DISEASE ECOLOGY OF RABIES IN THE GREAT PLAINS: SYNTHESIZING THE EFFECTS OF VIRAL PROPERTIES, HOST ATTRIBUTES, AND LANDSCAPE ON DISEASE EMERGENCE

by

HEATHER D. BARTON

B.S., Grove City College, 2005

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2010

Abstract

Emerging infectious diseases play an increasingly critical role in many biological fields, including conservation biology and public health. Many emerging diseases originate in wildlife, most are caused by viruses, and often emergence is due to adaptation to and amplification in a new host, frequently in areas where ecological transformation is occurring. These emergence patterns suggest that the complex interactions among host, virus, and landscape drive disease emergence. Terrestrial rabies in striped skunks (*Mephitis mephitis*) in the central Great Plains is an excellent model system to investigate the interactions among the components of disease emergence: host ecology, pathogen properties, and landscape features. Striped skunks are not only numerous in the central Great Plains, they are also the reservoir for two genetically distinct rabies strains that co-occur in the region. Additionally, the landscape in the central Great Plains has undergone significant land use change over the last 70 years through increased urbanization and industrial agriculture practices. I used a combination of molecular and spatial techniques to investigate the interactions among host, pathogen, and landscape. Molecular epidemiology results indicated that rabies strains in the central Great Plains exhibit different epidemiological properties, while population genetic analyses indicated that striped skunks in the region are highly admixed and comprise a single population. Spatial analysis revealed that landscape features such as rivers are not a barrier to striped skunk dispersal, but differentially influence the movement of the two rabies strains. Because striped skunks are reservoirs for many diseases other than rabies and are ubiquitous throughout North America, I also examined the historical movements and distribution of striped skunks in North America using a phylogeographic approach. Results revealed that a combination of multiple Pleistocene dispersal events and

Holocene admixture are responsible for the contemporary population structure of striped skunks in North America, and allowed me to place my regional-scale striped skunk rabies study into a larger biogeographic context. My results support the use of a holistic approach for studying emerging infectious diseases that includes studies of viral characteristics, host ecology and biogeography, and spatial features.

DISEASE ECOLOGY OF RABIES IN THE GREAT PLAINS: SYNTHESIZING THE EFFECTS OF VIRAL PROPERTIES, HOST ATTRIBUTES, AND LANDSCAPE ON DISEASE EMERGENCE

by

HEATHER D. BARTON

B.S., Grove City College, 2005

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2010

Approved by:

Major Professor
Samantha M. Wisely

Abstract

Emerging infectious diseases play an increasingly critical role in many biological fields, including conservation biology and public health. Many emerging diseases originate in wildlife, most are caused by viruses, and often emergence is due to adaptation to and amplification in a new host, frequently in areas where ecological transformation is occurring. These emergence patterns suggest that the complex interactions among host, virus, and landscape drive disease emergence. Terrestrial rabies in striped skunks (*Mephitis mephitis*) in the central Great Plains is an excellent model system to investigate the interactions among the components of disease emergence: host ecology, pathogen properties, and landscape features. Striped skunks are not only numerous in the central Great Plains, they are also the reservoir for two genetically distinct rabies strains that co-occur in the region. Additionally, the landscape in the central Great Plains has undergone significant land use change over the last 70 years through increased urbanization and industrial agriculture practices. I used a combination of molecular and spatial techniques to investigate the interactions among host, pathogen, and landscape. Molecular epidemiology results indicated that rabies strains in the central Great Plains exhibit different epidemiological properties, while population genetic analyses indicated that striped skunks in the region are highly admixed and comprise a single population. Spatial analysis revealed that landscape features such as rivers are not a barrier to striped skunk dispersal, but differentially influence the movement of the two rabies strains. Because striped skunks are reservoirs for many diseases other than rabies and are ubiquitous throughout North America, I also examined the historical movements and distribution of striped skunks in North America using a phylogeographic approach. Results revealed that a combination of multiple Pleistocene dispersal events and

Holocene admixture are responsible for the contemporary population structure of striped skunks in North America, and allowed me to place my regional-scale striped skunk rabies study into a larger biogeographic context. My results support the use of a holistic approach for studying emerging infectious diseases that includes studies of viral characteristics, host ecology and biogeography, and spatial features.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I relied on the advice, support, and encouragement of many people over the course of my doctoral program. First, I would like to thank my major advisor, Dr. Samantha Wisely, for her guidance and support over the past 5 years. Without her mentoring, the research presented in my dissertation would not have been possible. Sam was always available to answer questions, throw around ideas, or to just talk about whatever was on my mind. Her positive attitude and encouragement was a blessing during the times when I doubted myself, and I could not have asked for a better mentor. I would also like to thank the other members of my advisory committee: Dr. Jack Cully, Dr. Carolyn Ferguson, and Dr. Karen Garrett. Jack gave me a sense of how to look at diseases from both an ecological and an evolutionary perspective, which helped me to see the larger scope of my research. Carolyn provided me with a solid understanding of the basics of phylogenetics and systematics, and I am grateful for her advice and guidance on phylogenetic tree interpretation during the course of my research. Karen provided insight from an epidemiological perspective, and sparked my interest in using a modeling approach to look at rabies from a different angle. I appreciate the unique perspectives that each committee member has brought to my research, and I feel that they have all helped to prepare me for a successful career.

I would like to thank several family members, friends, and colleagues for their support. First, I would like to thank my family for their long-distance support and encouragement when things were stressful. They listened to all of my troubles and told me that everything would work out fine. Next, I would like to thank Alyssa Riley for the four and a half years of good conversations, venting sessions, support and encouragement when I needed it, and random trips around northeast Kansas. I would also like to thank my fellow Biology graduate students and my non-Biology friends for the good times and support during graduate school. Finally, I would like to thank the Molecular Ecology and Conservation Genetics lab group for the good times, good discussions, and support that they provided during my research program.

Dr. Cathleen Hanlon and Rolan Davis at the Kansas State University College of Veterinary Medicine Rabies Diagnostic Lab provided technical assistance and laboratory space. Lab assistance was provided by Claudia Ganser and Hailey Petersen, and Andy Gregory was

instrumental in conducting the spatial analyses. Feedback on research ideas and manuscripts was provided by current and former members of the Wisely lab. Funding for the research was provided by an Ecological Genomics Foundation seed grant, a Berryman Institute graduate student fellowship, the EPSCoR Ecoforecasting Program, and from the Wisely Molecular Ecology and Conservation Genetics Lab.

CHAPTER 1 - Introduction

Emerging infectious diseases (EIDs) play an increasingly critical role in many biological fields, including conservation biology and public health; many EIDs are zoonotic (originate in wildlife), about 75% are caused by viruses (Chomel, Belotto, & Meslin, 2007), and emergence is often the result of adaptation to and amplification in a novel host (Daszak, Cunningham, & Hyatt, 2001). Transmission to a diversity of potential hosts via host switches among wildlife reservoirs or spillover from wildlife to domestic animals leads to pathogen strains of increased fitness that are better adapted to novel environments (Daszak, et al., 2001; McCallum & Dobson, 2002). RNA viruses in particular are especially well-adapted for the invasion of novel environments.

RNA viral pathogens have enormous adaptive potential, and often generate suites of viral strains. However, the success of a given strain depends on a variety of factors, including evolutionary history, which shapes the adaptive potential of a viral strain. For a majority of RNA viruses, including retroviruses and segmented RNA viruses, mutation and environmental selection are the antagonistic mechanisms that most influence evolutionary potential (Moya, Holmes, & Gonzalez-Candelas, 2004). Because RNA viruses have no proofreading mechanism, they cannot repair transcription errors. Thus, most RNA viruses have mutation rates on the order of 10^{-3} to 10^{-5} mutations per site per year. These mutation rates are very near the error threshold, which is the limit at which lethal mutations rapidly accumulate, decrease individual fitness, and drive a population to extinction; the high mutation rates of RNA viruses, which are maintained just below this limit of critical mutagenesis, results in rapid sequence variation (Domingo & Holland, 1997).

The evolutionary success of RNA viruses is therefore the result of high mutation rates mediated by selective pressures imposed by the host environment. The high mutation rate of RNA viruses allows for maximum adaptability to new host environments; as host selection pressures change, RNA viruses can rapidly adapt (Holland & Domingo, 1998). The adaptive potential of RNA viruses is exemplified in virulence differences in HIV. HIV consists of two main strains, but through extremely high mutation rates those strains have diversified into a dozen different subtypes, each of which contains a myriad of unique viral sequences and virulence levels (Holmes, 2001). Ultimately, the increased genetic diversity of HIV has allowed for a spectrum of virulence levels across viral subtypes, and likely facilitated its successful switch from simians to humans (Van Heuverswyn & Peeters, 2007). Thus, RNA viral strains with a history of different host species should also have differences in virulence, pathogenesis, and transmission efficiency, which result in differences in disease emergence patterns.

The complex interactions among host, virus, and landscape, which influence disease emergence patterns, necessitate the use of well-established model disease systems to comprehensively investigate disease ecology. I used the striped skunk-rabies system to investigate rabies epidemiology, striped skunk host ecology and phylogeography, and landscape epidemiology. This research provides a comprehensive look at the three major components in rabies disease ecology, and emphasizes how molecular and spatial methods can be used to study infectious disease systems. For the remainder of the introduction, I will explore the three components involved in rabies disease ecology: the virus, the host, and the landscape.

The Pathogen: Rabies

The rabies virus is a member of the lyssavirus genus in the family rhabdoviridae (Krebs, Wilson, & Childs, 1995). Rabies is a single-stranded, antisense (i.e. the nucleotide sequence is read in the 3' to 5' direction), non-segmented RNA virus. The rabies virus genome is approximately 12,000 bp long, containing a leader sequence, 4 non-coding regions, and 5 genes that code for nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and polymerase (L; Meng, et al., 2007). Nucleoprotein encases the genomic RNA, preserving the integrity of the genetic material; matrix protein is associated with both the nucleoprotein and the virus envelope, making it critical for virion assembly; glycoprotein is found on the outer surface of the virus envelope, and is responsible for the formation of the viral antigenic sites on the surface of the envelope; polymerase is responsible for assembling new rabies RNA strands during transcription; and phosphoprotein is a jack-of-all-trades, playing roles in replication, transcription, transmission, and assembly (Meng, et al., 2007). The viral RNA and all of the proteins are contained in a bullet-shaped virion of approximately 180 nm in length by 75 nm in width (Tordo, 1996).

The potential for transmission begins when the virus induces behavioral changes like fearlessness, aggressiveness, and restlessness, in an infected host (Niezgoda, Hanlon, & Rupprecht, 2002). The infected animal bites a susceptible animal, directly transmitting virus-laden saliva into the muscle of the naïve animal (Lafon, 2005). If transmission is successful, the virus passes from the muscle into the peripheral nerves, and travels through the nervous system toward the brain. During this latent period, which can last from several days to several months, there are no signs that the animal has been infected. After the virus reaches the brain, the infected animal enters the prodromal phase and begins to show symptoms, which generally

manifest in one of two ways: furious rabies or paralytic rabies (Hemachudha, et al., 2003). Furious rabies is the most familiar form, in which behavioral changes cause infected animals to become fearless and highly aggressive, and initiate the bite reflex in carnivores (Mitrabhakdi, et al., 2005). Paralytic rabies, on the other hand, generates symptoms of lethargy and depression (Mitrabhakdi, et al., 2005); this manifestation is more common in spillover hosts, like livestock. Aggressive and paralytic rabies do share several symptoms, including the presence of thick, ropey, virus-filled saliva, and hydrophobia. Death generally occurs within two weeks after the infected animal becomes symptomatic (Krebs, et al., 1995).

Over 70,000 people die each year around the world as a result of rabies, usually domestic dog rabies (Dietzschold, Li, Faber, & Schnell, 2008). However, under-reporting is common in much of the developing world, so this value likely underestimates the actual number of human deaths due to rabies each year. Developing countries rarely have the funding to maintain education, surveillance, and mandatory dog vaccination (Centers for Disease Control and Prevention, 2010a). As a result, the large numbers of unvaccinated stray and feral dogs facilitate continued transmission among canines, leading to increased exposure to the virus in humans. An additional problem in developing nations is that even if individuals recognize that they may have been bitten by a rabid animal, post-exposure vaccination is too costly to obtain. Conversely, in the United States alone, over $300 million is spent each year on rabies prevention and treatment, which includes pre- and post-exposure vaccination for humans, education and surveillance programs, and vaccination of pets (Smith, 1996).

At the turn of the last century, death rates in the United States were similar to other parts of the world; over 100 people died each year mostly as a result of domestic dog rabies (Centers for Disease Control and Prevention, 2010b). However, mandatory dog vaccination programs

that began in the 1940's and 1950's led to the eradication of domestic dog rabies in the U.S. (Smith, 1996), and drastically reduced rabies-related deaths. As domestic dog rabies declined though, cases of wildlife rabies began to increase; while only 1 to 2 human rabies deaths now occur each year in the United States, they are usually as a result of wildlife rabies, and almost always insectivorous bat rabies (Centers for Disease Control and Prevention, 2010b). Despite the presence of rabies in other wildlife species, bats are responsible for virtually all rabies-related deaths in the United States because rabies education in this country has advanced to the point that signs of rabies in wild carnivores and companion animals are easily recognizable, so people are more likely to seek medical attention if they think exposure has occurred. Bat exposure, on the other hand, is often much more subtle. Most reports of bat rabies exposure indicate that the person awoke and noticed a bat in the room; bat bites are quite small and often look similar to insect bites, so if a person is bitten while asleep, bite marks may go unnoticed or be mistaken for an insect bite; the person assumes that exposure has not occurred, and fails to get post-exposure vaccination or have the bat tested. By the time symptoms emerge, treatment is no longer an option, and the person dies.

In addition to the many rabies strains circulating among bat species, several different strains of terrestrial wildlife rabies circulate in the United States (Figure 1.1; Blanton, Hanlon, & Rupprecht, 2007). Terrestrial rabies strains tend to be geographically partitioned according to the range of a particular terrestrial host species. The species-specific nature of these strains arises through a combination of rapid replication and the lack of a proofreading mechanism to repair transcription errors, leading to rapid mutation (Domingo & Holland, 1997). The high mutation rate of rabies allows the virus to exist as a quasispecies, a group of related viral sequences that differ by one or a few mutations within a single host individual (Morimoto, et al.,

1998). Because a majority of the sequences are closely related, they are well-adapted to a particular host species, and the slight mutational variation among sequences allows rabies to quickly adapt to differences among individuals of the reservoir species. The quasispecies nature of rabies also leads to variability in viral characteristics like virulence, transmissibility, and incubation period, which are all crucial for host-switching. For instance, a rabies strain that exhibits high transmission efficiency is more likely to survive through transmission to new hosts, while a strain with poor transmission efficiency has a higher probability of extinction. Similarly, a low virulence strain or a high virulence strain with a long incubation period allows an infected host to survive longer. Ultimately, the longer a host can survive, the higher the probability that it will come into contact with other susceptible hosts. Conversely, a highly pathogenic strain with a short incubation period incapacitates its host in a short time, before the host has a chance to transport or transmit the virus. Thus, when a novel species becomes infected, it is probable that a potentially advantageous mutation exists somewhere within the spectrum of possible viral sequences that will enable the rabies virus to persist in the new species long enough for transmission to occur among individuals, eventually leading to a rabies strain that is better adapted to the new species, and establishing a new reservoir species.

The Host: Striped Skunk

Striped skunks (*Mephitis mephitis*) are a native North American species, with a distribution ranging from northern Mexico through the entire continental United States and into southern Canada (Hall, 1981). These mesocarnivores are the quintessential generalist species; they can utilize a variety of habitats throughout their range to forage for a variety of food types. In general, skunks are insectivores, but will also feed on berries and nuts, small mammals, and

bird eggs (Azevedo, et al., 2006). As a result of their preference for insects, they also tend to prefer forest-field edge habitat (Bixler & Gittleman, 2000) and agricultural land (Larivière, Walton, & Messier, 1999) where these prey items are especially abundant. Striped skunks are also an urban-adapted species (Gehrt, 2005), easily adapting to life in cities and suburban areas.

The generalist life history characteristics of striped skunks make them an important ecosystem component. They can fill a variety of niches, and they affect both predator and prey distributions (Roemer, Gompper, & Van Valkenburgh, 2009). However, these same characteristics also make striped skunks a nuisance species. Striped skunks dig up crops and gardens in search of grubs, and their propensity for using human habitation as shelter and food sources puts them in direct conflict with humans and companion animals (Broadfoot, Rosatte, & O'Leary, 2001). The increasing proximity of striped skunks to humans and domestic animals is also a public health concern because striped skunks are a major disease reservoir, harboring a variety of viral and bacterial diseases including rabies (Krebs, Smith, Rupprecht, & Childs, 1999), canine distemper (Williams, Thorne, Appel, & Belitsky, 1988), and tularemia (Berrada, Goethert, & Telford, 2006), among others. Striped skunks also have a high dispersal capacity (Bixler & Gittleman, 2000), utilize communal dens (Sunquist, 1974), have a high reproduction rate (Rosatte, 1987), and are ubiquitous throughout much of North America (Hall, 1981). These demographic and behavioral traits in combination with their generalist life history traits make them an ideal disease reservoir. The suitability of striped skunks as model hosts can be seen in the fact that they are the reservoir for four of the seven main terrestrial rabies strains in the United States: California skunk strain, Arizona skunk strain (found in the northern part of Arizona), and North Central and South Central skunk strains, which are both found in the central Great Plains (Figure 1.1).

The Landscape

Ecologists have long known that landscape features influence plant and animal distribution and movement in an ecosystem. Landscape features can act either as barriers or corridors, shaping population structure and dispersal; when the landscape changes, the ecosystem also changes as species leave or arrive. Because landscape features affect host species and pathogens to shape disease emergence, a better understanding of how ecological transformation alters the landscape to aid in disease emergence is also of vital importance.

Ecological transformation falls into two general categories: global climate change and anthropogenic influence. Global climate change is predicted to dramatically alter precipitation and temperature patterns around the world (Raisanen, 2002). As weather patterns change, plants and animals that are better adapted to the altered environments will increase in abundance, bringing new pathogens with them. Ultimately, global climate change may lead to an increase in disease emergence. This phenomenon is already apparent in the massive rate of amphibian mortality in Central and South America due to infection by an emerging chytrid fungus, *Batrachochytrium dendrobatidis* (Rohr & Raffel, 2010). The other category of ecological transformation, anthropogenic influence, includes increased global connectivity, translocation and transport of animals and plants, and land use change (e.g. urbanization, industrial agriculture practices). As global connectivity increases, formerly remote regions of the world become more accessible, presenting more opportunities for humans to interact with novel pathogens. SARS exemplifies the results of increased global connectivity. SARS emerged from the crowded markets of Hong Kong where it was transmitted from chickens to civets (Greger, 2007). The infected civets transmitted the virus to humans, and rapid air travel allowed infected humans to

carry SARS from southeast China to Europe and the Americas, leading to the first pandemic of the twenty-first century (Lee & Krilov, 2005).

