UNRAVELING PLAGUE ECOLOGY THROUGH VECTOR AND HOST GENETICS

by

Rachael M Giglio

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Biological Sciences

at

The University of Wisconsin-Milwaukee

August 2019

ABSTRACT

UNRAVELING PLAGUE ECOLOGY THROUGH VECTOR AND HOST GENOMICS

by

Rachael M Giglio

The University of Wisconsin-Milwaukee, 2019
Under the Supervision of Professor Emily Latch

The transmission of vector-borne diseases involves complex interactions between vectors and their host species. These complex host-parasite interactions can be difficult to study with traditional, field-based methods. My dissertation aims to use a population genomics approach to elucidate transmission pathways of plague among prairie dog colonies. Plague is a flea-borne, zoonotic disease caused by the bacterium *Yersinia pestis*. It is infamous for causing the Black Death (1347-1353), one of the most devastating pandemics in human history. Since its emergence in North America around 1900, plague has spread to native rodents, thus creating a sylvatic cycle. Prairie dogs (*Cynomys* spp.) are highly susceptible to the disease, experiencing >90% mortality during outbreaks. Further, prairie dogs exacerbate the spread of plague by acting as an amplifying host, initiating epizootic events. In the first chapter of my dissertation, I examine how the landscape influences the connectivity of black-tailed prairie dog colonies in order to better understand the role of prairie dogs in plague transmission. I found that slope and bodies of water explain effective dispersal better than geographic distance alone. My second chapter describes patterns of connectivity for *Oropsylla hirsuta*, the main flea species found on prairie dogs and a known vector for plague. I compare those patterns of vector-mediated plague transmission to the host (prairie dogs) and potential alternative hosts [Northern grasshopper mice (*Onychomys leucogaster*) and deer mice (*Peromyscus maniculatus*)] to uncover alternative

modes of transmission. I found that the best performing model used patterns of connectivity for prairie dogs and deer mice to explain patterns of connectivity for *O. hirsuta*. My third chapter uses both neutral and putatively adaptive loci to characterize patterns of genetic variation for the threatened Utah prairie dog in order to improve recovery efforts for this threatened species. I found low species-wide genetic variation and high population divergence among sampling sites, which suggests that this species is highly vulnerable to the effects of genetic drift. Overall, this dissertation not only improves the conservation and management of prairie dogs in light of devastating plague outbreaks, but also provides a more general population genomics framework suitable for elucidating transmission pathways of wildlife diseases.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (SUPPLEMENTARY)

distribution of p-values compared to an expected uniform distribution of p-values (dark line) and (c) Manhattan plot showing the p-values for each locus).

LIST OF TABLES

LIST OF TABLES (SUPPLEMENTARY)

ACKNOWLEDGEMENTS

The list of people that have supported me throughout my graduate tenure is endless. However, I'd like to start by thanking my advisor, Dr. Emily Latch. I have been incredibly fortunate to have had the opportunity to learn from her expertise through my M.S. as well as my Ph.D. She has truly guided me to become the scientist I am today, and words cannot express my gratitude to her. I would also like to thank Dr. Filipe Alberto for his continued support and for sharing his knowledge of statistics, programming, and population genetics. I also thank Drs. Linda Whittingham and Peter Dunn for always being a source of sound advice and support. I would also like to thank Dr. Rafael Rodriguez who always encouraged me to see the bigger picture and to look at my research from a different perspective. I also thank Drs. Tonie Rocke and Jorge Osorio for accepting me into their labs and sharing their expertise in disease epidemiology.

I was able to meet and work with a number of amazing people throughout my dissertation research. I would like to thank the Latch lab (Elizabeth Kierepka, Maggie Haines, Samantha Hauser, Gennelle Uhurig, Xueling Yi, Ona Alminas, and Bennett Hardy) as well as the Rocke lab (particularly Carly Malave, Gebbiena Bron, and Elizabeth Falendysz) for their support and friendship through the years. Graduate school has been a learning process, and I'd like to thank those who helped prepare me for the journey (Dr. Janet Rachlow, Dr. Lisette Waits, Dr. Caren Goldberg, Dr. Jennifer Adams, Dr. Matthew Mumma, Dr. John Byers, Dr. Erin Clancey, and Dr. Stacey Dunn). I would also like to thank Dr. Jamie Ivy for teaching me to program. This skill has been invaluable throughout my dissertation. I also thank Dr. Chuck Wimpee for sharing his knowledge of bacteria and inspiring me with his passion for science.

I am fortunate to have an amazing family that has supported me throughout my dissertation. Thank you to my parents, Scott and Sheila Toldness and Laura and Gary Manning, for instilling in me a sense of adventure and passion for the outdoors that has driven me to pursue a Ph.D. in Biology. I would like to thank my sister, Carly Woolf, for being a constant pillar of support and friendship. I thank my mother and father-in-law, Joe and Nora Giglio, for their support and much needed humor. I would also like to thank my friends, Katie Barry, Ambi Henschen, and Brooke Berger for their camaraderie and support during my dissertation. Most of all, I'd like to thank my spouse, Daniel Giglio. He inspires me every day to be a better person and a better scientist. His support has been immeasurable and I'm sure I would have famished from starvation while writing my dissertation without him.

I would also like to thank my funding sources. This research would not have been possible without the funding provided by the Morris Animal Foundation in the form of a research grant and a fellowship. I also appreciate the support from the University of Wisconsin-Milwaukee through grants and travel awards (Ruth Walker travel award, Ruth Walker Grant-in-Aid, Ivy-Balsam- Milwaukee Audubon Society award, James D. Anthony award, and a Distinguished Dissertation Fellowship). I also extend my gratitude to The Wildlife Society for providing funding through the Aldo Leopold Scholarship as well as through travel awards.

Chapter 1. Landscape features explain functional connectivity of black-tailed prairie dogs better than distance alone

RACHAEL M. GIGLIO[1], TONIE E. ROCKIE[2], JORGE E. OSORIO[3], EMILY K. LATCH[1]

[1]*University of Wisconsin-Milwaukee, Milwaukee, WI*

[2]*United States Geological Survey-National Wildlife Health Center, Madison, WI*

[3]*University of Wisconsin-Madison, Madison, WI*

Abstract

The degree of connectivity among wildlife populations influences the retention of genetic diversity and has implications for the spread of disease. Connectivity is determined not only by the distance among populations, but also by characteristics of the landscape that an animal has to traverse during dispersal events. To quantify how the landscape influences connectivity of black-tailed prairie dog populations, we generated a population genomic dataset of 3,416 single nucleotide polymorphisms (SNPs) from 300 individuals from 10 sites within the Charles M. Russell National Wildlife Refuge in northern Montana. Prairie dogs suffer from severe population declines (typically killing >90% of a colony) during plague outbreaks. These outbreaks drive metapopulation dynamics, which are characterized by local extirpation followed by recolonization. Recolonization events are crucial to population viability, yet we lack information about how the landscape influences connectivity among colonies. We used a circuit theory-based approach to characterize the influence of the landscape on prairie dog dispersal. We determined that landscape features are better predictors of prairie dog dispersal than distance alone. The best performing model suggests that high slopes and bodies of water impede prairie dog dispersal, potentially shielding populations with such intervening habitat from plague exposure. Incorporating landscape features improves our predictions of prairie dog movement, advancing our understanding of recolonization dynamics and enhance our ability to predict where and why plague outbreaks occur.

Introduction

Wildlife populations often occur in a mosaic of heterogenous habitat across the landscape with differing degrees of connectivity. Different landscape features have diverse impacts on connectivity. The field of landscape genetics aims to quantify these differences by determining their effect on spatial genetic structure (Manel et al. 2003; Storfer et al. 2007). The ability to use landscape data to predict population connectivity is important for conservation planning. The success of conservation programs is often dependent on accurately characterizing and encouraging connectivity among small populations. These small and isolated populations are often characterized by low genetic variation, making them vulnerable to extinction and limiting their potential to adapt to future environmental change (Frankham 1996; Barrett and Schluter 2008; Savolainen et al. 2013; Tigano and Friesen 2016). By promoting or re-establishing connectivity among these vulnerable populations, we can mitigate the loss of genetic variation and maintain long-term viability.

Species with limited to moderate dispersal are predicted to follow an isolation-by-distance (IBD) model, where genetic structure can be explained by geographic distance (Aguillon et al. 2017). However, the dispersal of organisms is also affected by variables such as behavior, physical barriers, and landscape configuration and heterogeneity (Vallinoto et al. 2006; Hollatz et al. 2011). To test different hypotheses of connectivity, we can use a circuit theory approach to quantify the movement of individuals through the landscape (McRae 2006). In this approach, different landscapes features are assigned a resistance value based on the individual's inability to move through that feature in what is called an isolation-by-resistance (IBR) model. By using a quantitative framework to investigate and compare these different models (IBD and IBR), we

can infer patterns of dispersal that can guide management decisions to encourage natural gene flow (Ma et al. 2018; Epps et al. 2007; Cleary et al. 2017; McRae et al. 2008). The inferences we can make depend on the level of resolution in our data; dense sampling across the landscape and across the genome, using high resolution markers such as single nucleotide polymorphisms (SNPs), permit fine-scale analysis and ultimately improve our predictions of connectivity pathways (Luikart et al. 2003; Morin et al. 2004).

In this study, we used black-tailed prairie dogs (*Cynomis ludovicianus*) to test the competing hypotheses of IBD and IBR on population divergence in order to understand what drives patterns of connectivity – and thus population viability – in this species. Black-tailed prairie dogs are highly social burrowing mammals that are considered ecosystem engineers, as well as a keystone species of grassland ecosystems. Their burrows create habitat that benefits numerous species, including burrowing owls (*Speotyto cuniculaira*) and mountain plovers (*Eupoda montana*) (Tipton et al. 2008). Further, prairie dog declines have hindered the recovery of the endangered black-footed ferret (*Mustela nigripes*), which preys almost exclusively on prairie dogs, shelters in their burrows, and is susceptible to the diseases they carry (Matchett et al. 2010).

Population connectivity of black-tailed prairie dogs is likely dependent on a mixture of factors such as disease, behavior, and landscape. Prairie dogs are highly susceptible to plague, a disease caused by the bacterium *Yersinia pestis*. Once plague enters a prairie dog colony, it moves from coterie (small family group of prairie dogs) to coterie until nearly the entire colony is extirpated (near 100% mortality rate; Cully and Williams 2001). Further, these outbreaks (referred to as epizootics) drive metapopulation dynamics since they cause local extirpation events, which are

followed by recolonization events (Roach et al. 2001; Antolin et al. 2006; Augustine et al. 2007; Sackett et al. 2013). Dispersal events are rare because of the highly social nature of prairie dogs; when they do occur, dispersal to neighboring colonies is more likely to occur than long-distance dispersal (Hoogland 2013). These dispersal dynamics suggest that patterns of connectivity are driven primarily by geographic distance (IBD). However, prairie dogs are grassland specialists and may follow patterns of IBD in contiguous areas of grassland, but other landscape features may impede connectivity.

Our main objective was to understand how the landscape influences connectivity in terms of population divergence in black-tailed prairie dogs. To address this objective, we had two core aims. The first aim was to characterize genetic variation and spatial genetic structure using a population genomics dataset, to quantify connectivity within and among prairie dog colonies. Second, we evaluated landscape features to determine their impact on connectivity, to identify the model that best predicts prairie dog population structure. By understanding how the landscape influences the connectivity of colonies, we may begin to predict its contribution to metapopulation dynamics, especially relevant in the context of plague-mediated extinctions.

Methods

Study Area

To evaluate different models of connectivity, we sampled 10 sites of black-tailed prairie dogs within the Charles M. Russel Wildlife Refuge located in North-central Montana (Fig. 1.1). Sampling within the wildlife refuge encompassed roughly 110.2 hectares within an extent of 106,966.8 ha of heterogeneous habitat with an elevation gradient of 700 m to 867 m. The refuge

also contained bodies of water, such as the Missouri River south of the sampling locations. The uplands of the Missouri River, where the study sites are located, are primarily composed of rolling prairies. The distance between sites varied with a maximum of 34 km and a minimum of 0.86 km.

Generating the SNP Dataset

The hair and whiskers used for this study were collected from black-tailed prairie dogs during an oral sylvatic plague vaccine field trial as described in Rocke et al. (2017). DNA was extracted from hair and whisker follicles using a spin-column based method using the Zymo Research Universal Quick-DNA miniprep kit (Zymo Research) following the manufacturer's feather extraction protocol. The quantity of DNA was measured using a Qubit (Invitrogen) with the high sensitivity kit for double-stranded DNA. We used double digest restriction-site associated DNA (ddRAD) sequencing (Peterson et al. 2012) to identify genotypes for a total of 300 individuals (mean per site=30). DNA was sheared using *HindIII* and *NlaIII* restriction enzymes followed by a 300-600bp size selection step (Pippin Prep, Sage Sciences). Raw sequences were aligned to a Gunnison's prairie dog (*Cynomys gunnisoni*) genome (Sackett pers comm) using the BWA short-read aligner with default parameters using the BWA-MEM alignment algorithm (Li and Durbin 2009). Contigs were then assembled and filtered using the program STACKS v1.48 (Catchen et al. 2011) following the protocol outlined in Rochette and Catchen (2017).

Single nucleotide polymorphisms (SNPs) were identified after a series of filtering steps. First, in order the reduce linkage disequilibrium, we only used one SNP per contig. If more than one SNP was identified in a contig, only the first in the contig (5' most SNP) was retained for downstream

analyses. Second, we filtered loci with minor allele frequencies <0.05, as they may represent PCR errors. Third, to reduce the presence of multi-copy loci and paralogs, which can affect population genetic analyses, we removed loci with an observed heterozygosity above 0.70 and an average depth greater than twice the mode (90.93) (Willis et al. 2017; O'Leary et al. 2018). Fourth, only loci that were present in all sites in at least 80% of individuals per site were retained for analyses. Fifth, individuals with greater than 30% of missing data were removed from analyses. Filters to keep only a single SNP per contig, remove loci with minor alleles, remove loci with excess heterozygosity, and keep only loci found in all sampled sites in at least 80% of individuals were executed in the populations module of STACKS. The average depth of coverage for each locus and the amount of missing data per individual were calculated using VCFtools (v.0.1.16; Danecek et al. 2011) and removed using the R package adegenet v.2.1.1 (Jombart and Ahmed 2011).

Characterizing Genetic Variation

We characterized genetic variation in terms of observed and expected heterozygosity (H_O and H_E), allelic richness (A_R), number of private alleles (P_A), and inbreeding (F_{IS}). Heterozygosity (both H_O and H_E) were calculated using the R package adegenet (Jombart and Ahmed 2011). Allelic richness, or the average number of alleles per locus, was calculated using the R package hierfstat v.0.04-22 with rarefaction to account for uneven sample sizes (Goudet and Jombart 2015). The number of private alleles, or alleles unique to that sampling site, was calculated using the R package poppr v.2.8.2 (Kamvar et al. 2014). We calculated F_{IS} (a measure of heterozygosity excess) to characterize the level of inbreeding within each sampling site. We

generated a 95% confidence interval of F_{IS} using the R package diveRsity with 1,000 bootstraps (Keenan et al. 2013).

Characterizing Genetic Structure, Patterns of Gene Flow, and IBD

First, we used a principal component analysis (PCA) to visualize patterns of genetic variation using the R package ade4 v.1.7-13 (Chessel et al. 2004; Dray & Dufour 2007). Individuals were assigned to genetic clusters using an ordination-based method (discriminant analysis of principal components; DAPC) performed in the R package adegenet (Jombart et al. 2010, Jombart and Ahmed 2011) and a Bayesian program (STRUCTURE) (Pritchard et al. 2000). For STRUCTURE we ran 10 iterations of 100,000 burn-in iterations followed by 100,000 MCMC repetitions for each value of K (the number of genetic clusters) from 1 to 7. We ran STRUCTURE using the admixture model with correlated allele frequencies. We used STRUCTURE HARVESTER (Earl and Vonholdt 2012) to choose the optimal K, assessing both the mean LnP(K) and the ΔK methods (Evanno et al. 2005). Individuals were assigned to genetic clusters using the program CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) based on the proportional assignments (q values) generated from STRUCTURE. We performed the DAPC unsupervised (which is based on K-means) and supervised (based on the Bayesian information criterion (BIC)) to identify the optimal K.

We used two pairwise statistics to measure gene flow among sites, F_{ST} (calculated in the R package hierfstat; Goudet and Jombart 2015) and the proportion of shared alleles (D_{PS}, calculated in the R package PopGenReport v.3.0.4; Adamack and Gruber 2014). The significance of pairwise gene flow measurements was assessed by randomly reassigning a

population to each individual for 1,000 bootstraps. We used the false discovery rate (FDR) to correct p-values for multiple comparisons (conducted in R; R Core Team 2018; Benjamini and Hochberg 1995). Further, we measured the degree of migration (N_M) among sites using the divMigrate function in the diveRsity v1.9.89 package in R (Keenan et al. 2013; Alcala et al. 2014) and generated 95% confidence intervals for 1,000 bootstraps. Migration networks were created using the R program qgraph (Epskamp et al. 2012). To detect patterns of IBD, we performed a mantel test in the R package ecodist (Goslee and Urban 2007) using D_{PS} and geographic distance, calculated using the R core function dist (R Core Team 2018).

Isolation by Resistance

In order to elucidate the role of different landscape features in shaping population genetic structure in black-tailed prairie dogs, we used 8 landcover types from the National Landcover Database (NLCD; Jin et al. 2013) that were found within the geographic extent of the study (Fig. 1.2). The landcover types included barren land, cultivated cropland, evergreen forests, grassland/herbaceous, emergent herbaceous wetland, shrubland/scrubland, open bodies of water, and woody wetland at a resolution of 30 meters × 30 meters. Each of the 8 landcover types were assigned a range of resistance values (0.01, 0.1, 10, 100) to assess whether it hinders (resistance>1) or promotes (resistance<1) connectivity while all other features within each resistance surface (rs) were given a uniform value of 1 to represent the cost of distance. Resistance values were evaluated based on their ability to explain population divergence. Additionally, we assessed the influence of elevation and slope on genetic structure. Elevation data was obtained from the National Elevation Database (NED; Archuleta et al. 2017) at a resolution of 1 arc-second (approximately 30 meters). Slope was assessed by converting the

elevation raster to slope using the R package raster v2.8-19 following the 8-neighbor rule (Hijmans 2019). All IBR models were compared to a null hypothesis of IBD by generating a resistance surface with uniform resistance (resistance=1) for all pixels.

To compare IBD and IBR models, we used the program CIRCUITSCAPE v.4.0 (McRae and Shah 2011) to calculate the effective resistance for each landcover feature, distance, elevation, and slope. The effective resistance for each landcover feature was scaled and compared to genetic distance (calculated as the shared proportion of alleles) using linear mixed effect models with the R package lme4 (Bates et al. 2015). In order to correct for dependency among sites, we used a maximum-likelihood population-effect (MLPE) method (Clarke et al. 2002; van Strien et al. 2012), in which sampling sites were introduced as random effect terms and explanatory variables as fixed effect terms. Further, parameters within the MLPE model were fitted with the restricted maximum-likelihood (REML). The best univariate models were selected based on Akaike Information Criteria (AIC) and whether they performed better than the null (IBD) model. We also calculated the marginal coefficient of determination for each model (R^2_M) using the R package MuMIn v.1.43.6 (Bartoń 2019). To calculate R^2_M, models with fixed and random effects were compared to a null model (we used the IBD model). We determined the optimal resistance value of each landscape feature based on their model performance. In order to understand how complex landscapes influence gene flow, we used the top performing model (optimal resistance value) for each landscape feature to build multivariate models. Explanatory variables were evaluated for collinearity using their variance inflation factors (VIF) using the R package usdm v.1.1-18 (Naimi et al. 2014). Explanatory variables with high VIF (multicollinearity was assumed if VIF >4) were excluded in reduced models.

Results

Generating the SNP Dataset

After filtering using the populations module of STACKS, 3,771 SNPs were identified for 300 individuals from 10 sampling sites. An additional 355 loci were removed due to high depth of coverage (greater than two times the mode=90.93), leaving 3,416 loci in the final analysis. A total of 4 individuals were removed from subsequent analysis due to a high proportion of missing data (>30%). The final number of individuals per site ranged from 13 to 40 (mean=29.6).

Characterizing Genetic Variation

The average observed heterozygosity ranged from 0.07 to 0.15 (mean=0.12) while the expected heterozygosity ranged from 0.12 to 0.37 (mean=0.19) (Table 1.1). Allelic richness ranged from 1.32 to 1.91 (mean=1.57). Overall, eastern sites (3A,3B,4A,4B,5A, and 5B) had higher H_O (mean=0.14) compared to western sites (1A, 1B, 2A, and 2B; mean=0.09; p= 0.01); however, western sites had higher levels of H_E (mean=0.26) and A_R (mean=1.79) than eastern sites (p= 0.04). The exception to this pattern was site 3A, an eastern site with elevated H_O (0.14) but that also showed the elevated H_E (0.24) and A_R (1.86) more characteristic of western sites. Western sites had positive values of F_{IS}, while eastern sites had negative values (indicating assortative mating); site 3A was again atypical of eastern sites and exhibited a positive F_{IS} like other western sites.

Characterizing Genetic Structure, Patterns of Gene Flow, and IBD

We observed population divergence between western (1A, 1B, 2A, and 2B) and eastern (3A, 3B, 4A, 4B, 5A, and 5B) sites using PCA (Fig. 1.3). Further, we found that the presence of two genetically distinct clusters (K=2) was most supported by both the program STRUCTURE (based on both Ln|P and ΔK) and using the DAPC analysis (using unsupervised clustering and based on BIC) (Fig. 1.4; Fig. S1.1, S1.2, and S1.3). Genetic differentiation was detected between all sites using F_{ST}; however, paired A and B sites generally showed less differentiation (Fig. 1.5; Table S 1.2). Genetic differentiation was higher between the western-most sites (1A, 1B, 2A, and 2B) than between sites in the east (3A, 3B, 4A, 4B, 5A, and 5B). The highest degree of differentiation was between sites 1A and 4A in terms of F_{ST} (0.80) and between sites 1A and 5B in terms of D_{PS} (0.37). We detected significant IBD using the mantel test comparing D_{PS} and geographic distance based on 999 replicates (observation= −0.76; p-value=0.003). Relative migration based on N_M revealed high migration among the eastern sites (3A, 3B, 4A, 4B, 5A, and 5B) with relatively low migration occurring between eastern and western sites (Fig. 1.6). The little migration that occurred between regions was unidirectional, from eastern sites to site 2B.

Isolation by Resistance

Univariate MLPE models showed that slope was the single variable that best explained genetic distance (R^2_M=0.74), followed by elevation (R^2_M=0.61). The optimal resistance values for evergreen forests, woody wetlands, herbaceous wetlands, open bodies of water, and cropland were 100, suggesting that they represent significant obstacles for prairie dogs to traverse (Table S1.1). Conversely, the optimal resistance values for grassland and barren land were 0.1 and 0.01, respectively, suggesting that they are conducive to prairie dog dispersal. Shrubland was the only

landscape feature that did not explain genetic distance better than the null (IBD) model (optimal resistance =10).

We compared 1,014 MLPE multivariate models and found that the best performing model included slope, elevation, forest (rs=100), woody wetland (rs=100), open water (rs=100), cropland (rs=100), barren land (rs=0.01), herbaceous wetland (rs=100), grassland (rs=0.1), and shrubland (rs=10) (AIC_C =−59.01 and R^2_M = 0.60). However, we detected high multicollinearity within the fixed variables (VIF values ranged from 6.89 to 263961.60). We created 8 reduced MLPE multivariate models using only variables that showed no indication of multicollinearity. All eight reduced models included as fixed variables slope and one of eight other landcover variables (shrubland excluded). All reduced models had ΔAIC_C values under two, suggesting that they all had strong performance in their ability to explain patterns of connectivity. The best performing of the reduced models had slope and open water as fixed variables (AIC_C=−43.57 and R^2_M = 0.71; Fig. 1.7; Table 1.2).

Discussion

Maintaining viable prairie dog populations is essential for the health of grassland ecosystems as they provide a host of valuable ecological functions (Martínez-Estévez et al. 2013). However, prairie dog populations are in decline due to a number of factors, including plague outbreaks. These epizootic events drive metapopulation dynamics, which are characterized by local extirpation events followed by recolonization events. Recolonization is dependent on prairie dog dispersal to inactive burrows, which in turn is dependent upon features of the landscape that restrict or promote movement. Connectivity among populations is also important for mitigating

the erosion of genetic variation by drift within isolated populations. By characterizing patterns of genetic structure at high resolution, and incorporating landscape information in a circuit theory approach, we were able to 1) describe patterns of genetic diversity across prairie dog populations, 2) quantify the extent of prairie dog dispersal and the degree of connectivity between populations, and 3) tie connectivity to characteristics of the landscape that influence prairie dog movement. We observed strong population divergence between eastern and western sites, with evidence of limited long-distance dispersal among these two regions. Our study shows that although prairie dogs exhibit significant IBD, the inclusion of landscape features in an IBR model improves our ability to explain connectivity based on effective dispersal.

We detected a high degree of population divergence between eastern and western sites (measured as F_{ST} and D_{PS}). Further, our estimates of pairwise F_{ST} between these regions were much higher than previously reported in other black-tailed prairie dog populations (our highest F_{ST}=0.80; highest from previous studies =0.24; Pigg 2014; Antolin et al. 2006; Jones and Britten 2010; Sackett et al. 2012). However, these studies were based on microsatellite loci rather than SNPs, which may result in different patterns of differentiation due to different mutation rates, resolution, and biases (Morin et al. 2004). Regardless of the methodology, it is clear that long-distance prairie dog movement is highly restricted. In addition to low connectivity between regions, we detected a high degree of connectivity within regions. Specifically, connectivity among eastern sites was higher than anticipated, especially given that these sites are located roughly 18 km apart (compared to ~21 km between the closest eastern and western sites). This high degree of connectivity among eastern sites could be explained by their proximity to natural drainages. Previous studies have shown that prairie dogs are more likely to encounter a colony

along drainage systems and that they can act as dispersal corridors (Garrett and Franklin 1988; Roach et al. 2001). These drainages are also represented as areas of low slope, the best predictor of dispersal in our models.

Although population divergence is high between western and eastern sites overall, our results showed low but significant migration between specific sites (significant rates of N_M ranged from 0.03 to 0.10). This may explain the mixed genetic ancestry shown in some western sites. These migrant individuals from eastern colonies could be a result of long-distance dispersal events. Eastern and western sites are separated, for the most part, by grasslands. Grasslands, as long as vegetation is not too tall, has been shown to be highly conducive to prairie dog dispersal (Pigg 2014). Our resistance models also showed that grasslands (and barren lands) were conducive to prairie dog dispersal. However, based on our best resistance model, slope, particularly when combined in a reduced model with open water, is a far superior indicator of prairie dog dispersal. The current map of this best performing reduced model suggests that individuals could be moving along the drainages around the Missouri river. This pattern agrees with the results from Roach et al. (2001), which show that natural drainages may act as connectivity corridors among colonies. Another possibility is that the mixed genetic ancestry in some sites is a signature of translocation. In 1999, an experiment was conducted in the CMR wildlife refuge to determine the best translocation practices for prairie dog (Dullum et al. 2006). Individuals were moved from private land north of the refuge to colonies within the refuge. Up to 120 prairie dogs were moved to new and existing colonies, creating the potential to significantly alter the genetic composition of sites. However, with a generation time of 1.5 years (Hoogland 1995), approximately 10 generations had passed between the translocation and the time of sampling for this study. Barring

assortative mating, any remaining signature of the translocation would mostly be observed in admixed individuals. We did not detect evidence of highly admixed individuals from our STRUCTURE or DAPC analyses, making it more likely that migrants are from more recent, undocumented translocations or natural dispersal.

Incorporating landscape features as resistance surfaces improves our ability to explain connectivity. The MLPE model paired with D_{PS} has been shown to work well in identifying patterns of IBD (Shirk et al. 2017) and avoids the elevated type I error found in other common approaches like partial Mantel tests (Raufaste and Rousset 2001; Legendre and Fortin 2010; Kierepka and Latch 2015). This approach has the added benefit of allowing for model comparison via AIC. For this study, all univariate models (with optimal cost values) performed better at explaining population divergence (measured as D_{PS}) than geographic distance alone, with the exception of shrubland. However, it is important to note that the resistance surfaces have the optimal cost value assigned to pixels where the target landscape feature is present and a cost value of 1 for all non-target pixels. In doing this, we make distance implicit within each resistance surface. Therefore, landscape features with limited representation within the study extent (cropland, barren land, and herbaceous wetland had >1% representation) will have more representation of IBD within the model. Since IBD alone was significant, these models would perform well simply because they have a similar cost as IBD. However, if distance alone was the main driver of genetic differentiation, we would expect the furthest sites away to have the greatest differentiation (sites 2A and 5B). Instead, we found that sites 1A and 5B were the most differentiated.