Along with increased connectivity, translocation of species brings new pathogens into new potentially ideal habitats through the introduction of novel hosts. Raccoon rabies expanded up the east coast of the United States through a translocation event. Prior to the 1970's, rabies was confined to the southeast United States (Florida, Georgia, South Carolina, and Alabama); translocation of infected, but asymptomatic, raccoons to West Virginia and Virginia resulted in the first cases of rabies in these two states in the late 1970's, leading to an explosion of raccoon rabies up the east coast of the United States (Jenkins, Perry, & Winkler, 1988). Finally, land use change in the form of urbanization and industrial agriculture alters the physical landscape by fragmenting, degrading, or completely changing habitat. Species remaining in these habitats are increasingly stressed, which makes them more susceptible to the novel pathogens arriving with novel hosts that are better adapted to these fragmented or degraded habitats. Nipah virus (Mackenzie, 2005), Lyme disease (Allan, Keesing, & Ostfeld, 2003), and Rift Valley Fever (Greger, 2007) are all examples of diseases that expanded their ranges due to land use change.

Implications for Disease Ecology

The striped skunk-rabies system in the central Great Plains is an excellent model system to investigate the interactions among the three components of disease emergence patterns: host ecology, pathogen properties, and landscape features. Striped skunks are not only prolific in the central Great Plains, they are also the reservoir for two evolutionarily independent rabies strains, North Central and South Central skunk rabies (Favoretto, de Mattos, Morais, Araújo, & de Mattos, 2001; Real, Russell, Waller, Smith, & Childs, 2005). Additionally, the landscape in the

central Great Plains has undergone a significant amount of land use change over the last 70 years through increased urbanization and industrial agriculture practices (Laude, 1958; Ostlie, Schneider, Aldrich, Faust, & Chaplin, 1997; Samson, Knopf, & Ostlie, 2004).

I used a combination of molecular and spatial techniques to investigate the interactions among host, pathogen, and landscape. First, I used molecular techniques to examine the epidemiological properties and phylogeography of North Central and South Central skunk rabies over a five-state region in the central Great Plains. I also used molecular techniques to investigate the population structure and host ecology of corresponding striped skunk hosts in the same study area. Finally, I used a combination of molecular and spatial analyses to examine the interactions between landscape features and striped skunks, and between landscape features and rabies strains in the same study area. The results of these studies provided information on the epidemiological properties of North Central and South Central skunk rabies, the population structure of striped skunks in the central Great Plains, and the influence of landscape features on rabies and striped skunk movement. These results are described in Chapter 2 of this dissertation.

The generalist nature of striped skunks combined with their propensity for carrying diseases also reinforces the need to examine the species' historical and contemporary population characteristics in a biogeographic context. Because striped skunks are found throughout North America, not just in the central Great Plains, I examined the historical movements and distribution of striped skunks in North America using a phylogeographic approach. While phylogeography studies are most often used to determine the conservation necessity of endangered species (Andrus, et al., 2009; Arshad, Gonzalez, El-Sayed, Osborne, & Wink, 2009; Earl, Louie, Bardeleben, Swift, & Jacobs, 2010; Terrasa, et al., 2009), they are being used in studies of disease reservoirs and pathogens with increasing frequency (Cullingham, Kyle, Pond,

& White, 2008; Dragoo, et al., 2006; Llewellyn, et al., 2009). By examining the phylogeography of striped skunks on a continental scale, I was able to determine how geological landscape changes influenced the current structure and movements of striped skunks in North America. The results of this study provide a larger context in which to place the population structure of striped skunks in the central Great Plains. My findings are described in the third chapter of this dissertation. Finally, in the fourth chapter of my dissertation, I draw conclusions from the landscape epidemiology study and from the study of the phylogeography of striped skunks. From these conclusions, I make inferences regarding both the importance of using a holistic approach when investigating disease ecology and the value of examining comparative phylogeography between pathogens and their host reservoirs.

References

Allan, B. F., Keesing, F., & Ostfeld, R. S. (2003). Effect of forest fragmentation on Lyme disease risk. *Conservation Biology, 17*, 267-272.

Andrus, N., Tye, A., Nesom, G., Bogler, D., Lewis, C., Noyes, R., et al. (2009). Phylogenetics of Darwiniothamnus (Asteraceae: Astereae) - molecular evidence for multiple origins of the endemic flora of the Galapagos Islands. *Journal of Biogeography, 36*, 1055-1069.

Arshad, M., Gonzalez, J., El-Sayed, A. A., Osborne, T., & Wink, M. (2009). Phylogeny and phylogeography of critically endangered Gyps species based on nuclear and mitochondrial markers. *Journal of Ornithology, 150*, 419-430.

Azevedo, F. C. C., Lester, V., Gorsuch, W., Larivière, S., Wirsing, A. J., & Murray, D. L. (2006). Dietary breadth and overlap among five sympatric prairie carnivores. *Journal of Zoology, 269*, 127-135.

Berrada, Z. L., Goethert, H. K., & Telford, S. R. (2006). Raccoons and skunks as sentinels for enzootic tularemia. *Emerging Infectious Diseases, 12*, 1019-1021.

Bixler, A., & Gittleman, J. L. (2000). Variation in home range and use of habitat in the striped skunk (*Mephitis mephitis*). *Journal of Zoology, London, 251*, 525-533.

Blanton, J. D., Hanlon, C., & Rupprecht, C. E. (2007). Rabies surveillance in the United States during 2006. Retrieved from http://www.cdc.gov/rabies/docs/rabies_surveillance_us_2006.pdf

Broadfoot, J. D., Rosatte, R. C., & O'Leary, D. T. (2001). Raccoon and skunk population models for urban disease control planning in Ontario, Canada. *Ecological Applications, 11*, 295-303.

Centers for Disease Control and Prevention (2010a). Rabies around the world. Retrieved June 24, 2010, from http://www.cdc.gov/rabies/location/world/index.html.

Centers for Disease Control and Prevention (2010b). Rabies in the U.S. Retrieved June 24, 2010, from http://www.cdc.gov/rabies/location/usa/index.html.

Chomel, B. B., Belotto, A., & Meslin, F.-X. (2007). Wildlife, exotic pets, and emerging zoonoses. *Emerging Infectious Diseases, 13*, 6-11.

Cullingham, C. I., Kyle, C. J., Pond, B. A., & White, B. N. (2008). Genetic structure of raccoons in eastern North America based on mtDNA: Implications for subspecies designation and rabies disease dynamics. *Canadian Journal of Zoology, 86*, 947-958.

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2001). Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Tropica, 78*, 103-116.

Dietzschold, B., Li, J., Faber, M., & Schnell, M. (2008). Concepts in the pathogenesis of rabies. *Future Virology, 3*, 481-490.

Domingo, E., & Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annual Review of Microbiology, 51*, 151-178.

Dragoo, J. W., Lackey, J. A., Moore, K. E., Lessa, E. P., Cook, J. A., & Yates, T. L. (2006). Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a predictive framework for research on hantaviruses. *Journal of General Virology, 87*, 1997-2003.

Earl, D. A., Louie, K. D., Bardeleben, C., Swift, C. C., & Jacobs, D. K. (2010). Rangewide microsatellite phylogeography of the endangered tidewater goby, *Eucyclogobius newberryi* (Teleostei: Gobiidae), a genetically subdivided coastal fish with limited marine dispersal. *Conservation Genetics, 11*, 103-114.

Favoretto, S. R., de Mattos, C. C., Morais, N. B., Araújo, F. A. A., & de Mattos, C. A. (2001). Rabies in Marmosets (*Callithrix jacchus*), Ceará, Brazil. *Emerging Infectious Diseases, 7*, 1062-1065.

Gehrt, S. D. (2005). Seasonal survival and cause-specific mortality of urban and rural striped skunks in the absence of rabies. *Journal of Mammalogy, 86*, 1164-1170.

Greger, M. (2007). The Human/Animal Interface: Emergence and Resurgence of Zoonotic Infectious Diseases. *Critical Reviews in Microbiology, 33*, 243-299.

Hall, E. R. (1981). *The Mammals of North America* (2nd ed.). New York: John Wiley & Sons.

Hemachudha, T., Wacharapluesadee, S., Lumlertdaecha, B., Orciari, L. A., Rupprecht, C. E., La-Ongpant, M., Juntrakul, S., & Denduangboripant, J. (2003). Sequence analysis of rabies virus in humans exhibiting encephalitic or paralytic rabies. *Journal of Infectious Diseases, 188*, 960-966.

Holland, J., & Domingo, E. (1998). Origin and Evolution of Viruses. *Virus Genes, 16*, 13-21.

Holmes, E. C. (2001). On the origin and evolution of the human immunodeficiency virus (HIV). *Biological Reviews, 76*, 239-254.

Jenkins, S. R., Perry, B. D., & Winkler, W. G. (1988). Ecology and epidemiology of raccoon rabies. *Reviews of Infectious Diseases, 10*, S620-S625.

Knobel, D. L., Fooks, A. R., Brookes, S. M., Randall, D. A., Williams, S. D., Argaw, K., et al. (2008). Trapping and vaccination of endangered Ethiopian wolves to control an outbreak of rabies. *Journal of Applied Ecology, 45*, 109-116.

Krebs, J. W., Smith, J. S., Rupprecht, C. E., & Childs, J. E. (1999). Rabies surveillance in the United States during 1998. *Journal of the American Veterinary Medical Association, 215*, 1786-1789.

Krebs, J. W., Wilson, M. L., & Childs, J. E. (1995). Rabies - epidemiology, prevention, and future research. *Journal of Mammalogy, 76,* 681-694.

Lafon, M. (2005). Modulation of the immune response in the nervous system by rabies virus. *Current Topics in Microbiology and Immunology, 289,* 239-258.

Larivière, S., Walton, L. R., & Messier, F. (1999). Selection by striped skunks (*Mephitis mephitis*) of farmsteads and buildings as denning sites. *American Midland Naturalist, 142,* 96-101.

Laude, H. H. (1958). The fruitful plains. *Transactions of the Kansas Academy of Science, 4,* 19-22.

Lee, P. J., & Krilov, L. R. (2005). When animal viruses attack: SARS and Avian Influenza. *Pediatric Annals, 34,* 42-52.

Llewellyn, M. S., Lewis, M. D., Acosta, N., Yeo, M., Carrasco, H. J., Segovia, M., Vargas, J., Torrico, F., Miles, M. A., & Gaunt, M. W. (2009). *Trypanosoma cruzi* IIc: Phylogenetic and phylogeographic insights from sequence and microsatellite analysis and potential impact on emergent Chagas disease. *PLOS Neglected Tropical Diseases, 3,* 1-10.

Mackenzie, J. S. (2005). Emerging zoonotic encephalitis viruses: Lessons from Southeast Asia and Oceania. *Journal of Neurovirology, 11,* 434-440.

McCallum, H., & Dobson, A. (2002). Disease, habitat fragmentation and conservation. *Proceedings of the Royal Society of London, Series B, 269,* 2041-2049.

Meng, S.-L., Yan, J.-X., Xu, G.-L., Nadin-Davis, S. A., Ming, P.-G., Liu, S.-Y., Wu, J., Ming, H.-T., Zhu, F.-C., Zhou, D.-J., Xiao, Q.-Y., Dong, G.-M., & Yang, X.-M. (2007). A molecular epidmiological study targeting the glycoprotein gene of rabies virus isolates from China. *Virus Research, 124,* 125-138.

Mitrabhakdi, E., Shuangshoti, S., Wannakrairot, P., Lewis, R. A., Susuki, K., Laothamatas, J., & Hemachudha, T. (2005). Difference in neuropathogenetic mechanisms in human furious and paralytic rabies. *Journal of the Neurological Sciences, 238*, 3-10.

Morimoto, K., Hooper, D. C., Carbaugh, H., Fu, Z. F., Koprowski, H., & Dietzschold, B. (1998). Rabies virus quasispecies: Implications for pathogenesis. *Proceedings of the National Academy of Sciences of the USA, 95*, 3152-3156.

Moya, A., Holmes, E. C., & Gonzalez-Candelas, F. (2004). The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews Microbiology, 2*, 279-288.

Niezgoda, M., Hanlon, C., & Rupprecht, C. E. (2002). Animal Rabies. In A. C. Jackson & W. H. Wunner (Eds.), *Rabies, 2nd edn.* (pp. 163-218). Boston: Academic Press.

Ostlie, W. R., Schneider, J. M., Aldrich, T. M., Faust, R. L. B., & Chaplin, S. J. (1997). *The status of biodiversity in the Great Plains*. Minneapolis: The Nature Conservancy.

Raisanen, J. (2002). CO2-induced changes in interannual temperature and precipitation variability in 19 CMIP2 experiments. *Journal of Climatology, 15*, 2395-2411.

Real, L. A., Russell, C., Waller, L., Smith, D., & Childs, J. (2005). Spatial dynamics and molecular ecology of North American rabies. *Journal of Heredity, 96*, 253-260.

Roemer, G. W., Gompper, M. E., & Van Valkenburgh, B. (2009). The ecological role of the mammalian mesocarnivore. *BioScience, 59*, 165-173.

Rohr, J. R., & Raffel, T. R. (2010). Linking global climate change and temperature variability to widespread amphibian declines putatively caused by disease. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 8269-8274.

Rosatte, R. C. (1987). Striped, Spotted, Hooded, and Hog-nosed skunk. *Wild Furbearer Management and Conservation in North America* (pp. 598-613). Toronto, Ontario: Ontario Trappers Association.

Samson, F. B., Knopf, F. L., & Ostlie, W. R. (2004). Great Plains ecosystems: past, present, and future. *Wildlife Society Bulletin, 32*, 6-15.

Smith, J. S. (1996). New aspects of rabies with emphasis on epidemiology, diagnosis, and prevention of the disease in the United States. *Clinincal Microbiology Reviews, 9*, 166-176.

Sunquist, M. E. (1974). Winter activity of striped skunks (*Mephitis mephitis*) in east-central Minnesota. *American Midland Naturalist, 92*, 434-446.

Terrasa, B., Perez-Mellado, V., Brown, R. P., Picornell, A., Castro, J. A., & Ramon, M. M. (2009). Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*. *Diversity and Distributions, 15*, 207-221.

Tordo, N. (1996). Characteristics and molecular biology of the rabies virus. In F.-X. Meslin, M. M. Kaplan & H. Koprowski (Eds.), *Laboratory Techniques in Rabies* (4th ed.) (pp. 28-51). Geneva: WHO.

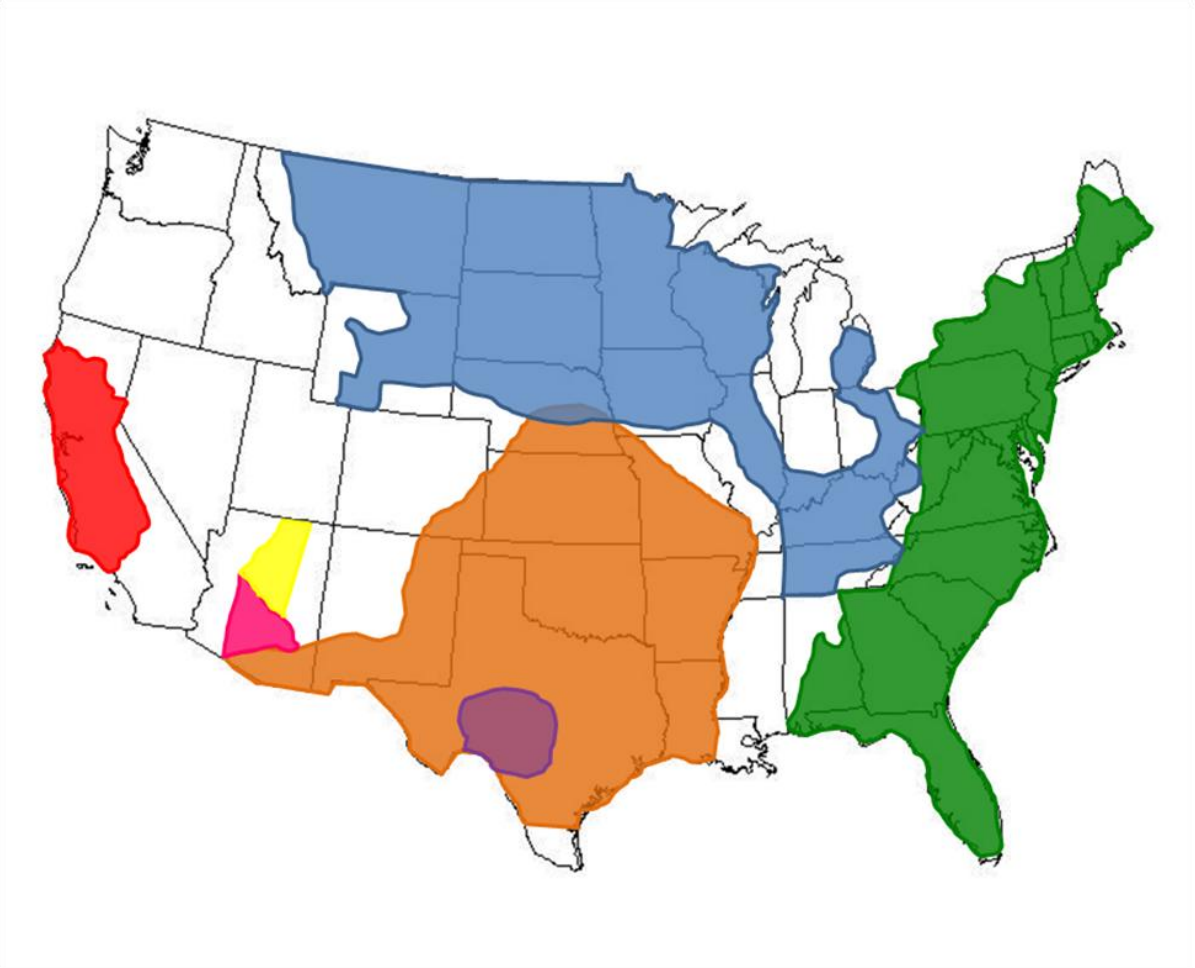
Van Heuverswyn, F., & Peeters, M. (2007). The origins of HIV and implications for the global epidemic. *Current Infectious Disease Reports, 9*, 338-346.

Vial, F., Cleaveland, S., Rasmussen, G., & Haydon, D. T. (2006). Development of vaccination strategies for the management of rabies in African wild dogs. *Biological Conservation, 131*, 180-192.