Conservation implications

Improving our ability to predict prairie dog movement advances our understanding of where and why plague outbreaks occur. In turn, this knowledge can be used to improve prairie dog conservation and enhance plague mitigation efforts. Vaccination efforts and insecticide application (e.g., flea dusting) can be used to protect individual prairie dog colonies from plague (Rocke et al. 2017; Tripp et al. 2017). Sites can be prioritized for these mitigation efforts based on genetic variation and degree of connectivity. For instance, populations with high genetic variation could be prioritized for outbreak mitigation efforts to maximize the conservation of genetic variation, adaptive potential, and the odds of long-term species viability. Our results as well as those of Proctor (1998) suggest that physical barriers, such as steep slopes, may help safeguard colonies from plague by reducing colony connectivity. However, a trade-off exists; while decreased connectivity among prairie dog populations may safeguard populations from plague, the isolation also contributes to loss of genetic variation via genetic drift, which could be important in building resistance to the disease. Therefore, maintaining some degree of connectivity among prairie dog colonies is essential to conserve genetic variation and facilitate recolonization of extirpated colonies.

Figure 1.1 Map of sampling sites of black-tailed prairie dog within the Charles M. Russell Wildlife Refuge in Phillips county, MT.

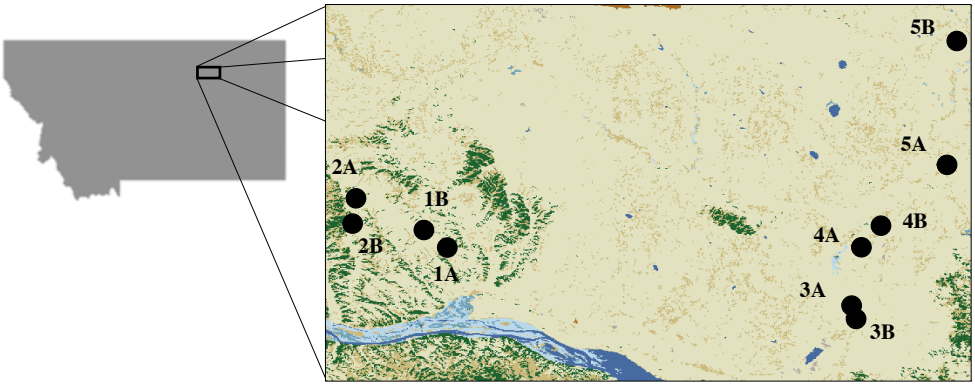


Figure 1.2. Eight landcover features as well as elevation and slope were used in resistance models to determine their effect on population divergence. The black plus signs depict the locations of the black-tailed prairie dog sampling sites.

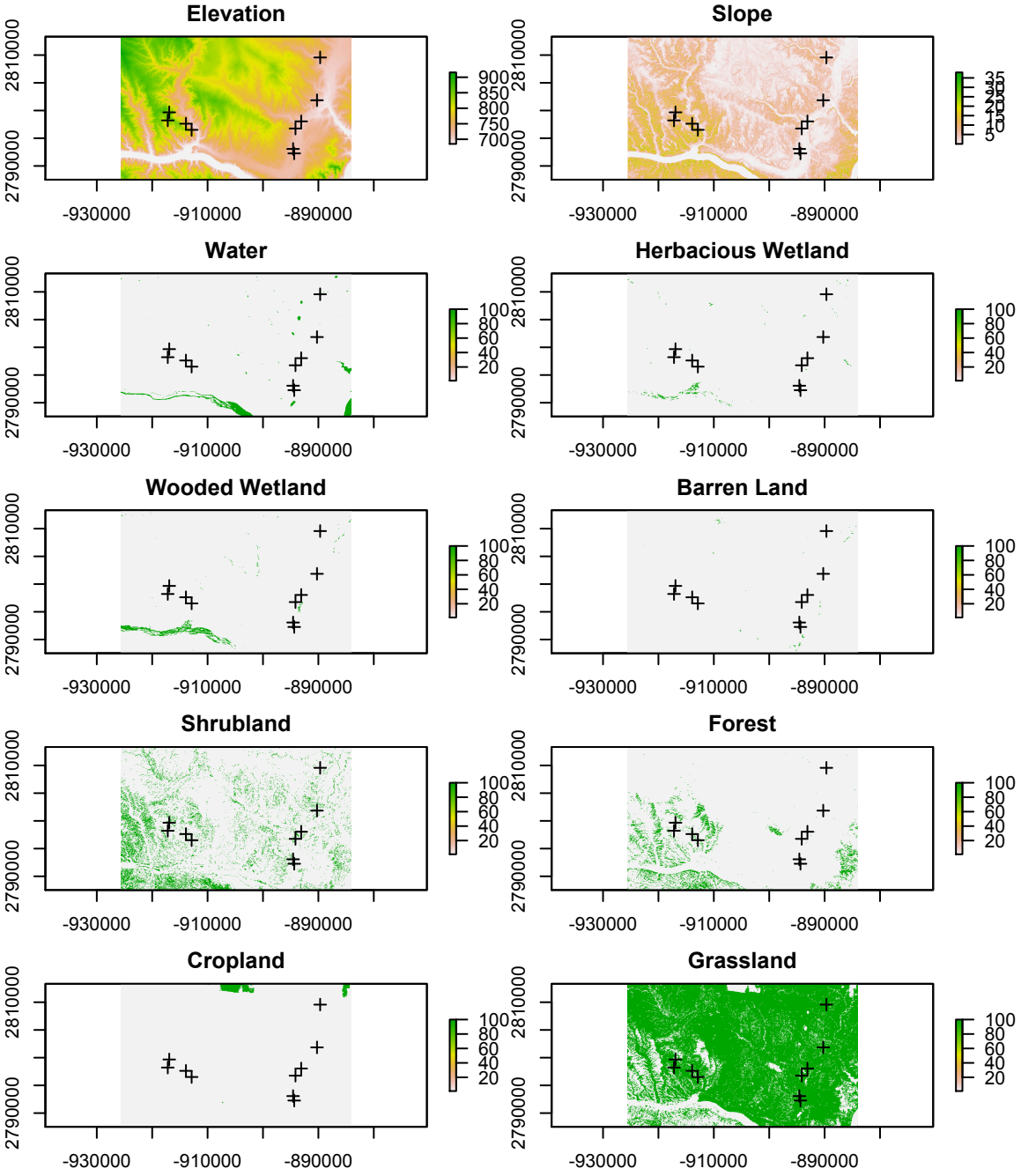


Figure 1.3. Principal component analysis (PCA) shows genetic differentiation among sampling sites of black-tailed prairie dogs using SNP loci. Colors and shapes distinguish the different sampling sites.

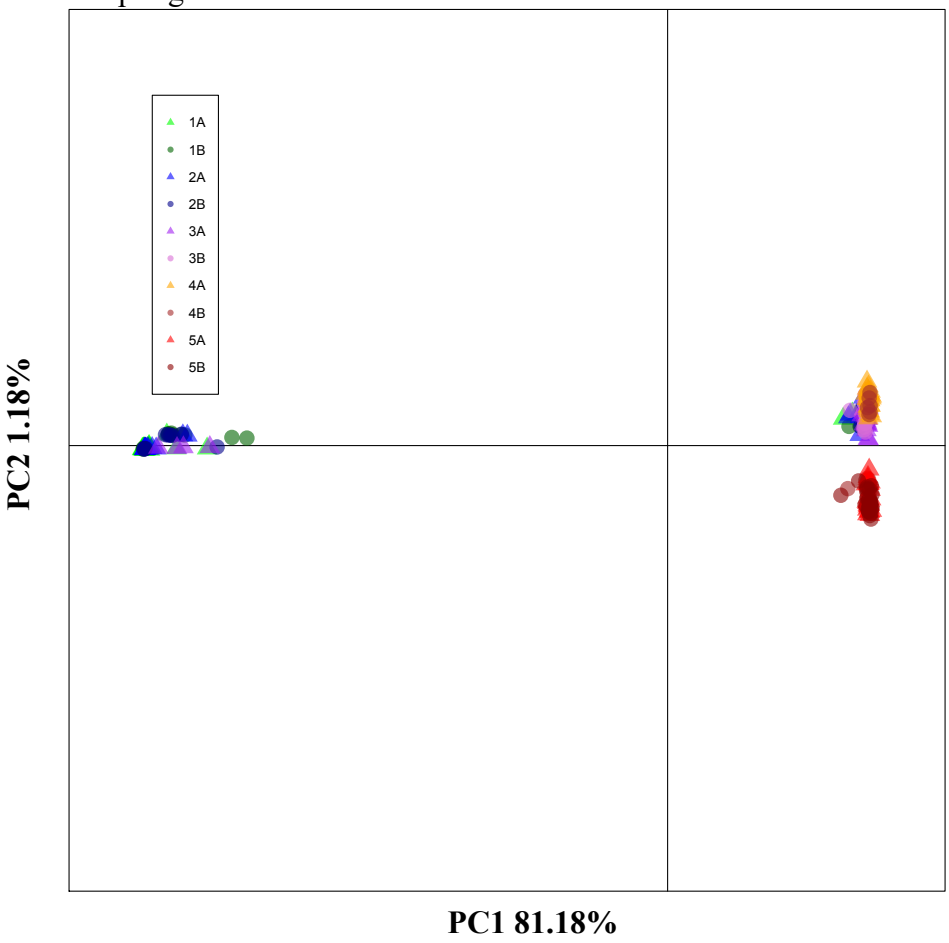


Figure 1.4. Results from the Bayesian clustering with STRUCTURE for k= 2. Individual bars depict the probability of membership of each individual to one of two distinct genetic clusters. Individuals are grouped based on sampling site with sample sizes shown in parentheses.

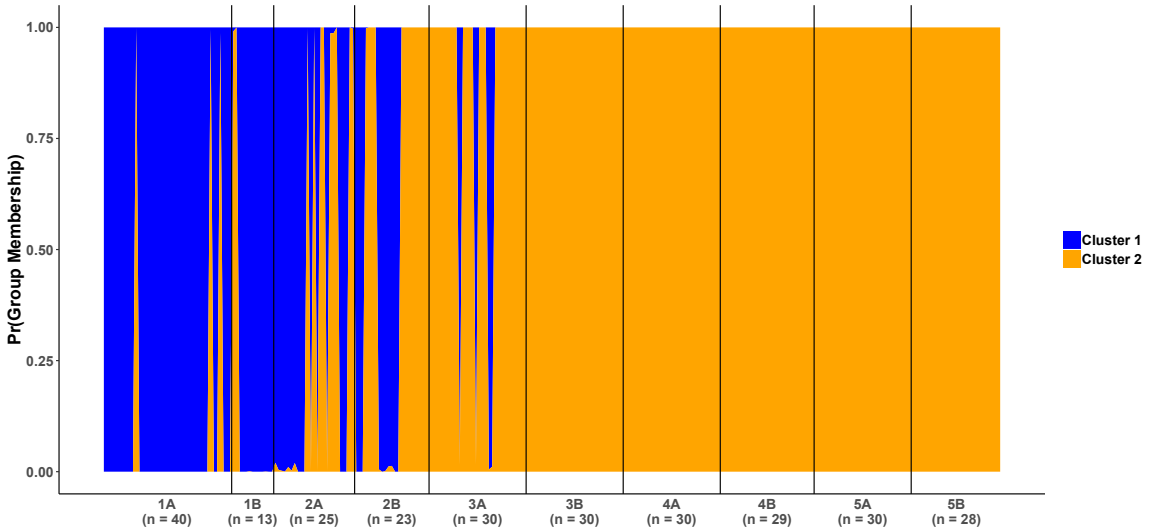


Figure 1.5. Pairwise genetic distance (F_{ST} and D_{PS}) among all sampling sites for black-tailed prairie dogs. Values closer to 0 are represented by darker colors.

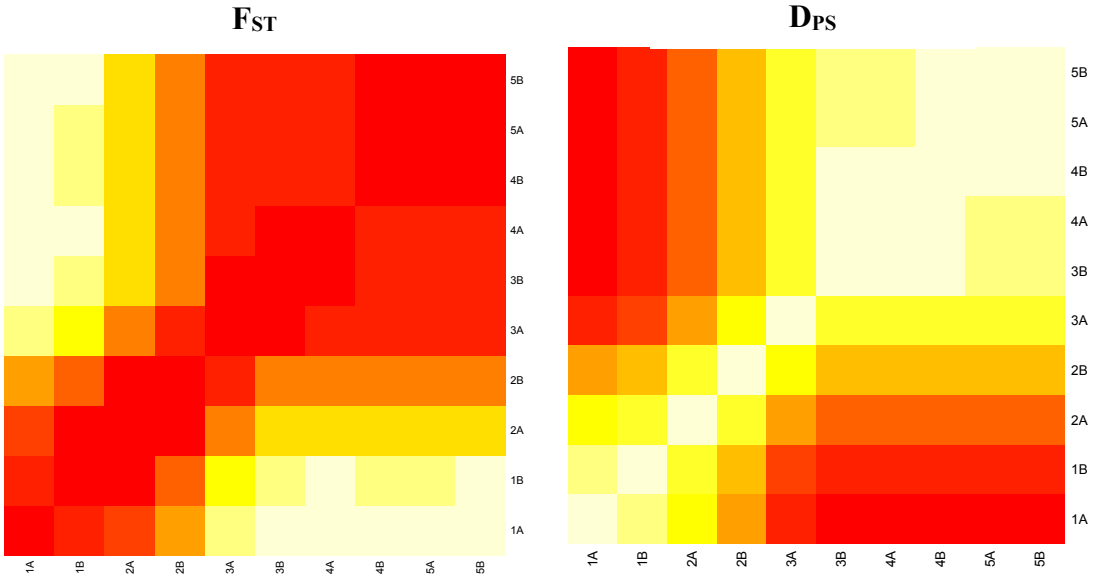


Figure 1.6. Network plot of relative migration (N_M) levels between sampling sites. The arrows refer to the direction of gene flow among nodes (sampling sites) whereas the edge values represent the degree of gene flow. Higher values of N_M mean a higher proportion of shared migrants.

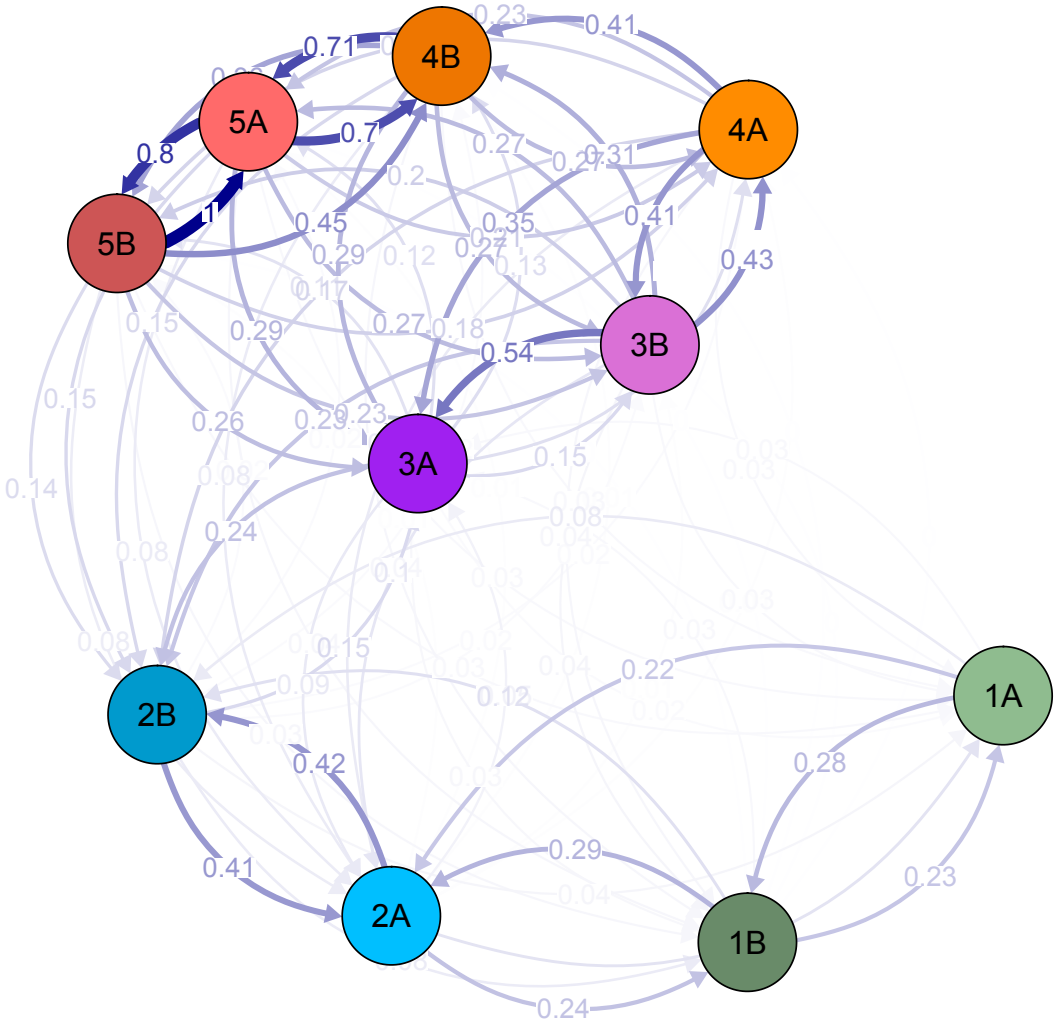
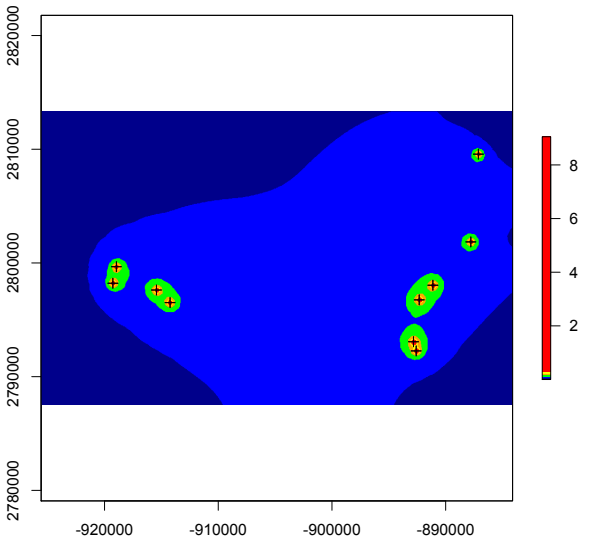


Figure 1.7. Current maps depicting connectivity of back-tailed prairie dogs based on: A) the null, IBD model and B) the best performing reduced model (slope and open bodies of water with resistance values of 100) generated in Circuitscape. Warmer colors indicate a higher degree of connectivity among sampling sites.

A.



B.

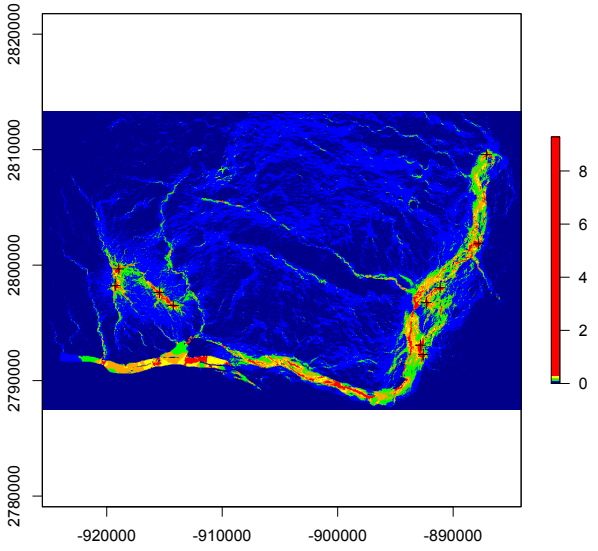


Table 1.1 Number of individuals (N), observed heterozygosity (H_O), expected heterozygosity (H_E), allelic richness (A_R), and inbreeding (F_{IS}) were calculated for 10 sites using 3,416 loci.

Site	N	H_O	H_E	A_R	F_{IS}
1A	40	0.07	0.13	1.59	0.45 (-0.29 – 0.71)
1B	13	0.07	0.21	1.80	0.65 (-0.21 – 0.85)
2A	25	0.10	0.34	1.87	0.71 (0.66 – 0.73)
2B	23	0.11	0.37	1.91	0.70 (0.61 – 0.74)
3A	30	0.14	0.24	1.86	0.42 (0.09 – 0.64)
3B	30	0.14	0.12	1.34	-0.17 (-0.20 –-0.15)
4A	30	0.14	0.12	1.32	-0.17 (-0.21 –-0.15)
4B	29	0.15	0.13	1.34	-0.16 (-0.19 – -0.14)
5A	30	0.15	0.12	1.34	-0.19 (-0.21 – -0.16)
5B	28	0.14	0.12	1.32	-0.20 (-0.24 – -0.17)

Table 1.2. Reduced multivariate linear mixed effects models using maximum likelihood population effects (MLPE) parameterization were built using explanatory variables with variance inflation cofactors less than four. Models were evaluated based on overall corrected AIC values from the best model (ΔAIC_C), AIC weights (AIC_{ew}), and the marginal coefficient of determination (R^2_M). Lower values of ΔAIC_C indicates a better model performance of explaining genetic structure (calculated using the proportion of shared alleles (D_{PS})).

Model	AIC_C	ΔAIC_C	AIC_{ew}	R^2_M
D_{PS} ~Slope + Water100 + (1|pop1)	-43.57	0.00	0.15	0.71
D_{PS} ~Slope + Wwet100 + (1|pop1)	-43.53	0.05	0.14	0.71
D_{PS} ~Slope + Crops100 + (1|pop1)	-43.52	0.06	0.14	0.71
D_{PS} ~Slope + Elevation + (1|pop1)	-43.45	0.12	0.14	0.71
D_{PS} ~Slope + Barren0.01 + (1|pop1)	-43.41	0.16	0.14	0.71
D_{PS} ~Slope + Hwet100 + (1|pop1)	-43.41	0.16	0.14	0.71
D_{PS} ~Slope + Forest100 + (1|pop1)	-42.78	0.79	0.10	0.70
D_{PS} ~Slope + Grass0.1 + (1|pop1)	-41.72	1.85	0.06	0.71

Figure S1.1 Results from STRUCTURE supported that two distinct genetic clusters are present among the black-tailed prairie dog sites based on the ΔK method.

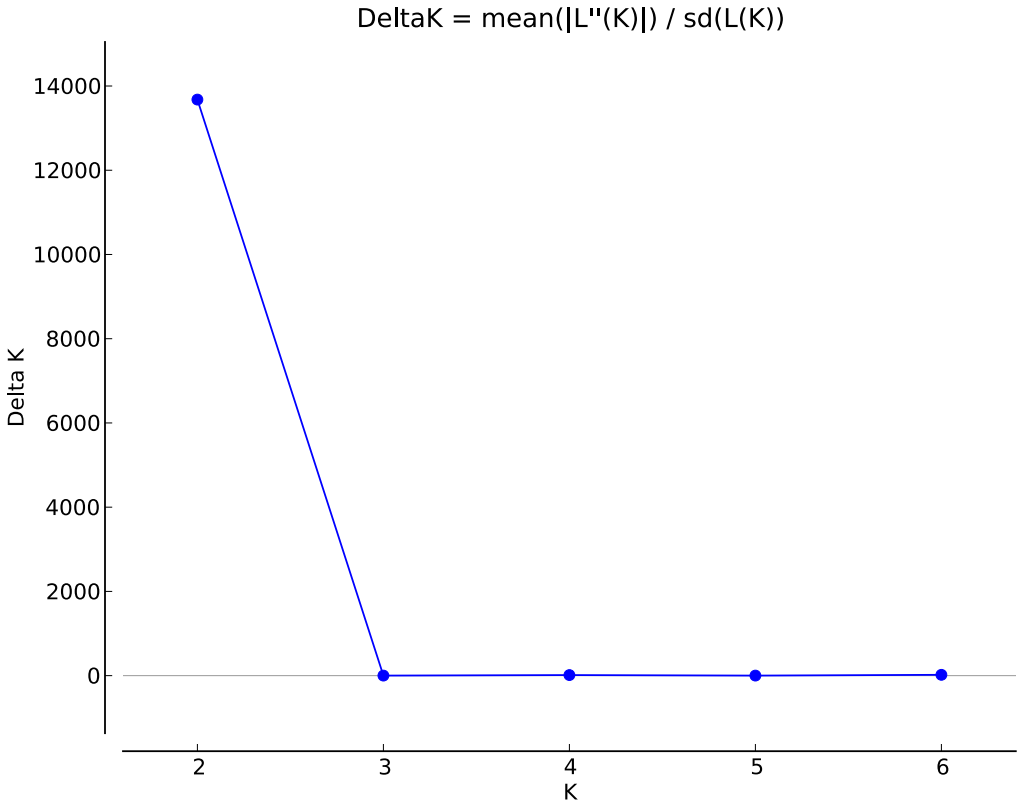


Figure S1.2. Results from STRUCTURE supported that two distinct genetic clusters are present among the black-tailed prairie dog sites based on the Ln(K) method.

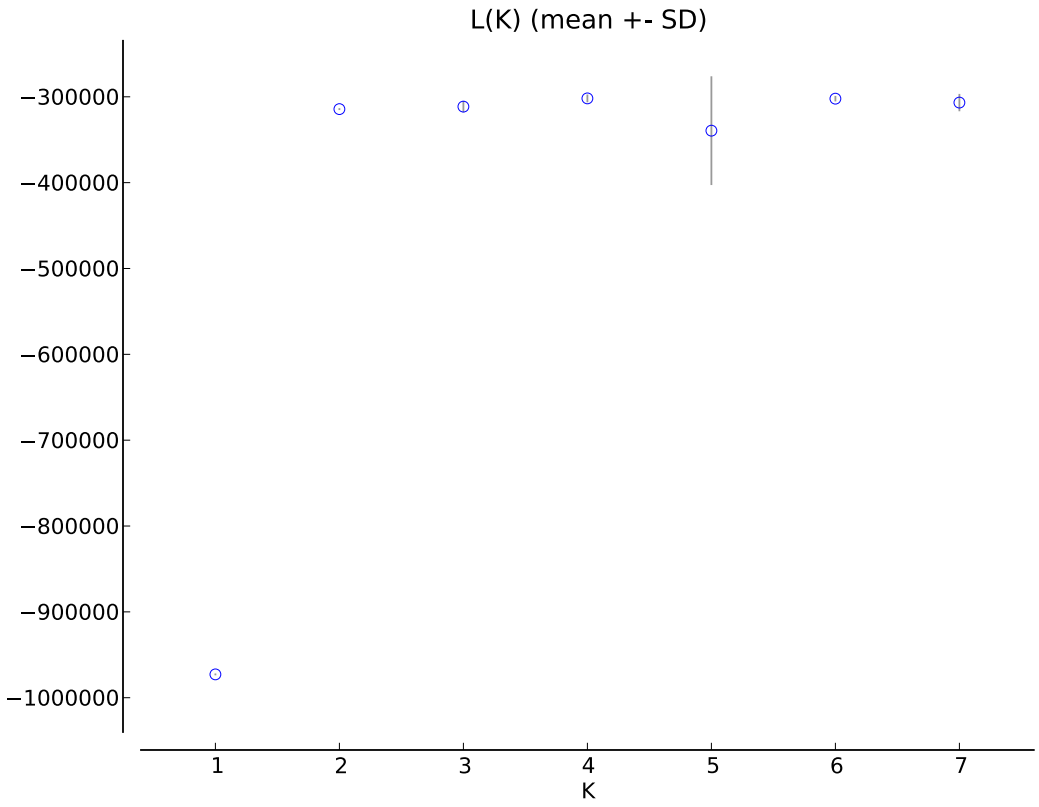


Figure S1.3. Results from discriminant analysis of principal components (DAPC) supported two distinct genetic clusters among the black-tailed prairie dog sites based on the Bayesian information criterion (BIC).