Williams, E. S., Thorne, E. T., Appel, M. J. G., & Belitsky, D. W. (1988). Canine distemper in black-footed ferrets (*Mustela nigripes*) from Wyoming. *Journal of Wildlife Diseases, 24,* 385-398.

Figures and Tables

Figure 1.1 Distribution of terrestrial rabies virus strains in the United States. Blue, orange, red, and yellow areas represent striped skunk strains; green represents raccoon strain; purple and pink represent gray fox strains; red fox strain is also present in Alaska (not shown). Adapted from Blanton et al., 2007.



CHAPTER 2 - Contrasting Landscape Epidemiology of Two Sympatric Rabies Virus Strains

Heather D. Barton, Andrew J. Gregory, Rolan Davis, Cathleen A. Hanlon, and Samantha M. Wisely

Abstract

Viral strain evolution and disease emergence are influenced by anthropogenic change to the environment. We investigated viral characteristics, host ecology, and landscape features in the rabies-striped skunk disease system of the Central Great Plains to determine how these factors interact to influence disease emergence. We amplified portions of the N and G genes of rabies viral RNA from 269 samples extracted from striped skunk brains throughout the distribution of two different rabies strains for which striped skunks were the reservoir. Because the distribution of these 2 strains overlapped on the landscape and were present in the same host population, we could evaluate how viral properties influenced epidemiological patterns in the area of sympatry. We found that South Central Skunk rabies (SCSK) exhibited intense purifying selection and high infectivity, which are both characteristics of an epizootic virus. Conversely, North Central Skunk rabies (NCSK) exhibited relaxed purifying selection and comparatively lower infectivity, suggesting the presence of an enzootic virus. The host population in the area of sympatry was highly admixed, and skunks among allopatric and sympatric areas had similar effective population sizes. Spatial analysis indicated that landscape features had minimal influence on NCSK movement across the landscape, but those same features were partial barriers to the spread of SCSK. We conclude that NCSK and SCSK have different epidemiological properties that interact differently with both host and landscape features to influence rabies

spread in the Central Great Plains. We suggest a holistic approach for future studies of emerging infectious diseases that includes studies of viral properties, host characteristics, and spatial features.

Introduction

Disease emergence and distribution are intimately connected with the landscape on which they emerge. Many emerging infectious diseases (EIDs) are caused by viruses that circulate in wildlife, and emerge due to adaptation to and amplification in a new host (Daszak *et al.*, 2001). Host switches among wildlife reservoirs or from wildlife to domestic animals facilitate new adaptive landscapes on which pathogens evolve, and each adaption event can generate suites of viral strains in the environment (McCallum & Dobson, 2002). Given that each viral strain must adapt to a unique set of hosts and environments, each strain also has the potential to exhibit unique epidemiological properties.

As viruses spread on the landscape aided by their host population, viral characteristics influence the pace and tempo of emergence on the landscape. Transmission efficiency, virulence, and incubation period are all viral characteristics that can interact with landscape features to influence virus emergence. A strain that exhibits high transmission efficiency is more likely to persist via transmission to new hosts, which gives it the potential to move further across the landscape, while a strain with poor transmission efficiency has a higher probability of extinction. Similarly, a strain with low virulence or a highly virulent strain with a long incubation period allows infected hosts to survive, disperse, migrate, or come into contact with other susceptible hosts. Conversely, a highly pathogenic strain with a short incubation period incapacitates its host before the host has a chance to transport or transmit the virus. Finally, once the host has passed through the incubation period, the severity of clinical symptoms to the viral

infection may limit the mobility of the host. Thus, the intrinsic properties of a virus coupled with the ability of the host to move through the landscape influence the emergent properties of a disease.

Rabies provides a study system to investigate how viral, host, and landscape characteristics work together to influence the emergence and expansion of disease outbreaks. Rabies is a negative sense, non-segmented RNA virus that epitomizes the characteristics of RNA viruses that are a source of many EID's. RNA viruses have mutation rates on the order of 10^{-3} to 10^{-5} mutations per site per year, very near the error threshold, which results in rapid sequence variation (Domingo & Holland, 1997). The evolutionary success of these mutations is tied closely to the selective pressures imposed by the host environment. The high mutation rate of RNA viruses allows for maximum adaptability; as host selection pressure increases, RNA viruses respond rapidly (Holland & Domingo, 1998).

Transmission of rabies is direct from host to host and the disease is fatal with very rare exception (Willoughby *et al.*, 2005). Increased aggression and bite reflex coupled with virion-filled saliva facilitates the transmission (Lafon, 2005); however, the probability of successful transmission and the incubation time (time from exposure to infective) of each host is highly variable. Viral dose, pathogenic properties of each viral variant, susceptibility of the species in general, susceptibility of the individual, severity of exposure (i.e. through multiple deep bites), and location of the exposure (proximity to the brain) are major factors which contribute to the outcome of exposure and the timing of disease onset (Rupprecht *et al.*, 2002; Shankar *et al.*, 2004). Differences in the pathogenicity of viral strains, such as duration of viral shedding, amount of virus in the saliva, and cell invasiveness (Finke & Conzelmann, 2005), likely are the result of rapid, and fine-tuned adaptation to different host species. Thus, strains with different

histories of host species reservoirs may evolve different virulence, transmissibility, or incubation periods, and these changes may manifest themselves differently on the landscape.

Terrestrial rabies in the Central Great Plains of North America presents a unique opportunity to study differences in adaptive potential and patterns of emergence of two strains of rabies with different host histories. These two strains currently share a common host reservoir, the striped skunk (*Mephitis mephitis*) and co-circulate in an area of sympatry. South Central Skunk rabies strain (SCSK), however, evolved from a North American bat virus (Real *et al.*, 2005b), while North Central Skunk rabies strain (NCSK) is most closely related to the globally distributed strain of dog rabies (Favoretto *et al.*, 2001). These two strains have been present in the Central Great Plains at least since surveillance began in 1969, and the leading edges of the two strains have been present in Nebraska since at least 2003 (Kansas State University Rabies Diagnostic Lab). The sympatric distribution of these two strains of skunk rabies allows us to test the hypothesis that different host histories and therefore different evolutionary histories have created adaptive differences in each strain, which translate into different emergent properties as rabies spreads across the landscape.

In the current study, we compared how viral characteristics, host ecology, and landscape features interact to influence viral emergence in two different strains of skunk rabies. Our study had three objectives. First, we compared the molecular epidemiological patterns of NCSK and SCSK. We hypothesized that due to the different adaptive pressures placed on these strains by their unique host histories, each strain would have a resulting unique pattern of emergence in the zone of sympatry. Because the strains were compared on the same landscape in the same host population, we could ascribe differences in viral demography and phylogeography to intrinsic differences in the viral strains. Second, we hypothesized that differences in demography of the

virus would be due in part to differences in the pathogenesis of each strain of the rabies virus as measured by antigenic diversity. The ability to invade host cells is partially controlled by the rabies virus glycoprotein, in which the glycoprotein ectodomain contains antigenic sites responsible for host cell recognition and neuroinvasiveness. Diversity in this region allows for the invasion of novel hosts and cell types. Finally, we hypothesized that intrinsic differences in the strains would translate to differences in how landscape heterogeneity influenced the pattern of disease emergence among virus strains. Studies of raccoon rabies have shown that landscape features such as large bodies of water and rivers are semi-permeable barriers to the spread of rabies on the East Coast (Real *et al.*, 2005a; Biek *et al.*, 2007; Rees *et al.*, 2008). Clinically normal raccoons and skunks are highly mobile mesocarnivores (Bixler & Gittleman, 2000) capable of swimming for long periods of time (Verts, 1967); however, the clinical manifestations of rabies include hydrophobia, disorientation, and paralysis, reducing the ability of these animals to navigate difficult terrain. If the pathogenic properties, such as incubation time or transmissibility, of each skunk rabies strain were different, we hypothesized that these differences would translate into different abilities of the virus population to expand past a river barrier.

Materials and Methods

Study Area and Surveillance History

Our study area consisted of a 5 state region of the American Central Great Plains ranging from North Dakota to Oklahoma. Rabies has been present in the Central Great Plains for at least 40 years; records of rabies-positive striped skunks exist for both Kansas (Kansas State University College of Veterinary Medicine [KSUCVM]) and South Dakota (South Dakota Department of Health [SDDH]) dating back to 1969, but rabies was likely present in those states prior to that

time. Surveillance maps show a steady increase in the range of both strains beginning in 2001, the first year for which a rabies distribution map exists (Centers for Disease Control and Prevention, 2007). The most recent map available (2007) shows an area of overlap between NCSK and SCSK in the state of Nebraska. However, based on data from this study, the actual area of sympatry between the two strains is located from northern Kansas (approximately 40º latitude) to mid-South Dakota (approximately 44º latitude), encompassing the entire state of Nebraska (Fig. 2.1A). Diagnostic data (KSUCVM Rabies Diagnostic Lab, unpublished data) indicate that both strains were present at low levels in Nebraska since 1998 but a lack of adequate surveillance and reporting contributed to underestimation of rabies cases in Nebraska until 2003, when an epidemic of striped skunk rabies occurred in the town of Grand Island. Surveillance indicates that NCSK and SCSK were sympatric by 2003, but the strains may have had shared distributions for at least the prior 5 years (KSUCVM Rabies Diagnostic Lab, unpublished data). Thus allopatric areas of the viruses are presumed to have older, more established viral populations, while the sympatric area is an emerging wavefront of the viruses (NCSK is moving south and SCSK is moving north).

Sample Collection

We analyzed a total of 269 brain samples from striped skunks that tested positive for either NCSK or SCSK. Animals were determined positive for rabies using a direct fluorescent antibody (DFA) test using a set of 7 fluorescently labeled anti-rabies monoclonal antibodies. Following incubation, samples were examined under a fluorescence microscope to visualize the presence of fluorescent green foci. We analyzed 244 samples that were submitted to the KSUCVM Rabies Diagnostic Laboratory between 2003 and 2006. In addition, the North Dakota Department of Health provided 7 samples, and the Oklahoma Department of Health provided 18

samples. For the molecular epidemiological analysis of rabies, we chose samples from North Dakota, South Dakota, Nebraska, Kansas, and Oklahoma ranging from approximately 35º to 48º latitude (Fig. 2.1A). Sample sizes per county ranged from 1 to 2 for North Dakota, 1 to 8 in South Dakota, 1 to 19 for Nebraska, 2 to 16 in Kansas, and 1 to 9 in Oklahoma. We chose a random subset of 86 samples from South Dakota, Kansas and Nebraska to analyze population structure in the host, striped skunks. This dataset was composed of 29 samples from Kansas and 29 samples from South Dakota which were outside the zone of sympatry; we also sampled 28 samples from the sympatric area of NCSK and SCSK rabies strains.

Laboratory Procedures

We isolated rabies RNA from each sample using a Trizol-Chloroform technique (Invitrogen, Inc.). We amplified approximately half of the rabies nucleoprotein (N) gene cDNA using primers 550F (5'-ATGTGYGCTAAYTGGAGYAC-3') and 304 (5'-TTGACGAAGATCTTGCTCAT-3'; Trimarchi & Nadin-Davis, 2007), and the ectodomain of rabies glycoprotein (G) gene cDNA using degenerate primers 93Gdeg (5'-ATYTACRCRATACYAGACAA-3') and 989Gdeg (5'-CTKAGACGTCTRAARCTYAC-3') through reverse transcription-PCR (RT-PCR). 5 µl of template RNA was combined with 1 µl of the RT primer (5 µM 550F), heated to 94^{o} C for 90 seconds, and immediately cooled on ice to allow for polymerization. 14 µl of RT reaction buffer (RTRX) was added to each tube and incubated at 42^{o} C for 90 minutes to complete RT. The RTRX mix contained 100 mM Tris (pH 8.3), 10 mM MgCl, 0.5 mM dNTP, 10 units of reverse transcriptase and 16 units of RNase inhibitor. Following RT, 80 µl of PCR pre-mix, which consisted of 69 µl of distilled water, 8 µl of Tris (1 M, pH 8.3), 0.5 µl Taq DNA, 1 µl of primer 550F (20 µM) and 1.5 µl of primer 304 (20 µM), was added the 20 µl RT product for a total volume of 100 µl. Cycler conditions for

PCR were 94° C for 60 seconds, 39 cycles of 94° C for 30 seconds, 42° C for 30 seconds, 72° C for 90 seconds, and a final denaturation/annealing cycle followed by a 7 minute elongation period, ending with a 4° C hold. The same procedure was followed for glycoprotein using 10 µM 93Gdeg as the RT primer, and 40 µM 93Gdeg and 40 µM 989Gdeg as the PCR primers. The resulting PCR products were purified and bi-directionally sequenced at the Advanced Genetics Technology Center (AGTC, University of Kentucky). We had one anomalous sequence, SDSK030554, which we re-extracted and re-sequenced. The resulting sequence was the same as when sequenced previously, and was therefore included in our analyses.

We extracted striped skunk DNA from each brain tissue sample using a DNeasy Blood and Tissue extraction kit (Qiagen, Inc.). Extracted DNA was amplified at eight striped skunk-specific microsatellite loci (Dragoo *et al.*, 2009) at a total volume of 10 µl, which consisted of 1 µl DNA extract, 1x PCR buffer, 2.7 mM $MgCl_2$, 0.3 mM dNTPs, 0.1 µg/µl BSA, 0.8 M Betaine, 0.2 µM forward primer, 0.5 µM reverse and dye-tagged primers, and 0.5 unit Taq. Cycler conditions were 94º C for 5 minutes, followed by 30 cycles of 94º C for 30 sec, 54º C for 45 sec, and 72º C for 45 sec, followed by 10 cycles of 94º C for 30 sec, 53º C for 45 sec, and 72º C for 45 sec, followed by 72º C for 10 minutes, ending with a 4º C hold. PCR product was visualized on a Li-Cor 4300 AutoAnalyzer (Li-Cor, Inc.) and scored at all 8 loci using the SAGA GT fragment analysis program (v.3.3; Li-Cor, Inc.). Each locus was independently amplified and visualized an average of four times.

Rabies Genetic Analysis

Sequences were edited using Sequencher 4.7 (Gene Codes Corporation, Ann Arbor, MI). Rough sequence alignment was conducted using ClustalW 2.0.10 (Larkin *et al.*, 2007) and fine alignment was conducted using Se-Al v.2.0a11 (Rambaut, 2002). The final sequence size

following alignment was 821bp for N gene and 888bp for G gene ectodomain. We determined the placement of each N gene and G gene sequence into either NCSK or SCSK using minimum-spanning haplotype networks constructed using Arlequin 3.1 (Excoffier *et al.*, 2005). To confirm the placement of SDSK030554 within NCSK, we generated a phylogenetic tree for NCSK in PAUP* (Swofford, 2003). We used a neighbor-joining method, with a tree bisection-reconnection (TBR) branch-swapping algorithm. The tree consisted of 98 NCSK samples and was rooted with four of 160 SCSK samples.

We estimated all genetic parameters separately for NCSK and SCSK N gene and G gene sequences. We analyzed NCSK and SCSK in two different ways. First, we used datasets for NCSK and SCSK that included both allopatric and sympatric samples. Then, we used only samples that were collected from the sympatric area in order to compare the properties of molecular epidemiology.

The average number of skunks infected with a unique viral isolate (a unique genetic sequence) was calculated from minimum-spanning networks as a measure of infectivity. Infectivity can be sensitive to sample bias; however, samples from each strain were collected by the same agency (USDA APHIS) personnel at the same localities, over the same period of time. We assessed nucleotide diversity, the probability that any two randomly chosen homologous DNA sites in a population will be different, and the mean number of pairwise differences, the average number of nucleotide differences among all pairs of DNA sequences, using Arlequin 3.1. To determine the amount and type of selection occurring in NCSK and SCSK, we estimated the ratio of nonsynonymous to synonymous substitutions (dN/dS) using DnaSP 5.0 (Rozas *et al.*, 2003). We also estimated the dN/dS ratio for antigenic site II (aa34-42 and 198-200) of the G gene ectodomain separately.

We conducted four different neutrality tests for N gene: Tajima's D, Fu's Fs, Fu and Li's D*, and Fu and Li's F*. Tajima's D compares the number of polymorphic sites to the pairwise number of nucleotide differences, so positive test statistics indicate diversifying selection and negative test statistics indicate purifying selection (Tajima, 1989) or selective sweeps (Braverman *et al.*, 1995). Fu's F_S compares the number of polymorphic sites to the total number of nucleotide differences, therefore positive test statistics indicate positive selection and negative test statistics indicate balancing selection (Fu, 1997). By using both methods, we could infer the mode of selection is acting on NCSK and SCSK. Fu & Li's D* and Fu & Li's F* assisted in determining if background selection was occurring in a population. These two methods compare singletons and non-singletons to determine the presence of background selection, which is the removal of deleterious mutations along with the random removal of non-deleterious mutations near the deleterious mutations (Hartl & Clark, 2007). If Fu & Li's D* and F* are significant and Fu's F_S is not, then background selection is operating in the system. If, however, Fu's F_S is significant and Fu & Li's D* and F* are not, then purifying or diversifying selection is occurring (Fu & Li, 1993). Tajima's D and Fu's Fs were determined using Arlequin 3.1 and Fu and Li's D* and F* were determined using DnaSP 5.0.

We estimated the average patristic distance between isolates, i.e. the sum of the branch lengths between the tree tips, for the N gene of each strain in both the allopatric and sympatric zones. In our study, patristic distance is a measure of within strain divergence, and takes viral evolution into account. We also estimated rabies virus demographic history and time to most recent common ancestor (TMRCA) for allopatric and sympatric samples combined using a Bayesian coalescent-based approach implemented in BEAST 1.4.8 (Drummond & Rambaut, 2007), and specified a Bayesian skyline plot as the model for both NCSK and SCSK (Drummond

et al., 2005). We analyzed each strain separately, and analyzed NCSK both with and without SDSK030554. We re-ran the analyses for 1,000,000 iterations until each scale factor was optimized to the criteria accepted by the program. The final run consisted of 10,000,000 iterations with the first 1,000,000 iterations discarded as burn-in. Bayesian skyline analysis uses an MCMC approach (Drummond *et al.*, 2002), allowing for simultaneous estimation of genealogy, nucleotide substitution rate, and demography. Bayesian skyline analysis, therefore, does not impose a specific demographic model *a priori*, as demography is one of the fitted components. We visualized the results with Tracer 1.4.1 (available from http://tree.bio.ed.ac.uk/software/tracer/).

Striped Skunk Population Analysis

To assess striped skunk population structure in the geographic distribution of each virus strain, we grouped skunk samples in two different ways. First, we grouped skunk samples according to whether they occurred in the range of NCSK or SCSK. We also grouped skunk samples according to whether they were found in the sympatric zone, the northern allopatric zone, or the southern allopatric zone. We tested for HWE using GENEPOP 3.4 (Raymond & Rousset, 1995); and to ensure that microsatellite loci were randomly associated with one another, we tested for linkage disequilibrium (GENEPOP 3.4). We estimated mean allelic richness for each population using FSTAT 2.9 (Goudet *et al.*, 2002). We also determined the global F_{ST} value with FSTAT 2.9 to determine the amount of gene flow among skunks throughout the Central Great Plains. We repeated our analyses in GENEPOP 3.4 and FSTAT 2.9 with a pooled dataset. Finally, we determined the number of segregating populations in our host population using STRUCTURE 2.3 (Pritchard *et al.*, 2000).

Because striped skunk populations are reservoirs for a contagious and fatal disease, we tested all populations for the signature of a bottleneck using BOTTLENECK 1.2.02 (Cornuet & Luikart, 1996). Because our sampling area spanned five states with considerable climatological, topographical, and epidemiological variation, we tested whether the population sizes of striped skunks were equivalent from north to south and in the area containing different viral strains. We used MIGRATE 2.3 (Beerli & Felsenstein, 1999) to estimate the relative effective population size (θ) of striped skunks in the Central Great Plains by calculating $\theta = 4N_e\mu$ for skunk populations within the range of NCSK and SCSK. We also used N_eEstimator 1.3 (Peel *et al.*, 2004) to independently estimate striped skunk effective population size.

Spatial Analysis of Host and Virus

Using microsatellite data, we conducted an analysis of molecular variance (AMOVA; Excoffier *et al.*, 2005) using Arlequin 3.1 to test whether rivers structured striped skunk populations. We defined populations as all of the samples within a county. We tested the influence of rivers to partition genetic variation in striped skunks by grouping skunks populations north or south of major rivers ($\Phi_{GROUP\text{-}RIVER}$). We tested for partitioning among populations of skunks within river-separated groups ($\Phi_{POP\text{-}GROUP}$), and within skunk populations. To define groups, only major rivers (drainage basins over 233,000 km^2) were included, resulting in 3 skunk groups separated by 2 river systems: the Missouri River (1,371,000 km^2) and the Platte River (233,099 km^2).

To investigate the influence of different habitat types on the ability of the two rabies virus strains to permeate the landscape, we compared spatially explicit measures of landscape permeability with sequence divergence estimates among isolates to determine which habitat features influenced the spread of rabies on the landscape. Because we hypothesized that host

ecology and viral properties influenced viral spread, we investigated the permeability of habitats important to striped skunks (i.e. edge habitat, urban areas [Verts, 1967; Gehrt, 2005]) and important to rabies virus (i.e. water bodies [e.g. Rees *et al.*, 2008]) to determine how the epidemiology of each strain was influenced by the landscape of the Central Great Plains.

Spatially explicit, landscape resistance to gene flow was modeled for a suite of landscape permeability hypotheses using circuit theory to connect habitat types with habitat-specific resistance values. We used CircuitScape v3.4 (McRae, 2006) with the GAP, 30 m, 2005 land cover data (United States Geological Survey, GLS Database) to generate rasterized permeability maps. We converted the 30 m resolution land cover data to 1,000 m grid data using the majority rule in Arc Info 9.3 (Environmental Systems Research Incorporated, Redlands, CA). Pairwise land cover resistance values were generated for pairs of viral populations (viral isolates pooled by county [Fig. 2.1B]). Rabies-positive sample locations were recorded only for the county level; therefore, we used the county centroid from the county in which samples were obtained. In cases where multiple samples were obtained from the same county, sample locations were adjusted so that they differed by a distance of 1 m.

We tested 4 hypothesized models of landscape resistance by assigning each habitat type a resistance value that ranged from 0 to 10. Values closer to 0 represented habitats through which the virus spread more easily and therefore had less resistance, and values closer to 10 indicated habitats with greater resistance to viral spread (McRae, 2006). When modeling isolation by resistance, absolute resistance values are less important than the relative ratio between values (McRae, 2006). The first model tested rivers as barriers to dispersal (RB). This model assumed that all land cover types were equally suited for rabies spread, but imposed a moderate resistance, except for rivers which acted as strong barriers (resistance = 10). Studies of raccoon

rabies indicate that bodies of water appear to impede the spread of raccoon rabies (Biek *et al.*, 2007; Rees *et al.*, 2008; Cullingham *et al.*, 2009), thus rivers may also impede the spread of striped skunk rabies. The second model took host habitat affiliation into account. Striped skunks forage in agricultural fields and grasslands, but use forest-grassland edge habitat for dispersal and movement between areas (Bixler & Gittleman, 2000). Striped skunks are also an urban-adapted species, and are quite abundant in urban and suburban landscapes where resources are abundant (Randa & Yunger, 2006). Therefore, this model assumed that the landscape was uniformly difficult to cross, but that dispersal corridors existed along edge habitat (Forest x Grassland, Urban x Grassland, and Row Crop Agriculture x Grassland). This edge model (EM) assumed a resistance of 4 for all non-edge habitats and resistance of 1 for all edges. The third model combined the RB model and the EM model. This was our combined model (CM), which assumed that all continuous habitat patches had a resistance of 4, edges had a resistance of 1, and rivers posed barriers and had a resistance of 10. The fourth model was a simple isolation by resistance (SIBR) model. SIBR was used to characterize the effect that habitat shape and distance had on model output and imposed an intermediate resistant value (4) uniformly on all habitat types. In biological terms, SIBR has similar assumptions to isolation by distance (IBD) and assumed that all land cover types were equal in their ability to enhance or inhibit dispersal. SIBR differed from IBD in that the circuit theory approach averaged the resistance across all possible paths (created by the 1 km habitat grid) between demes and accounted for some hindrance of dispersal imposed on rabies by the shape of the surrounding landscape. We tested these hypothesized models against the null model of IBD (Slatkin, 1993) which assumed that genetic differentiation was a function of Euclidean distance alone, and that habitat had a negligible impact on viral spread. We inferred that if viral emergence was not affected by

habitat specific landscape permeability and composition (RB, EM, or CM models), or by landscape shape (SIBR model), then differentiation of rabies sequences would simply increase with Euclidean distance (IBD).

We assessed the influence of these spatial characteristics of the landscape on viral genetic distance using Mantel tests (SAS Institute Incorporated, SAS Campus Drive, Cary, NC, USA); we used partial Mantel tests with IBD as a covariate for the RB, EM, and CM models, and Mantel tests for SIBR and IBD models. To select the best Mantel model, we used AICc with correction for small sample size bias (Burnham & Anderson, 1998). AICc values were calculated from maximum likelihood estimates extracted from SAS output. The number of parameters (K) included in model selection was determined by the number of land cover resistance classes included in the GIS land cover ecological resistance model for RB, EM, and CM models, or in the case of IBD and SIBR by the number of distance matrices included in the Mantel test. For covariate models (IBD x RB and IBD x RB x EM), K was determined by summing the K values for all contributing models plus 1 for the regression intercept of the covariate models.

Using viral sequence data, we conducted an AMOVA to confirm the results of the IBR analysis. We conducted the AMOVA using Arlequin 3.1 using the same model definition as the AMOVA of skunk populations. As with skunks, virus populations were defined by samples within a county. Because of our larger sample size for virus sequence data than skunk microsatellite marker data, we included viral populations to the south of the Arkansas River (drainage basin = 505,000 km^2), for a total of 4 virus groups separated by 3 rivers.

Results

Virus Molecular Epidemiology

Comparison of N gene datasets (GenBank Accession Numbers: HM113685-HM113942) of each rabies strain using samples from both allopatric and sympatric areas revealed similar levels of nucleotide diversity between the two strains, although the mean number of pairwise differences was twice as high for NCSK than for SCSK (Table 2.1). Neutrality test results and the ratio of nonsynonymous to synonymous site substitutions indicated that purifying selection was acting on both strains. NCSK, however, exhibited relaxed purifying selection compared to SCSK (Table 2.1). Additionally, the number of infected skunks per rabies isolate was slightly higher for SCSK than for NCSK (Table 2.1), suggesting that SCSK had higher infectivity than NCSK.

Differences between the two strains were more apparent when we compared NCSK and SCSK N gene samples from the sympatric area. The mean number of pairwise differences in NCSK increased from twice that of SCSK to over three times larger than SCSK. Purifying selection became even more intense for SCSK and slightly more relaxed for NCSK (Table 2.1). Finally, while infectivity did not change for NCSK, it was higher in SCSK (Table 2.1).

Comparison of complete G gene datasets (GenBank Accession Numbers: HM113943-HM114211) of each rabies strain from allopatric and sympatric areas combined also revealed similar levels of genetic diversity and infectivity between the two strains (Table 2.2). The ratio of nonsynonymous to synonymous site substitutions indicated that purifying selection was acting on the ectodomain of the G gene. However, dN/dS values were opposite of those exhibited by the N gene; SCSK exhibited more relaxed purifying selection compared to NCSK (Table 2.2). When only samples from the sympatric zone were compared, genetic diversity and infectivity

between NCSK and SCSK G gene samples remained similar, but purifying selection became even more intense for NCSK and slightly more relaxed for SCSK (Table 2.2). Antigenic site II of the ectodomain showed positive selection for both strains (SCSK dN/dS = 1.97; NCSK dN/dS = 1.74).

When we compared patristic distance of isolate pairs between the allopatric zone and the sympatric zone for each strain, average branch length was longer in the sympatric zone than in the allopatric zone in both cases. However, average patristic distance in the sympatric zone was over 5 times longer for NCSK than for SCSK (NCSK: Allopatric = 5.9 mutations, Sympatric = 16.36 mutations; SCSK: Allopatric = 2.05 mutations, Sympatric = 3.4 mutations). Bayesian skyline analysis for allopatric and sympatric samples combined showed the TMRCA for NCSK including sample SDSK030554 as ranging from 450 to 2150 years ago, and NCSK without SDSK030554 as ranging from 50 to 550 years ago. The TMRCA for SCSK ranged from 75 to 205 years ago. The phylogenetic tree generated by PAUP indicated that SDSK030554 did belong in the NCSK clade, but was on a long branch separate from the rest of the sequences in the clade (Figure 2.2; Appendix A). While the skyline plots for NCSK and SCSK showed a steady number of infections over time (300 and 100 infections respectively), the confidence intervals for the plots overlapped considerably, suggesting no differences in the viral population size between NCSK and SCSK.

Striped Skunk Host Ecology

Analysis of microsatellite markers in striped skunks indicated that none of the loci exhibited linkage disequilibrium ($P > 0.05$ after Bonferroni correction). Mean allelic richness and heterozygosity were similar for skunks infected with NCSK or SCSK (Table 2.3). We found significant heterozygote deficiency at seven of eight loci (Table 2.3), which is expected in a

system where population density is periodically perturbed by epizootic events. In support of this inference, a Wilcoxon Sign-Rank test in program BOTTLENECK 1.2.02 under the two-phase mutation (TPM) model was significant for each host population affected by NCSK or SCSK. A Wilcoxon Sign-Rank test under the TPM model for the pooled dataset was also significant (Table 2.3) suggesting a past population reduction. The global F_{ST} was 0.006 ± 0.003, which indicated that there were high levels of connectivity among skunks throughout the Central Great Plains. θ (Table 2.3) for striped skunk populations in the northern allopatric area, the sympatric area, or the southern allopatric area was not significantly different among populations (Allopatric North θ = 5.49, Sympatric θ = 7.27, Allopatric South θ = 7.53).

Results of the AMOVA for striped skunks microsatellite markers indicated that rivers do not act as barriers to striped skunk gene flow. $\Phi_{GROUP\text{-}RIVER}$ accounted for virtually none of the variation (Φ_{CT} = 0.002, P = 0.15), and $\Phi_{POP\text{-}GROUP}$ accounted for approximately 2% of the variation (Φ_{SC} = 0.022, P = 0.17). The majority of the variation (Φ_{ST} = 0.023, P = 0.14) was explained by within population variation. Pairwise F_{ST} for groups on either side of the Missouri river was 0.005 (P = 0.29), and the pairwise F_{ST} for groups on either side of the Platte river was 0.011 (P = 0.03). These values translate to 23 and 51 effective migrants per generation respectively, so while the F_{ST} for the Platte River was statistically significant, we found a large level of gene flow between groups on either side of the river. Results from program STRUCTURE indicated K=1 as the number of populations in the area sampled. This interpretation is based on the ΔK distribution, the modal value of which tends to be located at the actual K (Evanno *et al.*, 2005). Results of population structure analyses of striped skunks suggested no population sub-structuring, large levels of gene flow, and genetic admixture in the Central Great Plains.

Landscape Epizootiology

NCSK and SCSK responded differently to landscape heterogeneity in the zone of sympatry. In our landscape resistance models, the greatest model support for the cause of genetic differentiation among SCSK rabies isolates was found to be a function of an interaction between rivers as barriers and Euclidean distance (Table 2.4). Conversely, NCSK seemed to move independently of rivers and showed only a weak indication of Euclidean distance as a barrier to dispersal (Table 2.4). Edge habitat did not appear to influence rabies emergence on the landscape for either strain.

The AMOVA confirmed that each strain possessed different abilities to cross rivers. The AMOVA for NCSK revealed no significant partitioning of viral isolates by rivers ($\Phi_{GROUP\text{-}RIVER}$; Φ_{CT} = -0.007, P = 1.00). Results for SCSK, however, revealed that rivers were a significant barrier and accounted for approximately 26% of the variation in that strain ($\Phi_{GROUP\text{-}RIVER}$; Φ_{CT} = 0.26, P = 0.08).

Discussion

Comparisons of NCSK and SCSK in the area of sympatry indicated substantial differences in molecular epidemiology, evolutionary trends, and geographic patterns of emergence between the two strains. These differences were masked when allopatric datasets were included. In host populations, we found high levels of admixture and similar effective population sizes among allopatric and sympatric zones, suggesting that differences between NCSK and SCSK emergence patterns are driven by strain-specific epidemiology. Our results suggest that differences in viral properties between the two strains are interacting with host attributes and landscape features to shape unique disease emergence patterns.

Landscape Influence on Pathogen and Host

Landscape heterogeneity has been shown to influence the epidemiology of a variety of pathogens and parasites, including rabies (Real & Biek, 2007; Rees *et al*., 2008; Wheeler & Waller, 2008), influenza (Matthews & Haydon, 2007), infestations of the fungus *Phytophthora lateralis* (Kauffman & Jules, 2006), and infections by trematodes of the genus *Microphallus* (Lively, 1999). Rivers represent a defining geographical feature in the Central Great Plains, but the influence of rivers as a barrier to the emergence of rabies has been debated. Some studies of host ecology show that rivers do (Cullingham *et al.*, 2009) or do not (Arjo *et al.*, 2008) influence the population structure of raccoons, the main host of rabies on the east coast of the United States. However, spatial modeling predicts that river crossings should slow rabies transmission (Real *et al.*, 2005b), and case data show that river crossings can slow the spread of rabies across rivers by a range of 11 to 16 months (Lucey *et al.*, 2002; Smith *et al.*, 2002). The effect of rivers on rabies may also be scale-dependent. On the east coast of the United States, the Susquehanna River and Chesapeake Bay have been shown to confine and direct the flow of raccoon rabies on a local scale, but have a minimal effect at larger scales (Biek *et al.*, 2007). We found that rivers minimally influenced striped skunk population structure and gene flow, but differentially affected the movement of each rabies strain. Although rivers presented no barrier to the spread of NCSK, they were a semi-permeable barrier to the spread of SCSK in the Central Great Plains. Because both host and landscape were the same in the area of sympatry, we suspect that differences in patterns of viral spread among the two strains are a result of differences in host-pathogen interactions of each strain.

Differences in the intrinsic properties of each strain can interact with the host in unique ways. Strain-specific differences in virulence, transmission efficiency, or incubation period could explain the differences in the ability of each strain to spread across the rivers. Differences

in virulence would allow a more pathogenic strain to kill the host prior to transmission, while a more attenuated strain would allow a higher proportion of the population to recover. We found fewer cases of SCSK crossing rivers, which could be explained by differences in virulence. While this is a plausible hypothesis, there is no evidence of differences in virulence among strains; however, pathogenic properties of wild type strains have rarely been studied. Alternatively, differences in transmission efficiency may drive differences in the emergent properties of each strain. Many directly transmitted pathogens, including rabies, rely on transmission of infectious particles between host individuals in order to propagate (Merrell & Falkow, 2004). As transmission efficiency increases, the rate of successful secondary infections per infected individual also increases. If transmission was more efficient in NCSK than in SCSK, it would also explain the observation that more individuals infected with NCSK cross rivers. Our data, however, suggest the contrary; unique isolates of SCSK infected more hosts than did NCSK.

We also considered that the duration of the asymptomatic period in each strain may drive differences in landscape epidemiology. The incubation period of an incapacitating disease influences how far an infected host is able to travel. Incubation period varies both among rabies strains (Carey & McLean, 1983), and within strains (Parker & Wilsnack, 1966). Incubation period for rabies can be variable by strain due to adaptations to its host (Carey & McLean, 1983) or due to inoculant load (Parker & Wilsnack, 1966). The fact that rivers are barriers to SCSK but not NCSK could be explained by a longer incubation period in NCSK than SCSK. Thus, if SCSK-infected skunks have a shorter incubation period than NCSK-infected skunks, they will have fewer opportunities to cross rivers prior to presenting clinical signs of illness. Although experimental transmission studies are needed to definitively determine differences in viral

pathogenesis among strains, the comparative molecular epidemiology of these strains provides some additional insight into the pathogenic properties of each strain.

Molecular Epidemiology of the Pathogen

SCSK and NCSK viral characteristics are the product of their unique evolutionary histories as they pass from one reservoir host to another; these species-specific environments lead to strains that exhibit unique viral characteristics. In molecular epidemiology, the N gene has been used to represent phylogenetic history of viral emergence (Davis *et al.*, 2006). The N gene is highly conserved and thus thought to be under constant purifying selection (Holmes *et al.*, 2002). Therefore, we interpret the low diversity coupled with short patristic distances of the N gene in the SCSK strain as evidence of serial selective sweeps that allowed for a rapid spread of the virus throughout the sympatric zone. The rapid expansion of similar isolates is further supported by the infectivity data. Unique SCSK isolates (as determined from N gene sequences) infected twice as many skunks as viral isolates of NCSK, suggesting that SCSK is more transmissible than NCSK.

Molecular clock data suggest that NCSK has been on the landscape for at least 100 years longer than SCSK. If we include the anomalous sequence from NCSK, then our data suggest that the strain has been present for much longer. This isolate may be a remnant haplotype which has persisted from a past lineage. The demographic and molecular clock data suggest that SCSK has higher infectivity and transmissibility, which characterizes an epizootic strain (one that exceeds the number of expected cases in a given time period) that quickly moves through the landscape. NCSK, by contrast, appears to be less transmissible, with longer persistence in the host population and behaves like a more enzootic strain of rabies (one that is maintained in a population at sustainably low levels).