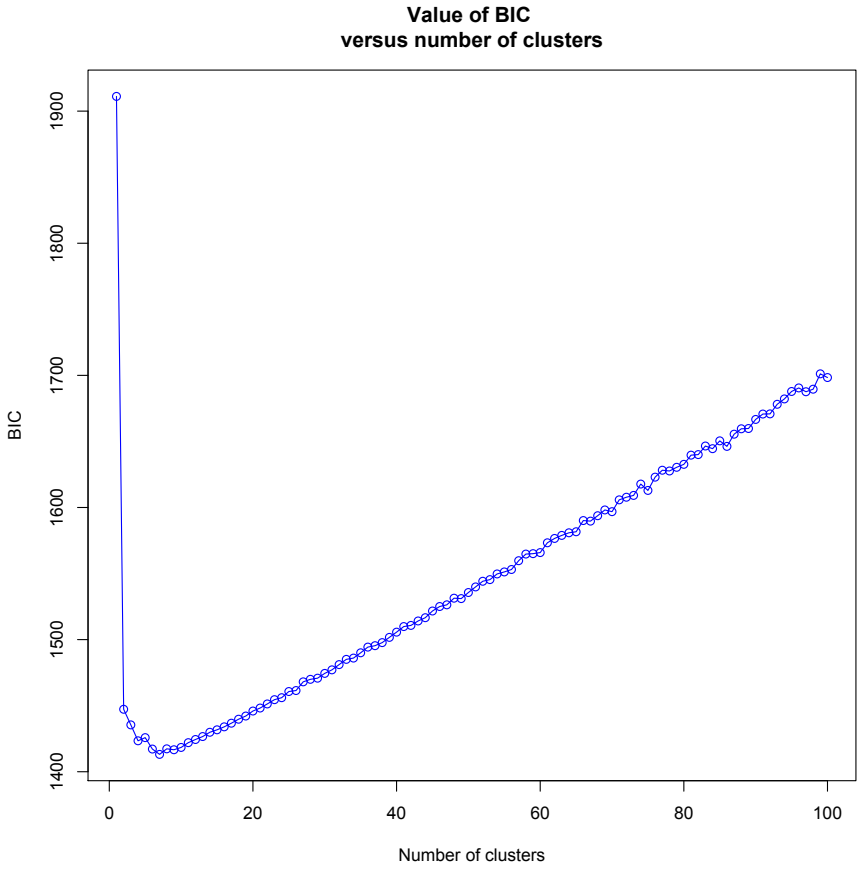


Figure S1.4. Membership of individuals based on the discriminant analysis of principal components (DAPC). The size of the squares depicts the number of individuals from each site that assigned to each of the two distinct genetic clusters (inf 1 and inf 2).

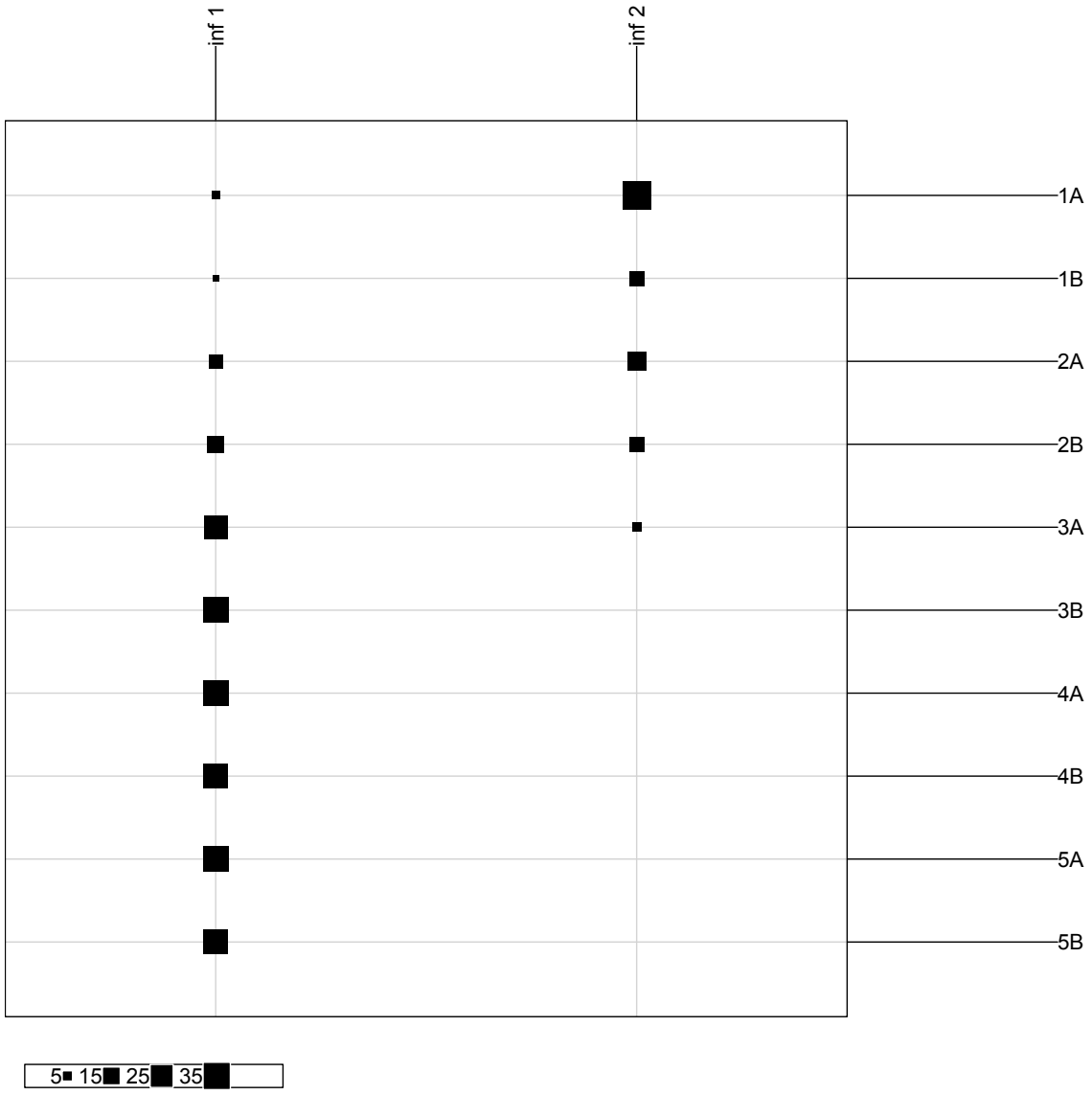


Table S1.1. Each landscape feature was assigned four resistance values (0.01, 0.1, 10, 100) and compared using a corrected Akaike Information Criteria (AIC_C). Lower values of ΔAIC_C indicates a better model performance of explaining genetic structure (calculated using the proportion of shared alleles (D_{PS})). Further, the marginal coefficient of determination for each model was calculated to indicate the relative amount of variation in genetic structure explained by each model (R^2_M). Models in bold indicate those that performed best for each landscape variable. Distance is italicized to indicate that it is the null model to which we based our resistance model comparisons from.

Model	**AIC_C**	**ΔAIC_C**	**AIC_{ew}**	**R^2_M**
Slope	**-44.72**	**0.00**	**0.23**	**0.74**
Elevation	**-43.00**	**1.72**	**0.10**	**0.61**
Forest100	**-41.59**	**3.13**	**0.05**	**0.57**
Forest10	-41.27	3.45	0.04	0.57
Wwet100	**-40.77**	**3.95**	**0.03**	**0.56**
Water100	**-40.73**	**3.99**	**0.03**	**0.56**
Crops100	**-40.65**	**4.07**	**0.03**	**0.56**
Water10	-40.57	4.15	0.03	0.56
Barren0.01	**-40.57**	**4.15**	**0.03**	**0.56**
Crops10	-40.56	4.16	0.03	0.56
Wwet10	-40.53	4.19	0.03	0.56
Grass0.1	**-40.48**	**4.24**	**0.03**	**0.59**
Barren0.1	-40.44	4.28	0.03	0.56
Hwet100	**-40.43**	**4.29**	**0.03**	**0.56**
Hwet10	-40.36	4.36	0.03	0.56
Grass0.01	-40.29	4.43	0.02	0.60
Distance	***-40.26***	***4.46***	***0.02***	***0.56***
Hwet0.01	-40.25	4.47	0.02	0.56
Hwet0.1	-40.22	4.50	0.02	0.56
Barren10	-40.19	4.53	0.02	0.56
Barren100	-40.17	4.55	0.02	0.55
Wwet0.1	-39.79	4.93	0.02	0.55
Crops0.1	-39.76	4.96	0.02	0.55
Crops0.01	-39.52	5.20	0.02	0.55
Water0.1	-39.38	5.34	0.02	0.55
Wwet0.01	-39.33	5.39	0.02	0.55
Shrub10	**-38.58**	**6.14**	**0.01**	**0.58**
Forest0.1	-38.26	6.46	0.01	0.53
Water0.01	-38.01	6.71	0.01	0.54
Shrub0.1	-37.87	6.85	0.01	0.54
Shrub100	-37.46	7.26	0.01	0.58

Forest0.01	-36.45	8.26	0.00	0.52
Shrub0.01	-34.84	9.88	0.00	0.55
Grass10	-33.79	10.93	0.00	0.52
Grass100	-27.72	17.00	0.00	0.53

Table S1.2. Pairwise values of F_{ST} are below the diagonal with p-values based on 1,000 iteration are above the diagonal. Italicized values refer to those that were statistically significant ($p < 0.05$) before correction using the false discovery rate (FDR) and bolded if remained significant after correction.

	1A	**1B**	**2A**	**2B**	**3A**	**3B**	**4A**	**4B**	**5A**	**5B**
1A	–	0.040	0.006	0.002	0.002	0.002	0.002	0.002	0.002	0.002
1B	***0.089***	–	0.089	0.009	0.002	0.002	0.002	0.002	0.002	0.002
2A	***0.174***	0.052	–	0.103	0.002	0.002	0.002	0.002	0.002	0.002
2B	***0.394***	***0.204***	0.042	–	0.019	0.002	0.004	0.002	0.002	0.006
3A	***0.673***	***0.546***	***0.333***	***0.129***	–	0.081	0.059	0.058	0.063	0.085
3B	***0.797***	***0.730***	***0.511***	***0.304***	0.060	–	0.081	0.051	0.058	0.062
4A	***0.800***	***0.736***	***0.519***	***0.316***	*0.080*	0.054	–	0.062	0.040	0.500
4B	***0.796***	***0.727***	***0.510***	***0.308***	*0.078*	*0.076*	*0.069*	–	0.169	0.123
5A	***0.797***	***0.731***	***0.515***	***0.314***	*0.084*	*0.084*	***0.099***	0.029	–	0.223
5B	***0.800***	***0.734***	***0.515***	***0.315***	0.090	*0.098*	*0.114*	0.053	0.020	–

Literature Cited

Adamack, A. T., and Gruber, B., 2014. PopGenReport: simplifying basic population genetic analyses in R. Methods Ecol. Evolut. 5, 384–387. doi: 10.1111/2041-210X.12158

Aguillon, S.M., Fitzpatrick, J.W., Bowman, R., Schoech, S.J., Clark, A.G., Coop, G., Chen, N., 2017. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. PLoS Genet. 13, 1–27. https://doi.org/10.1371/journal.pgen.1006911

Alcala, N., Goudet, J., Vuilleumier, S., 2014. On the transition of genetic differentiation from isolation to panmixia: What we can learn from GST and D. Theor. Popul. Biol. 93, 75–84. https://doi.org/10.1016/j.tpb.2014.02.003

Antolin, M.L., Savage, L., Eisen, R., 2006. Landscape features influence genetic structure of black-tailed prairie dogs (Cynomys ludovicianus). Landsc Ecol 21:867–875

Archuleta, C.M., Constance, E.W., Arundel, S.T., Lowe, A. J., Mantey, K. S., & Phillips, L. A., 2017. The National Map seamless digital elevation model specifications: U.S. Geological Survey Techniques and Methods. https://doi.org/10.3133/tm11b9

Augustine, D.J., Matchett, M.R., Toombs, T.P., Cully, J.F., Johnson, T.L., Sidle, J.G., 2007. Spatiotemporal dynamics of black-tailed prairie dog colonies affected by plague. Landsc. Ecol. 23, 255–267. https://doi.org/10.1007/s10980-007-9175-6

Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Bartoń, K., 2019. MuMIn: Multi-Model Inference. R package version 1.43.6. https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1-48. doi:10.18637/jss.v067.i01.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. Stacks : Building and Genotyping Loci De Novo From Short-Read Sequences. Genes|Genomes|Genetics 1, 171–182. https://doi.org/10.1534/g3.111.000240

Chessel, D., Dufour, A., Thioulouse, J., 2004. The ade4 Package - I: One-Table Methods. R News, 4, 5-10. https://cran.r-project.org/doc/Rnews.

Clarke, R. T., Rothery, P., Raybould, A. F., 2002. Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. Journal of Agricultural, Biological and Environmental Statistics, 7, 361–372. https://doi. org/10.1198/108571102320

Cleary, K.A., Waits, L.P., Finegan, B., 2017. Comparative landscape genetics of two frugivorous bats in a biological corridor undergoing agricultural intensification. Mol. Ecol. 26, 4603–4617. https://doi.org/10.1111/mec.14230

Cully, J.F., Williams, E.S., 2001. Interspecific Comparisons of Sylvatic Plague in Prairie Dogs. J. Mammal. 82, 894–905. https://doi.org/10.1644/1545-1542(2001)082<0894:ICOSPI>2.0.CO;2

Danecek, P., Auton, A., Abecasis, G. et al. & 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics (Oxford, England), 27, 2156–2158

Dray, S., Dufour, A., 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical Software, 22, 1-20. doi: 10.18637/jss.v022.i04

Dullum, J.A.L.D., Foresman, K.R., Matchett, M.R., 2006. Efficacy of translocations for restoring populations of black-tailed prairie dogs. Wildl. Soc. Bull. 33, 842–850. https://doi.org/10.2193/0091-7648(2005)33[842:eotfrp]2.0.co;2

Earl, D. A., vonHoldt, B. M.,2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 4, 359-361. doi: 10.1007/s12686-011-9548-7

Epps, C.W., Wehausen, J.D., Bleich, V.C., Torres, S.G., Brashares, J.S., 2007. Optimizing dispersal and corridor models using landscape genetics. J. Appl. Ecol. 44,714-724

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., Borsboom, D., 2012. qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software 48, 1-18. URL http://www.jstatsoft.org/v48/i04/.

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

Frankham, R., 1996. Relationship of Genetic Variation to Population Size in Wildlife. Conserv. Biol. 10, 1500–1508. https://doi.org/10.1046/j.1442-9071.2001.00397.x

Garrett, M.G., Franklin, W.L., 1988. Ecology of Dispersal in the Black-Tailed Prairie Dog. J. Mammal. 69, 236–250.

Goslee, S.C., Urban, D.L., 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22, 1-19.

Goudet, J., Jombart, T., 2015. hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.04-22. https://CRAN.R-project.org/package=hierfstat

Hijmans, R.J., 2019. raster: Geographic Data Analysis and Modeling. R package version 2.8-19. https://CRAN.R-project.org/package=raster

Hollatz, C., Vilaca, S. T., Redondo, R. A. F., Marmontel, M., Baker, C. S., Santos, F. R., 2011. The Amazon River system as an ecological barrier driving genetic differentiation of the pink dolphin (Inia geoffrensis). Biological Journal of the Linnean Society, 102, 812-827.
Hoogland, J., 1995. The black-tailed prairie dog: Social life of a burrowing mammal. University of Chicago Press, Chicago, IL.

Hoogland, J.L., 2013. Prairie Dogs Disperse When All Close Kin Have Disappeared. Science 339, 1205–1207. https://doi.org/10.1126/science.1231689

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment, 132, 159 – 175.

Jombart, T., Ahmed, I., 2011. adegenet 1 . 3-1 : new tools for the analysis of genome-wide SNP data 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jombart, T., Devillard, S., Balloux, F., Falush, D., Stephens, M., et al., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94

Jones, P.H., Britten, H.B., 2010. The absence of concordant population genetic structure in the black-tailed prairie dog and the flea, Oropsylla hirsuta, with implications for the spread of Yersinia pestis. Mol. Ecol. 19, 2038–2049. https://doi.org/10.1111/j.1365-294X.2010.04634.x

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2, e281. doi: 10.7717/peerj.281

Keenan, K., Mcginnity, P., Cross, T.F., Crozier, W.W., Prodöhl, P.A., 2013. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol. Evol. 4, 782–788. https://doi.org/10.1111/2041-210X.12067

Kierepka, E.M., Latch, E.K., 2015. Performance of partial statistics in individual-based landscape genetics. Mol. Ecol. Resour. 15, 512-525 doi: 10.1111/1755-0998.12332

Legendre, P., Fortin, M.J., 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol. Ecol. Resour. 10, 831–844. https://doi.org/10.1111/j.1755-0998.2010.02866.x

Li, H., Durbin, R., 2009. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp698

Luikart, G., England, P. R., Tallmon, D., Jordan, S., Taberlet, P., 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Rev. Genet. 4, 981–994

Ma, T., Hu, Y., Russo, I.-R.M., Nie, Y., Yang, T., Xiong, L., Ma, S., Meng, T., Han, H., Zhang, X., Bruford, M.W., Wei, F., 2018. Walking in a heterogeneous landscape: dispersal, gene-flow and conservation implications for the giant panda in the Qinling Mountains, Evolutionary Applications. https://doi.org/10.1111/eva.12686

Manel, S., Schwartz, M.K., Luikart, G., Taberlet, P., 2003. Landscape genetics: Combining landscape ecology and population genetics. Trends Ecol. Evol. 18, 189–197. https://doi.org/10.1016/S0169-5347(03)00008-9

Martínez-Estévez, L., Balvanera, P., Pacheco, J., Ceballos, G., 2013. Prairie dog decline reduces the supply of ecosystem services and leads to desertification of semiarid grasslands. Plos One 8, e75229. doi:10.1371/journal.pone.0075229

Matchett, M.R., Biggins, D.E., Carlson, V., Powell, B., Rocke, T., 2010. Enzootic Plague Reduces Black-Footed Ferret (*Mustela nigripes*) Survival in Montana. Vector-Borne Zoonotic Dis. 10, 27–35. https://doi.org/10.1089/vbz.2009.0053

McRae, B.H., 2006. Isolation by resistance. Evolution 60, 1551-1561.

McRae, B.H., Dickson, B.G., Keitt, T.H., Shah, V.B., Mcrae, B.H., Dickson, B.G., Keitt, T.H., Shah, V.B., 2008. Using Circuit Theory to Model Connectivity in Ecology , Evolution , and Conservation. Ecology 89, 2712–2724.

McRae, B.H., and Shah, V.B., 2011. Circuitscape User Guide. ONLINE. The University of California, Santa Barbara. Available at: http://www.circuitscape.org.

Morin, P.A., Luikart, G., Wayne, R.K., 2004. SNPs in ecology, evolution and conservation. Trends Ecol. Evol. 19, 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Naimi, B., Hamm, N.A., Groen, T.A., Skidmore, A.K., Toxopeus, A.G., 2014. Where is positional uncertainty a problem for species distribution modelling. *Ecography*, 37, pp. 191-203. doi: 10.1111/j.1600-0587.2013.00205.x.

O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M., Portnoy, D.S., 2018. These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 0–3. https://doi.org/10.1111/mec.14792

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7. https://doi.org/10.1371/journal.pone.0037135

Pigg, R.M., 2014. *A multi-scale investigation of movement patterns among black-tailed prairie dog colonies.* Dissertation, Kansas State University, Manhattan, Kansas.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Proctor, J., 1998. *A GIS model for identifying potential black-tailed prairie dog habitat in the Northern Great Plains shortgrass prairie*. Thesis, University of Montana, Missoula, USA.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raufaste, N., Rousset, F.,2001). Are partial mantel tests adequate? Evolution 55, 1703–1705.

Roach, J., Stapp, P., Van Horne, B., MF, A., 2001. Genetic structure of a metapopulation of black-tailed prairie dogs. J. Mammal. 82, 946–959.

Rochette, N.C., Catchen, J.M., 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Publ. Gr. 12, 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rocke, T.E., Tripp, D.W., Russell, R.E., Abbott, R.C., Richgels, K.L.D., et al., 2017. Sylvatic Plague Vaccine Partially Protects Prairie Dogs (Cynomys spp.) in Field Trials. Ecohealth 14, 438–450. https://doi.org/10.1007/s10393-017-1253-x

Sackett, L.C., Collinge, S.K., Martin, A.P., 2013. Do pathogens reduce genetic diversity of their hosts? Variable effects of sylvatic plague in black-tailed prairie dogs. Mol. Ecol. 22, 2441–2455. https://doi.org/10.1111/mec.12270

Sackett, L.C., Cross, T.B., Jones, R.T., Johnson, W.C., Ballare, K., Ray, C., Collinge, S.K., Martin, A.P., 2012. Connectivity of prairie dog colonies in an altered landscape: Inferences from analysis of microsatellite DNA variation. Conserv. Genet. 13, 407–418. https://doi.org/10.1007/s10592-011-0293-y

Savolainen, O., Lascoux, M., Merilä, J., 2013. Ecological genomics of local adaptation. Nat. Rev. Genet. 14, 807–820. https://doi.org/10.1038/nrg3522

Shirk, A.J., Landguth, E.L., Cushman, S.A., 2017. A comparison of individual-based genetic distance metrics for landscape genetics. Mol Ecol Resour 17, 1308–1317. https://doi.org/10.1111/1755-0998.12684

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L., Waits, L.P., 2007. Putting the "landscape" in landscape genetics. Heredity (Edinb). 98, 128–142. https://doi.org/10.1038/sj.hdy.6800917

Tigano, A., Friesen, V.L., 2016. Genomics of local adaptation with gene flow. Mol. Ecol. 25, 2144–2164. https://doi.org/10.1111/mec.13606

Tipton, H.C., Dreitz, V.J., Doherty, P.F., 2008. Occupancy of mountain plover and burrowing owl in Colorado. J. of Wildl. Managem. 72, 1001-1006. Doi: 10.2193/2007-168.

Tripp, D.W., Rocke, T.E., Runge, J.P., Abbott, R.C., Miller, M.W., 2017. Burrow dusting or oral vaccination prevents plague-associated prairie dog colony collapse. Ecohealth 14, 451-462.

Vallinoto, M., Araripe, J., do Rego, P. S., Tagliaro, C. H., Sampaio, I., Schneider, H., 2006. Tocantins river as an effective barrier to gene flow in Saguinus niger populations. Genetics and Molec. Biol. 29, 215-219.

van Strien, M. J., Keller, D., Holderegger, R., 2012. A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. Mol. Ecol., 21, 4010-4023.

Willis, S.C., Hollenbeck, C.M., Puritz, J.B., Gold, J.R., Portnoy, D.S., 2017. Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Mol. Ecol. Resour. 17, 955–965. https://doi.org/10.1111/1755-0998.12647

Chapter 2. Elucidating dispersal patterns of the prairie dog flea, *Oropsylla hirsuta*, via alternative hosts: Implications for the spread of plague

RACHAEL M. GIGLIO[1], TONIE E. ROCKIE[2], JORGE E. OSORIO[3], EMILY K. LATCH[1]

[1]*University of Wisconsin-Milwaukee, Milwaukee, WI*

[2]*United States Geological Survey-National Wildlife Health Center, Madison, WI*

[3]*University of Wisconsin-Madison, Madison, WI*

Abstract

Plague is a flea-mediated disease caused by the bacterium *Yersinia pestis*. It moves through prairie dog colonies quickly, causing >90% mortality. However, the transmission pathways among prairie dog colonies remain unclear. Prairie dogs live in highly structured social groups with limited gene flow among colonies. Yet, the primary flea vector for plague found on black-tailed prairie dogs, *Oropsylla hirsuta*, has a higher degree of gene flow among prairie dog colonies than the prairie dogs themselves. This suggests that an alternative host may serve to move fleas among colonies. To evaluate the potential for small rodents to carry plague-infected fleas among prairie dog colonies, we used cutting-edge genomic techniques to characterize patterns of gene flow for two candidate small rodent species [deer mouse (*Peromyscus maniculatus*) and northern grasshopper mouse (*Onychomys leucogaster*)] and the flea vector (*O. hirsuta*) most prominently found in prairie dog colonies. We found concordant patterns of gene flow in prairie dog fleas and their prairie dog hosts in areas where prairie dogs exhibit high connectivity. However, in areas where prairie dog gene flow is restricted, flea dispersal is better explained by models that incorporate *O. hirsuta* movement by deer mice. These results, paired with field observations, indicate that deer mice are a viable alternative host for moving prairie dog fleas among colonies, thereby exacerbating the spread of plague.

Introduction

Vector-borne diseases are challenging to study for a number of reasons, including the complex interactions that exist between vector species and their hosts. As a result, much of the basic ecology of these disease systems remains unknown. Recently, tools from the field of molecular ecology have aided our quest to elucidate vector-host interactions and pathways of transmission (De MeeÛs et al. 2007; McCoy 2008; Lymbery and Thompson 2012). Specifically, population genetic approaches can be used to understand the impact of host population genetic structure on that of the parasite. Many parasites, including fleas, have limited movement without the aid of their hosts (Tripet et al. 2002). In these systems, hosts dictate the dispersal dynamics of their parasites, resulting in concordant genetic structure between host and parasite (Little et al. 2006; McCoy et al. 2005; Whiteman et al. 2007). For parasites that occupy a suite of hosts, concordant genetic structure is expected between the parasite and the host with the largest dispersal (Criscione et al. 2005). We can compare the population genetic structure of parasites and hosts to elucidate dispersal patterns and determine the host specificity of vector species (Whiteman et al. 2007; Criscione 2008; Stefka et al. 2011; Gómez-Díaz et al. 2012; McCoy et al. 2013). This molecular approach affords new insights into how population dynamics drive the spread of vector-borne diseases, and could aid in the identification of cryptic, alternative hosts (Prugnolle et al. 2005; Brouat et al. 2011).

Plague is a flea-mediated mammalian disease caused by the bacterium *Yersinia pestis*. Since its emergence in North America around 1900 (Gage and Kosoy 2005), plague has spread to native rodents, thus creating a sylvatic cycle. Previous work suggests that *Y. pestis* is maintained through hematophagous adult fleas and rodent hosts (Gage and Kosoy 2005). Virulence is

variable among mammals, with some species such as prairie dogs (*Cynomys* spp.) and the endangered black-footed ferret (*Mustela nigripes)* more severely affected than others (Antolin et al. 2002; Gage and Kosoy 2005). Specifically, outbreaks of plague are devastating to prairie dog colonies (>90% mortality), frequently leading to colony extirpation. Plague outbreaks in North America occur over the entire range of prairie dogs, including all five extant species: black-tailed prairie dog (*Cynomys ludovicianus*), white-tailed prairie dog (*C. leucurus*), Gunnison's prairie dog (*C. gunnisoni*), Utah prairie dog (*C. parvidens*), and Mexican prairie dog (*C. mexicanus*). For the threatened Utah prairie dog, plague-associated declines put the species at risk of extinction (Biggins and Kosoy 2001; Hoogland et al. 2004).

The flea species *Oropsylla hirsuta* is the primary flea species associated with prairie dogs and is implicated in the spread of plague, particularly with respect to supporting the fast-moving epizootics that occur in prairie dog colonies (Ubico et al. 1988; Cully et al. 1997; Cully and Williams 2001). However, previous studies have shown higher levels of gene flow for *O. hirsuta* than their prairie dog hosts, suggesting that the fleas are moving between prairie dog colonies on an alternative host (Jones and Britten 2010; Brinkeroff et al 2011). The alternative host species may serve as an enzootic host of *Y. pestis* or an unaffected host that shares fleas with infected prairie dogs and moves them from one colony to another. Two small rodent species, the deer mouse (*Peromyscus maniculatus*) and the northern grasshopper mouse (*Onychomys leucogaster*), have been found to harbor *O. hirsuta* fleas (with flea loads >2; Bron et al. 2019), and are thus feasible alternative hosts to move *Oropsylla* fleas among prairie dog colonies.

Our main objective was to elucidate the dispersal patterns of the prairie dog flea, *Oropsylla hirsuta*, and evaluate the feasibility of deer mice and northern grasshopper mice as facilitators of *Oropsylla* dispersal. To address this objective, we had two aims. The first aim was to assess the specificity of *O. hirsuta* to black-tailed prairie dogs as hosts. To address this aim, we compared patterns of genetic structure to quantify concordance. High specificity of the flea to its host is hypothesized to yield highly concordant patterns of genetic structure. The second aim was to determine if *O. hirsuta* dispersal is mediated by an alternative host. For this aim, we quantified concordance in the patterns of genetic structure between *O. hirsuta* and each of the two most abundant rodent species found in prairie dog colonies harboring *O. hirsuta* (i.e., deer mice and northern grasshopper mice). The importance of *O. hirsuta* as a vector for plague transmission makes understanding how it moves among colonies critical for improving plague mitigation efforts. Further, by identifying alternative hosts to transmission we may shed light on how plague is introduced to healthy prairie dog colonies, thereby initiating epizootic events.