These inferences are supported by molecular epidemiological data from the G gene. Although isolates of SCSK were similar within a given geographic region (as inferred by N gene data), the antigenic diversity was very high. The G gene in SCSK had a five-fold higher amino acid substitution rate in the antigenic region of the glycoprotein than NCSK. A high dN/dS ratio in RNA viruses is indicative of high viral adaptability, increased cell invasiveness, and increased transmission potential as seen in the antigenic regions of HIV glycoproteins (Rong *et al.*, 2007). In other viruses, a high frequency of nonsynonymous changes is positively correlated with increased passage among hosts (Cuevas *et al.*, 2009). The higher antigenic diversity of SCSK provides some indirect evidence for our hypothesis that SCSK has higher infectivity and a shorter incubation period (due to more efficient cell-to-cell transfer). Though this is not definite evidence of differential pathogenesis among strains, it does support our hypotheses.

These intrinsic differences between rabies viral strains translate into differences at the landscape scale of rabies epidemiology. SCSK is partially restricted by landscape heterogeneity due to a hypothesized short incubation period; however, once the viral strain crosses rivers it rapidly infects many host animals with similar isolates. NCSK, with its potentially longer incubation period and lower infectivity, moves slowly across the landscape without habitat barriers and persists on the landscape for a longer time period.

Population Structure of the Reservoir Host

Striped skunks are the major rabies reservoir in the Central Great Plains, and have been interacting with one or both rabies strains for decades. Despite these long-term interactions, we found that host population structure was not predictive of rabies population structure. The low pairwise F_{ST} values and lack of evidence that rivers were barriers in our study demonstrate virtually unrestricted striped skunk movement, unlike our findings for SCSK strain. Skunks in

the Great Plains appear to be highly admixed, likely a result of high dispersal capacity, high reproductive output, and the confounding influence of demographic history. Indeed, striped skunks are highly mobile mesocarnivores, especially during mating season (Greenwood *et al.*, 1997; Bixler & Gittleman, 2000), and are capable of swimming for hours (Verts, 1967) and crossing bodies of water when necessary (Wilber & Weidenbacher, 1961; Olson & Lewis, 1999). Our findings indicate that a large, admixed striped skunk population exists in the Central Great Plains, and harbors two distinct strains of rabies virus. Populations in both the allopatric and sympatric zones were situated in a variety of habitat types; yet all three populations were of similar size, suggesting that population size was not affected by ecological differences associated with latitude, or that those differences were obscured by high levels of dispersal. The highly mobile, large, and admixed nature of this host population makes it an excellent host for this directly transmitted and fatal virus.

Conclusions

Many host-parasite studies emphasize using host data to predict parasite or pathogen spread (McCoy *et al.*, 2005; Berrada *et al.*, 2006; Kauffman & Jules, 2006). We, however, took a holistic approach to determine how virus properties, host attributes, and landscape features interact to influence disease emergence. The high infectivity and intense purifying selection in SCSK N gene suggests a more transmissible strain, whereas the lower infectivity and comparatively relaxed purifying selection in NCSK N gene is indicative of an enzootic strain. The relaxed purifying selection in SCSK G gene compared to the intense purifying selection in NCSK G gene is indicative of a more transmissible strain in the south.

While molecular characteristics of each strain point to differences in pathogenic properties, definitive comparative studies of transmission are needed. Nonetheless, it appears

that viral properties, host attributes, and landscape features may be interacting with one another to shape strain-specific rabies emergence patterns throughout the Central Great Plains.

References

Arjo, W.M., Fisher, C.E., Armstrong, J., Boyd, F. & Slate, D. (2008) Effects of natural barriers and habitat on the western spread of raccoon rabies in Alabama. *Journal of Wildlife Management,* 72**,** 1725-1735.

Beerli, P. & Felsenstein, J. (1999) Maximum likelihood estimation of migration rates and population numbers of two populations using a coalescent approach. *Genetics,* 152**,** 763-773.

Berrada, Z.L., Goethert, H.K. & Teleford, S.R. (2006) Raccoons and skunks as sentinels for enzootic tularemia. *Emerging Infectious Diseases,* 12**,** 1019-1021.

Biek, R., Henderson, J.C., Waller, L., Rupprecht, C.E. & Real, L.A. (2007) A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proceedings of the National Academy of Sciences of the USA,* 104**,** 7993-7998.

Bixler, A. & Gittleman, J.L. (2000) Variation in home range and use of habitat in the striped skunk (*Mephitis mephitis*). *Journal of Zoology, London,* 251**,** 525-533.

Braverman, J.M., Hudson, R.R., Kaplan, N.L., Langley, C.H. & Stephan, W. (1995) The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics,* 140**,** 783-796.

Burnham, K.P. & Anderson, D.R. (1998) *Model Selection and Inference: A practical information-theoretic approach,* New York, Springer-Verlag.

Carey, A.B. & McLean, R.G. (1983) The Ecology of Rabies - Evidence of Co-Adaptation. *Journal of Applied Ecology,* 20**,** 777-800.

Centers for Disease Control and Prevention (2007) *Rabies Epidemiology*. [Online] Available from: http://www.cdc.gov/rabies/epidemiology.html [Accessed 5 November 2009]

Cornuet, J.M. & Luikart, G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics,* 144**,** 2001-2014.

Cuevas, J.M., Moya, A. & Sanjuan, R. (2009) A genetic background with low mutational robustness is associated with increased adaptability to a novel host in an RNA virus. *Journal of Evolutionary Biology,* 22**,** 2041-2048.

Cullingham, C.I., Kyle, C.J., Pond, B.A., Rees, E.E. & White, B.N. (2009) Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Molecular Ecology,* 18**,** 43-53.

Daszak, P., Cunningham, A.A. & Hyatt, A.D. (2001) Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Tropica,* 78**,** 103-116.

Davis, P.L., Bourhy, H. & Holmes, E.C. (2006) The evolutionary history and dynamics of bat rabies virus. *Infection, Genetics and Evolution,* 6**,** 464-473.

Domingo, E. & Holland, J.J. (1997) RNA virus mutations and fitness for survival. *Annual Review of Microbiology,* 51**,** 151-178.

Dragoo, J.W., Coan, K.E., Moore, K.A., *et al.* (2009) Polymorphic microsatellite markers for the striped skunk, *Mephitis mephitis,* and other mephitids. *Molecular Ecology Resources,* 9**,** 383-385.

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G. & Solomon, W. (2002) Estimating mutation parameters, population history, and genealogy simultaneously from temporally spaced sequence data. *Genetics,* 161**,** 1307-1320.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology,* 7:214.

Drummond, A.J., Rambaut, A., Shapiro, B. & Pybus, O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution,* 22**,** 1185-1192.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,* 14**,** 2611-2620.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* 1**,** 47-50.

Favoretto, S.R., De Mattos, C.C., Morais, N.B., Araújo, F.A.A. & De Mattos, C.A. (2001) Rabies in Marmosets (*Callithrix jacchus*), Ceará, Brazil. *Emerging Infectious Diseases,* 7**,** 1062-1065.

Finke, S. & Conzelmann, K.-K. (2005) Replication strategies of rabies virus. *Virus Research,* 111**,** 120-131.

Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics,* 147**,** 915-925.

Fu, Y.-X. & Li, W.-H. (1993) Statistical tests of neutrality of mutations. *Genetics,* 133**,** 693-709.

Gehrt, S.D. (2005) Seasonal survival and cause-specific mortality of urban and rural striped skunks in the absence of rabies. *Journal of Mammalogy*, 86, 1164-1170.

Goudet, J., Perrin, N. & Waser, P. (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology,* 11**,** 1103-1114.

Greenwood, R.J., Newton, W.E., Pearson, G.L. & Schamber, G.J. (1997) Population and movement characteristics of radio-collared striped skunks in North Dakota during an epizootic of rabies. *Journal of Wildlife Diseases,* 33**,** 226-241.

Hartl, D.L. & Clark, A.G. (2007) *Principles of Population Genetics,* 4th edn. Sinauer Associates, Inc., Sunderland, Massachussets.

Holland, J. & Domingo, E. (1998) Origin and Evolution of Viruses. *Virus Genes,* 16**,** 13-21.

Holmes, E.C., Woelk, C.H., Kassis, R. & Bourhy, H. (2002) Genetic constraints and the adaptive evolution of rabies virus in nature. *Virology,* 292**,** 247-257.

Kauffman, M.J. & Jules, E.S. (2006) Heterogeneity shapes invasion: Host size and environment influence susceptibility to a nonnative pathogen. *Ecological Applications,* 16**,** 166-175.

Lafon, M. (2005) Modulation of the immune response in the nervous system by rabies virus. *Current Topics in Microbiology and Immunology,* 289**,** 239-258.

Larkin, M.A., Blackshields, G., Brown, N.P., *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics,* 23, 2947-2948.

Lively, C.M. (1999) Migration, virulence, and the geographic mosaic of adaptation by parasites. *The American Naturalist,* 153**,** S34-S37.

Lucey, B.T., Russell, C.A., Smith, D., *et al.* (2002) Spatiotemporal analysis of epizootic raccoon rabies propagation in Connecticut, 1991-1995. *Vector Borne and Zoonotic Diseases,* 2**,** 77-86.

Matthews, L. & Haydon, D. (2007) Introduction. Cross-scale influences on epidemiological dynamics: from genes to ecosystems. *Journal of the Royal Society Interface,* 4**,** 763-765.

McCallum, H. & Dobson, A. (2002) Disease, habitat fragmentation and conservation. *Proceedings of the Royal Society of London, Series B,* 269**,** 2041-2049.

McCoy, K.D., Boulinier, T. & Tirard, C. (2005) Comparative host-parasite population structures: Disentangling prospecting and dispersal in the black-legged kittiwake *Rissa tridactyla*. *Molecular Ecology,* 14**,** 2825-2838.

McRae, B.H. (2006) Isolation by resistance. *Evolution,* 60**,** 1551-1562.

Merrell, D.S. & Falkow, S. (2004) Frontal and stealth attack strategies in microbial pathogenesis. *Nature,* 430**,** 250-256.

Olson, R., & Lewis, A.M. (1999) Skunk ecology and damage management techniques for homeowners. Bulletin B-1075, Cooperative Extension Service, University of Wyoming, 16 pp.

Parker, R.F. & Wilsnack, R.F. (1966) Pathogenesis of skunk rabies virus: quantitation in skunks and foxes. *American Journal of Veterinary Research,* 27**,** 33-38.

Peel, D., Ovenden, J.R. & Peel, S.L. (2004) NeEstimator: software for estimating effective population size, Version 1.3., Queensland Government, Department of Primary Industries and Fisheries.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure from multilocus genotype data. *Genetics,* 155**,** 945-959.

Rambaut, A. (2002) SE-AL Sequence Alignment Editor v.2.0a11. University of Oxford, Oxford, UK.

Randa, L.A. & Yunger, J.A. (2006) Carnivore occurrence along an urban-rural gradient: A landscape-level analysis. *Journal of Mammalogy,* 87**,** 1154-1164.

Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity,* 86**,** 248-249.

Real, L.A. & Biek, R. (2007) Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. *Journal of the Royal Society Interface,* 4**,** 935-948.

Real, L.A., Henderson, J.C., Biek, R., *et al.* (2005a) Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. *Proceedings of the National Academy of Sciences of the USA,* 102**,** 12107-12111.

Real, L.A., Russell, C., Waller, L., Smith, D. & Childs, J. (2005b) Spatial dynamics and molecular ecology of North American rabies. *Journal of Heredity,* 96**,** 253-260.

Rees, E.E., Pond, B.A., Cullingham, C.I., *et al.* (2008) Assessing a landscape barrier using genetic simulation modelling: Implications for raccoon rabies management. *Preventative Veterinary Medicine,* 86**,** 107-123.

Rong, R., Gnanakaran, S., Decker, J.M., *et al.* (2007) Unique mutational patterns in the envelope alpha 2 amphipathic helix and acquisition of length in gp120 hypervariable domains are associated with resistance to autologous neutralization of subtype C human immunodeficiency virus type 1. *Journal of Virology,* 81**,** 5658-5668.

Rozas, J., Sánchez-Del Barrio, J.C., Messeguer, X. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics,* 19**,** 2496-2497.

Rupprecht, C.E., Hanlon, C.A. & Hemachudha, T. (2002) Rabies re-examined. *The Lancet Infectious Diseases,* 2**,** 327-343.

Shankar, V., Bowen, R.A., Davis, A.D., Rupprecht, C.E. & O'Shea, T.J. (2004) Rabies in a captive colony of big brown bats (*Eptesicus fuscus*). *Journal of Wildlife Diseases,* 40**,** 403-413.

Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution,* 47, 264-279.

Smith, D.L., Lucey, B., Waller, L.A., Childs, J.E. & Real, L.A. (2002) Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 3668-3672.

Swofford, D.L. (2003) *PAUP 4.0 beta version for the Macintosh PowerPC - Phylogenetic analysis using parsimony (and other methods),* New York, NY, Sinauer Associates.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* 123**,** 585-595.

Trimarchi, C.V. & Nadin-Davis, S.A. (2007) Diagnostic Evaluation. In: *Rabies,* 2nd edn. (eds Jackson, A.C. & Wunner, W.H.), pp. 411-469. Academic Press, New York.

Verts, B.J. (1967) *The biology of the striped skunk,* Urbana, Illinois, University of Illinois Press.

Wheeler, D.C. & Waller, L.A. (2008) Mountains, valleys, and rivers: The transmission of raccoon rabies over a heterogeneous landscape. *Journal of Agricultural, Biological, and Environmental Statistics,* 13**,** 388-406.

Wilber, C.G. & Weidenbacher, G.H. (1961) Swimming capacity of some wild mammals. *Journal of Mammalogy,* 42**,** 428-429.

Willoughby, R.E., Tieves, K.S., Hoffman, G.M., *et al.* (2005) Survival after treatment of rabies with induction of coma. *New England Journal of Medicine,* 352**,** 2508-2514.

Figures and Tables

Figure 2.1 Map of the central Great Plains study area. A. Sampled counties are patterned according to which rabies strain is present. Counties where both strains were present are shaded gray. The leading edge of each strain is indicated by black dashed lines, which also delineate the zone of sympatry. B. CircuitScape study area representing the River Barrier Model. Open circles represent SCSK samples and closed circles represent NCSK samples.

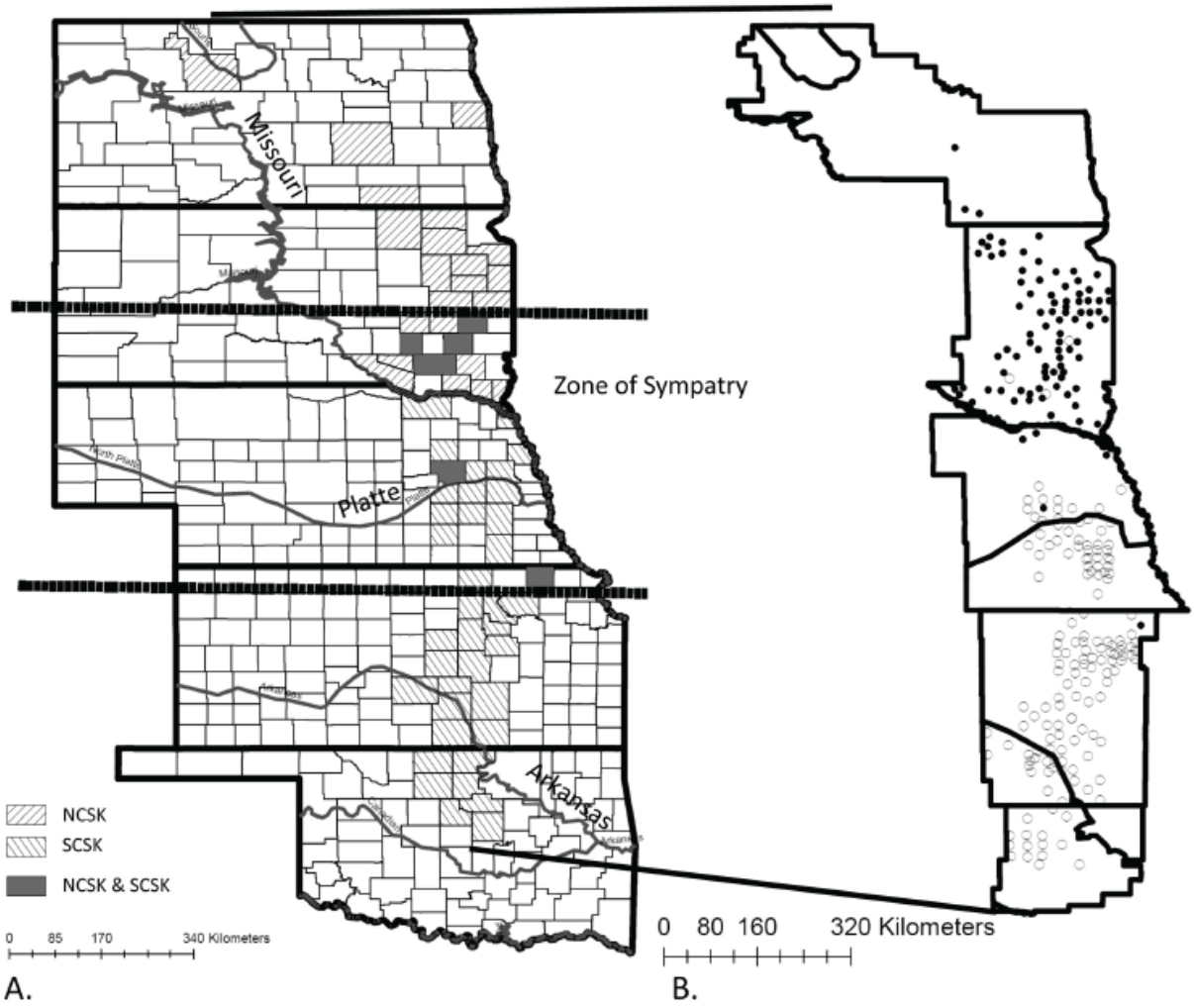


Figure 2.2 Phylogram of NCSK N gene implemented using a neighbor-joining method in PAUP* and rooted with four SCSK N gene samples. Terminal triangles indicate collapsed branches. Numbers in parentheses indicate the number of samples associated with each geographic state found at a branching location. Numbers on branches indicate bootstrap support for the branch. NCSK and SCSK are separated into two different clades. One sequence, SDSK030554, while a member of the NCSK clade, is on a long branch and distant from the rest of the samples, indicating that it may be a remnant of a past rabies epizootic.

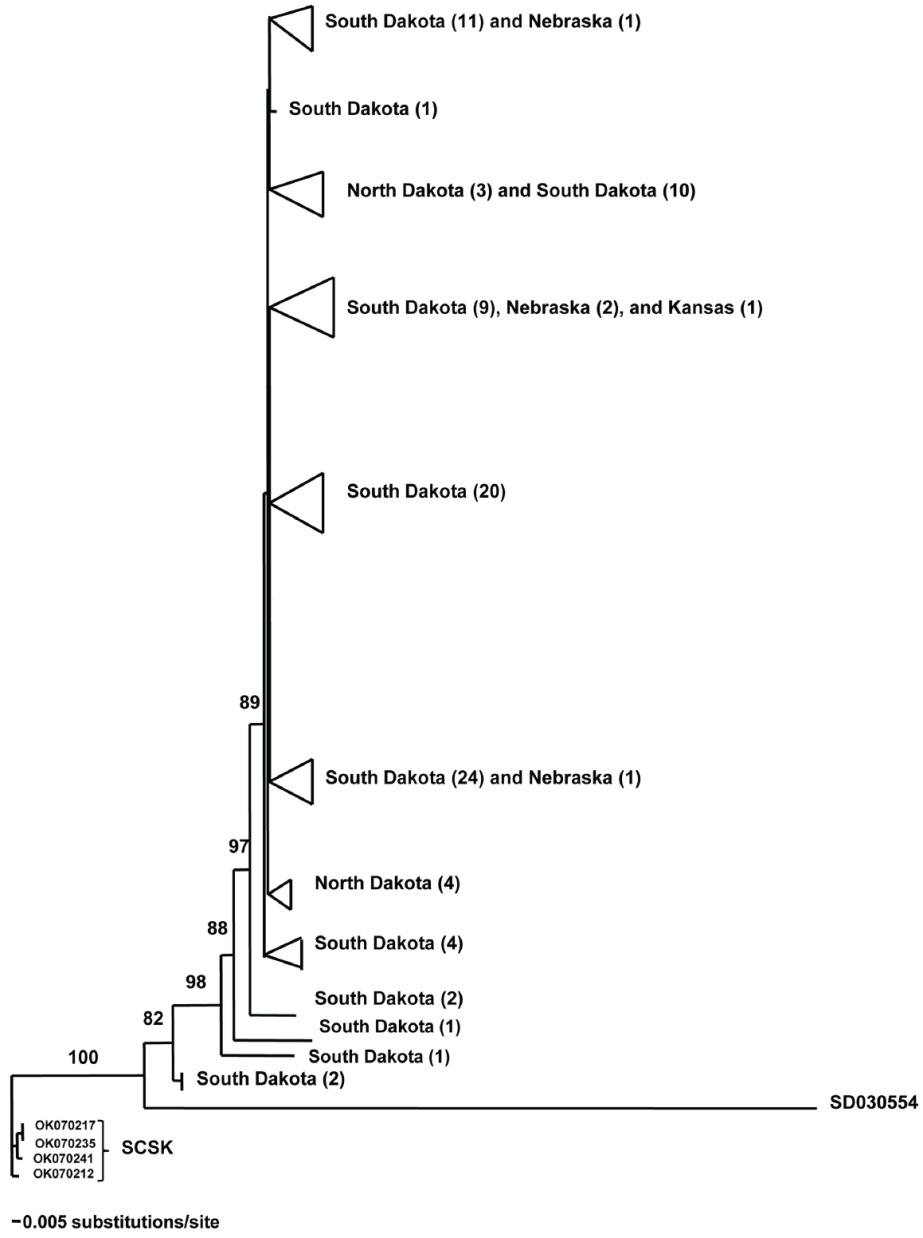


Table 2.1 Diversity and neutrality statistics for NCSK and SCSK using complete N gene datasets containing both allopatric and sympatric samples for each strain suggested the presence of similar levels of diversity in both strains, but slightly higher infectivity and more purifying selection in SCSK. The same analysis using only samples from the sympatric zone for each strain revealed marginal differences in diversity between the two strains, but more intense purifying selection and twice as much infectivity in SCSK when compared to NCSK.

	Sympatric and Allopatric		Sympatric Only	
	NCSK (N=98)	SCSK (N=160)	NCSK (N=51)	SCSK (N=60)
Nucleotide Diversity	0.03 ± 0.02	0.02 ± 0.01	0.05 ± 0.02	0.02 ± 0.01
Pairwise Differences	26.26 ± 11.61	11.93 ± 5.42	37.00 ± 16.36	10.86 ± 5.01
Fu & Li's D*	-7.11**	-4.17**	-5.56*	-3.39*
Fu & Li's F*	-6.22**	-3.99**	-5.35*	-3.61*
Tajima's D	-2.56***	-2.27***	-2.52***	-2.36***
Fu's Fs	-23.95**	-24.12**	-23.79***	-24.52***
dN/dS	0.57	0.37	0.66	0.30
# skunks/isolate	1.2	1.8	1.2	2.4

*P<0.05 **P<0.02 ***P<0.001

Table 2.2 Diversity and neutrality statistics for NCSK and SCSK using complete G gene datasets containing both allopatric and sympatric samples for each strain suggested the presence of similar levels of diversity and infectivity in both strains, but more purifying selection in NCSK. The same analyses using only samples from the sympatric zone for each strain revealed more intense purifying selection in the NCSK when compared to SCSK.

	Sympatric and Allopatric		Sympatric Only	
	NCSK (N=103)	SCSK (N=166)	NCSK (N=51)	SCSK (N=60)
Nucleotide Diversity	0.02 ± 0.01	0.04 ± 0.02	0.02 ± 0.01	0.03 ± 0.01
Pairwise Differences	17.68 ± 7.92	35.49 ± 15.53	16.66 ± 7.54	21.09 ± 9.44
Fu & Li's D*	-4.76**	-7.08**	-3.61**	-4.90**
Fu & Li's F*	-4.49**	-5.91**	-3.64**	-4.79**
Tajima's D	-2.34***	-2.32***	-2.02**	-2.07**
Fu's Fs	-24.01***	-23.80***	-24.22***	-24.11***
dN/dS	0.088	0.22	0.077	0.56
# skunks/isolate	0.86	0.94	1.13	1.07

*P<0.05 **P<0.02 ***P<0.001

Table 2.3 Population genetic analyses of striped skunk populations divided by rabies strain revealed slightly higher genetic diversity among skunks in the south than in the north. Analysis of sympatric and allopatric populations revealed similar effective population sizes for all three populations. Analysis of the populations combined into a global population showed the presence of a population bottleneck. Seven of eight loci were out of Hardy-Weinberg equilibrium in the global population.

		AR	H_o	H_e	θ	N_e	Bottleneck	HWE
2 Strains	NCSK	12.57 ± 0.89	0.767 ± 0.01	0.877 ± 0.01	---	---	P = 0.01	---
	SCSK	14.50 ± 0.80	0.789 ± 0.01	0.896 ± 0.01	---	---	P = 0.02	---
3 Groups	Allopatric North	13.25 ± 0.77	0.772 ± 0.04	0.893 ± 0.01	5.49	42.7	P = 0.03	---
	Sympatric	11.65 ± 0.79	0.754 ± 0.04	0.875 ± 0.01	7.27	68.4	P = 0.006	---
	Allopatric South	12.13 ± 0.76	0.808 ± 0.02	0.883 ± 0.01	7.53	76.2	P = 0.01	---
Global	Pooled Data	15.63 ± 1.05	0.778 ± 0.03	0.889 ± 0.01	---	---	P = 0.01	---
	22-70	17.00	0.849	0.902	---	---	---	P = 0.01
	22-67	20.00	0.733	0.916	---	---	---	P << 0.001
	22-14	15.00	0.791	0.893	---	---	---	P = 0.009
	42-26	19.00	0.837	0.916	---	---	---	P <<

							0.001
42-15	14.00	0.744	0.850	---	---	---	P = 0.006
42-25	16.00	0.721	0.901	---	---	---	P = 0.001
22-19	12.00	0.674	0.856	---	---	---	P << 0.001
42-73	12.00	0.872	0.877	---	---	---	P = 0.4

Table 2.4 Results of AIC model selection for all landscape permeability models evaluated for skunk rabies across the central Great Plains.

Model	r-square	-2LN(P)	ΔAIC	Wi	C	C-P
North						
$IBD*	0.03	4099	0	0.25	0.5	0.71
IBD x RB x EM	0.01	4112	6.3	< 0.01	0.5	0.13
EM*	0.04	4094	1.9	0.10	0.5	0.08
RB*	0.16	4106	2.2	0.08	0.5	0.40
SIBR*	0.12	4061	3.1	0.05	0.5	0.01
CM*	0.13	4057	2.3	0.08	0.5	0.34
IBD x RB	0.18	4034	7.1	< 0.01	0.6	0.09
South						
IBD*	0.45	132	0.9	0.38	0.6	0.01
IBD x RB x EM	0.15	138	13.2	< 0.01	0.6	0.01
EM	0.06	137	9.0	< 0.01	0.5	0.20
RB*	0.28	134	5.7	0.03	0.6	< 0.01
SIBR	0.19	138	11.9	< 0.01	0.5	0.19
CM	0.63	129	12.0	< 0.01	0.6	0.04
$IBD x RB*	0.63	129	0.0	0.59	0.6	0.09

* Indicates potentially good models with ΔAIC_c values $\leq$ 6.0. $ Indicates the AIC_c model with the highest weight. All models were run independently for NCSK and SCSK as molecular data suggest that the two rabies strains are distinct. Variables are as follows: r-square is the r^2 value from regression analysis, K = the number of parameters included in the model determined by the number of different land cover and distance values included in the GIS land cover ecological

resistance model + 1 for the regression intercept parameter, -2Ln(K) is the negative log likelihood for the model with K parameters, ΔAIC is the standardized difference in AIC values for the given model, W_i is the AIC weight for the given model and can be thought of as the degree of certainty that it is the best model given the data, C is the C value from Hosemer-Lemeshow goodness of fit tests for the regression model, and the C-P value is the associated P value for the goodness of fit test. Model variables are defined as follows: IBD = isolation by distance, CM = combined model enhanced edge effects and river barriers, EM = edge model, RB = rivers barrier model, SIBR =simple isolation by resistance, IBD x RB = interaction between IBD and RB model, IBD x RB x EM = interaction between IBD, RB, and EM models.

CHAPTER 3 - Phylogeography of the striped skunk (*Mephitis mephitis*) in North America: Pleistocene dispersal and contemporary population structure

Heather D. Barton and Samantha M. Wisely

Abstract

Pleistocene climate fluctuations influenced the distribution of many species. While specialist species were often restricted to isolated refugia by Pleistocene climate change, generalist species may have been less constrained in their distribution and movements. We used a combination of genetic data and previously published fossil data to investigate the phylogeography and contemporary population structure of a generalist species, the striped skunk (*Mephitis mephitis*). We sequenced portions of the mitochondrial cytochrome *b* gene (601 bp) and control region (D-loop; 381 bp) and amplified 8 microsatellite loci from 314 striped skunk specimens. Phylogenetic analysis of the cytochrome *b* gene revealed the presence of 4 distinct phylogroups, and microsatellite data indicated a pattern of secondary contact among several of these phylogroups. We infer from these data that during the Rancholabrean stage prior to the Illinoian glaciation, striped skunks emerged from a southern refugium in the Texas-Mexico region and colonized the southeastern United States, forming a second refugium in the east. This colonization was followed by a second dispersal event from the southern refugium to the west of the Rocky Mountains during the Illinoian glacial period. During the Sangamonian interglacial stage, two distinct subclades formed on either side of the Sierra Nevada Mountains. During the Holocene, the subclade that colonized the Great Basin then expanded across the northern Rockies and recolonized the Great Plains to create an area of secondary contact. Secondary

contact occurred to a lesser extent with individuals from the eastern refugium east of the Mississippi River. It appears that periodic Pleistocene glacial expansions and retreats caused a series of range expansions and secondary contact events in this native North American species to create a complex pattern of population structure today.

Introduction

The Pleistocene was a time of periodic climate fluctuations that led to the formation and subsequent retraction of large ice sheets over much of the Northern Hemisphere. The cyclical expansion and contraction of these ice sheets throughout the Pleistocene affected the environment and landscape, and influenced the distribution of continental biota. At the end of the Pleistocene (12,000 BP), temperatures yet again increased, glaciers retreated, and flora and fauna that had been confined to refugia during the Wisconsinan glacial period expanded as ecosystems developed throughout the Holocene and into the Recent periods into their contemporary forms (Graham et al. 1996; Hewitt 2000). On a species level of organization, Pleistocene glacial cycles and subsequent Holocene climatic warming produced patterns of vicariance and dispersal which, in turn, generated modern patterns of genetic diversity and population structuring in many species of mammals (Hewitt 2004).

Phylogeographic studies, which can combine fossil, genetic, climatic, and biogeographic data, allow us to infer the vicariance and dispersal of plant and animal species from Pleistocene refugia into what are now their contemporary distributions. Many of these studies describe the Holocene expansion of endemic montane or boreal species out of low-elevation or low-latitude refugia to high latitude or altitude modern distributions (Lessa et al. 2003; Lunt et al. 1998; Santucci et al. 1998; Taberlet and Bouvet 1994). The specialist nature of these species means that they were dependent on a narrow range of resources, and as glaciers advanced and

ecosystems changed, boreal species like the black bear (*Ursus americanus*) and the pine marten (*Martes martes*) were forced out of high-elevation or high-latitude boreal forests as they became uninhabitable, and into smaller pockets of lower-elevation, lower-latitude refugia. In North America, these refugia were located in the southeastern and eastern United States (Soltis et al. 2006), the southwestern United States and Mexico (Olah-Hemmings et al. 2010), and along the Pacific coast (Aubry et al. 2009); for some species, eastern Beringia (central Alaska and western Yukon) also functioned as a Pleistocene refugium (Lessa et al. 2003). These refugia served as origins of future range expansion for endemic and specialist species into uninhabited, developing ecosystems following the final retreat of the Laurentide ice sheet at the end of the Wisconsinan glaciation. Generalist species, which were not subject to the same habitat restrictions as specialist species during the Pleistocene, were also relegated to lower latitudes, but they could traverse a broader spectrum of habitats in the unglaciated areas. The greater freedom of movement afforded to generalist species during the Pleistocene may have allowed them to utilize different colonization routes or to colonize new areas during the Pleistocene (Dragoo et al. 2006; Lister 2004). We investigated the effects of Pleistocene climate change on the distribution and movements of one such North American generalist species, the striped skunk (*Mephitis mephitis*).

Striped skunks are small-bodied mesocarnivores, whose modern distribution ranges from northern Mexico through the continental United States and into southern Canada. All extant skunk species (Mephitidae) are confined to the New World, and evolved from a primitive skunk genus (*Martinogale*) that migrated to the New World in the late Miocene as determined by fossil data (Wang et al. 2005); however, two species that are basal in the skunk phylogenetic tree, *Mydaus javanensis* and *Mydaus marchei*, are found in Java and the Philippines, respectively

(Dragoo and Honeycutt 1997). Phylogenetic evidence suggests that striped skunks diverged from a common ancestor with their sister genus, *Conepatus* (Dragoo et al. 1993) in the southern part of North America (Wang and Carranza-Castañeda 2008). The earliest fossil evidence of the genus *Mephitis* is from the early Pleistocene (~1.8 million BP) and comes from the Broadwater site in Nebraska (Kúrten and Anderson 1980). Late Pleistocene (70,000-14,500 BP) fossil records suggest that striped skunks were broadly distributed across much of the southern half of the United States by the end of the Pleistocene, and fossil records from the Holocene (~10,000-4,500 BP) provide evidence that striped skunks expanded into the upper Midwest and Northeast regions of the United States following the retreat of the Wisconsinan glacier (Faunmap Working Group 1994).

In the current study, we used a combination of genetic, fossil, and biogeographic data to develop a phylogeography of striped skunks in North America. We hypothesized that, unlike boreal species, the life history traits of generalist species afforded striped skunks the ability to move longitudinally during glacial stages when high latitudes and high elevations were uninhabitable, but southern latitudes provided habitat that could be inhabited or provide a route for dispersal, as seen in the phylogeographies of such habitat generalists as raccoons (*Procyon lotor*; Cullingham et al. 2008), deer mice (*Peromyscus maniculatus*; Dragoo et al. 2006), and five-lined skinks (*Eumeces fasciatus*; Howes et al. 2006). We hypothesized that geographic features such as mountains and major river drainages would have acted as isolating mechanisms which caused striped skunk populations to diverge into separate lineages. Studies of another highly mobile generalist species, the raccoon, have shown that rivers and mountains are barriers to movement, generating population structure (Biek et al. 2007; Jenkins et al. 1988). Specifically, we hypothesized that the Mississippi River was a biogeographic barrier to striped

skunk dispersal. One phylogeographic study revealed that the Mississippi River in particular was a barrier to raccoon dispersal (Cullingham et al. 2008); the Mississippi River was also an important barrier to the expansion of other mammals like the northern short-tailed shrew (*Blarina brevicauda*; Brant and Orti 2003), and for reptiles like the five-lined skink (Howes et al. 2006). Because the Mississippi River has been an isolating mechanism for an array of species, it may have isolated populations of striped skunks as well. We further hypothesized that climatic warming during the Holocene would have allowed this highly mobile species to expand latitudinally across North America, and for populations which diverged during the Pleistocene to admix. Studies of striped skunk ecology have shown that they are true habitat generalists (Larivière and Messier 1998; Larivière et al. 1999; Verts 1967) and are highly mobile with a large dispersal capacity (Bixler and Gittleman 2000). Thus, we hypothesized that the biogeographic history of this species would provide evidence for both vicariance and population admixture during the dynamic climatic events of the last 2 million years.

Methods

Sample collection

Because striped skunks are a native North American species (Hall 1981), we wanted to ensure that we adequately sampled the species range to obtain a broad picture of striped skunk phylogeography. Sampling included ear tissue samples of 314 striped skunks collected by USDA-APHIS and the Texas State Health Department from a total of 20 U.S. states, which were sent to Kansas State University for analysis. Geographic sampling units were designated according to six major regions of the United States: West (California, Oregon, and Nevada); Plains (Montana, Wyoming, North Dakota, and Nebraska); Great Lakes (Michigan, Illinois, Indiana, and Ohio); South (Texas, New Mexico, Arizona, and Louisiana); East (Georgia,

Virginia, and West Virginia); and New England (Maine and Vermont). Each geographic sampling unit was chosen based on a combination of geographically and politically established regions of the United States (e.g. the West region was all of the states west of the Rocky Mountains, and the New England region was the collection of states north of New York). Within each geographic sampling unit, sample sizes ranged from 5 to 49 striped skunk ears per state.