Methods

Sample sites and collection

Black-tailed prairie dogs and small rodents were trapped in 2014 from the Charles M. Russell National Wildlife Refuge (CMR) in Montana, USA as part of a sylvatic plague vaccine trial conducted by the USGS National Wildlife Health Center (Fig. 2.1; Rocke et al 2017; Bron et al. 2018). The distance between sites ranged from 0.86 km to 34 km. Hair and whisker samples were collected from black-tailed prairie dogs, deer mice, and northern grasshopper mice for DNA extraction. Fleas, including the *O. hirsuta* fleas used in this study, were collected from each trapped animal within the study area and identified to species.

DNA extraction and sequencing

Prairie dogs

DNA was extracted from black-tailed prairie dog samples using a spin-column kit (*Quick*-DNA Miniprep kit; Zymo Research), following the manufacturer's feather extraction protocol. The quantity of DNA was estimated using a Qubit 3.0 fluorometer (Invitrogen) with the high sensitivity kit for double-stranded DNA. Samples with >300 ng of DNA were used for double digest restriction site associated DNA (ddRAD) sequencing (Peterson et al. 2012). DNA was sheared using the restriction enzymes *HindIII* and *NlaIII* followed by a 250–500 bp size selection step using a Pippin Prep (Sage Sciences). Sequences were generated on a NovaSeq6000 with 150 bp paired-end reads at Texas A&M AgriLife Genomics. Sequences were aligned to a Gunnison's prairie dog genome (Sackett pers comm) using the BWA short-read aligner with the BWA-MEM alignment algorithm (Li and Durbin 2009). We used the program *Stacks* v1.48 (Catchen et al. 2011) to assemble contigs and filter single nucleotide polymorphisms (SNPs) following the guidelines outlined in Rochette and Catchen (2017). Specifically, we used the *populations* module in *Stacks* to limit one SNP per contig, filter loci with minor allele frequencies <0.05, remove loci with an observed heterozygosity >0.70, and reduce loci to those present in all sites and in at least 80% of individuals per site. We then used *VCFtools* v.0.1.16 (Danecek et al. 2011) and the R package *adegenet* (Jombart et al. 2011) to remove individuals with >30% missing data and loci with an average depth of coverage exceeding 2X the mode. We sequenced 300 black-tailed prairie dogs to identify SNPs. We then subsampled 10 individuals per sampling location for downstream analyses, to equalize sample size across species.

Small rodents and fleas

DNA was extracted from hair and whiskers from individual deer mice and northern grasshopper mice using the *Quick*-DNA Miniprep kit (Zymo Research). DNA was extracted from individual whole fleas that were homogenized via bead beating using a *Quick*-DNA Microprep kit (Zymo Research). Low DNA yield from fleas and from rodent whiskers required pooling samples from multiple individuals of the same species prior to sequencing. Pools consisted of 3–10 individuals from the same sampling site, pooled in equimolar amounts as quantified using a Qubit fluorometer. Pooled samples were sequenced using a ddRAD protocol similar to that used for black-tailed prairie dogs. DNA was sheared using *SpeI* and *MbolI* for deer mice and *SpeI* and *NlaIII* for fleas and northern grasshopper mice followed by a 300–600 bp size selection step (Pippin Prep, Sage Sciences) and paired-end sequencing on a NovaSeq6000 at Texas A&M AgriLife Genomics. We used the genomic pipeline *dDocent* (Puritz et al. 2014a, b) to conduct a *de novo* assembly for each species. We used *VCFtools* to filter loci with a minimum quality score of < 30, a minor allele count of < 3, a minimum depth of coverage of < 10, a maximum depth of coverage of < 100, minor allele frequencies < 0.05, or with > 50% missing data. Further, we used *VCFtools* to remove indels and identify SNPs within contigs.

Genetic structure

A principal component analysis (PCA) was used to visualize patterns of genetic variation using the R package *ade4* v.1.7-13 (Chessel et al. 2004; Dray and Dufour 2007). We calculated pairwise F_{ST} for prairie dogs using the R package *hierfstat* (Goudet and Jombart 2015) and for pooled samples (fleas, deer mice, and northern grasshopper mice) using *popoolation2* (Kofler et al. 2011).

Linear models

To investigate the influence of known and candidate hosts on vector dispersal, we built models in a two-step process. First, we used basic linear models to quantify the explanatory power of each host's dispersal on flea dispersal, and then we used multivariate models to identify the set of hosts that best predicted flea dispersal. Basic linear models were built using data from all host species (prairie dog, deer mouse, and northern grasshopper mouse) and geographic distance (calculated as Euclidean distance using the *dist* function in R (R Core Team 2018). Northern grasshopper mice were only found at 3 of the 5 sites (CMR2, CMR3, and CMR5), so linear models were built using data from these 3 sites only. The basic linear models were evaluated based on corrected AIC (AIC_C) and R^2, which were calculated using R core functions (R Core Team 2018).

To account for non-independence among pairwise points, we built linear mixed effects models with maximum likelihood population effects parameterization (MLPE) using the *lme4* package in R (Bates et al. 2015; Clarke et al. 2002). Although it has not previously been used to evaluate host-vector dispersal patterns, this MLPE method has been shown to perform well in landscape genetic studies to evaluate how the landscape influences species dispersal (Shirk et al. 2017). For our host-vector application of the MLPE approach, we used patterns of pairwise genetic differentiation (measured as F_{ST}) for prairie dogs, deer mice, and northern grasshopper mice as explanatory variables (introduced as fixed effects), and flea F_{ST} as the response variable. Differences among sampling sites were introduced as random effect terms. We built both univariate and multivariate models and included a null model with geographic distance as the

only predictor of flea F_{ST} to test for isolation-by-distance (IBD). Northern grasshopper mice were omitted from these analyses because they did not occur at all sites. Multivariate models with explanatory variables containing variance inflation factors (VIF) > 4 were excluded as high VIF values indicate multicollinearity (calculated using *usdm* package in R; Naimi et al. 2014). All models were evaluated using corrected Akaike Information Criterion (AIC_C) values, weights (AIC_{EW}), and the calculated marginal coefficient of determination (R^2) (R package *MuMIn* v. Bartoń 2019).

Results

DNA extraction and sequencing

After filtering, we identified 15,507 SNPs for *O. hirsuta* fleas from 5 pools from 5 different sites. Each flea pool contained DNA from 5–10 individuals from a single site (mean=9 individuals per site). We identified 3,416 SNPs for black-tailed prairie dogs from a set of 300 individuals (before subsampling 10 individuals per site for all analyses to equalize sampling effort among species). We identified 80,558 SNPs for deer mice from 5 pools from 5 sites. Each pool of deer mice contained DNA from 5–10 individuals from a single site (mean=8.4 individuals per site). We identified 19,619 SNPs for northern grasshopper mice from 3 pools from 3 different sites (CMR 2, 3, and 5; northern grasshopper mice were not found at CMR1 or CMR4). Each pool of northern grasshopper mice contained DNA from 3–6 individuals from a single site (mean=5 individuals per site).

Genetic structure

In *O. hirsuta* fleas, the greatest degree of population divergence occurred between western (CMR1 and CMR2) and eastern (CMR3, CMR4, and CMR5) sites based on PCA (Fig. 2.2). However, fleas exhibited low genetic divergence across all sampling sites based on F_{ST} (global F_{ST}=0.10). The lowest measures of pairwise F_{ST} for *O. hirsuta* were observed across the eastern sampling sites (CMR 3, 4, and 5; Fig. 2.3). CMR1 was more differentiated from all other sites, though F_{ST} remained < 0.14.

In black-tailed prairie dogs, a high degree of population divergence between western (CMR1 and CMR2) and eastern (CMR3, CMR4, and CMR5) sites was observed for black-tailed prairie dogs based on PCA (Fig. 2.2) and F_{ST} (Fig. 2.3). Like the fleas, black-tailed prairie dogs showed low F_{ST}, but only across eastern sites. The global F_{ST} was much higher (F_{ST}=0.41) in black-tailed prairie dogs than in fleas, due to the strong differentiation of western sites (CMR1 and CMR2) from each other (F_{ST}=0.36) and from eastern sites (F_{ST}=0.60).

Based on PCA, deer mice from CMR 1, CMR 2, and CMR 4 were the most genetically similar (Fig. 2.2). However, F_{ST} was low for all sites (deer mouse global F_{ST}= 0.14; Fig. 2.3). Compared to *O. hirsuta*, deer mice exhibited more genetic structure among eastern sites (deer mouse F_{ST}=0.17 and *O. hirsuta* F_{ST}=0.06) but similar levels of genetic structure between eastern and western sites (deer mouse F_{ST}=0.13 and *O. hirsuta* F_{ST}=0.11), and between western sites (deer mouse F_{ST}=0.09 and *O. hirsuta* F_{ST}=0.11).

A slightly higher degree of population divergence was detected for northern grasshopper mice between western (CMR2) and eastern (CMR3 and CMR5) sites based on PCA (Fig. 2.2);

however northern grasshopper mouse global F_{ST} was low (global F_{ST}= 0.17). Like the deer mice, northern grasshopper mice showed a lower degree of genetic structure between western and eastern sites (F_{ST}= 0.15) than among the eastern sites (F_{ST}=0.19). Both potential alternative hosts of *O. hirsuta* showed lower genetic structure than black-tailed prairie dogs (deer mouse global F_{ST}= 0.14 and northern grasshopper mouse global F_{ST}= 0.17), but a slightly higher degree of genetic structure compared to *O. hirsuta*.

Linear models

Patterns of genetic structure in *O. hirsuta* fleas were best explained using a linear model that included black-tailed prairie dog F_{ST} (R^2=0.99; Table 2.1). The next best performing model was Euclidean distance (IBD) (R^2=0.91). Models using the proposed alternative hosts performed the worst (deer mouse R^2=0.88 and northern grasshopper mice R^2=0.21). None of the linear models were statistically significant ($p<0.05$), largely because these models were based on data from only 3 sites (CMR 2, 3, 5).

The linear mixed effects models yielded similar results and were more powerful in that we could utilize data from all 5 sampled sites. Univariate models including black-tailed prairie dog F_{ST} (R^2_M=0.80), and to a lesser degree IBD (R^2_M=0.35), best predicted flea differentiation (Table 2.2). However, the multivariate model using deer mice and prairie dogs as explanatory variables of flea differentiation outperformed all univariate models and was the best performing linear mixed effects model overall (R^2_M=0.89). This multivariate model also exhibited low values of VIF (VIF=1.15).

Discussion

Vector-borne diseases involve complex interactions between hosts and vector species that make elucidating transmission pathways challenging. However, transmission pathways are driven by patterns of dispersal for both vectors and hosts, which leave genetic signatures that can be detected using molecular approaches. Here, we found high dispersal of the plague vector, *O. hirsuta*, throughout our study area. We also found concordant patterns of population divergence among black-tailed prairie dogs and their flea specialist *O. hirsuta* within our eastern sites. This indicates that black-tailed prairie dog dispersal has a large influence on the patterns of genetic structure for their flea specialist. However, at a broader scale, high gene flow in *O. hirsuta* was not reflected in prairie dog genetic structure patterns, suggesting that an alternative host contributes to the long-distance dispersal of *O. hirsuta*.

Our models signal the close relationship between *O. hirsuta* and black-tailed prairie dogs. Parasites tend to show concordant patterns of genetic structure with their hosts, because host behavior drives patterns of parasite dispersal (Little et al. 2006; Martinu et al. 2018). For example, Calhoun (2015) found concordant patterns of genetic structure between a species of chewing louse and two subspecies of their pocket gopher host, thereby illustrating the point that parasite dispersal is highly dependent on host behavior. Further, concordance in host-parasite genetic structure is dependent on the host specificity of the parasite (Barrett et al. 2008; Martinu et al. 2018). The concordant pattern of genetic structure we observed for *O. hirsuta* and black-tailed prairie dogs, combined with the fact that *O. hirsuta* was the dominant flea species found on prairie dogs at our sampling locations (Russell et al. 2018), shows that *O. hirsuta* is highly specific to prairie dogs. In accordance with these findings, factors such as geographic distance

that influence dispersal in prairie dogs (Jones and Britten 2010, Sackett et al. 2012) were also found to be important predictors of *O. hirsuta* gene flow.

Our results show that prairie dog dispersal alone does not account for the low degree of genetic structure of *O. hirsuta*. Previous studies have found similar results of low genetic structure in *O. hirsuta* higher genetic structure in black-tailed prairie dogs (Jones and Britten 2010; Jones et al. 2011). This suggests an alternative host may move prairie dog fleas among colonies. Even if the dispersal of fleas via an alternative host is rare, only a few migrants are necessary to prevent population differentiation from occurring (Spieth 1974). Across our entire study area, the model that best explained patterns of genetic structure in *O. hirsuta* included both prairie dog and deer mice as the explanatory variables. Whereas local genetic structure of the flea is likely driven by black-tailed prairie dog dispersal, long-distance flea dispersal is likely facilitated by an alternative host that moves beyond the dispersal range of prairie dogs. While numerous species with high dispersal capabilities exist on prairie dog colonies, few have been found with *O. hirsuta* fleas on them. However, both deer mice and northern grasshopper mice have been found carrying *O. hirsuta* fleas, and previous studies have implicated both mouse species in the spread of plague (Stapp et al. 2009; Salkeld et al. 2010; Kraft and Stapp 2013; Danforth et al. 2018; Foley and Foley 2010; Thiagarajan et al. 2008; Bron et al. 2019). Our results lend support to the hypothesis that deer mice are important facilitators of long-distance dispersal of *O. hirsuta*.

A large body of prior evidence exists to support the hypothesis that northern grasshopper mice play a substantial role in moving prairie dog fleas (Stapp et al. 2008; Stapp et al. 2009; Stapp and Salkeld 2009; Thiagarajan et al. 2008). For example, Kraft and Stapp (2013) found high flea

burdens including *O. hirsuta*, and heavy burrow use by northern grasshopper mice in northern Colorado. Further, they showed that grasshopper mice are wide-ranging and capable of moving relatively long distances. However, models in our study that used the northern grasshopper mouse as a predictor for *O. hirsuta* genetic structure performed poorly compared to models that included deer mice or black-tailed prairie dogs as predictors. This could have been due to low detection of northern grasshopper mice within our study. Only a few individuals were found on three of our study sites (Bron et al. 2018). We were able to gauge long-distance connectivity from these three sites, though not at the resolution we were able to achieve for the other species in our study. It is possible that in portions of prairie dog range where northern grasshopper mice occur in greater densities, they have a greater impact on the dispersal of *O. hirsuta*.

Comparative population genetic studies such as ours can be informative in elucidating transmission pathways of vector-borne diseases. They are but a first step, however, because gene flow is not the only determinant of population divergence (measured as F_{ST} in our study). The life history of a species also plays a decisive role in population genetic structure (Barrett et al. 2008; Dharmarajan et al. 2016; Martinu et a. 2018). Accordingly, direct comparisons of genetic structure patterns across species should be interpreted in light of life history differences. For example, shorter generation times of parasites compared to their hosts might result in more rapid divergence, and a signature of population structure in the parasite that is not reflected in the host (Whiteman and Parker 2005). Because of this, studies have used parasites to identify fine-scale host movements. For example, tick species were used to give insight to intercolony movements of their hosts, the black-legged kittwake (McCoy et al. 2005). In disease transmission systems, parasites could be used to track fine-scale transmission of diseases that may not be reflected in

the host. Parasite specificity has also been found to impact host-parasite population structure comparisons, with highly specific parasites exhibiting a higher degree of genetic structure than their host (Barrett et al. 2008; Martinu et al. 2018). Little is known about the life history of *O. hirsuta*, but under the assumption that *O. hirsuta* is fairly host-specific, we would expect *O. hirsuta* to evolve more rapidly than black-tailed prairie dogs. Therefore, the two most likely scenarios resulting in lower genetic structure of *O. hirsuta* compared to black-tailed prairie dogs are that *O. hirsuta* dispersal occurs during non-reproductive movements of prairie dogs (dispersal without gene flow), or on an alternative host. Since prairie dog dispersal is uncommon and typically occurs over distances of less than 5 km (Garrett and Franklin 1988; Hoogland 2013), using an alternative host for the long-distance dispersal is the most likely scenario.

Prairie dogs are keystone species within prairie ecosystems and are highly susceptible to plague. Understanding how plague is spread among prairie dog colonies is vital to the conservation of not only prairie dogs, but also the numerous species that depend on them. In this study, we used population genetic theory to investigate potential transmission pathways of plague. We showed that alternative hosts move among prairie dog colonies more frequently than the prairie dogs themselves. This information, paired with field studies that found *O. hirsuta* on these alternative hosts, suggests that deer mice and northern grasshopper mice are viable alternative hosts and may serve to carry fleas beyond prairie dog dispersal range. Further, plague-positive fleas were found on deer mice and northern grasshopper mice prior to plague-induced die-offs of prairie dogs, including at our sampling sites (Bron et al. 2019). This suggests that these alternative hosts may also introduce plague to prairie dog colonies. By characterizing the transmission pathways

that include all potential hosts, we may improve our ability to predict disease outbreaks and develop the proper mitigation efforts.

Figure 2.1. Map of sampling sites within the Charles M. Russell Wildlife Refuge in Phillips county, MT. For each site, samples were collected for black-tailed prairie dogs, deer mice, and norther grasshopper mice. Fleas were collected off all trapped mammals and identified to species.

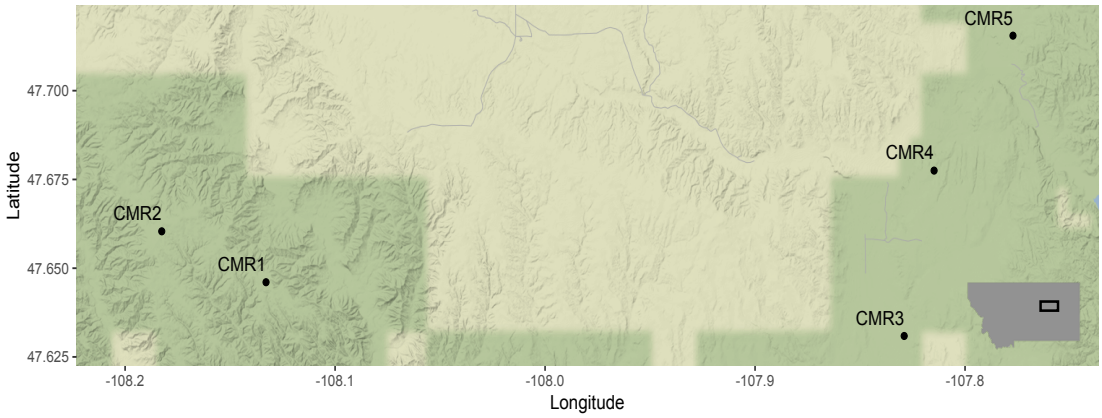


Figure 2.2. Principal component analysis (PCA) shows genetic differentiation among the five sampling locations for *O. hirsuta*, black-tailed prairie dogs, deer mice, and northern grasshopper mice. Population divergence was most evident among eastern and western sites for *O. hirsuta* and black-tailed prairie dogs.

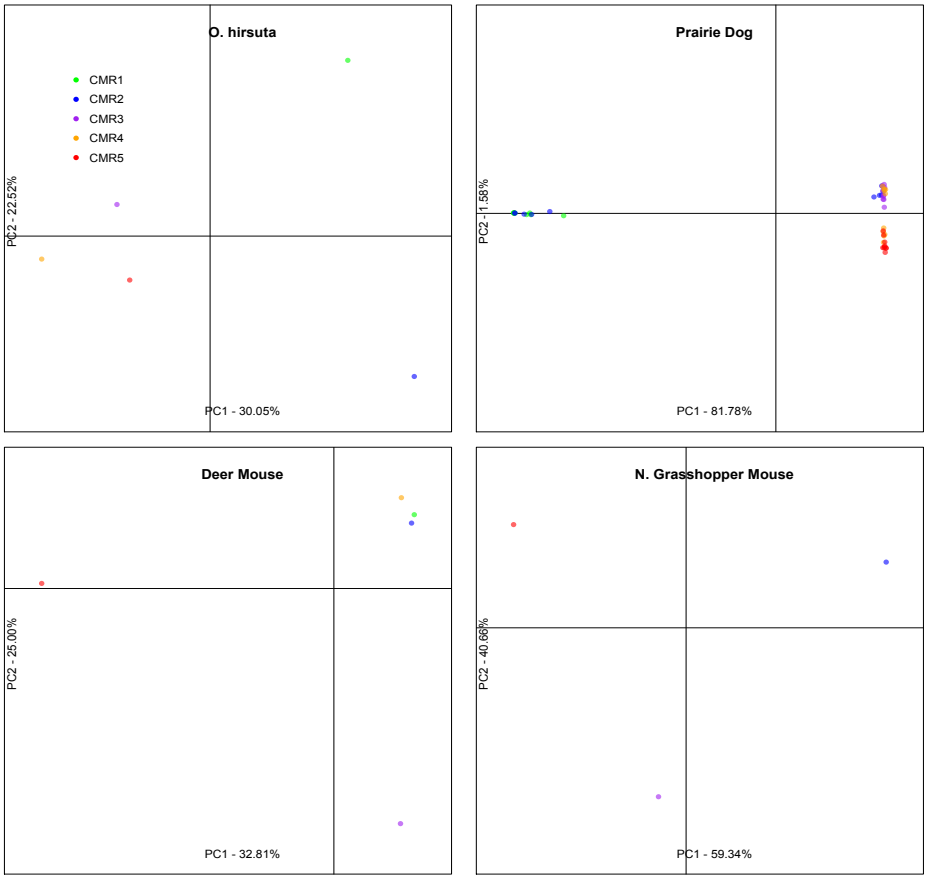


Figure 2.3. Pairwise F_{ST} of the flea *O. hirsuta* (A), black-tailed prairie dog (B), deer mouse (C), and northern grasshopper mouse (D). Warmer colors indicate lower F_{ST} values (less population differentiation) while cooler colors indicate higher F_{ST}. Grey blocks indicate no data for that site.

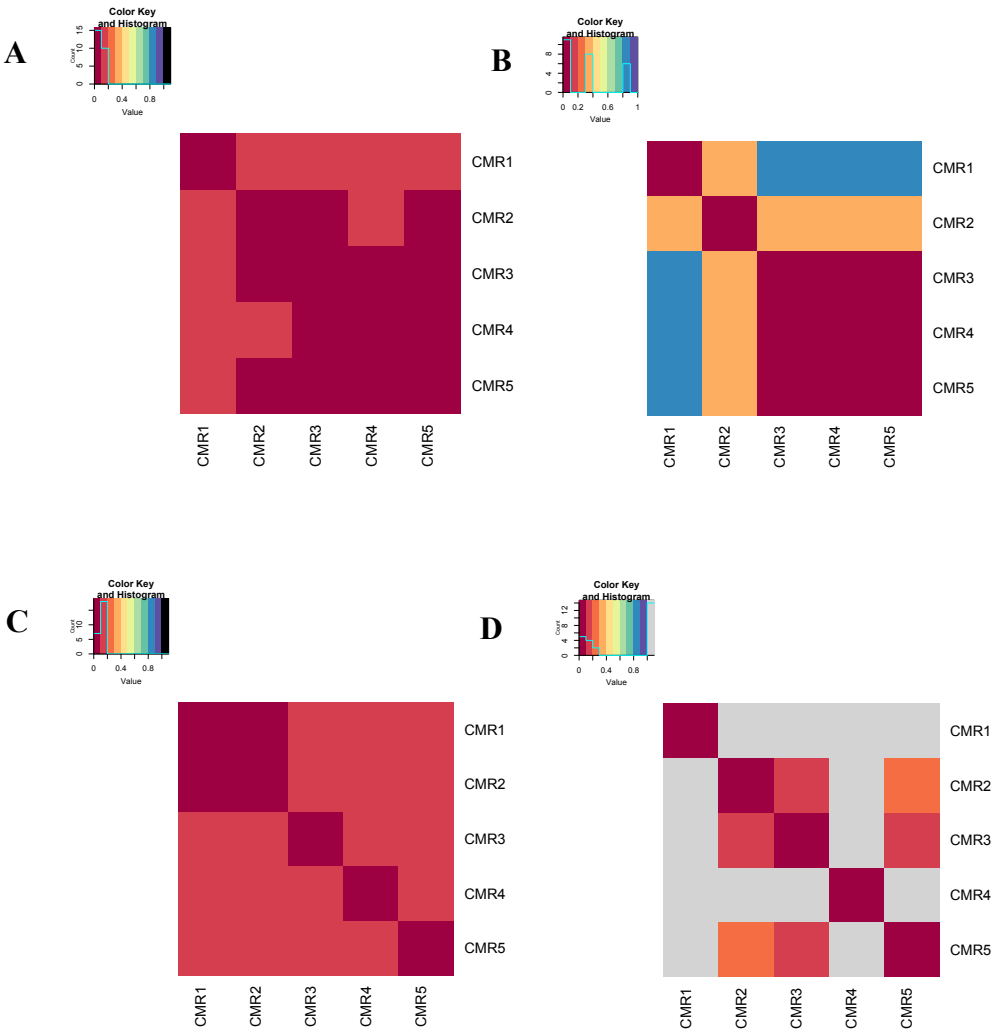


Table 2.1. Results of basic linear models. Models were evaluated based on overall log-likelihood (LogLik), change in corrected AIC values from the best model (ΔAIC_C), AIC weights (AIC_{ew}), the amount of variation explained by each model (R^2), and whether the model was statistically significant (P-value).

Model	**LogLik**	**ΔAIC_C**	**AIC_{ew}**	**R^2**	**P-value**
Prairie Dog	15.39	0.00	0.92	0.99	0.07
Distance	12.36	6.05	0.04	0.91	0.20
Deer Mouse	12.00	6.77	0.03	0.88	0.22
Grasshopper Mouse	9.17	12.43	0.00	0.21	0.70

Table 2.2. Results from linear mixed effects model using maximum likelihood population effects (MLPE) parameterization. Models were evaluated based on overall log-likelihood (LogLik), change in corrected AIC values from the best model (ΔAIC_C), AIC weights (AIC_{ew}), and the marginal coefficient of determination (R^2_M).

Model	**LogLik**	**ΔAIC_C**	**AIC_{ew}**	**R^2_M**
Deer Mouse + Prairie Dog	33.81	0.00	0.86	0.89
Prairie Dog	30.91	3.74	0.13	0.80
Distance	28.51	8.55	0.01	0.35
Deer Mouse	25.80	13.96	0.00	0.44

Literature Cited

Antolin, M.F., Gober, P., Luce, B., Biggins, D.E., Van Pelt, et al., 2002. The influence of sylvatic plague on North American wildlife at the landscape level, with special emphasis on black-footed ferret and prairie dog conservation. US Fish Wildl. Publ. 57.

Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Bartoń, K., 2019. MuMIn: Multi-Model Inference. R package version 1.43.6. https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1-48. doi:10.18637/jss.v067.i01.

Biggins, D.E., Kosoy, M.Y., 2001. Influences of Introduced Plague on North American Mammals: Implications From Ecology of Plague in Asia. J. Mammal. 82, 906–916. https://doi.org/10.1644/1545-1542(2001)082<0906:IOIPON>2.0.CO;2

Brinkerhoff, R.J., Martin, A.P., Jones, R.T., Collinge, S.K., 2011. Population genetic structure of the prairie dog flea and plague vector, Oropsylla hirsuta. Parasitology 138, 71–79. https://doi.org/10.1017/S0031182010001046

Bron, G.M., Malavé, C.M., Boulerice, J.T., Osorio, J.E., Rocke, T.E., 2019. Plague-Positive Mouse Fleas on Mice Before Plague Induced Die-Offs in Black-Tailed and White-Tailed Prairie Dogs. Vector-Borne Zoonotic Dis. XX, vbz.2018.2322. https://doi.org/10.1089/vbz.2018.2322

Bron, G.M., Richgels, K.L.D., Samuel, M.D., Poje, J.E., Lorenzsonn, F., Matteson, J.P., Boulerice, J.T., Osorio, J.E., Rocke, T.E., 2018. Impact of Sylvatic Plague Vaccine on Non-target Small Rodents in Grassland Ecosystems. Ecohealth 1–11. https://doi.org/10.1007/s10393-018-1334-5

Calhoun, C.S., 2015. Host behaviour drives parasite genetics at multiple geographic scales : population genetics of the chewing louse , Thomomydoecus minor 4129–4144. https://doi.org/10.1111/mec.13306

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. Stacks : Building and Genotyping Loci De Novo From Short-Read Sequences. Genes|Genomes|Genetics 1, 171–182. https://doi.org/10.1534/g3.111.000240

Chessel, D., Dufour, A., Thioulouse, J., 2004. The ade4 Package - I: One-Table Methods. R News, 4, 5-10. https://cran.r-project.org/doc/Rnews.

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. Journal of Agricultural Biological & Environmental Statistics, 7(3), 361.

Criscione, C.D., 2008. Parasite Co-Structure: Broad and Local Scale Approaches. Parasite 15, 439–443. https://doi.org/10.1051/parasite/2008153439

Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: Elucidating ecological and microevolutionary processes. Mol. Ecol. 14, 2247–2257. https://doi.org/10.1111/j.1365-294X.2005.02587.x

Cully, A.M. Barnes, T.J. Quan and G. Maupin. 1997. Dynamics of plague in a Gunnison's prairie dog colony. Journal of Wildlife Diseases 33: 706–718.

Cully, J.F., Williams, E.S., 2001. Interspecific Comparisons of Sylvatic Plague in Prairie Dogs. J. Mammal. 82, 894–905. https://doi.org/10.1644/1545-1542(2001)082<0894:ICOSPI>2.0.CO;2

Danecek, P., Auton, A., Abecasis, G. et al. & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics (Oxford, England), 27, 2156–2158

Danforth M, Tucker J, Novak M. The deer mouse (Peromyscus maniculatus) as an enzootic reservoir of plague in California. EcoHealth 2018; 15:566–576.

De MeeÛs, T., McCoy, K.D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-bousse, S., Renaud, F., 2007. Population genetics and molecular epidemiology or how to "‘ d´ busquer la b ˆ te ."' Infect. Genet. Evol. 7, 308–332. https://doi.org/10.1016/j.meegid.2006.07.003

Dharmarajan, G., Beasley, J.C., Beatty, W.S., Olson, Z.H., Fike, J.A., Jr, O.E.R., 2016. Genetic co- structuring in host- parasite systems : Empirical data from raccoons and raccoon ticks. Ecosphere 7, 1–15.

Dray, S., Dufour, A., 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical Software, 22, 1-20. doi: 10.18637/jss.v022.i04

Foley, P., Foley, J., 2010. Modeling susceptible infective recovered dynamics and plague persistence in California rodent-flea communities. Vector borne zoonotic Dis. 10, 59–67. https://doi.org/10.1089/vbz.2009.0048

Gage, K.L., Kosoy, M.Y., 2005. Natural History of Plague: Perspectives from more than a Century of Research. Annu. Rev. Entomol 50, 505–28. https://doi.org/10.1146/annurev.ento.50.071803.130337

Garrett, M.G., Franklin, W.L., 1988. Ecology of Dispersal in the Black-Tailed Prairie Dog. J. Mammal. 69, 236–250.

Gomez-Diaz, E., Morris-Pocock, J. a., Gonzalez-Solis, J., McCoy, K.D., 2012. Trans-oceanic host dispersal explains high seabird tick diversity on Cape Verde islands. Biol. Lett. 8, 616–619. https://doi.org/10.1098/rsbl.2012.0179

Goudet, J., Jombart, T., 2015. hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.04-22. https://CRAN.R-project.org/package=hierfstat

Hoogland, J.L., 2013. Prairie dogs disperse when all close kin have disappeared. Science (80-.). 339, 1205–1207. https://doi.org/10.1126/science.1231689

Hoogland, J.L., Davis, S., Benson-Amram, S., Labruna, D., Goossens, B., Hoogland, M.A., 2004. Pyraperm Kills Fleas and Halts Plague among Utah Prairie Dogs. Southwest. Nat. 49, 376–383.

Jombart, T., Ahmed, I., 2011. adegenet 1 . 3-1 : new tools for the analysis of genome-wide SNP data 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, P.H., Britten, H.B., 2010. The absence of concordant population genetic structure in the black-tailed prairie dog and the flea, Oropsylla hirsuta, with implications for the spread of Yersinia pestis. Mol. Ecol. 19, 2038–2049. https://doi.org/10.1111/j.1365-294X.2010.04634.x

Jones, P.H., Washburn, L.R., Britten, H.B., 2011. Gene flow in a Yersinia pestis vector, Oropsylla hirsuta, during a plague epizootic. J. Vector Borne Dis. 48, 125–132.

Kofler, R., Pandey, R.V., Schlötterer, C., 2011. PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27, 3435–3436. https://doi.org/10.1093/bioinformatics/btr589

Kraft, J.P., Stapp, P., 2013. Movements and burrow use by northern grasshopper mice as a possible mechanism of plague spread in prairie dog colonies. J. Mammal. 94, 1087–1093. https://doi.org/10.1644/12-MAMM-A-197.1

Li, H., Durbin, R., 2009. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp698

Little, T. J., K. Watt, and D. Ebert. 2006. Parasite-host specificity: Experimental studies on the basis of parasite adaptation. Evolution 60:31-38.

Lymbery, A.J., Thompson, R.C.A., 2012. The molecular epidemiology of parasite infections : tools and applications. Mol. Biochem. Parasitol. 181, 102–116.

Martinů, J., Hypša, V., Štefka, J., 2018. Host specificity driving genetic structure and diversity in ectoparasite populations : Coevolutionary patterns in Apodemus mice and their lice 10008–10022. https://doi.org/10.1002/ece3.4424

McCoy, K.D., 2008. What can the population structure of vectors tell us? Parasite 15, 444–448.

McCoy, K.D., Boulinier, T., Tirard, C., 2005. Comparative host-parasite population structures: Disentangling prospecting and dispersal in the black-legged kittiwake Rissa tridactyla. Mol. Ecol. 14, 2825–2838. https://doi.org/10.1111/j.1365-294X.2005.02631.x

Mccoy, K.D., Léger, E., Dietrich, M., 2013. Host specialization in ticks and transmission of tick-borne diseases: a review. Front. Cell. Infect. Microbiol. 3, 57. https://doi.org/10.3389/fcimb.2013.00057

Naimi, B., Hamm, N.A., Groen, T.A., Skidmore, A.K., Toxopeus, A.G., 2014. Where is positional uncertainty a problem for species distribution modelling. *Ecography*, 37, pp. 191-203. doi: 10.1111/j.1600-0587.2013.00205.x.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7. https://doi.org/10.1371/journal.pone.0037135

Puritz, J.B., Hollenbeck, C.M., Gold, J.R., 2014a. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2, e431. https://doi.org/10.7717/peerj.431

Puritz, J.B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., Bird, C. E., 2014b. Demystifying the RAD fad. Molecular Ecology 23, 5937–5942.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rochette, N.C., Catchen, J.M., 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Publ. Gr. 12, 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rocke, T.E., Tripp, D.W., Russell, R.E., Abbott, R.C., Richgels, K.L.D., et al., 2017. Sylvatic Plague Vaccine Partially Protects Prairie Dogs (Cynomys spp.) in Field Trials. Ecohealth 14, 438–450. https://doi.org/10.1007/s10393-017-1253-x

Russell, R.E., Abbott, R.C., Tripp, D.W., Rocke, T.E., 2018. Local factors associated with on-host flea distributions on prairie dog colonies. Ecol. Evol. https://doi.org/10.1002/ece3.4390

Sackett, L.C., Cross, T.B., Jones, R.T., Johnson, W.C., Ballare, K., Ray, C., Collinge, S.K., Martin, A.P., 2012. Connectivity of prairie dog colonies in an altered landscape: Inferences from analysis of microsatellite DNA variation. Conserv. Genet. 13, 407–418. https://doi.org/10.1007/s10592-011-0293-y

Salkeld, D.J., Salathé, M., Stapp, P., Jones, J.H., 2010. Plague outbreaks in prairie dog populations explained by percolation thresholds of alternate host abundance. Proc. Natl. Acad. Sci. U. S. A. 107, 14247–50. https://doi.org/10.1073/pnas.1002826107

Shirk, A.J., Landguth, E.L., Cushman, S.A., 2017. A comparison of individual-based genetic distance metrics for landscape genetics. Mol Ecol Resour 17, 1308–1317. https://doi.org/10.1111/1755-0998.12684

Spieth, P.T., 1974, Gene flow and genetic differentiation. Genetics 78, 961-965.

Stapp, P., Salkeld, R.J., 2009. Inferring host-parasite relationships using stable isotopes: implications for disease transmission and host specificity. Ecology 90, 3268–3273.

Stapp, P., Salkeld, D.J., Eisen, R.J., Pappert, R., Young, J., et al., 2008. Exposure of small rodents to plague during epizootics in black-tailed prairie dogs. Journal of Wildlife Diseases 44:724–730

Stapp, P., Salkeld, D.J., Franklin, H.A., Kraft, J.P., Tripp, D.W., Antolin, M.F., Gage, K.L., 2009. Evidence for the involvement of an alternate rodent host in the dynamics of introduced plague in prairie dogs. J. Anim. Ecol. 78, 807–817. https://doi.org/10.1111/j.1365-2656.2009.01541.x

Stefka, J., Hoeck, P. E. A., Keller, L. F. and Smith, V. S., 2011. A hitchhikers guide to the Galapagos: co-phylogeography of Galapagos mockingbirds and their parasites. BMC Evolutionary Biology 11, 284.

Thiagarajan, B., Bai, Y., Gage, K.L., Cully, J.F., 2008. Prevalence of Yersinia pestis in rodents and fleas associated with black- tailed prairie dogs (Cynomys ludovicianus) at Thunder Basin National Grassland, Wyoming. J Wildl Dis 44:731– 736

Tripet, F., Christe, P., Moller, A.P., 2002. The importance of host spatial distribution for parasite specialization and speciation: a comparative study of bird fleas (Siphonaptera: Ceratophyllidae). J. Anim. Ecol. 71, 735-748. https://doi.org/10.1046/j.1365-2656.2002.00639.x

Ubico, S.R., Maupin, G.O., Fogerstone, K.A., McLean, R.G., 1988. A plague epizootic in the white- tailed prairie dogs (Cynomys leucurus) of meeteetse, Wyoming. J. Wildl. Dis. 24, 399-406.

Whiteman, N.K., Parker, P.G., 2005. Using parasites to infer host population history: a new rationale for parasite conservation. Anim. Conserv. 8, 175–181. doi: https://doi.org/10.1017/S1367943005001915

Whiteman, N.K., Kimball, R.T., Parker, P.G., 2007. Co-phylogeography and comparative population genetics of the threatened Galápagos hawk and three ectoparasite species: Ecology shapes population histories within parasite communities. Mol. Ecol. 16, 4759–4773. https://doi.org/10.1111/j.1365-294X.2007.03512.x

Chapter 3. Characterizing patterns of genomic variation in the threatened Utah prairie dog: Implications for conservation and management

RACHAEL M. GIGLIO[1], TONIE E. ROCKIE[2], JORGE E. OSORIO[3], EMILY K. LATCH[1]

[1]*University of Wisconsin-Milwaukee, Milwaukee, WI*

[2]*United States Geological Survey-National Wildlife Health Center, Madison, WI*

[3]*University of Wisconsin-Madison, Madison, WI*

Abstract

Utah prairie dogs (*Cynomys parvidens*) are federally threatened due to persecution, habitat destruction, and outbreaks of plague. Today, Utah prairie dogs exist in small, isolated populations, making them less demographically stable and more susceptible to erosion of genetic variation by genetic drift. We characterized patterns of genetic structure at neutral and putatively adaptive loci in order to evaluate the relative effects of genetic drift and local adaptation on population divergence. We sampled individuals across the Utah prairie dog species range and generated 2,955 single nucleotide polymorphisms (SNPs) using double digest restriction site associated DNA sequencing (ddRAD). Genetic diversity was lower in low elevation sites compared to high elevation sites. Population divergence was high among sites and followed an isolation-by-distance (IBD) model. Our results suggest that genetic drift plays a substantial role in the population divergence of the Utah prairie dog and colonies would likely benefit from translocation of individuals from HE sites to CC sites despite the detection of environmental associations with outlier loci. By understanding the processes that shape genetic structure, we can make better informed decisions with respect to the management of threatened species to ensure that adaptation is not stymied.

Introduction

As a result of habitat loss and fragmentation, threatened species often occur in small, isolated populations. Genetic variation can be rapidly eroded in these small populations by genetic drift, a process that goes unmitigated in isolated populations without gene flow (Frankham 1996). A lack of genetic variation weakens species viability, stifling the evolutionary potential of a species and constraining adaptation to local environmental conditions (Barrett and Schluter 2008; Savolainen et al. 2013). As global environmental change intensifies (Urban 2015; Wiens 2016), the ability of a species to adapt to changing local conditions will be increasingly central to its long-term viability.

Conservation and management activities that facilitate the retention of genetic variation, evolutionary potential, and adaptability will further help species avoid extinction. For example, translocations, used to bolster declining populations, can be improved by incorporating genetic data to tailor translocation actions to outcomes that boost genetic variation. Selecting source populations that are genetically appropriate for the target population (e.g., Johnson et al. 2010), prioritizing target populations with low genetic variation (e.g., Whiteley et al. 2015), or gauging incorporation of source genotypes (e.g., Latch and Rhodes 2005, Bateson et al. 2014, Mulder et al. 2017) can improve efforts to retain genetic variation. Conservation and management actions could further improve the evolutionary potential and adaptive capacity of populations by incorporating data from studies of adaptive variation, especially if those actions include translocations, genetic rescue, or assisted gene flow (Funk et al. 2018, Flanagan et al. 2018). Advances in genomics for non-model species means that we can generate broad coverage and high-resolution genomic data for an increasing number of species. Genomic data can be used to

survey both adaptive and neutral genetic variation and can be incorporated into conservation policy to improve long-term species viability against changing environments and exposure to new diseases (Funk et al. 2018, Flanagan et al. 2018).

In this study, we use a population genomics approach to understand the maintenance of genetic variation in the threatened Utah prairie dog (*Cynomys parvidens*). The Utah prairie dog is one of five species of prairie dog found in North America and is listed as Threatened under the United States' Endangered Species Act. Due to a history of heavy range-wide eradication campaigns during the 19th and 20th centuries as well as ongoing habitat loss and epizootic outbreaks of plague, the Utah prairie dog has been reduced from roughly 95,000 individuals range-wide in the 1920s to approximately 14,000 today (Collier and Spillet 1973; Brown et al. 2016). In 1972, a recovery plan for the Utah prairie dog was enacted that focused on translocating individuals from private land to protected public lands (McDonald 1993; United States Fish and Wildlife Service 2012). Today, Utah prairie dogs exist in small, isolated populations, making them less demographically stable and more susceptible to the erosion of genetic variation through genetic drift (Wright 1931, Gilpin and Soule 1986). Further, Utah prairie dogs are highly social mammals that live in small family groups called coteries (consisting of 1 or 2 unrelated adult males, a group of related females, and their young; Hoogland 2006), with adjacent coteries forming a colony. In contrast to many mammalian species that exhibit natal dispersal, Utah prairie dogs rarely leave their natal coteries unless nearly all of the individuals in a coterie are gone (Hoogland 2013), for example following an outbreak of plague. Any such dispersal that does occur is likely male-biased and to nearby coteries, often within the same colony (Hoogland 2013).

Translocations are a common practice in the management of prairie dogs (United States Fish and Wildlife Service 2012). This is an effective tool to combat the loss of genetic variation experienced through a history of population bottlenecks, eradication campaigns and disease, and limited natural gene flow. However, a large degree of variation in habitat (in terms of land cover, elevation, and climate) exists across the Utah prairie dog species range. Combined with a lack of natural gene flow, this habitat heterogeneity could encourage local adaptation (Blanquart et al. 2013). Under this scenario, translocating individuals from different areas of the species range could introduce maladaptive genes to other colonies, leading to outbreeding depression. Even though fears of outbreeding depression may be inflated (Frankham et al. 2011, Ralls et al. 2018), translocations between locally adapted populations could still have potentially disastrous consequences for the evolution of plague resistance in prairie dogs. Resistance to plague has been shown in populations of both black-tailed prairie dogs (Rocke et al. 2012) and Gunnison's prairie dogs in areas that have a history of plague exposure (Busch et al. 2013), which suggests that prairie dogs are under high selective pressure and can become locally adapted to their environment. Further, differences in climate across prairie dog colonies can change flea densities (Eads and Hoogland 2017), thus impacting prairie dog exposure to plague and their ability to build resistance. Utah prairie dogs, like other prairie dog species, are affected by plague. Although plague resistance has not been detected in this species, they may have the potential to become resistant through strong selective pressures. Thus, translocating plague-naïve individuals to areas where plague-resistant individuals exist could lead to a less resilient population due to a swamping effect of plague-resistant genes. By characterizing not only neutral genetic markers,

but also those under selection, we may avoid stifling the evolutionary potential of Utah prairie dogs.

The objective of this study was to use a population genomics approach to investigate the impact of gene flow, genetic drift, and divergent selection on the maintenance of genetic variation in Utah prairie dogs across their species range. To accomplish this objective, we carried out five aims. First, we characterized patterns of genetic structure and gene flow among Utah prairie dog populations. Second, we evaluated the impact of sex-biased dispersal on the maintenance of genetic variation by identifying differences in patterns of genetic structure and gene flow for females and males separately. Third, we characterized the impact of genetic drift on the erosion of genetic variation and genetic differentiation. Fourth, we identified loci under divergent selection and compared patterns of population divergence at these loci against neutral loci. Fifth, we examined how the environment might influence local adaptation by identifying genotype-by-environment associations (GEAs). Our genome-wide approach allows us to harness information in both neutral and adaptive loci to tailor conservation activities to maximize the success of recovery efforts without incurring the potentially substantial costs that could result from translocating locally adapted individuals.

Methods

Generating the SNP Dataset

We trapped Utah prairie dogs and collected hair and whiskers during a field trial of a sylvatic plague vaccine (Rocke et al. 2017). Samples were collected from paired sites, with adjacent coteries located in close proximity (0.25 – 8km). Individuals were sampled throughout the Utah

prairie dog range at three paired sites near Cedar City, UT (CC) and four paired high-elevation sites within the Awapa Plateau (HE) (Fig. 3.1). DNA was extracted from the hair and whiskers using the Zymo universal spin column-based tissue extraction kit (Zymo Scientific) following the manufacturer's feather and hair follicle protocol. We sequenced 4–43 individuals (mean = 22.86 individuals, total=320) per site (Table 3.1).

To generate single nucleotide polymorphisms (SNPs), samples with greater than 300 ng of total genomic DNA, quantified using a Qubit 2.0 Fluorometer (Invitrogen), were used for double digest RAD sequencing (ddRAD) (Peterson et al. 2012). Genomic DNA was digested using the restriction enzymes *HindIII* and *NlaIII*, barcoded, and size selected for 300–600 bp fragments using a Pippin Prep (Sage Sciences). Fragments were paired-end sequenced on an Ilumina HiSeq at Texas A&M AgriLife Genomics. We aligned sequences to a Gunnison's prairie dog (*Cynomys gunnisoni*) genome (Sackett pers comm) using the BWA short-read aligner with default parameters and the MEM alignment algorithm (Li and Durbin 2009). Contigs were assembled using the program STACKS v.1.48 software (Catchen et al. 2011,2013), following the proposed workflow outlined by Rochette and Catchen (2017). Preliminary sequence alignment revealed high sequencing error in 83 individuals, including all individuals found in sites HE2B and HE3A. We removed these individuals from alignment and subsequent analyses.

After calling SNPs, several additional quality control measures were taken. First, in cases where more than 1 SNP per contig was present, only the first (most 5') SNP was used. Second, only loci represented in 80% or more of individuals were retained. Third, only loci present in all 12 sampling locations were retained. Fourth, individuals missing greater than 30% of data (n=4)

were removed (calculated using VCFtools v.0.1.16; Danecek et al. 2011). Fifth, since low-frequency alleles may represent PCR errors, we removed loci with minor allele frequencies <0.05. Sixth, we removed potential paralogs by excluding loci with an observed heterozygosity exceeding 0.7 using the populations module of STACKS, and loci with a depth of coverage greater than twice the mode of the depth of coverage for each locus using R (Willis et al. 2017; O'Leary et al. 2018). Paralogous loci can skew common downstream analyses for population genomics by artificially inflating levels of heterozygosity (Willis et al. 2017).

Characterizing the patterns of genetic structure, gene flow, and sex-biased dispersal

To visualize genetic divergence, we used a principal component analysis (PCA). We used two methods to determine the number of genetic clusters (k) present in our sampling sites: the Bayesian clustering program STRUCTURE (Pritchard et al. 2000) and a multivariate approach using discriminant analysis of principal components (DAPC) in the R package adegenet (Jombart et al. 2010, Jombart and Ahmed 2011). We ran STRUCTURE with an MCMC burn-in of 100,000 steps followed by 100,000 steps for inference clustering using the admixture model with correlated allele frequencies. For each value of k, we completed 10 replicates. For STRUCTURE, we used the alternative prior for population specific ancestry (α=1) since we had unequal sampling among sites (Wang 2017). In order to accurately resolve the number of genetic clusters (k) using STRUCTURE, we used a combination of the LnP(D) and delta K as outlined in (Janes et al. 2017) calculated using STRUCTURE HARVESTER v.0.6.93 (Earl and vonHoldt 2012). We also adopted a 'hierarchical STRUCTURE analysis' approach where each genetic cluster was analyzed iteratively in a new STRUCTURE run in order to gauge substructure (Vaha et al. 2007). We used the program CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) to assign

individuals to genetic clusters using q values from STRUCTURE. For the DAPC, the analysis was first performed unsupervised (no prior knowledge of groups) using the sequential K-means clustering algorithm executed through the find.clusters function in adegenet (Jombart et al. 2010, Jombart and Ahmed 2011). We then performed the DAPC analysis supervised by using the Bayesian information criterion (BIC) to determine the final value of k.

To determine patterns of gene flow, we calculated the pairwise genetic differentiation between sites using Weir and Cockerham's F_{ST} and Jost's D using the R packages hierfstat (Goudet and Jombart 2015) and mmod (Winter 2012), respectively. We tested for significant population divergence using both F_{ST} and Jost's D with 1,000 random permutations and corrected p-values for multiple comparisons using a false discovery rate (FDR) of 0.05 (conducted in R; R Core Team 2018). To test for sex-biased dispersal, we repeated the above analyses on males and females separately. We tested relative migration rates among sampling sites using N_M with the divMigrate function in the R package diveRsity and identified significant migration rates using 10,000 bootstrap iterations (Keenan et al. 2013; Alcala et al. 2014; Sundqvist et al. 2016). Migration networks were created using the R package qgraph (Epskamp et al. 2012).

Characterizing the Effect of Genetic Drift

Genetic drift erodes standing genetic variation where rare alleles face a greater chance of being lost due to random chance. As populations decrease in size and become isolated, the effects of genetic drift grow in significance, which accelerates the erosion of genetic variation. Thus, populations in which genetic drift is the primary driver of genetic structure are predicted to have 1) less genetic variation, 2) fewer private alleles, and 3) a higher degree of inbreeding than

populations in which genetic structure is shaped by other mechanisms. We characterized genetic variation in each site as well as each genetic cluster by measuring expected heterozygosity (H_e), observed heterozygosity (H_o), allelic richness (A_R), the number of private alleles (A_P), and inbreeding (F_{IS}). Since the effects of genetic drift are stronger in more isolated populations, we compared the average pairwise F_{ST} for each site with H_e, H_o, and A_R using linear regression in R (R Core Team 2018). We calculated H_e, H_o, and A_P using the R package poppr v.2.8.2 (Kamvar et al. 2014) and A_R and F_{IS} using the R package diveRsity, after which 95% confidence intervals were calculated using 1000 bootstrap iterations (Keenan et al. 2013).

We estimated the effective population size (N_e) and tested for evidence of recent genetic bottlenecks for each site and genetic cluster, as these factors significantly impact the rate at which genetic variation is lost as well as the rate of increase of inbreeding and genetic drift (Charlesworth 2009; Banks et al. 2013; Gasca-Pineda et al. 2013). We estimated N_e for each genetic cluster using the linkage disequilibrium (LD) method implemented in the program NeEstimator (Do et al. 2014). One assumption to calculate N_e using the LD method is that linkage disequilibrium at independent loci in randomly mating, closed populations comes exclusively from genetic drift (Hill 1981). To meet this assumption, we created a neutral set of loci by excluding loci potentially under the influence of natural selection; those identified as high F_{ST} outliers by either Bayescan or PCAdapt for this analysis (see next section for details).

We tested each of the 12 prairie dog sampling sites as well as each genetic cluster for evidence of genetic bottlenecks using the program Bottleneck v. 1.2.02 (Cornuet and Luikart 1997; Piry et al. 1999). Like the estimation of N_e, only neutral loci were used to test for evidence of bottlenecks.

We used the infinite alleles model (IAM) and tested for significant heterozygosity excess compared to the level predicted under mutation-drift equilibrium using standardized differences tests.

Neutral Loci vs. Loci Under Selection

Outlier loci were detected via a Bayesian method conducted in Bayescan 2.1 (Foll and Gagiotti 2008) and a non-constrained ordination method executed in the R package PCAdapt (Luu et al 2019). For the Bayescan method of outlier detection, an FDR of 0.05 was used. For the PCAdapt method, outlier loci with p-values less than 0.05 were kept. To generate the set of outlier loci, we retained loci that were identified as under divergent selection in both Bayescan and PCAdapt methods. We compared this outlier dataset to the neutral set of loci to assess the relative contribution of genetic drift and selection in shaping patterns of genetic structure. Specifically, we compared patterns of genetic structure for outlier and neutral loci using the program STRUCTURE and by using DAPC in adegenet (Jombart et al. 2010, Jombart and Ahmed 2011). Gene flow was also estimated for each set of loci using pairwise Jost's D and F_{ST}. To determine if geographic distance alone drives patterns of genetic structure, we tested for isolation by distance (IBD) in the full genetic dataset (neutral and outlier loci), neutral loci only, and outlier loci only using Mantel tests between genetic and geographic distance carried out in the R package adegenet (Jombart et al. 2010, Jombart and Ahmed 2011). Significance was assessed based on 999 replicates.

Selection and Environmental Associations

Genotype-environment associations were characterized using a multivariate ordination method, Redundancy Analysis (RDA) (Forester et al. 2018). For the environmental comparisons, we used land cover data from the National Landcover Database (NLCD; resolution= 30 m × 30 m; Jin et al. 2013) and climatic variables from Worldclim (resolution = 30 arc-seconds; Hijmans et al. 2005). Since prairie dog sampling coordinates represent the center of a larger sampling area, we characterized the most represented land cover types by the proportion of each landcover type within a 5 km buffer around the coordinate point (Fig. S3.1). To avoid multicollinearity, we used Pearson's correlations to remove variables that were correlated with an |r| of 0.7 or higher prior to performing the RDA. Predictor variables in the RDA model were further pruned based on their variance inflation factors (VIF; multicollinearity was assumed if VIF > 4). To determine if associations exist between environmental variables and the environment, we converted the full SNP dataset to allele frequencies. Using the RDA, we identified outlier loci and the environmental variables most associated with those loci (Forester et al. 2018, Capblanco et al. 2018). We tested each RDA as a full model to quantify how well it explained genetic variation with 999 permutations. We then used an analysis of variance (ANOVA) to see if some axes performed better than others in terms of explaining genetic variation.

Results

Generating the SNP Dataset

After initial filtering steps, a total of 3,549 variable SNP loci were retained. An additional 594 loci were removed due to exceptionally high depth of coverage, suggesting potential paralogs (2X the mode of depth of each locus averaged for all individuals; mode=17.55). The final genomic dataset contained 2,955 variable SNP loci with a mean depth of coverage of 20.08

(averaged across all individuals and ranged from 7.68 to 35.58; Fig. S3.2). Four individuals with a high amount of missing data (>30%) were removed from the dataset, leaving a total of 233 individuals used for subsequent analyses (4–40 individuals per site, mean=19.42). These 233 individuals used for analyses had an average of 4.2% missing data (Fig. S3.3).

Characterizing the patterns of genetic structure, gene flow, and sex-biased dispersal

Differentiation among sampling sites was observed using PCA; however, paired A and B sites were indistinguishable from one another (Fig. 3.2). The most supported number of genetic clusters (K) was two, using both the program STRUCTURE (based on Ln(K) and the ΔK method) and the unsupervised clustering method using DAPC (Figs. S3.4 and S3.5). In the two clusters, all individuals from the CC sites grouped to make one genetic cluster and all individuals from HE sites grouped to make the second genetic cluster. When using a supervised clustering approach for DAPC, additional clustering solutions (K=2–4) were informative for describing genetic structure based on BIC (Fig. S3.6). For example, under a K of four for DAPC, sites from CC3 formed a separate genetic cluster and an additional genetic cluster included all individuals from HE2A and 11 individuals from HE4 A and B (Fig. S3.7). This pattern was also observed using the hierarchical clustering approach in STRUCTURE (Fig. S3.5).