Laboratory procedures

We extracted DNA from 314 ear tissue samples at Kansas State University using a PrepGem Blood and Tissue extraction kit (Zygem, Inc., Solana Beach, CA). We amplified a 601 bp segment of the cytochrome *b* gene using two primer sets: 31CB (5'-TGAAACTTCGGTTCCTTG-3') and 728CB (5'-TTCAGTGGATTGGCTGGAGT-3'), and 48CB (5'-GCTCGGAATTTGCTTGATTC-3') and 745CB (5'-TAATATGGGGTGGGGTGTTC-3') at a total volume of 10 µl, which consisted of 2 µl DNA extract (4.5 – 52.0 ng/µl), 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 1.0 µg/µl BSA, 0.5 µM of each primer, and 0.3 unit Taq. PCR conditions were 55º C for 2 minutes, followed by 30 cycles of 94º C for 15 sec, 50º C for 15 sec, and 72º C for 30 sec, followed by 72º C for 1 minute, ending with a 4º C hold. We also amplified a 372 bp section of the control region (D-loop) using primers 15D (5'-TAAACAACCCCGCCATCA-3') and 474D (5'-CCAAATGTATGACGCCGCAGTTATG-3'); reaction mixtures and cycler conditions were the same as for cytochrome *b*. PCR product was purified and sequenced at the Advanced Genetic Technology Center (AGTC, University of Kentucky, Lexington, KY). Sequences were edited using Sequencher 4.7 (Gene Codes Corporation, Ann Arbor, MI). Post-alignment, sequences were deposited in the GenBank database.

We also amplified the extracted DNA at 8 microsatellite loci (Dragoo et al. 2009) at a total volume of 10 µl, which consisted of 2 µl DNA extract (4.5 – 52.0 ng/µl), 1x PCR buffer, 2.7 mM $MgCl_2$, 0.3 mM dNTPs, 0.1 µg/µl BSA, 0.8 M Betaine, 0.2 µM forward primer, 0.5 µM reverse and four dye-tagged primers (FAM and HEX, Integrated DNA Technologies, Coralville, IA; PET and NED, Applied Biosystems, Carlsbad, CA), and 0.5 unit Taq. PCR conditions for primer sets 22-70, 42-15, 22-19, and 42-73 were 94º C for 5 minutes, followed by 30 cycles of 94º C for 30 sec, 51º C for 45 sec (primer sets 22-14 and 22-26: 54º C, primer set 42-25: 53º C), and 72º C for 45 sec, followed by 10 cycles of 94º C for 30 sec, 53º C for 45 sec, and 72º C for 45 sec, followed by 72º C for 10 minutes. Primer set 22-67 required the use of a touchdown PCR, with conditions as follows: 94º C for 5 minutes, followed by 8 cycles of 94º C for 30 sec, 60º C for 30 sec, and 72º C for 30 sec, with the annealing temperature decreasing every two cycles to 58º C, 56º C, and 52º C respectively, followed by 30 cycles of 94º C for 30 sec, 50º C for 30 sec, and 72º C for 30 sec, followed by 72º C for 10 minutes. Samples that failed to amplify were re-extracted using a DNeasy Blood and Tissue extraction kit (Qiagen Inc., Valencia, CA) and re-amplified. PCR product was visualized on an ABI 3730 DNA analyzer (Applied Biosystems, Carlsbad, CA) and scored at all 8 loci using GeneMarker v.1.85 (SoftGenetics LLC, State College, PA). Each locus was independently amplified and visualized an average of 4 times.

Data analyses

We based our analyses on a 601 bp section of the cytochrome *b* gene, a 381 bp aligned section of the D-loop, and concatenated cytochrome *b* and D-loop sequences, and used both graphical and statistical methods of data analysis. Traditional graphical approaches, like nested clade analysis, allow us to determine whether population separations exist (Templeton 2008).

Striped skunks are relatively long-distance dispersers for a small-bodied carnivore (Bixler and Gittleman 2000), and our sampling was representative of the majority of the modern range of striped skunks; our extensive sampling, therefore, minimizes the risk of type I errors that are often associated with traditional graphical analyses. In order to account for stochastic effects on evolutionary history, we used a model-based statistical approach to obtain maximum-likelihood estimates of divergence times (Hey and Nielson 2007).

Phylogenetic and geographical analyses

We conducted separate phylogenetic analyses for cytochrome *b*, D-loop, and concatenated data sets. We used Arlequin 3.11 to estimate nucleotide and haplotype diversity, and to generate a matrix of pairwise Φ_{ST} values based on pairwise differences between haplotypes (Nei and Li 1979). We evaluated statistical significance ($\alpha = 0.05$) based on 1000 permutations, then sequential Bonferroni-corrected for multiple tests (Rice 1989).

We determined the relationships among haplotypes using median-joining networks in Network 4.5.1.6 (www.fluxus-engineering.com). Mutation rates for cytochrome *b* in striped skunks have not been previously calculated; therefore, we relied on known transition and transversion fractions for the Mephitinae superfamily (Marmi et al. 2004). We used these fractions and the date for the first known striped skunk fossil (1.8 MBP; Kúrten and Anderson 1980) to calculate a mutation rate of 6.20×10^{-8} substitutions per site per year. Assuming a generation time of 1 year in striped skunks, we expected a mutation to occur every 26,837 years on average for a 601 bp sequence. We used this mutation rate to estimate divergence times in terms of average number of mutations (rho) separating ancestral and descendent haplotypes in Network 4.5.1.6 (Forster et al. 1996).

We constructed phylogenetic trees using maximum likelihood (ML) in PAUP* 4.0 Beta v.10 (Swofford 2003), and calculated bootstrap support from 1000 replicates using fast-stepwise addition. We used a general time-reversible (GTR) model of DNA substitution, proportion of invariable sites, and shape of the gamma distribution [GTR+I+G model, base frequencies of A=0.2961, C=0.2519, G=0.1353, T=0.3167, rate matrix = (4.2082, 17.7883, 1.8460, 1.0952, 44.5668), I=0.4611, G=0.6328] determined using MrModeltest 2.3 (Nylander 2004), to analyze the cytochrome *b* data set in ML analyses. For D-loop haplotypes, which included insertions and deletions, we used the Hasegawa-Kishino-Yano (HKY), proportion of invariable sites, and shape of the gamma distribution (HKY+I+G) model of DNA substitution, base frequencies of A=0.2589, C=0.2627, G=0.1550, T=0.3235, transition to transversion ratio=6.3678, I=0.0961, G=0.4131 determined using MrModeltest 2.3. We used the eastern spotted skunk [*Spilogale putorious*; cytochrome *b*, X94928 (Ledje and Arnason 1996); D-loop, AY587076, (Dragoo et al. 2004)] as our outgroup. We repeated these analyses using the maximum parsimony method in PAUP* 4.0 Beta, and calculated bootstrap support from 10,000 bootstrap replicates using the full heuristic method. We used as-is addition, and a tree-bisection reconnection (TBR) swapping strategy. We also calculated consistency indices (CI) to evaluate homoplasy in our data sets.

We conducted nested clade analysis (NCA; Templeton 1998) using the median-joining networks generated in Network 4.5.1.6, except that the network was generated from a concatenated dataset of cytochrome *b* and D-loop to resolve some of the ambiguous links. We defined nested clades based on the rules of Templeton et al. (1987), and conducted the NCA using GeoDis 2.4 (Posada et al. 2000) and Templeton's inference key (Templeton 2004).

For geographical analyses, we used SAMOVA (v.1.0) to independently evaluate population structure (Dupanloup et al. 2002). We assumed that the number of geographical

groupings (K) ranged from 2 to 5 for cytochrome *b* and concatenated data sets, based on haplogroups from the median-joining networks. Using multiple geographic groupings allowed us to test hypotheses about the relationships among geographic sampling units to determine patterns of population structure among striped skunks. We evaluated the degree to which population genetic differences could be explained by isolation by distance in Arlequin 3.11 using Mantel tests between geographic sampling unit pairwise Φ_{ST} and geographical distances.

We used MCMC-based simulations in program IMa2 to assess an isolation-with-migration demographic model for striped skunks, and to produce maximum-likelihood estimates and confidence intervals for divergence times among haplogroups. We used the HKY model of substitution and the cytochrome *b* data set for this analysis. We began with multiple runs of 1,000 steps (following 100,000 iterations as burn-in) to assess mixing and to fine-tune the parameter space. We then conducted two independent runs of 1,000,000 steps in MCMC mode. Consistent marginal peak locations with unimodal likelihood curves approaching zero on either end of the distribution indicated reasonable sampling of trees, which were then used in "LoadTree" mode to estimate joint distributions, final parameter estimates, and credibility intervals (Hey and Nielson 2007). To verify divergence times, we calculated generation time to most recent common ancestor (T_{MRCA}) using a Bayesian coalescent-based approach implemented in BEAST 1.4.8 (Drummond and Rambaut 2007), and specified a Bayesian skyline plot as the demographic model (Drummond et al. 2005). Bayesian skyline analysis uses an MCMC approach (Drummond et al. 2002), allowing for simultaneous estimation of genealogy, nucleotide substitution rate, and demography. Bayesian skyline analysis is a coalescent model and does not impose a specific demographic growth pattern *a priori*, as demography is one of the fitted components. T_{MRCA} is different from divergence time in that T_{MRCA} represents the time (in

generations) that lineages shared a common relative, while divergence time represents when populations genetically diverged from one another. T_{MRCA} will be older than divergence time; however, the high dispersal capacity of striped skunks combined with Pleistocene climate fluctuations may have led to rapid population structuring, generating divergence time estimates that are close to T_{MRCA} in time. Thus, we felt that it was appropriate to use T_{MRCA} to confirm our divergence time estimates. We ran the analysis for 10^6 iterations, discarding the first 10^5 iterations as burn-in; we re-ran the analysis until each scale factor was optimized. Our final run consisted of 10^7 iterations, with the first 10^6 iterations discarded as burn-in. We visualized the results with Tracer 1.4.1 (available from http://tree.bio.ed.ac.uk/software/tracer/). We also tested for population growth by calculating Fu's F_S in DnaSP 5.5 (Rozas et al. 2003) for different phylogenetic subunits as indicated by our haplotype networks and phylogenetic trees. Fu's F_S (Fu 1997) compares the number of polymorphic sites to the total number of nucleotide differences to detect population growth; populations with recent expansion show statistically significant negative values.

We used 8 microsatellite loci to assess the contemporary population structure of striped skunks. First, we determined the number of segregating populations in our host population using STRUCTURE 2.3 (Pritchard et al. 2000). We then grouped striped skunks according to the populations determined in STRUCTURE. We tested for Hardy-Weinberg equilibrium using GENEPOP 3.4 (Raymond and Rousset 1995); to ensure that microsatellite loci were randomly associated with one another, we tested for linkage disequilibrium (GENEPOP 3.4). We estimated mean allelic richness for each population using FSTAT 2.9 (Goudet et al. 2002). We also determined the global F_{ST} value with FSTAT 2.9 and population pairwise F_{ST} values to determine the amount of gene flow among skunks throughout North America. We repeated our

analyses in GENEPOP 3.4 and FSTAT 2.9 with a pooled dataset. We used MIGRATE 2.3 (Beerli and Felsenstein 1999) to estimate the relative effective population size (θ) of striped skunks throughout North America by calculating $\theta = 4N_e\mu$ for skunk populations. Because striped skunks are hosts to a variety of diseases, we also tested all populations for the signature of a bottleneck using BOTTLENECK 1.2.02 (Cornuet and Luikart 1996).

Finally, we investigated the fine-scale population structure of striped skunks using a combination of linear regression and population structure analyses for the geographic sampling units in the west, central, and eastern U.S. Program STRUCTURE generates the proportion of each individual's genome that can be assigned to a given population (hereafter, q-value). For the central U.S., we calculated the average q-value for all individuals in southern Texas, central Texas, northern Texas, Nebraska, Wyoming, Montana, and North Dakota, and conducted a linear regression of average q-values against latitude to test the hypothesis that secondary contact between the Intermountain West clade and South stock population was occurring in the Great Plains; we conducted linear regression analysis using SAS (SAS Institute Incorporated, SAS Campus Drive, Cary, NC, USA). We also tested for secondary contact between the Intermountain West clade and the East clade within the Great Lakes/East geographic sampling unit by regressing the average q-values for Georgia, Virginia, West Virginia, Indiana, Illinois, and Michigan against latitude. If secondary contact were occurring in either of these geographic sampling units, linear regression analysis of q-value by latitude would result in a significant r^2 value, indicating that latitude is an important predictor of gene flow.

Results

We obtained complete cytochrome *b* sequences (601 bp) from 269 of the 314 samples; we also obtained complete D-loop sequences (381 bp) from 304 of the 314 samples. The

concatenated dataset consisted of 267 sequences (982 bp). Although we found high levels of homoplasy among D-loop sequences (CI = 0.26), we used concatenated sequences of cytochrome *b* and D-loop to resolve ambiguities in the cytochrome *b* data, in addition to using a data set of only cytochrome *b* sequences (CI = 0.59).

Striped Skunk Phylogenetics and Nested Clade Analysis

The cytochrome *b* network revealed that haplotypes were divided into 3 distinct clades and 1 stock population. The NCA revealed a pattern of contiguous range expansion for the network as a whole (Figure 3.1). The Pacific clade contained only samples from California; the South stock population contained mostly samples from the southwestern states; the Intermountain West clade consisted mostly of samples from the northern Great Plains, the northern Rocky Mountains and the Pacific Northwest; finally, the East clade contained mostly samples from the southeastern states and states bordering the Great Lakes (Figure 3.2). Some samples from Montana and Wyoming were also present in the South stock population, and some samples from Texas and New Mexico were present amongst haplotypes found in the Intermountain West clade. Additionally, samples from Illinois were also found in both the South and Intermountain West clades. The East clade showed evidence of allopatric fragmentation; the South stock population showed evidence of restricted gene flow and isolation by distance. Analysis of the concatenated data set also resulted in 3 distinct clades and a stock population (Appendix B) that were concordant with their cytochrome *b* counterparts (Figure 3.1). Mantel tests using the Φ_{ST} values from the concatenated data set and geographic distance confirmed the presence of a pattern of isolation by distance for striped skunks throughout North America (r^2 = 0.16, P = 0.03).

The clade patterning seen in the cytochrome *b* median-joining network was supported by a ML phylogenetic tree (Figure 3.3). Using eastern spotted skunk (*S. putorius*) as the outgroup, the phylogram showed well-formed clades except for the basal group, the South stock population, which is paraphyletic. Because the stock population is large, it has had ample time to diversify, and we did not have the resources to sample it well enough; however, it is clear that the South stock population is ancestral to all other clades. The split of the East clade from the South stock population is the oldest divergent event, followed by the split of the Intermountain West clade from the South stock population. Finally, the Pacific clade split from the Intermountain West clade most recently, and represents the most derived clade. While the tree bifurcation points representing these splits had less than 50% bootstrap support, trees built using maximum parsimony had similar topologies (results not shown). Phylogenetic trees generated using D-loop data resulted in branches with no resolution due to high levels of homoplasy (CI = 0.26), and trees generated using concatenated sequences did not increase resolution (CI = 0.33); therefore, we did not include phylogenetic trees for D-loop or concatenated sequences in our results.

Population Structure of Striped Skunks in North America

Results of the SAMOVA indicated significant P-values for both K = 3 and K = 4 for cytochrome *b*. In general, the cytochrome *b* SAMOVA showed a pattern of separation by longitude (Table 3.1). The cytochrome *b* Φ_{CT} values were similar for all four levels of K, but the Φ_{CT} value for K = 4 was higher than for K = 3 and was an order of magnitude more significant, suggesting that K = 4 is most likely the correct number of groupings.

We found a significant moderate to high level of differentiation (pairwise Φ_{ST}) among all pairs of geographic sampling units for cytochrome *b* (Table 3.2). Based on the cytochrome *b*

data, the East geographic sampling unit was most differentiated from both the South geographic sampling unit and the West geographic sampling unit; the East geographic sampling unit was also highly differentiated from the Plains geographic sampling unit. Finally, the New England geographic sampling unit was highly differentiated from both the East and the South geographic sampling units (Table 3.2). The concatenated data exhibited the same patterns seen in the cytochrome *b* data, but at a more moderate level (results not shown).

Striped Skunk Demography

We used IMa2 to assess the splitting time between the South stock population and East clade, between the South stock population and Intermountain West clade, and between the Pacific and Intermountain West clades under an isolation-with-migration demographic model, and designated the South stock population as the ancestral clade. Results of the isolation-with-migration model indicated that the East clade split from the stock population an estimated 209,000 BP, the Intermountain West clade split from the stock population an estimated 149,000 BP, and the Pacific clade split from the Intermountain West clade an estimated 132,000 BP. However, the credibility intervals were broad for all divergence time estimates (Table 3.3). We used two other methods, the generation time to most recent common ancestor (T_{MRCA}) for each haplogroup in BEAST using a skyline model, and the estimation of rho statistics in Network, to corroborate the splitting times generated by IMa2. The T_{MRCA} values estimated using the cytochrome *b* data set were similar to the splitting time estimates for the isolation-with-migration model generated using IMa2, as were the rho estimates generated in program Network (Table 3.3). Scale factors for Bayesian skyline analysis were optimized after 5-7 runs. The demographic model constructed from the entire cytochrome *b* data set corroborated our interpretation of the NCA, indicating steady population size throughout much of the Pleistocene,

followed by extensive range expansion by striped skunks in North America following the retreat of the Wisconsin glacier (Figure 3.5).

Fu's F_S for cytochrome *b* revealed a signature of expansion (Fu's F_S = -6.403, P << 0.001); the South stock population, and the East and Intermountain West clades all displayed significant signatures of population expansion (Fu's F_S = -7.083, P = 0.001; Fu's F_S = -4.319, P = 0.008; Fu's F_S = -2.748, P = 0.035 respectively), while the Pacific clade did not (Fu's F_S = 1.137, P = 0.388). Concatenated data demonstrated the same patterns seen in cytochrome *b*, so results were not included. These values indicated a steady expansion of striped skunks throughout much of North America.

Population Genetics of Striped Skunks in North America

The combination of haplotype and nucleotide diversity patterns present in each geographic sampling unit provides some indication of demographic processes. For instance, high haplotype diversity results when a population has been long established, while low haplotype diversity indicates the presence of a more recent or isolated population. High nucleotide diversity also indicates the presence of an older population, whereas low nucleotide diversity generally indicates a selective sweep, population bottleneck, or the presence of a more recently established population. The low haplotype and nucleotide diversity in the cytochrome *b* data for the East geographic sampling unit suggest that the East was isolated from the rest of the U.S. after animals initially expanded east of the Mississippi River (Table 3.4). This trend disappears in the concatenated data, which includes D-loop, a marker commonly used to examine more recent phylogeographic patterns, suggesting that individuals from the East geographic sampling unit gradually expanded north to colonize the Great Lakes and New England geographic sampling units (Table 3.4). The combination of high haplotype diversity

and low nucleotide diversity in the Plains geographic sampling unit using the concatenated data set suggests that while the populations in the Great Plains are not the oldest, they may have been colonized by individuals from multiple Pleistocene refugia. Finally, the combination of high haplotype diversity and high nucleotide diversity for the Great Lakes geographic sampling unit using the concatenated data set suggests that a large amount of gene flow and admixture occurred around the Great Lakes geographic sampling unit.