We found similar patterns of population differentiation with pairwise values of F_{ST} and Jost's D (Table S3.1), so only values of F_{ST} are reported. Values of pairwise F_{ST} were low among paired A and B sites, suggesting little to no population differentiation, but we detected significant genetic differentiation among sampling locations (Fig. 3.3, Table S3.1). Particularly, a high degree of differentiation was observed between the CC sites and the HE sites. Among those CC

to HE comparisons, the lowest F_{ST} values were between the CC3 sites and the HE1 sites. Further, the CC3 sites showed high values of F_{ST} to CC1 and CC2 sites, compared to other CC-CC or HE-HE comparisons.

The relative migration analyses conferred with our estimates of population differentiation using pairwise F_{ST} in that we observed high migration rates (N_M) among paired A and B sites. We also observed high migration rates among CC1 and CC2 sites as well as among the HE sites (Fig. 3.4). The CC3 sites showed intermediate levels of N_M among both other CC sites as well as the HE sites. Overall, females and males showed comparable patterns of population differentiation (Fig. 3.3). We observed global F_{ST} values of 0.43 and 0.40 for males and females, respectively.

Characterizing the Effect of Genetic Drift

Using all 2,955 variable SNP loci, we determined the level of genomic diversity for all 12 sampling sites within CC and HE as well as each genetic cluster for K=2. Overall, levels of genomic diversity (as measured by H_O, H_E, and A_R) were higher in HE sites compared to CC sites (Table 3.1). Observed and expected heterozygosity were greatest in the HE sites, particularly in the HE 1 sites (CC average H_O= 0.12, H_E=0.10; HE average H_O= 0.27, H_E=0.26; Table 3.1). The highest levels of A_R were observed in the HE1 sites (A_R for HE1A=1.71 and HE1B= 1.70; Table 3.1). F_{IS} values ranged from −0.28 to 0.19 for CC sites and from −0.15 to 0.11 for HE sites (Table3.1). All sites had significant, negative F_{IS} values except the HE4 sites (F_{IS} for HE4A=0.11 and HE4B=0.04; Table 3.1), a deficit of heterozygotes that suggests individuals in these subpopulations are more related than expected. No alleles were unique to any one site; however, a large number of private alleles were detected when individuals were

grouped into genetic clusters (Table 3.1). The genetic group that contained individuals from the HE sites had the largest number of private alleles (A_P=1,326). This genetic cluster also had the highest amount of genomic variation in terms of H_O, H_E, and A_R (Table 3.1). We found a negative relationship between F_{ST} and measures of genomic variation, indicating that more divergent populations also have reduced genomic variation (Fig. 3.5).

Estimates of N_e ranged widely across sites (1.0–48.4) and had large confidence intervals (Table 3.1). We found that site CC3A had the largest effective population size at 48.4 individuals; however, confidence intervals went to infinity (Table 3.1). When genetic clusters were used to estimate N_e, values were much smaller (CC=2.1 and HE =4.1) and confidence intervals were tighter. We also detected significant bottlenecks (p-value < 0.000001) for all sites as well as for each genetic cluster.

Neutral Loci vs. Loci Under Selection

Using Bayescan, 303 loci were identified as under divergent selection compared to 531 loci identified as under selection using PCAdapt (Fig. S3.8). For our outlier dataset, only loci that were identified as potentially under divergent selection in both Bayescan and PCAdapt were used (n_{loci}=51). We found significant patterns of isolation by distance using the full SNP dataset (n_{loci}=2955; r=0.85; p=0.002), only neutral loci (n_{loci}=2904; r=0.84; p=0.002), and only outlier loci (n_{loci}=51; r=0.83; p=0.001) based on 999 replicates (Fig. 3.6).

The neutral loci were most similar to that of the full set, suggesting that neutral loci drive the overall pattern of population divergence (Fig. 3.2). The outlier loci explained the greatest total

variance in genetic structure compared to either the full set of loci or neutral loci. Overall, outlier loci depicted a high degree of divergence among CC and HE sites but less divergence within either the HE or CC sites.

Selection and Environmental Associations

We found that a large proportion of the climatic variables were highly correlated ($|r| > 0.70$) to elevation. Therefore, we conducted two different RDAs, one with landcover and elevation data, and another with only climatic data. For the first RDA using landcover and elevation data, cropland was removed as a predictor variable due to low variation across sites. For the second RDA with climate data, we reduced the 19 predictor variables to 2 – one temperature variable (BIO9- Mean temperature of the driest quarter) and one precipitation variable (BIO17- Precipitation of the driest quarter). For our landcover and elevation RDA, we found outlier loci associated with elevation (n=32), forests (n=42), and shrubland (n=20) (Fig. 3.7). Our full landcover and elevation model was significant when compared to the SNP data ($p= 0.002$). However, when we compared how well each component of the RDA explained genetic variation using an ANOVA, we found that only the first component was significant ($p=0.001$). For our RDA with climate data, the full model explained genetic variation well ($p=0.003$). However, the ANOVA analyses showed that only the first component (RDA1) was significant ($p=0.002$). We found SNPs associated with BIO9 (n=34) and BIO17 (n=4) (Fig. 3.7). All predictors within their respective RDA models had VIF <2. The two separate RDA models shared 34 outlier loci in common. These loci were associated with only one climate variable (BIO9) and both shrubland (n=12) and forests (n=22). However, higher correlations were observed with the BIO9 variable

compared to correlations with either shrubland or forests. No outlier loci had overlapping correlations with elevation or BIO17.

Discussion

Through the use of high-resolution, genome-wide markers we have demonstrated that we can characterize patterns of genetic structure and gene flow, elucidate demographic processes, compare the effects of drift and selection on population divergence, and identify genotype-by-environment associations in a genetically depauperate species. This information is invaluable in tailoring the conservation policy of threatened species to mitigate the loss of genetic variation and maintain evolutionary potential. We found that Utah prairie dogs show high genetic structure, limited gene flow, and a lack of genetic variation across their range. This pattern can be explained sufficiently well by demographic processes– population isolation and genetic drift. These processes seem to have had a greater impact on the CC sites, where genetic variation was lower than in the HE sites. However, selection also plays a role in the divergence of Utah prairie dog populations, primarily between low-elevation (CC) sites and high-elevation (HE) sites.

Limited genetic variation in Utah prairie dogs puts them at risk for further erosion of variation through genetic drift. Although Utah prairie dog population sizes have either stabilized or increased within their conservation units (United States Fish and Wildlife Service 2012), these positive demographic trends contrast with our observations of limited gene flow, low effective population size, and recent genetic bottlenecks. Utah prairie dogs are highly social and have been shown to rarely leave natal colonies (Hoogland 2013). This was reflected in both Jost's D and F_{ST} metrics. Higher rates of migration between paired A and B sites than between more distant

sites, combined with strong evidence of isolation-by-distance, suggest that when prairie dogs do disperse, it is likely to nearby colonies and that long-distance dispersal is rare or non-existent.

We did not detect evidence of sex-biased dispersal, which we expected based on the observations in Hoogland (2013). This could be explained by a number of factors. First, young males are likely to remain in their natal coterie until they mature. If a large proportion of the males we sampled were juveniles, then our estimates of male dispersal could be unrealistically low. Male prairie dogs also actively defend their coterie (and females) and have even been found to occasionally engage in infanticide and cannibalism (Hoogland 2007). This means it is possible that males disperse as observed by Hoogland (2013), but individuals that disperse experience high mortality by either other prairie dogs or by predators before they pass on their genes.

The effective population size (N_e) remains an important metric for characterizing the long-term evolutionary potential of a species, as it describes the ideal population size that will result in the same amount of genetic drift as the census population size (Jamieson and Allendorf 2012). Although confidence intervals for N_e values were large, it is clear that the N_e values for Utah prairie dog populations are remarkably low. Small N_e was also reported in another study of Utah prairie dog using different markers and methodology (Brown et al. 2016). Drift acts more rapidly in small vs. large populations (Lacy 1987), which means that without gene flow we should expect genetic diversity to decline quickly in Utah prairie dog populations, despite positive demographic trends. Facilitated gene flow (i.e., translocation) from populations with unique alleles and high genetic variation would increase N_e of small populations and slow drift-based erosion of diversity (Laikre et al. 2016). However, this strategy should be pursued with caution,

as maladapted individuals introduced to new environments could reduce population fitness, as shown by the low establishment success of poorly adapted seeds (Kulpa and Leger 2013) and reduced hatching success in non-native habitat in pike (Berggren et al. 2016).

Whereas restricted gene flow leaves genetic drift and inbreeding unmitigated, it also facilitates adaptation to local environments (Slatkin 1987; Barrett and Schluter 2008; Savolainen et al. 2013). A large proportion of our SNP loci in our study were identified as under divergent selection (10% using Bayescan and 18% using PCAdapt). These outlier loci reflected high divergence between high-elevation (HE) sites and low-elevation (CC) sites. Specifically, patterns of variation at outlier loci were associated with high elevation environments, a greater proportion of shrubland and forests, and a higher amount of precipitation in the HE sites. These genotype-by-environment associations may be important in terms of adaptation to plague as precipitation affects flea prevalence (Eads and Hoogland 2017) and temperature affects a flea's ability to transmit plague (Williams et al. 2013). Further, the HE sites harbor a greater diversity of flea and small mammal species compared to CC sites (Bron et al. 2018; Russell et al. 2018). This variation in species composition likely plays a role in plague dynamics. Differences in habitat and community composition between HE and CC sites suggest that local adaptation could occur, and our outlier and genotype-by-environment association analyses suggest that adaptation is ongoing. Experimental work such as common garden or reciprocal transplant experiments would help to validate the true connection between the genotypes and their associated environmental variables. However, in Utah prairie dogs, any local adaptation is likely diminished by stochastic processes such as genetic drift (Hereford 2009), suggesting that the benefit of translocations for overall genetic variation may outweigh the risk of incurring outbreeding depression.

Implications for Conservation

With continued change in land use and climate, maintaining genetic variation is of paramount importance for the survival of threatened species. Greater standing genetic variation allows for selection to act on a larger pool of phenotypic traits. Species that are most at risk of losing genetic variation and exhibiting reduced viability in the face of a changing environment are those that have 1) small, isolated populations, 2) a history of bottlenecks/founder events, 3) low effective population size, and 4) limited gene flow among populations. Our range-wide study on Utah prairie dogs, as well as another by Brown et al (2016), showed that Utah prairie dogs have all of the above characteristics and already exhibit limited genomic variation. We also observed signatures of selection via local adaptation that is contributing to divergence among CC and HE sites. Local adaptation reduces genome-wide variation, but also improves a population's ability to survive in their current environment.

Translocations are currently used for prairie dogs, with considerable research focused on improving success (Curtis 2014; Truett et al. 2001; United States Fish and Wildlife Service 2012). Translocations initiate gene flow and increase the genetic variation of a population (Whiteley et al. 2015) and have been successful in other threatened species (Johnson et al. 2010; Bateson et al. 2014). However, translocations could exacerbate population decline in threatened species by introducing maladaptive genes into populations that are locally adapted (Weeks et al. 2011; Frankham 2015). Incorporating selection into translocation strategies could help to mitigate the risks associated with translocations (Harrison et al. 2017; Flanagan et al. 2018). With low levels of species-wide genetic variation and limited gene flow in Utah prairie dogs, the

consequences of genetic drift may outweigh concerns of outbreeding depression from moving potentially locally adapted individuals (Ralls et al. 2018). Specifically, translocations from HE sites to CC sites may serve to increase genetic variation in CC sites, and, provided HE individuals could successfully integrate into CC sites, would not disrupt the local adaptation we identified in HE sites.

Figure 3.1. Map of Utah prairie dog sampling locations across the species range (shown in green on the Utah state inset map). Prairie dogs were sampled from 6 paired sites near Cedar City, UT (CC) and 8 paired, high-elevation sites on the Awapa Plateau (HE).

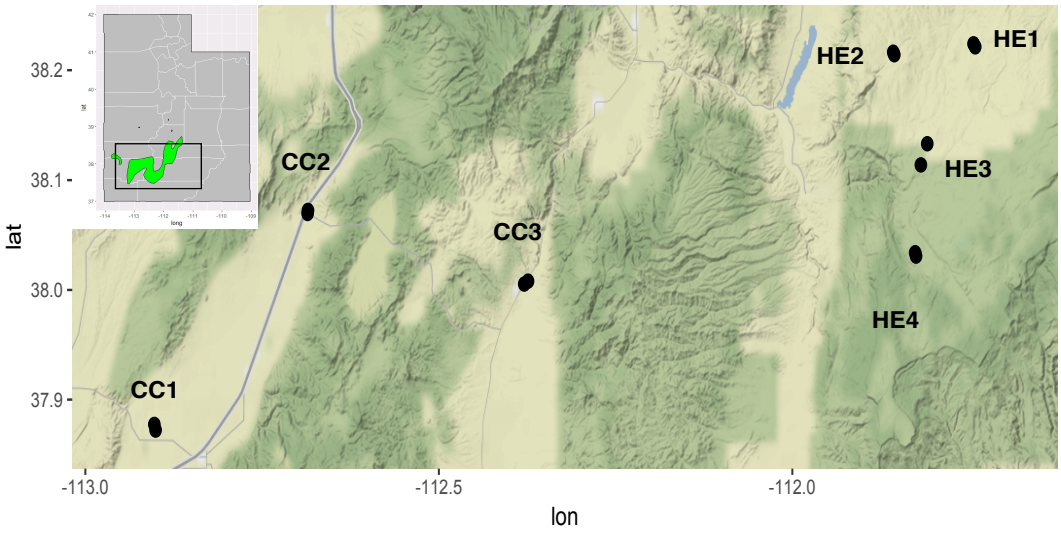


Figure 3.2. Principal component analysis (PCA) to characterize genetic differentiation among Utah prairie dogs using all SNP loci (Full), only SNP loci identified as outliers in both Bayescan and PCAdapt (Outlier), and only neutral SNP loci (Neutral). Colors correspond to site (CC1, CC2, CC3, HE1, HE2, HE3, and HE4) while symbols distinguish paired sites (A and B).

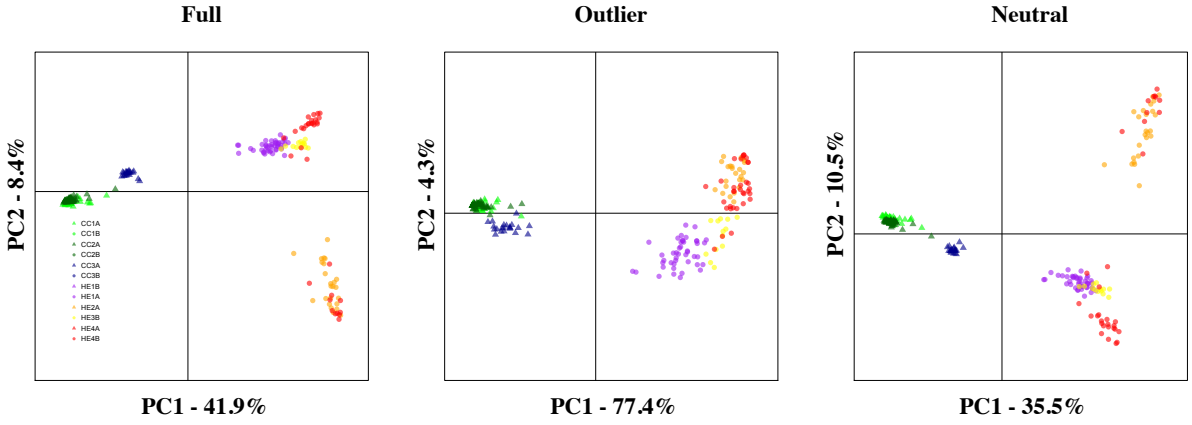


Figure 3.3. Population differentiation based on pairwise F_{ST} across Utah prairie dog sites for a) all individuals; b) only males; and c) only females. Warmer colors indicate lower F_{ST} (less differentiation) while cooler colors indicate higher F_{ST} (more differentiation).

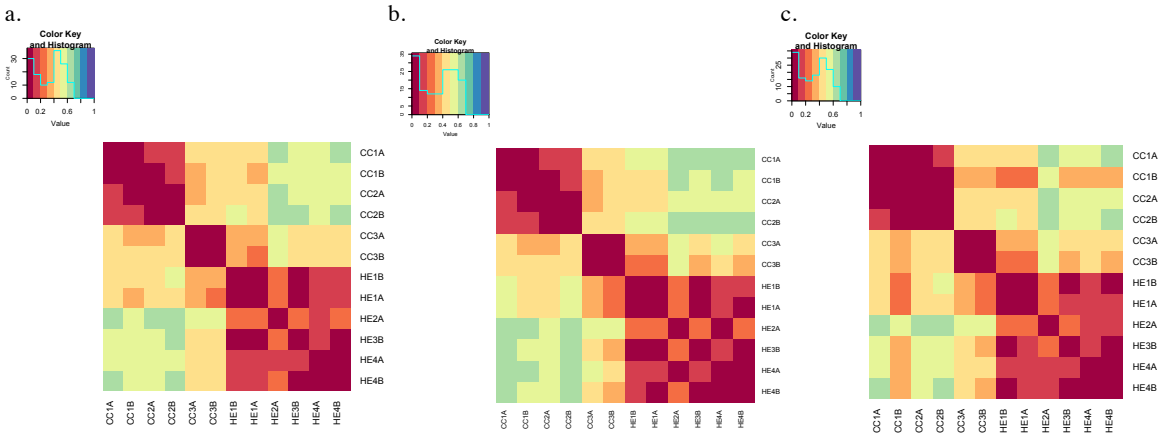


Figure 3.4. The relative rate of migration (N_M) among sampling sites. Arrows indicate direction and relative rates among sites with darker blue indicating higher migration rate than lighter blue.

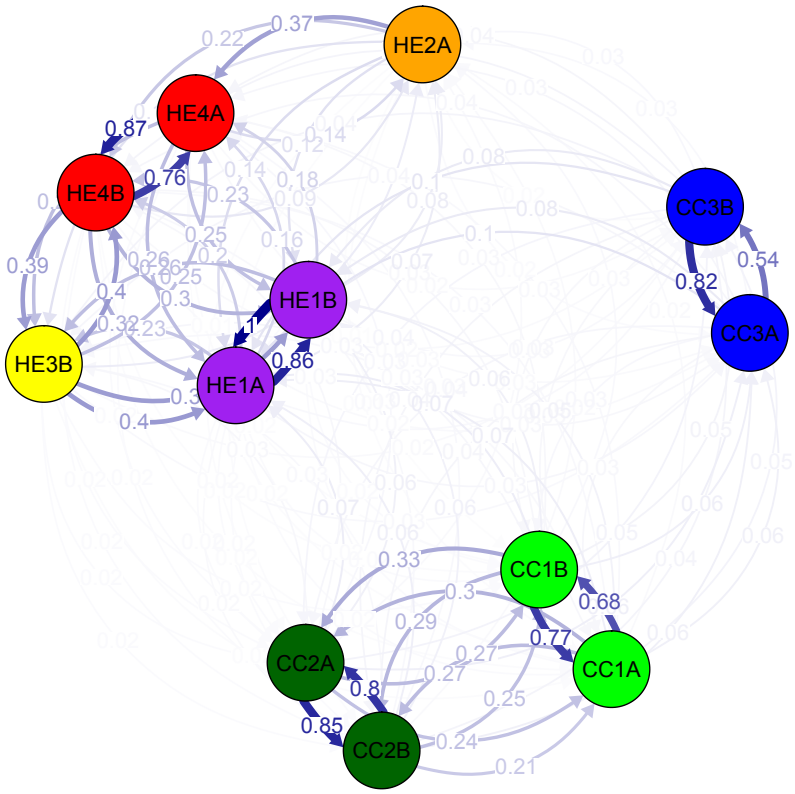


Figure 3.5. Regression plots relating population divergence (average pairwise F_{ST}) to genetic variation [observed heterozygosity (H_O), expected heterozygosity (H_E), and allelic richness (A_R)] in Utah prairie dogs.

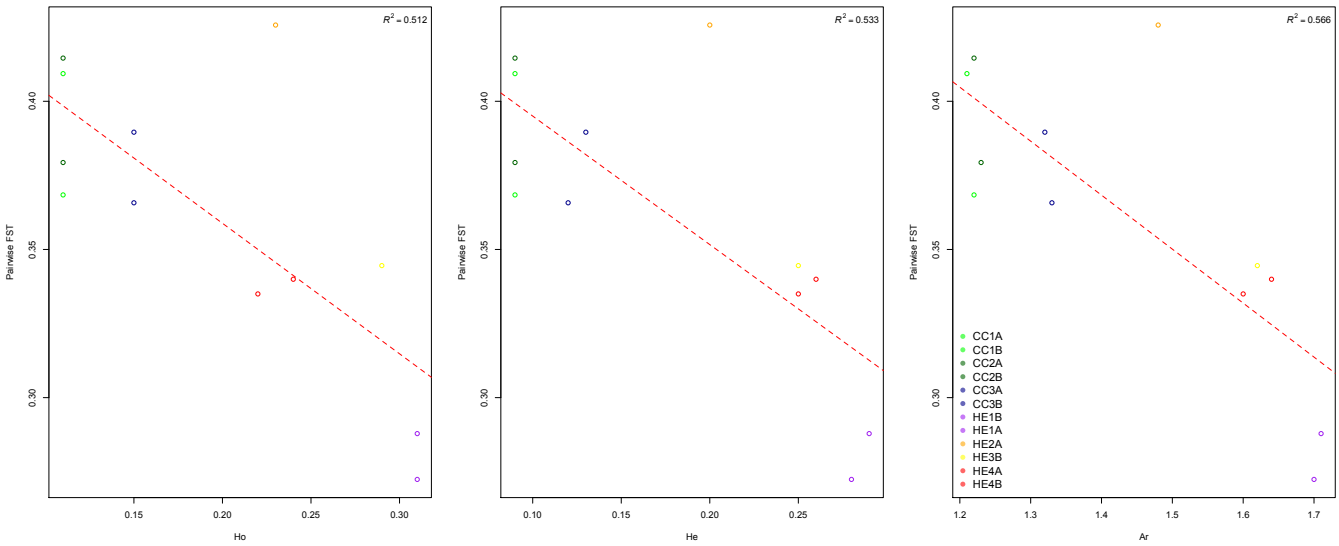


Figure 3.6. Correlation between genetic and geographic distance for a) all SNP loci; b) only neutral SNP loci; and c) only SNP loci identified as outliers.

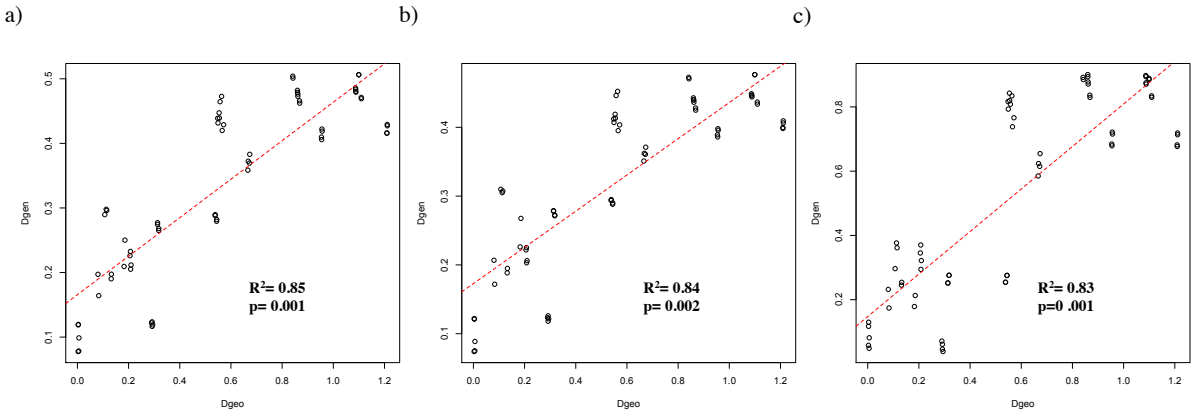
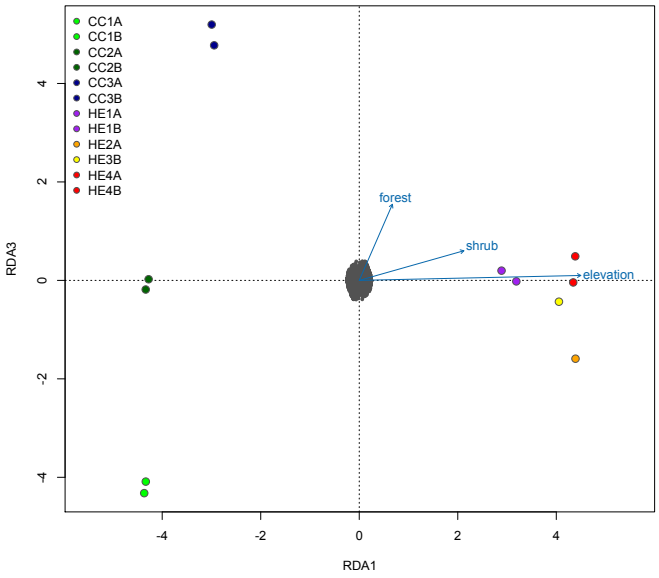


Figure 3.7. Redundancy Analysis (RDA) showing environmental associations with outlier loci. A) An RDA using only landcover data and elevation found outlier loci associated with elevation (n=32), forests (n=42), and shrubland (n=20). B) An RDA using climatic variables identified outlier loci associated with the mean temperature of the driest quarter (BIO9; n=34) and precipitation of the driest quarter (BIO17; n=4).

a)



b)

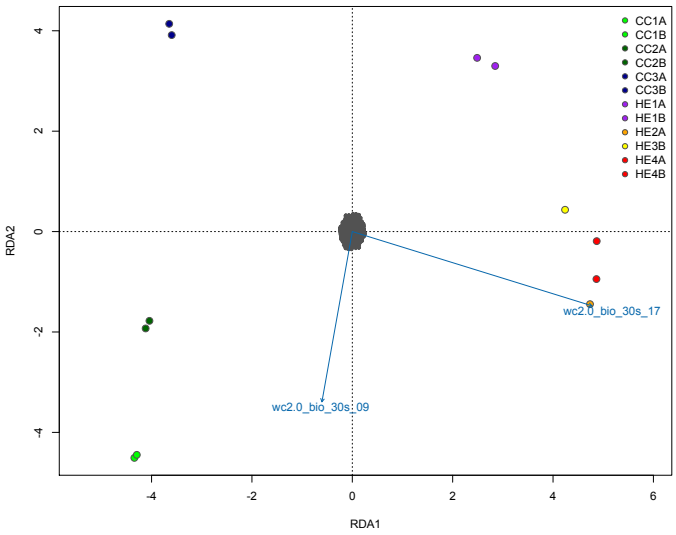


Table 3.1. Measures of genetic variability in Utah prairie dogs for a) each of the sampled sites and b) the two putative populations identified using both STRUCTURE v. and DAPC based on 2,955 variable SNP loci.

	N	*Nm*	*Nf*	*Ho*	*He*	*Ar*	*Ap*	*Fis (95%CI)*	*Ne (95%CI)*
a) Sampled Sites									
CC1A	28	17	11	0.11	0.09	1.21	0	-0.24 (-0.26 - -0.21)	25.4 (16.7 - 44.4)
CC1B	12	11	1	0.11	0.09	1.22	0	-0.27 (-0.34 - -0.21)	8.6 (2.7 - 44.7)
CC2A	18	8	10	0.11	0.09	1.23	0	-0.24 (-0.28 - -0.21)	23.1 (15.6 - 39.3)
CC2B	40	21	20	0.11	0.09	1.22	0	-0.24 (-0.26 - -0.22)	16.7 (9.2 - 32.6)
CC3A	17	12	5	0.15	0.13	1.32	0	-0.19 (-0.22 - -0.16)	48.4 (20.2 - infinity)
CC3B	4	2	2	0.15	0.12	1.33	0	-0.28 (-0.47 - -0.14)	infinity
HE1A	24	9	12	0.31	0.29	1.71	0	-0.09 (-0.13 - -0.06)	12.7 (6.2 - 30.3)
HE1B	20	11	9	0.31	0.28	1.70	0	-0.11 (-0.15 - -0.07)	10.6 (6.9 - 16.9)
HE2A	24	15	9	0.23	0.20	1.48	0	-0.14 (-0.19 - -0.09)	6.3 (3.0 - 11.4)
HE3B	13	6	8	0.29	0.25	1.62	0	-0.15 (-0.21 - -0.10)	24.2 (12.5 -91.6)
HE4A	22	7	15	0.22	0.25	1.60	0	0.11 (0.02 - 0.17)	1.0 (0.8 - 1.3)
HE4B	11	6	5	0.24	0.26	1.64	0	0.04 (-0.15 - 0.16)	2.2 (1.3 - 7.3)
b) Putative Populations									
CC	119	71	49	0.12	0.12	1.54	111	0.04 (-0.01 - 0.08)	2.1 (1.9 - 2.3)
HE	114	54	58	0.27	0.30	1.96	1,326	0.10 (0.07 - 0.12)	4.1 (3.5 - 6.8)

N=sample size; Nm= number of males sampled; Nf= number of females sampled; Ho= observed heterozygosity; He= expected heterozygosity; Ar= allelic richness; Ap= number of private alleles; Fis= inbreeding with 95% confidence intervals (95%CI); and Ne= estimated effective population size with 95% confidence intervals (95%CI).

Figure S3.1. Landcover for each sampling location was characterized using the proportion of each landcover type within a 5km buffer area around the center of the sampling area.

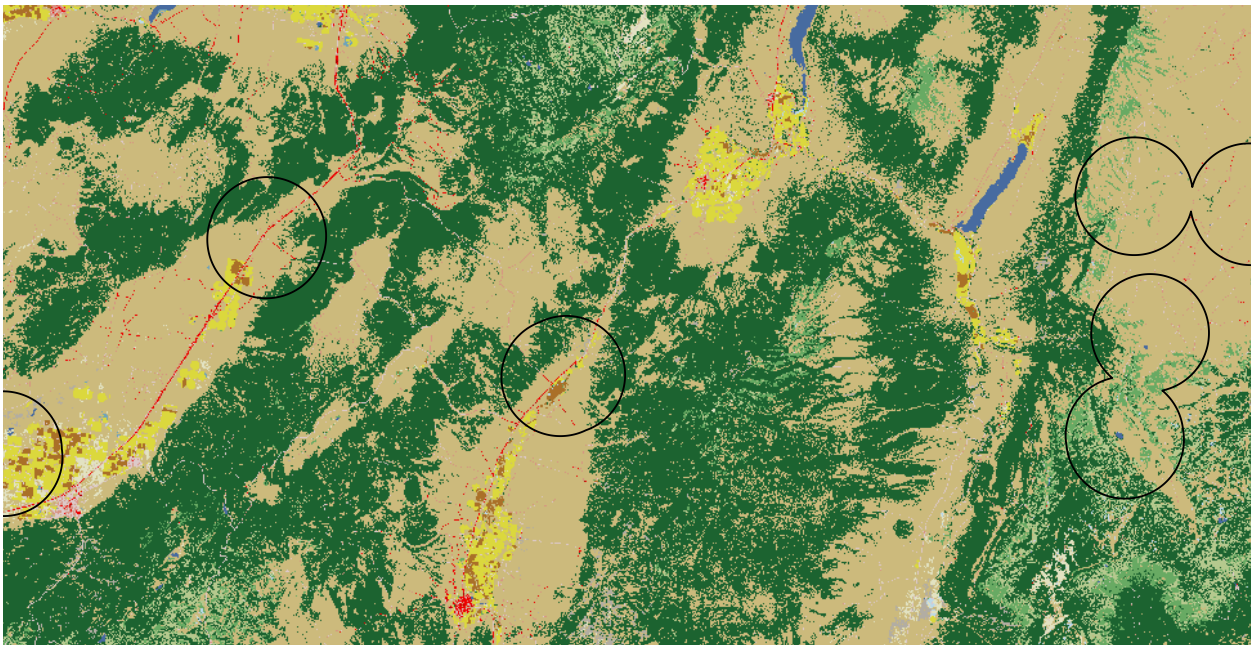


Figure S3.2. The mean depth of coverage for the 2,955 SNP loci used in this study was 20.08 (averaged across all individuals) and ranged from 7.68-35.58 after filtering high depth loci (based on 2X the mode of depth of each locus averaged for all individuals; mode=17.55).

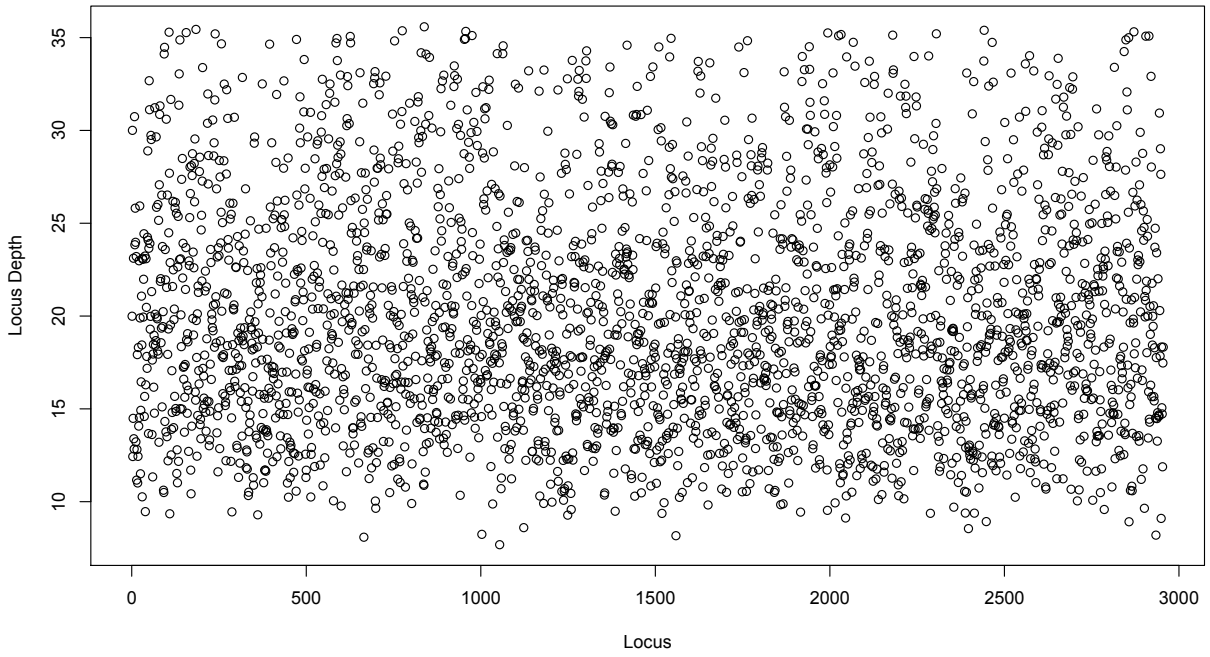


Figure S3.3. Average number of missing loci per individual after filtering individuals with a high amount of missing data (>30%). The minimum amount of missing data for an individual was 0.00% while the maximum was 26.16% (mean=4.20%).

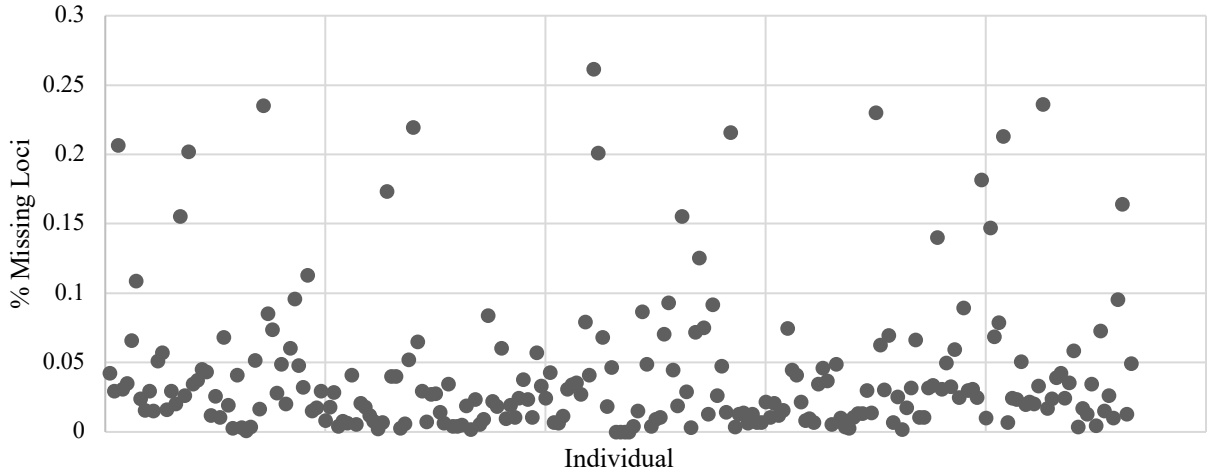


Figure S3.4. A K=2 was best supported in STRUCTURE based on the a) ΔK and b) Ln(K) method.

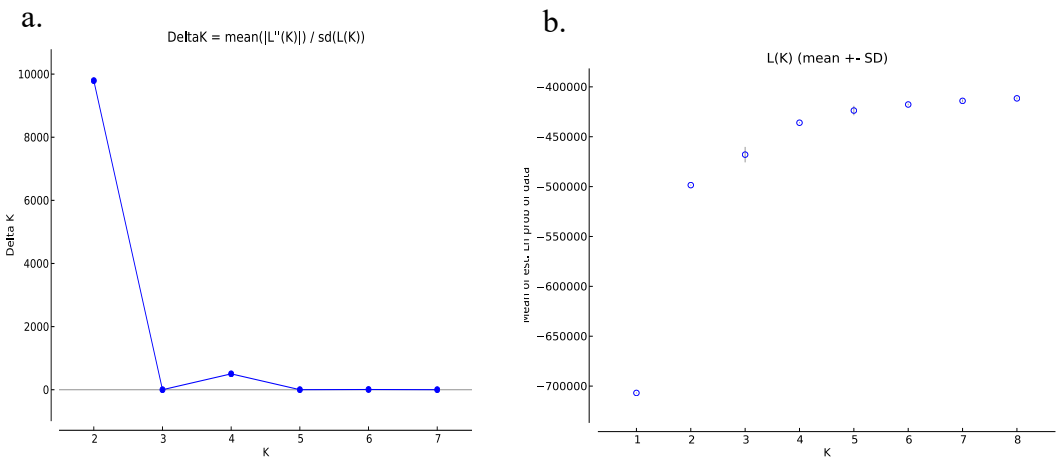


Figure S3.5. STRUCTURE plots for a) full run for K=2 (most supported based on Ln(K) and the deltaK method) b) hierarchical results for the CC and HE genetic clusters (K=2 for each) c) females only (K=2) d) males only (K=2) e) neutral loci only (K=2) and f) outlier loci only (K=2).

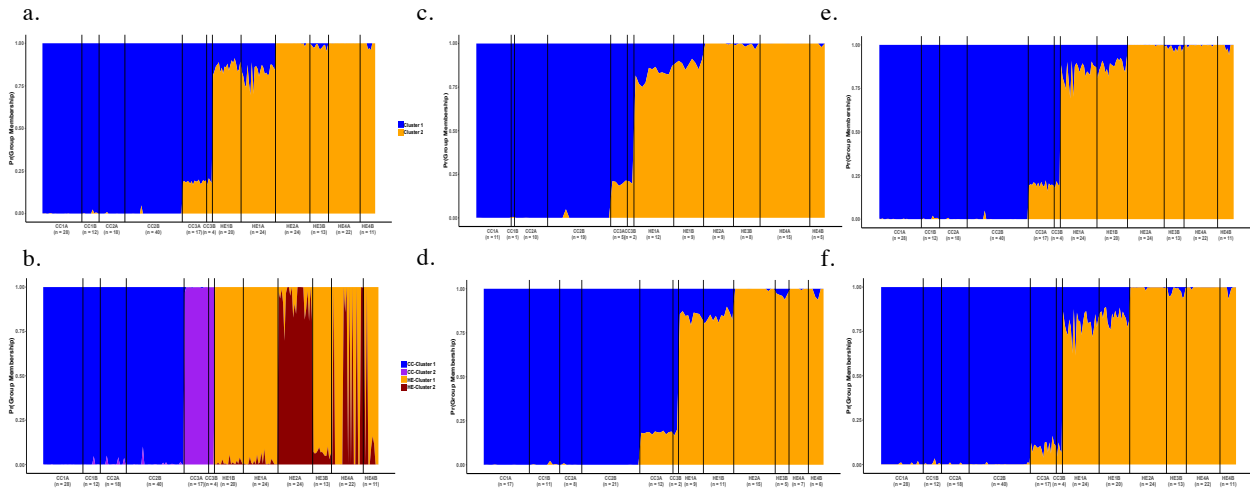
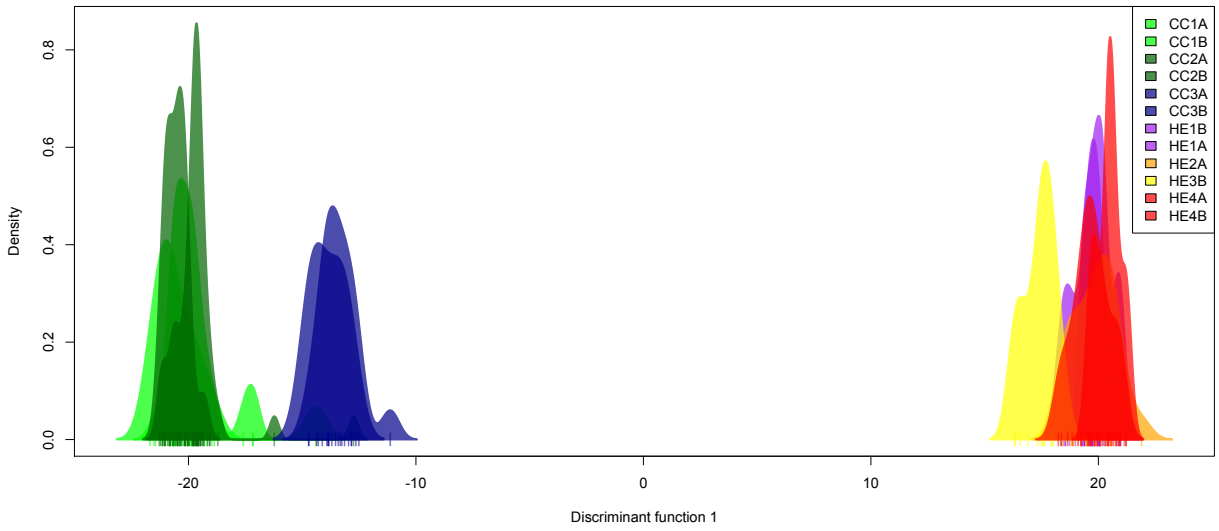


Figure S3.6. Discriminant analysis of principal components (DAPC) showing genetic structure in Utah prairie dogs. Solutions dividing samples into 2-4 genetic clusters were informative (a) based on the Bayesian information criterion (BIC) (b).

a.



b.

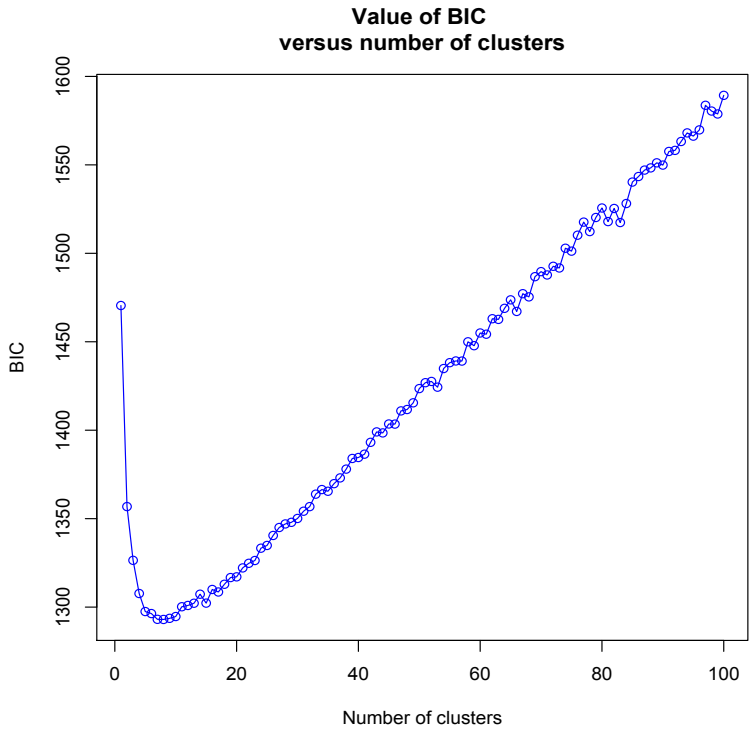
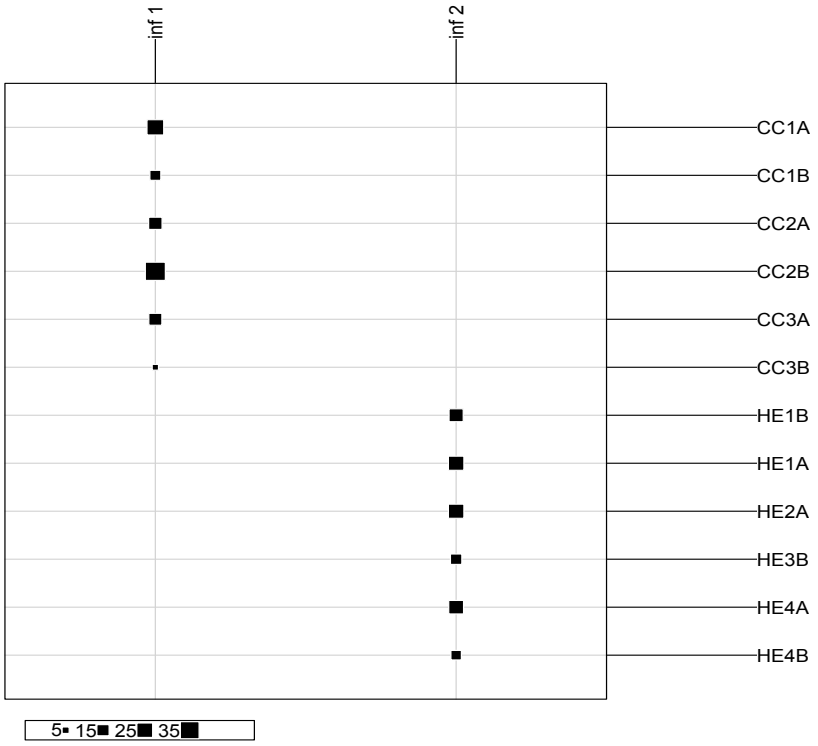


Figure S3.7. Sites from the Cedar City (CC) area made up one genetic cluster and the high elevation site (HE) made up another with K=2 in the DAPC analysis (a). When K=4, we saw CC3 sites form their own cluster and some further differentiation among the HE sites (b).

a.



b.

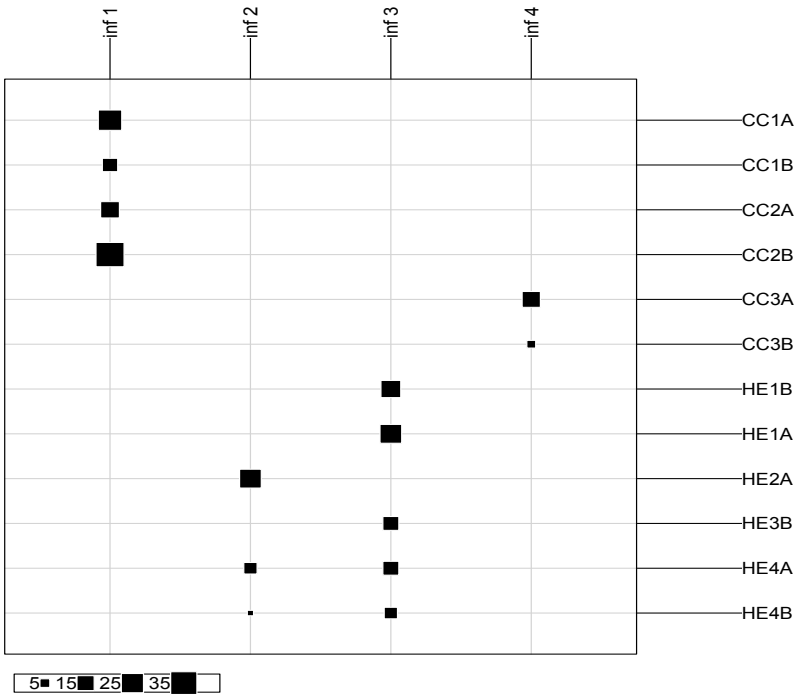


Figure S3.8. Two approaches were used to identify outlier loci. A Bayesian-based program (Bayescan; (a) showing the posterior odds (PO) compared to F_{ST} of each locus with a FDR threshold of 0.05) and an ordination-based method (PCAdapt; (b) Q-Q plot showing the distribution of observed p-values compared to an expected uniform distribution of p-values (dark line) and (c) Manhattan plot showing the p-values for each locus).

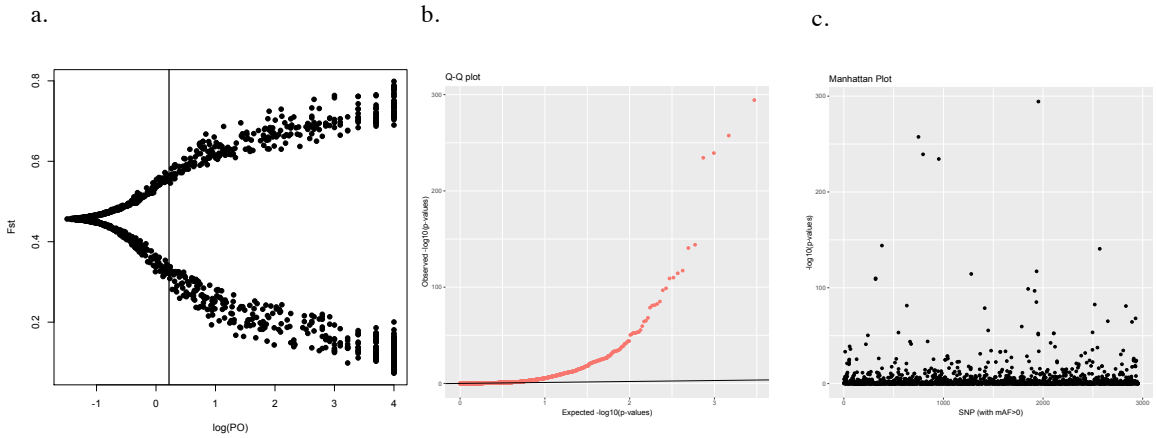


Table S3.1. Pairwise estimates of F_{ST} (below the diagonal) and Jost's D (above the diagonal) for all pairs of sampling sites of Utah prairie dog. Bolded and italicized numbers indicate measures that were statistically significant with a false discovery rate (FDR) correction using 999 permutations.

	CC1A	CC1B	CC2A	CC2B	CC3A	CC3B	HE1B	HE1A	HE2A	HE3B	HE4A	HE4B
CC1A	-	0.00	0.01	0.01	***0.08***	***0.08***	***0.20***	***0.19***	***0.29***	***0.25***	***0.26***	***0.26***
CC1B	0.03	-	0.01	0.01	***0.08***	***0.08***	***0.20***	***0.18***	***0.29***	***0.24***	***0.26***	***0.25***
CC2A	***0.10***	***0.09***	-	0.00	***0.08***	***0.07***	***0.20***	***0.18***	***0.29***	***0.24***	***0.26***	***0.25***
CC2B	***0.12***	***0.10***	0.04	-	***0.08***	***0.08***	***0.20***	***0.18***	***0.29***	***0.24***	***0.26***	***0.25***
CC3A	***0.42***	***0.39***	***0.39***	***0.41***	-	0.00	***0.16***	***0.15***	***0.25***	***0.20***	***0.22***	***0.22***
CC3B	***0.45***	***0.42***	***0.41***	***0.43***	0.00	-	***0.15***	***0.14***	***0.25***	***0.20***	***0.22***	***0.21***
HE1B	***0.49***	***0.43***	***0.45***	***0.51***	***0.37***	***0.31***	-	0.01	***0.11***	***0.04***	***0.06***	0.05
HE1A	***0.46***	***0.40***	***0.42***	***0.47***	***0.34***	***0.28***	0.03	-	***0.11***	***0.04***	***0.07***	0.05
HE2A	***0.64***	***0.60***	***0.61***	***0.65***	***0.55***	***0.51***	***0.25***	***0.25***	-	***0.10***	***0.05***	***0.07***
HE3B	***0.60***	***0.53***	***0.56***	***0.61***	***0.47***	***0.40***	***0.09***	***0.10***	***0.27***	-	***0.04***	0.02
HE4A	***0.58***	***0.52***	***0.54***	***0.60***	***0.47***	***0.42***	***0.14***	***0.15***	***0.14***	***0.11***	-	0.01
HE4B	***0.61***	***0.54***	***0.57***	***0.62***	***0.48***	***0.41***	0.10	***0.11***	***0.20***	0.06	0.02	-

Table S3.2. The Redundancy Analysis (RDA) with landcover and elevation variables identified 94 outlier loci-environment associations.

axis	snp	loading	shrub	forest	elevation	pred	correlation
3	65468_33.03	-0.179	-0.073	-0.521	0.212	forest	0.521
3	65468_33.02	0.179	0.073	0.521	-0.212	forest	0.521
3	73547_133.04	0.174	0.407	0.561	0.411	forest	0.561
3	73547_133.03	-0.174	-0.407	-0.561	-0.411	forest	0.561
3	74672_9.04	-0.145	-0.545	-0.462	-0.646	elevation	0.646
3	74672_9.03	0.145	0.545	0.462	0.646	elevation	0.646
3	75811_137.01	0.171	0.441	0.524	0.417	forest	0.524
3	75811_137.02	-0.171	-0.441	-0.524	-0.417	forest	0.524
3	79651_62.01	-0.138	0.053	-0.258	0.597	elevation	0.597
3	79651_62.02	0.138	-0.053	0.258	-0.597	elevation	0.597
3	80478_20.02	0.134	0.555	0.419	0.663	elevation	0.663
3	80478_20.03	-0.134	-0.555	-0.419	-0.663	elevation	0.663
3	83606_106.04	0.171	0.438	0.547	0.448	forest	0.547
3	83606_106.03	-0.171	-0.438	-0.547	-0.448	forest	0.547
3	88180_80.03	0.143	0.242	0.602	0.448	forest	0.602
3	88180_80.01	-0.143	-0.242	-0.602	-0.448	forest	0.602
3	89828_143.04	0.152	0.349	0.505	0.377	forest	0.505
3	89828_143.03	-0.152	-0.349	-0.505	-0.377	forest	0.505
3	91664_77.04	0.170	0.375	0.545	0.357	forest	0.545
3	91664_77.03	-0.170	-0.375	-0.545	-0.357	forest	0.545
3	93801_71.03	-0.181	-0.462	-0.469	-0.281	forest	0.469
3	93801_71.04	0.181	0.462	0.469	0.281	forest	0.469
3	98064_112.04	0.179	0.500	0.540	0.481	forest	0.540
3	98064_112.03	-0.179	-0.500	-0.540	-0.481	forest	0.540
3	123865_13.03	0.148	0.545	0.433	0.573	elevation	0.573
3	123865_13.01	-0.148	-0.545	-0.433	-0.573	elevation	0.573
3	123876_15.03	0.136	0.545	0.399	0.598	elevation	0.598
3	123876_15.01	-0.136	-0.545	-0.399	-0.598	elevation	0.598
3	126307_112.01	-0.140	-0.100	-0.326	0.249	forest	0.326
3	126307_112.03	0.140	0.100	0.326	-0.249	forest	0.326
3	126451_57.01	-0.140	-0.535	-0.467	-0.682	elevation	0.682
3	126451_57.03	0.140	0.535	0.467	0.682	elevation	0.682
3	126524_19.02	-0.145	-0.217	-0.116	0.481	elevation	0.481
3	126524_19.04	0.145	0.217	0.116	-0.481	elevation	0.481
3	126530_39.02	-0.147	-0.179	-0.147	0.502	elevation	0.502
3	126530_39.04	0.147	0.179	0.147	-0.502	elevation	0.502
3	127062_142.01	0.171	0.447	0.535	0.446	forest	0.535
3	127062_142.02	-0.171	-0.447	-0.535	-0.446	forest	0.535
3	129015_133.04	0.171	0.438	0.547	0.448	forest	0.547
3	129015_133.03	-0.171	-0.438	-0.547	-0.448	forest	0.547
3	131409_30.01	0.168	0.318	0.453	0.120	forest	0.453
3	131409_30.02	-0.168	-0.318	-0.453	-0.120	forest	0.453
3	137314_20.03	-0.177	-0.552	-0.511	-0.519	shrub	0.552
3	137314_20.02	0.177	0.552	0.511	0.519	shrub	0.552

3	139534_23.01	-0.171	-0.553	-0.488	-0.523	shrub	0.553
3	139534_23.03	0.171	0.553	0.488	0.523	shrub	0.553
3	140130_111.04	0.167	0.412	0.540	0.426	forest	0.540
3	140130_111.03	-0.167	-0.412	-0.540	-0.426	forest	0.540
3	144498_133.04	0.162	0.360	0.547	0.398	forest	0.547
3	144498_133.03	-0.162	-0.360	-0.547	-0.398	forest	0.547
3	163865_133.01	-0.144	-0.575	-0.445	-0.671	elevation	0.671
3	163865_133.02	0.144	0.575	0.445	0.671	elevation	0.671
3	170146_142.04	-0.174	-0.438	-0.585	-0.497	forest	0.585
3	170146_142.02	0.174	0.438	0.585	0.497	forest	0.585
3	171880_58.04	0.169	0.576	0.359	0.338	shrub	0.576
3	171880_58.03	-0.169	-0.576	-0.359	-0.338	shrub	0.576
3	184937_147.01	0.140	0.481	0.293	0.278	shrub	0.481
3	184937_147.04	-0.140	-0.481	-0.293	-0.278	shrub	0.481
3	185994_105.04	0.170	0.428	0.545	0.438	forest	0.545
3	185994_105.03	-0.170	-0.428	-0.545	-0.438	forest	0.545
3	193452_58.03	-0.135	0.127	-0.718	-0.142	forest	0.718
3	193452_58.02	0.135	-0.127	0.718	0.142	forest	0.718
3	193526_53.04	-0.160	-0.306	-0.214	0.279	shrub	0.306
3	193526_53.02	0.160	0.306	0.214	-0.279	shrub	0.306
3	193737_20.01	0.173	0.559	0.504	0.546	shrub	0.559
3	193737_20.04	-0.173	-0.559	-0.504	-0.546	shrub	0.559
3	195037_37.03	0.139	0.567	0.434	0.676	elevation	0.676
3	195037_37.01	-0.139	-0.567	-0.434	-0.676	elevation	0.676
3	195890_75.04	0.158	0.661	0.139	0.148	shrub	0.661
3	195890_75.03	-0.158	-0.661	-0.139	-0.148	shrub	0.661
3	199906_109.03	0.158	0.515	0.444	0.480	shrub	0.515
3	199906_109.04	-0.158	-0.515	-0.444	-0.480	shrub	0.515
3	215068_104.04	0.170	0.315	0.441	0.077	forest	0.441
3	215068_104.03	-0.170	-0.315	-0.441	-0.077	forest	0.441
3	222778_31.02	-0.144	-0.505	-0.128	0.008	shrub	0.505
3	222778_31.04	0.144	0.505	0.128	-0.008	shrub	0.505
3	223812_22.04	-0.154	-0.577	-0.459	-0.631	elevation	0.631
3	223812_22.01	0.154	0.577	0.459	0.631	elevation	0.631
3	230647_132.04	0.152	0.651	0.110	0.129	shrub	0.651
3	230647_132.03	-0.152	-0.651	-0.110	-0.129	shrub	0.651
3	233925_17.03	0.148	0.519	0.450	0.563	elevation	0.563
3	233925_17.04	-0.148	-0.519	-0.450	-0.563	elevation	0.563
3	234279_42.04	0.153	0.550	0.462	0.599	elevation	0.599
3	234279_42.02	-0.153	-0.550	-0.462	-0.599	elevation	0.599
3	234817_121.02	-0.134	-0.187	-0.137	0.414	elevation	0.414
3	234817_121.04	0.134	0.187	0.137	-0.414	elevation	0.414
3	268078_55.02	-0.150	-0.051	-0.313	0.419	elevation	0.419
3	268078_55.01	0.150	0.051	0.313	-0.419	elevation	0.419
3	271602_21.03	0.151	0.397	0.474	0.397	forest	0.474
3	271602_21.01	-0.151	-0.397	-0.474	-0.397	forest	0.474
3	314486_128.02	0.141	0.574	0.430	0.666	elevation	0.666
3	314486_128.01	-0.141	-0.574	-0.430	-0.666	elevation	0.666

3	359516_126.03	-0.167	-0.435	-0.534	-0.452	forest	0.534
3	359516_126.04	0.167	0.435	0.534	0.452	forest	0.534

Table S3.3. The Redundancy Analysis (RDA) with climatic variables (BIO9= Mean temperature of the driest quarter and BIO17=Precipitation of the driest quarter) identified 38 outlier loci-environment associations.

axis	snp	loading	wc_9	wc_17	predictor	correlation
2	73547_133.04	0.174	-0.879	-0.008	wc_9	0.879
2	73547_133.03	-0.174	0.879	0.008	wc_9	0.879
2	75811_137.01	0.172	-0.867	-0.008	wc_9	0.867
2	75811_137.02	-0.172	0.867	0.008	wc_9	0.867
2	83606_106.04	0.172	-0.872	0.020	wc_9	0.872
2	83606_106.03	-0.172	0.872	-0.020	wc_9	0.872
2	93801_71.03	-0.183	0.908	0.119	wc_9	0.908
2	93801_71.04	0.183	-0.908	-0.119	wc_9	0.908
2	98064_112.04	0.176	-0.896	0.060	wc_9	0.896
2	98064_112.03	-0.176	0.896	-0.060	wc_9	0.896
2	127062_142.01	0.170	-0.861	0.027	wc_9	0.861
2	127062_142.02	-0.170	0.861	-0.027	wc_9	0.861
2	129015_133.04	0.172	-0.872	0.020	wc_9	0.872
2	129015_133.03	-0.172	0.872	-0.020	wc_9	0.872
2	137314_20.03	-0.174	0.891	-0.113	wc_9	0.891
2	137314_20.02	0.174	-0.891	0.113	wc_9	0.891
2	140130_111.04	0.168	-0.851	0.005	wc_9	0.851
2	140130_111.03	-0.168	0.851	-0.005	wc_9	0.851
2	171880_58.04	0.184	-0.914	-0.133	wc_9	0.914
2	171880_58.03	-0.184	0.914	0.133	wc_9	0.914
2	185994_105.04	0.171	-0.865	0.013	wc_9	0.865
2	185994_105.03	-0.171	0.865	-0.013	wc_9	0.865
2	193526_53.04	-0.174	0.802	0.597	wc_9	0.802
2	193526_53.02	0.174	-0.802	-0.597	wc_9	0.802
2	193737_20.01	0.172	-0.882	0.125	wc_9	0.882
2	193737_20.04	-0.172	0.882	-0.125	wc_9	0.882
2	195890_75.04	0.185	-0.893	-0.323	wc_9	0.893
2	195890_75.03	-0.185	0.893	0.323	wc_9	0.893
2	215068_104.04	0.179	-0.861	-0.317	wc_9	0.861
2	215068_104.03	-0.179	0.861	0.317	wc_9	0.861
2	230647_132.04	0.182	-0.875	-0.348	wc_9	0.875
2	230647_132.03	-0.182	0.875	0.348	wc_9	0.875
2	260660_30.01	-0.170	0.754	0.800	wc_17	0.800
2	260660_30.03	0.170	-0.754	-0.800	wc_17	0.800
2	260661_119.03	-0.169	0.752	0.802	wc_17	0.802
2	260661_119.01	0.169	-0.752	-0.802	wc_17	0.802
2	359516_126.03	-0.168	0.853	-0.024	wc_9	0.853
2	359516_126.04	0.168	-0.853	0.024	wc_9	0.853

Literature Cited

Alcala, N., Goudet, J., Vuilleumier, S., 2014. On the transition of genetic differentiation from isolation to panmixia: What we can learn from GST and D. Theor. Popul. Biol. 93, 75–84. https://doi.org/10.1016/j.tpb.2014.02.003

Banks, S.C., Cary, G.J., Smith, A.L., Davies, I.D., Driscoll, D.A., Gill, A.M., Lindenmayer, D.B., Peakall, R., 2013. How does ecological disturbance influence genetic diversity? Trends Ecol. Evol. 28, 670–679. https://doi.org/10.1016/j.tree.2013.08.005

Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Bateson, Z.W., Dunn, P.O., Hull, S.D., Henschen, A.E., Johnson, J.A., Whittingham, L.A., 2014. Genetic restoration of a threatened population of greater prairie-chickens. Biol. Conserv. 174, 12–19. https://doi.org/10.1016/j.biocon.2014.03.008

Berggren, H., Nordahl, O., Tibblin, P., Larsson, P., Forsman, A., 2016. Testing for local adaptation to spawning habitat in sympatric subpopulations of pike by reciprocal translocation of embryos. PLoS One 11, 1–15. https://doi.org/10.1371/journal.pone.0154488

Blanquart, F., Kaltz, O., Nuismer, S.L., Gandon, S., 2013. A practical guide to measuring local adaptation. Ecol. Lett. 16, 1195–1205. https://doi.org/10.1111/ele.12150

Bron, G.M., Richgels, K.L.D., Samuel, M.D., Poje, J.E., Lorenzsonn, F., Matteson, J.P., Boulerice, J.T., Osorio, J.E., Rocke, T.E., 2018. Impact of Sylvatic Plague Vaccine on Non-target Small Rodents in Grassland Ecosystems. Ecohealth 1–11. https://doi.org/10.1007/s10393-018-1334-5

Brown, N.L., Peacock, M.M., Ritchie, M.E., 2016. Genetic variation and population structure in a threatened species, the Utah prairie dog Cynomys parvidens: The use of genetic data to inform conservation actions. Ecol. Evol. 6, 426–446. https://doi.org/10.1002/ece3.1874

Busch, J.D., Van Andel, R., Stone, N.E., Cobble, K.R., Nottingham, et al., 2013. the Innate Immune Response May Be Important for Surviving Plague in Wild Gunnison'S Prairie Dogs. J. Wildl. Dis. 49, 920–931. https://doi.org/10.7589/2012-08-209

Capblancq, T., Luu, K., Blum, M.G.B., Bazin, E., 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. Mol. Ecol. Resour. 18, 1223–1233. https://doi.org/10.1111/1755-0998.12906

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H., 2011. Stacks : Building and Genotyping Loci De Novo From Short-Read Sequences. Genes|Genomes|Genetics 1, 171–182. https://doi.org/10.1534/g3.111.000240

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

Charlesworth, B., 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10, 195–205. https://doi.org/10.1038/nrg2526

Collier, G. D., Spillett, J. J., 1973. The Utah prairie dog – decline of a legend. Utah Sci. 34:83–87.

Cornuet, J.M., Luikart, G., 1997 Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001–2014.

Curtis, R., Frey, S.N., Brown, N.L., 2014. The effect of coterie relocation on release-site retention and behavior of Utah prairie dogs. J. Wildl. Manage. 78, 1069–1077. https://doi.org/10.1002/jwmg.755

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol. Ecol. Resour. 14, 209–14. https://doi.org/10.1111/1755-0998.12157

Eads, D., Hoogland, J.L., 2017. Precipitation, climate change, and parasitism of prairie dogs by fleas that transmit plague. J. Parasitol. https://doi.org/10.1645/16-195

Earl, D. A., vonHoldt, B. M.,2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 4, 359-361. doi: 10.1007/s12686-011-9548-7

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., Borsboom, D., 2012. qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software 48, 1-18. URL http://www.jstatsoft.org/v48/i04/.

Flanagan, S.P., Forester, B.R., Latch, E.K., Aitken, S.N., Hoban, S., 2018. Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. Evol. Appl. https://doi.org/10.1111/eva.12569

Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180, 977–993.

Forester, B.R., Lasky, J.R., Wagner, H.H., Urban, D.L., 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol. Ecol. 27, 2215–2233. https://doi.org/10.1111/mec.14584

Frankham, R., 1996. Relationship of Genetic Variation to Population Size in Wildlife. Conserv. Biol. 10, 1500–1508. https://doi.org/10.1046/j.1442-9071.2001.00397.x

Frankham, R., 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Mol. Ecol. 24, 2610–2618. https://doi.org/10.1111/mec.13139

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R., Fenster, C.B., 2011. Predicting the probability of outbreeding depression. Con. Biol. 25, 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Funk, W.C., Forester, B.R., Converse, S.J., Darst, C., Morey, S., 2018. Improving conservation policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act decisions for conservation practitioners and geneticists. Conserv. Genet. 0, 0. https://doi.org/10.1007/s10592-018-1096-1

Gasca-Pineda, J., Cassaigne, I., Alonso, R.A., Eguiarte, L.E., 2013. Effective Population Size, Genetic Variation, and Their Relevance for Conservation: The Bighorn Sheep in Tiburon Island and Comparisons with Managed Artiodactyls. PLoS One 8, 20–22. https://doi.org/10.1371/journal.pone.0078120

Gilpin ME, Soulé ME, 1986. Minimum viable populations: process of species extinction. Conserv. Biol. Sci. Scarcity Divers. 19–34.

Goudet, J., Jombart, T., 2015. hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.04-22. https://CRAN.R-project.org/package=hierfstat

Harrisson, K.A., Amish, S.J., Pavlova, A., Narum, S.R., Telonis-Scott, M., et al., 2017. Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. Mol. Ecol. 26, 6253–6269. https://doi.org/10.1111/mec.14368

Hereford, J., 2009. A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. Am. Nat. 173, 579–588. https://doi.org/10.1086/597611

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276

Hill, WG., 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical research 38, 209-216.

Hoogland, J.L., 2006. Social behavior of prairie dogs. Pps 7-26, In Hoogland, J. (ed.) Conservation of the black-tailed prairie dog: Saving North America's Western grasslands. Island Press, Washington.

Hoogland, J.L., 2007. Alarm calling, multiple mating, and infanticide among black-tailed Gunnison's, and Utah prairie dogs. In Rodent societies (eds JO Wolff, PW Sherman), pp. 438 – 449. Chicago, IL: University Chicago Press

Hoogland, J.L., 2013. Prairie Dogs Disperse When All Close Kin Have Disappeared. Science 339, 1205–1207. https://doi.org/10.1126/science.1231689

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801-1806.

Jamieson, I.G., Allendorf, F.W., 2012. How does the 50/500 rule apply to MVPs? Trends Ecol. Evol. 27, 578–584. https://doi.org/10.1016/j.tree.2012.07.001

Janes, J.K., Miller, J.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., Cullingham, C.I., Andrew, R.L., 2017. The K = 2 conundrum. Mol. Ecol. 3594–3602. https://doi.org/10.1111/mec.14187

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment, 132, 159–175.

Johnson, W.E., Onorato, D.P., Roelke, M.E., Land, E.D., Cunningham, M., et al., 2010. Genetic restoration of the Florida panther. Science 329, 1641–1645. https://doi.org/10.1126/science.1192891

Jombart, T., Ahmed, I., 2011. adegenet 1 . 3-1 : new tools for the analysis of genome-wide SNP data 27, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jombart, T., Devillard, S., Balloux, F., Falush, D., Stephens, M., et al., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94

Kamvar ZN, Tabima JF, Grünwald NJ, 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ, 2, e281.

Keenan, K., Mcginnity, P., Cross, T.F., Crozier, W.W., Prodöhl, P.A., 2013. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol. Evol. 4, 782–788. https://doi.org/10.1111/2041-210X.12067

Kulpa, S.M., Leger, E.A., 2013. Strong natural selection during plant restoration favors an unexpected suite of plant traits. Evol. Appl. 6, 510–523. https://doi.org/10.1111/eva.12038

Lacy, Robert, C., 1987. Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. Conserv. Biol. 1, 143–158.

Laikre, L., Olsson, F., Jansson, E., Hössjer, O., Ryman, N., 2016. Metapopulation effective size and conservation genetic goals for the Fennoscandian Wolf (Canis lupus) population. Heredity (Edinb). 117, 279–289. https://doi.org/10.1038/hdy.2016.44

Latch, E.K., Rhodes, O.E., 2005. The effects of gene flow and population isolation on the genetic structure of reintroduced wild turkey populations: Are genetic signatures of source populations retained? Con. Gen. 6, 981-997. doi 10.1007/s10592-005-9089-2

Li, H., Durbin, R., 2009. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp698

Luu, K., Blum, M., Privé, F., 2019. pcadapt: Fast Principal Component Analysis for Outlier Detection. R package version 4.1.0. https://CRAN.R-project.org/package=pcadapt

McDonald, K. P, 1993. Analysis of the Utah prairie dog recovery program, 1972–1992. Publication No. 93–16. Utah Division of Wildlife Resources, Cedar City, UT. 81 pp.

Mulder, K.P., Walde, A.D., Boarman, W.I., Woodman, A.P., Latch, E.K., Fleischer, R.C., 2017. No paternal genetic integration in desert tortoises (Gopher agassizii) following translocation into an existing population. Biol. Conserv. 210, 318-324. https://doi.org/10.1016/j.biocon.2017.04.030

O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M., Portnoy, D.S., 2018. These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 0–3. https://doi.org/10.1111/mec.14792

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7. https://doi.org/10.1371/journal.pone.0037135

Piry, S., Luikart, G., Cornuet, J.M., 1999. bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity, 90, 502–503.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ralls, K., Ballou, J.D., Dudash, M.R., Eldridge, M.D.B., Fenster, C.B., Lacy, R.C., Sunnucks, P., Frankham, R., 2018. Call for a Paradigm Shift in the Genetic Management of Fragmented Populations. Conserv. Lett. 11, 1–6. https://doi.org/10.1111/conl.12412

Rochette, N.C., Catchen, J.M., 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nat. Publ. Gr. 12, 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rocke, T.E., Tripp, D.W., Russell, R.E., Abbott, R.C., Richgels, K.L.D., et al., 2017. Sylvatic Plague Vaccine Partially Protects Prairie Dogs (Cynomys spp.) in Field Trials. Ecohealth 14, 438–450. https://doi.org/10.1007/s10393-017-1253-x

Rocke, T.E., Williamson, J., Cobble, K.R., Busch, J.D., Antolin, M.F., Wagner, D.M., 2012. Resistance to Plague Among Black-Tailed Prairie Dog Populations. Vector-Borne Zoonotic Dis. 12, 111–116. https://doi.org/10.1089/vbz.2011.0602

Russell, R.E., Abbott, R.C., Tripp, D.W., Rocke, T.E., 2018. Local factors associated with on-host flea distributions on prairie dog colonies. Ecol. Evol. 2004–2016. https://doi.org/10.1002/ece3.4390

Savolainen, O., Lascoux, M., Merilä, J., 2013. Ecological genomics of local adaptation. Nat. Rev. Genet. 14, 807–820. https://doi.org/10.1038/nrg3522

Slatkin, M., 1987. the Geographic Structure of. Science (80-.). 236, 787–792. https://doi.org/10.4269/ajtmh.2010.09-0588

Sundqvist, L., Keenan, K., Zackrisson, M., Prodohl, P., Kleinhans, D., 2016. Directional genetic differentiation and relative migration. Ecol. Evol. 6, 3461–3475. https://doi.org/10.1002/ece3.2096

Truett, J.C., Dullam, J.A.L.D., Matchett, M.R., Owens, E., Seery, D., 2001. Translocating prairie dogs: a review. Wildl. Soc. Bull. 29, 863–872. https://doi.org/10.2307/3784413

United States Fish and Wildlife Service, 2012. Utah prairie dog (Cynomys parvidens) revised recovery plan. United States Fish and Wildlife Service. Denver, CO. 169 pp.

Urban, M.C., 2015. Accelerating extinction risk from climate change. Science (80-.). 348, 571–573. https://doi.org/10.1111/1467-8322.12302

Vähä, J.P., Erkinaro, J., Niemelä, E., Primmer, C.R., 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Mol. Ecol. 16, 2638–2654. https://doi.org/10.1111/j.1365-294X.2007.03329.x

Wang, J., 2017. The computer program Structure for assigning individuals to populations: easy to use but easier to misuse. Mol. Ecol. Res. 17, 981-990. https://doi.org/10.1111/1755-0998.12650

Weeks, A.R., Sgro, C.M., Young, A.G., Frankham, R., Mitchell, N.J., Miller, K.A., Byrne, M., Coates, D.J., Eldridge, M.D.B., Sunnucks, P., Breed, M.F., James, E.A., Hoffmann, A.A., 2011. Assessing the benefits and risks of translocations in changing environments: A genetic perspective. Evol. Appl. 4, 709–725. https://doi.org/10.1111/j.1752-4571.2011.00192.x

Whiteley, A.R., Fitzpatrick, S.W., Funk, W.C., Tallmon, D.A., 2015. Genetic rescue to the rescue. Trends Ecol. Evol. 30, 42–49. https://doi.org/10.1016/j.tree.2014.10.009

Wiens, J.J., 2016. Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. PLoS Biol. 14, 1–19. https://doi.org/10.1371/journal.pbio.2001104

Williams, S.K., Schotthoefer, a M., Montenieri, J. a, Holmes, J.L., Vetter, S.M., Gage, K.L., Bearden, S.W., 2013. Effects of Low-Temperature Flea Maintenance on the Transmission of Yersinia pestis by Oropsylla montana. Vector-Borne Zoonotic Dis. 13, 468–478. https://doi.org/DOI 10.1089/vbz.2012.1017

Willis, S.C., Hollenbeck, C.M., Puritz, J.B., Gold, J.R., Portnoy, D.S., 2017. Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Mol. Ecol. Resour. 17, 955–965. https://doi.org/10.1111/1755-0998.12647

Winter, D.J., 2012. mmod: an R library for the calculation of population differentiation statistics. Molecular Ecology Resources 12, 1158– 1160.

Wright, S., 1931. Evolution in mendelian populations. Genetics 97–159. https://doi.org/10.1007/BF02459575

RACHAEL M. GIGLIO

Curriculum Vitae

Education

PhD in Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI (In Progress)
Dissertation title: "Unraveling plague ecology through vector and host genetics"
Advisor: Dr. Emily K. Latch

MS in Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI (2014)
Thesis title: "Evaluation of alternative culling strategies on maintenance of genetic variation in bison (*Bison bison*)"
Advisor: Dr. Emily K. Latch

B.S. Ecology & Conservation Biology, University of Idaho, Moscow, ID (2009)
Thesis title: "Do pronghorn mothers shift their anti-predator behavior at night?"
Advisors: Dr. John A. Byers and Dr. Janet L. Rachlow

Positions/ Employment

Aug 2018 – Present	Distinguished Dissertator Fellow & Research/Teaching Assistant (Intro to Cons. and Env. Sci.) University of Wisconsin-Milwaukee (Milwaukee, WI)
Aug 2017 – May 2018	Research/Teaching Assistant (Foundation of Biology II) University of Wisconsin-Milwaukee (Milwaukee, WI)
June 2015 – June 2017	Research Assistant (Funded by the Morris Animal Foundation) University of Wisconsin-Milwaukee (Milwaukee, WI)
Aug 2011 – May 2015	Research/Teaching Assistant (Foundation of Biology II) University of Wisconsin-Milwaukee (Milwaukee, WI)
Jan 2010 – May 2011	GIS Technician, Habitat Use of Mountain Goats in WA, University of Idaho (Moscow, ID)
Sept 2009 – June 2011	Lab Technician, Conservation Genetics Lab, Dr. Lisette Waits, University of Idaho (Moscow, ID)
Aug 2009 – Dec 2009	Teaching Assistant, Introduction to Wildlife Management (NR 101), University of Idaho (Moscow, ID)

Publications *Surname changed from Toldness to Giglio in 2015

Giglio, R.M., J.A. Ivy, L. Jones, E.K. Latch. 2018. Pedigree-based genetic management improves bison conservation. Journal of Wildlife Management 82(4):766–774.

Giglio, R.M., J.A. Ivy, L. Jones, E.K. Latch. 2016. Evaluation of alternative management strategies for maintenance of genetic variation in wildlife populations. Animal Conservation 19:380–390.

Goldberg, C.S., K.Z. Woodruff, ***R. M. Toldness**, and L.P. Waits. 2011. Robust molecular sex identification of beaver (*Castor canadensis*) from non-destructive samples. Conservation Genetics Resources 3:729–731.

Grants

Research

Osorio, J.E., **R.M. Giglio –** Morris Animal Foundation, Grant-In-Aid (**$88,000** in 2015–2017)
Giglio, R.M., E.K. Latch – American Society of Mammalogists, Grant-in-Aid (**$1,500** in 2015)

Travel

Giglio, R.M. – The Wildlife Society, Travel Awards (**$250** in 2019)
Giglio, R.M. – UWM Graduate School Travel Award **($1,000** in 2018**)**
Giglio, R.M. – The Wildlife Society, Travel Award (**$350** in 2017)
Toldness, R.M. – Summer Institute for Statistical Genetics Scholarship (**$850** in 2013)
Toldness, R.M. – UWM Graduate School Travel Award (**$450** in 2012, **$400** in 2013)
Toldness, R.M. – UWM Ruth Walker Travel Award (**$200** in 2013)

Recent Honors/Awards

2019	Best Poster- UWM Bio Sci Symposium (**$200**)
2018–2019	Distinguished Dissertator Fellowship (UWM; **$16,500**)
2018	Best Presentation-Riveredge Nature Center Research Symposium (**$300**)
2018–2019	Ruth Walker Grant-in-Aid **(**UWM; Total **$3,000**)
2014–2018	Chancellor's Award (Total **$24,550**)
2017	Best Graduate Student Presentation – UWM Bio Sci Symposium (**$100**)
2016; 2012	Ivy Balsam- Milwaukee Audubon Society Award (UWM; Total **$750**)
2015	Leopold Scholarship (WI Chapter of The Wildlife Society; **$1,000**)
2014	James D. Anthony Award (UWM; **$200**)
2008	Lawrence and Eleanor Arneson Scholarship (UIdaho)
2007	John Vallega Scholarship (UIdaho)
2006, 2007	Idaho Robert Lee Promise B Scholarship (UIdaho)
2006	All Academic-Honor Division I NCAA (Women's Cross Country)

Presentations

Giglio, R.M. 2019. Unraveling plague ecology through vector and host genomics. Biological Sciences Colloquium, Milwaukee, WI. (Talk)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2019. Using landscape genomics to characterize potential transmission barriers to plague among prairie dog colonies. Biological Sciences Research Symposium, Milwaukee, WI. (Poster)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2018. Using landscape genomics to characterize potential transmission barriers to plague among prairie dog colonies. Riveredge Nature Center Research Symposium, Saukevill, WI. (Talk)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2018. Improving the conservation of the threatened Utah prairie dog through the evaluation of genomic variation and population structure. The Annual Conference of The Wildlife Society, Cleveland, OH. (Talk)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2017. Do small rodents act as vectors of plague among prairie dog colonies? Biological Sciences Research Symposium, Milwaukee, WI. (Talk)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2013. Evaluation of alternative culling strategies on maintenance of genetic variation in bison (Bison bison). Riveredge Research Symposium, Saukville, WI. (Talk)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2013. Evaluation of alternative culling strategies on maintenance of genetic variation in bison (*Bison bison*). The Annual Conference of The Wildlife Society, Milwaukee, WI. (Talk)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2013. Evaluation of alternative culling strategies on maintenance of genetic variation in bison (*Bison bison*). The American Society of Mammalogists, Philadelphia, PA. (Talk)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2012. Evaluation of alternative culling strategies on maintenance of genetic variation in bison (*Bison bison*). The Annual Conference of The Wildlife Society, Portland, OR. (Talk)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2012. Modeling genetic variation of bison based on current culling strategies. Evolution 2012, Ottawa, ON. (Poster)

Toldness, R.M., J.A. Ivy, L.C. Jones, E.K. Latch. 2012. Evaluation of alternative culling strategies on maintenance of genetic variation in bison (*Bison bison*). Biological Sciences Research Symposium, Milwaukee, WI. (Talk)

Invited Talks

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2019. Elucidating plague transmission among prairie dog colonies by using vector and host genomics. The Annual Conference of The Wildlife Society, Reno, NV. (Talk)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2019. Characterizing patterns of genomic variation in prairie dogs: Implications for the transmission of plague. USGS-National Wildlife Health Center, Madison, WI. (Talk)

Giglio, R.M., T. Rocke, J. Osorio, E.K. Latch. 2017. Optimizing disease management of prairie dogs with population genomics. The Annual Conference of The Wildlife Society, Albuquerque, NM. (Talk)

Giglio, R.M. 2017. Using molecular ecology to improve wildlife management. University of Wisconsin-Whitewater Colloquium, Whitewater, WI. (Talk)

Certificates

Introduction to Research Computing
The University of Wisconsin-Milwaukee, Milwaukee, WI (October 2016)

Elements of R for Genetics and Bioinformatics
The 18th Summer Institute in Statistical Genetics. University of Washington, Seattle, WA (July 2013).

Bayesian Statistics for Genetics-

The 18th Summer Institute in Statistical Genetics. University of Washington, Seattle, WA (July 2013).

Professional Societies

The Wildlife Society
American Society of Mammalogists
Society for the Study of Evolution

Outreach

Judge for the UW-Milwaukee Undergraduate Research Symposium, Milwaukee, WI (2017, 2018, 2019)

Co-presenter of "DNA Decoded" for UW-Milwaukee Science Bag, Milwaukee, WI (January 2018)