Contemporary Striped Skunk Population Genetics

Program STRUCTURE indicated 5 contemporary populations of striped skunks in North America, and those populations consisted of (i) the West geographic sampling unit, (ii) the Plains geographic sampling unit, (iii) the South geographic sampling unit, (iv) the New England geographic sampling unit, and (v) the Great Lakes and East geographic sampling units. This interpretation is based on the ΔK distribution where K = the number of populations, the modal value of which tends to be located at the actual K (Evanno et al. 2005). We assigned individual skunks to one of the 5 populations using individual q-values, and used these groupings to conduct population genetic analyses.

Analysis of microsatellite markers indicated that none of the loci exhibited linkage disequilibrium ($P > 0.05$ after Bonferroni correction). Mean allelic richness and heterozygosity were similar for all 5 populations (Table 3.5). We found heterozygote deficiency at 3 loci in the West population, at 5 loci in the Plains population, at 5 loci in the South population, at 3 loci in the New England population, and at 7 loci in the Great Lakes/East population (Table 3.5). A Wilcoxon Sign-Rank test in program BOTTLENECK 1.2.02 under the two-phase mutation (TPM) model was significant for the West and Plains populations, suggesting a past population reduction and subsequent expansion; when all populations were combined, the bottleneck

signature was only marginally significant (Table 3.5). The global F_{ST} was 0.046 ± 0.008, which indicated that there was a modest degree of divergence among populations of striped skunks across different regions of North America. Pairwise F_{ST} values among the 5 populations (Table 3.2) indicated a minor but significant amount of separation among striped skunks throughout North America. θ (Table 3.5) for striped skunk populations in the Plains, South, and Great Lakes/East were not significantly different among populations (Plains θ = 3.18, South θ = 3.41, Great Lakes/East θ = 3.19); however, populations in the West and New England were about two-thirds and half the size of the other three populations, respectively (West θ = 2.45, New England θ = 1.75). Because MIGRATE accounts for unequal sample sizes, we can infer that the small effective population sizes in the West and New England are due to demographic processes of recent colonization and founder effects.

We also examined fine scale population structure in three sampling units: within the West population, within the Great Lakes/East population, and within the Plains and South populations in the central Great Plains. Within the West population, samples on either side of the Sierra Nevada Mountains (California vs. Nevada and Oregon), which had divergent haplotypic signatures, all comprised a single population when microsatellite markers were analyzed in program STRUCTURE. The pairwise F_{ST} between populations on either side of the Sierra Nevada mountain range was 0.04 (P = 0.02). Population structure analysis within the Great Lakes/East population showed a slight gradient in individual q-value with latitude; however, linear regression indicated that this gradient was not significant (r^2 = 0.56; P = 0.09; Figure 3.4), and a pairwise F_{ST} = 0.018 (P = 0.01) between the East and Great Lakes geographic sampling units suggested that they are essentially a single population. Finally, the structure analysis using only Plains and South populations revealed a gradient in individual q-values with latitude,

suggesting a pattern of secondary contact between these two divergent populations throughout the Great Plains; latitude was a significant predictor of q-value in the Great Plains ($r^2 = 0.80$, P = 0.02; Figure 3.4).

Discussion

During the Pleistocene, fluctuating climates and ecosystems redistributed the majority of the North American biota; however, generalist species like the striped skunk (Rosatte 1987), which were better able to persist in a variety of environmental conditions, did not rely solely on interglacial or Holocene colonization out of refugia for population expansion. Indeed, population size of striped skunks throughout the Pleistocene appeared to remain constant until they increased during the Wisconsinan glaciation (Figure 3.5). The data from our study suggest that the biogeographic patterns of striped skunks represent multiple Pleistocene dispersal events (Figure 3.2, Figure 3.6) from the ancestral stock population in the south central United States, from which all other lineages of striped skunks descended. From Irvingtonian deposits (1.8 MBP to 0.5 MBP), fossil remains have been found from Florida, Colorado, and Arkansas (Anderson 2004) and suggest that striped skunks expanded across the Mississippi River and into the southeastern U.S. Our data estimate that the East clade split from the South stock population approximately 360,000 BP (Table 3.3) during the Rancholabrean when glacial melting repeatedly widened the Mississippi River and acted as a barrier to admixture between the two biogeographic regions. Ultimately, this vicariant event led to the generation of two separate refugia, one represented by the South stock population and one represented by the East clade.

The southern and eastern refugia represent the ancestral lineages from which all other lineages descended. The pattern of restricted gene flow with isolation by distance (RGF/IBD) in the South stock population revealed by NCA indicates stable population size and equilibrium

conditions persisted. This population was likely stable during the oscillating ice ages due to its distance from the effects of glaciers to the north. Fu's F_S estimate for the South stock population indicates a signature of population expansion, providing more evidence that this clade was the point of origin for Pleistocene and Holocene expansion across North America. The presence of allopatric fragmentation (AF) in the NCA for the East clade (Figure 3.1) and the combination of low haplotype and nucleotide diversity in the East sampling unit (Table 3.4) indicate that subsequent to the South stock population-East clade split, the eastern refugium was small and isolated for the remainder of the Pleistocene.

The retreat of the Wisconsinan glacier in the east precipitated the development of habitable ecosystems in the northeast, enabling animals in the East clade to expand north around the Great Lakes. The complete retreat of the Laurentide ice sheet ultimately provided habitat for striped skunks in the New England regions of the United States by the middle to late Holocene. The appearance of fossil remains from that time period (8,500 BP to 4,500 BP) from Tennessee, Kentucky, Ohio, Pennsylvania, and New York support the hypothesis that striped skunks expanded northward toward the Great Lakes and New England throughout the Holocene; by the late Holocene (~4,500 BP), striped skunks had expanded into much of their modern distribution (Faunmap Working Group 1994). Yet, we found no evidence that the East clade expanded westward across the Mississippi River at any point following the initial South stock population-East clade split, which suggests that the Mississippi River was an historical barrier, and continues to be a modern barrier to dispersal by striped skunks. Indeed, our microsatellite data reveal that modern striped skunk populations show evidence of very little admixture between clades on either side of the Mississippi River (Figure 3.4). Only in the northern half of the

Mississippi River do we find evidence that animals from both the stock population and Intermountain West clade crossed the Mississippi River.

The Mississippi River represents a considerable biogeographic barrier in the eastern United States, the clearest example of which can be seen in raccoon phylogeography in the southeastern U.S. It appears that distinct lineages of raccoons persist on either side of the Mississippi, with minimal admixture between the two populations (Cullingham et al. 2008). These findings suggest that the Mississippi River acts as a modern and historical barrier to dispersal. Additionally, phylogeographies for other mammals (northern short-tailed shrew, *Blarina brevicauda*; Brant and Orti 2003), reptiles (five-lined skink, *Eumeces fasciatus*; Howes et al. 2006; black rat snake, *Elaphe obsoleta*; Burbrink et al. 2000), and amphibians (northern leopard frog, *Rana pipiens*; Hoffman and Blouin 2004; tiger salamander, *Ambystoma tigrinum*; Templeton et al. 1995) all indicate that the Mississippi River was a major biogeographic barrier, responsible for vicariance among lineages, and ultimately generated distinct phylogroups within species. The frequency of this pattern across a broad range of species indicates the importance of the Mississippi River as a major biogeographic barrier to Pleistocene and Holocene dispersal; in some cases it is still a factor in contemporary population structuring.

The Pleistocene was not only a time of expansion to the east, but striped skunks also expanded west during this time. Divergence time estimates from IMa2 and Network, and T_{MRCA} estimates from BEAST indicate that the Intermountain West clade diverged from the South stock population about 200,000 BP during the Illinoian glaciation (300,000 BP to 130,000 BP; Table 3.3). We infer, based on patterns of haplotypic distribution and the presence of inhospitable habitat to the north, that striped skunks colonized the Great Basin from a southerly route. As extensive montane glaciers formed in the Rocky Mountains during the Illinoian glacial

maximum, the South and Intermountain West clades diverged. Subsequent clade formation occurred once the far west was colonized by striped skunks and then were isolated from the Great Basin population (the Intermountain West clade) by the Sierra Nevada mountain range. The resulting isolation of populations to the west of the Sierra Nevadas formed the Pacific clade, which split from the Intermountain West clade approximately 110,000 BP during the Sangamonian interglacial stage (125,000 BP to 75,000 BP; Table 3.3).

Fossil remains dated to the Wisconsinan stage (70,000 BP to 10,000 BP) from Idaho and central California suggest that striped skunks expanded north along both sides of the Sierra Nevada Mountains. During the Holocene warming trend, however, it appears that these isolated clades may have once again become admixed. Microsatellite markers suggest admixture of these two clades such that they are not distinguishable from one another. Thus, the signature of Sangamonian vicariance observed in the maternally inherited mitochondrial genome is obscured by contemporary gene flow as observed in the nuclear genome. This divergent pattern of spatial partitioning could be influenced by male-biased dispersal in this carnivore species (Sargeant et al. 1982). Nonetheless, it appears that these two clades were separated during the last interglacial period and have likely been admixing since the late Wisconsinan or early Holocene, however, a more detailed biogeographic study of striped skunks in the west is warranted.

The warming trend at the end of the Pleistocene led to the retreat of high elevation glaciers in the Rocky Mountains, and opened up high latitude colonization routes for striped skunks. The distribution of haplotypes from the Great Basin to the northern Great Plains suggests that during the late Pleistocene or early Holocene, individuals in the Intermountain West clade crossed the Continental Divide to the north, and then expanded east and south into the north and central Great Plains. Following the Wisconsinan glaciation, the Great Plains would

have initially been a mixture of grassland and forest, an ideal habitat of striped skunks (Bixler and Gittleman, 2000), allowing them to readily expand their distribution. The gradually improving habitat in the Great Plains during the Holocene also provided an opportunity for individuals in the southern refugium of the southern Great Plains to expand northward. Eventually, the South stock population and Intermountain West clade came into contact in the central Great Plains. The significant linear regression of q-value against latitude using microsatellite data (Figure 3.4) indicates a strong pattern of secondary contact among divergent populations from the northern (Intermountain West clade) and southern (South stock population) Great Plains that expanded towards the central Great Plains where they admixed. Middle Holocene (~8,500 BP) fossils appear in Nebraska and Kansas and late Holocene fossils (~5,000 BP) from several states (i.e. Minnesota, Iowa, Nebraska, Kansas, and Oklahoma) support the expansion of *Mephitis mephitis* across the Great Plains throughout the Holocene (Faunmap Working Group 1994).

The signature of a population bottleneck in the Western and Plains populations at first appears to contradict the results of the mitochondrial data, which indicate steady population expansion of striped skunks throughout North America. However, striped skunks are reservoirs for a multitude of diseases including rabies (Blanton et al. 2007), canine distemper (Gehrt et al. 2010) and tularemia (Berrada et al. 2006). California and the Great Plains are the foci of two strains of rabies for which striped skunks are the major disease reservoir (Blanton et al. 2007; Crawford-Miksza et al. 1999). Rabies is virtually always fatal (but see Willoughby et al. 2005), and routinely decimates striped skunk populations where the disease is endemic, leading to periodic fluctuations in population size that are evident in the bottleneck test. The majority of the Great Lakes/East population, and the New England population, on the other hand, are within

the distribution of a rabies strain that is endemic in raccoons (Blanton et al. 2007) and impact striped skunk populations much less than in California or the Great Plains.

Our findings provide the phylogenetic context needed to evaluate the importance of dispersal routes and biodiversity hotspots for striped skunks. Several native North American species, like the deer mouse and the raccoon, have comparable geographic distributions and generalist tendencies as striped skunks and exhibit similar phylogeographic patterns. The Rocky Mountains, Sierra Nevada Mountains, and Mississippi River have all been implicated as biogeographic barriers to Pleistocene dispersal for one or more of these species (Cullingham et al. 2008; Dragoo et al. 2006; Yang and Kenagy 2009). By the end of the Pleistocene, deer mice were using the same southern colonization route to expand from New Mexico into Texas (Dragoo et al. 2006) that striped skunks used during the Illinoian glacial stage to expand westward. Similarly, *P. maniculatus* populations west of the Rocky Mountains mirror *M. mephitis* populations in the same region and suggest that the Sierra Nevada mountains were a source of vicariance for multiple species, and provided separate colonization routes on either side of the range (Yang and Kenagy 2009). Finally, phylogeography of the deer mouse on a continental scale is similar to clade patterning of *M. mephitis* in the Great Plains and upper Midwest. Thus, for multiple generalist species, these regions represent intraspecific biodiversity hotspots where secondary contact has led to high levels of diversity and the admixture of previously separated lineages (Blackburn and Measey 2009, He et al. 2008, Hopper and Gioia 2004, Petit et al 2003).

Comparative phylogeography among these species implies that many different species shared similar dispersal and vicariance events, and suggests that dispersal routes in the southern and western U.S. were important during glaciation events, and that dispersal routes to the north

and east were important during interglacial periods for a number of generalist species. Additionally, the presence of distinctive phylogroups and contemporary populations for species like deer mice, raccoons, and striped skunks suggests that it may be worthwhile to further characterize gene flow among divergent phylogroups of these and other generalist species; assessing gene flow among divergent phylogroups allows us to explore current phylogeographic processes (e.g. vicariance, dispersal, admixture) in light of habitat modification, which will aid in determining the future distributions of these species as global climate change continues to alter ecosystems. Finally, comparative phylogeography is useful for studying mammalian disease reservoir species (Cullingham et al. 2008; Dragoo et al. 2006; Ngamprasertwong et al. 2008). Disease reservoir species, many of which are generalists, can respond rapidly to habitat alteration; as reservoir species shift ranges, so do their pathogens (Brooks and Hoberg 2007, Hoberg and Brooks, 2008, Holmes 2004). Therefore, pathogen range shifts should mirror range shifts in their host species, and by investigating the phylogeography of one or both, we can examine host-pathogen co-evolution and the potential for disease emergence or pandemics in new habitats. Our findings about the phylogeography of striped skunks on a continental scale allow us to make inferences about the Pleistocene and Holocene dispersal of generalist species, and to make inferences about the spread of diseases from other native North American generalist species with similar biogeographic and life histories to striped skunks. We know from studies of emerging diseases that directly transmitted pathogens need their hosts for survival and reproduction; we would predict that as the reservoir species expands into new habitats, diseases associated with that reservoir should emerge in those novel habitats. Striped skunks and other generalist species carry a wide variety of potentially zoonotic diseases; by studying the future range expansion of these generalist species, we can better control for and prevent zoonosis.

References

ANDERSON, E. 2004. The Carnivora from Porcupine Cave, Pp. 141-154 in Biodiversity Response to Climate Change in the Middle Pleistocene: The Porcupine Cave Fauna from Colorado (A. D. Barnosky, ed.). University of California Press, Berkeley.

BEERLI, P., AND J. FELSENSTEIN. 1999. Maximum likelihood estimation of migration rates and population numbers of two populations using a coalescent approach, Genetics 152**:**763-773.

BERRADA, Z. L., H. K. GOETHERT, AND S. R. TELFORD. 2006. Raccoons and skunks as sentinels for enzootic tularemia, Emerging Infectious Diseases 12**:**1019-1021.

BIEK, R., J. C. HENDERSON, L. WALLER, C. E. RUPPRECHT, AND L. A. REAL. 2007. A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus, Proceedings of the National Academy of Sciences of the USA 104**:**7993-7998.

BIXLER, A., AND J. L. GITTLEMAN. 2000. Variation in home range and use of habitat in the striped skunk (*Mephitis mephitis*), Journal of Zoology, London 251**:**525-533.

BLACKBURN, D.C. AND G.J. MEASEY. 2009. Dispersal to or from an African biodiversity hotspot?, Molecular Ecology 18:1904-1915.

BLANTON, J. D., C. HANLON, AND C. E. RUPPRECHT. 2007. Rabies surveillance in the United States during 2006, Pp. 540-556 in Journal of the American Veterinary Medical Association.

BRANT, S. V., AND G. ORTI. 2003. Phylogeography of the Northern short-tailed shrew, *Blarina brevicauda* (Insectivora: Soricidae): past fragmentation and postglacial recolonization, Molecular Ecology 12**:**1435-1449.

BROOKS, D.R. AND E.P. HOBERG. 2007. How will global climate change affect parasite-host assemblages?, Trends in Parasitology 23:571-574.

BURBRINK, F. T., R. LAWSON, AND J. B. SLOWINSKI. 2000. Mitochondrial DNA phylogeography of the polytopic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept, Evolution 54**:**2107-2118.

CORNUET, J. M., AND G. LUIKART. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data, Genetics 144**:**2001-2014.

CRAWFORD-MIKSZA, L. K., D. A. WADFORD, AND D. P. SCHNURR. 1999. Molecular epidemiology of enzootic rabies in California, Journal of Clinical Virology 14**:**207-219.

CULLINGHAM, C. I., C. J. KYLE, B. A. POND, AND B. N. WHITE. 2008. Genetic structure of raccoons in eastern North America based on mtDNA: Implications for subspecies designation and rabies disease dynamics, Canadian Journal of Zoology 86**:**947-958.

DRAGOO, J. W., R. D. BRADLEY, R. L. HONEYCUTT, AND J. W. TEMPLETON. 1993. Phylogenetic relationships among the skunks: a molecular perspective, Journal of Mammalian Evolution 1**:**255-267.

DRAGOO, J. W., K. E. COAN, K. A. MOORE, S. E. HENKE, R. C. FLEISCHER, AND S. M. WISELY. 2009. Polymorphic microsatellite markers for the striped skunk, *Mephitis mephitis*, and other mephitids, Molecular Ecology Resources 9**:**383-385.

DRAGOO, J. W., AND R. L. HONEYCUTT. 1997. Systematics of mustelid-like carnivores, Journal of Mammalogy 78**:**426-443.

DRAGOO, J. W., J. A. LACKEY, K. E. MOORE, E. P. LESSA, J. A. COOK, AND T. L. YATES. 2006. Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a

predictive framework for research on hantaviruses, Journal of General Virology 87**:**1997-2003.

DRAGOO, J. W., D. K. MATTHES, A. ARAGON, C. C. HASS, AND T. L. YATES. 2004. Identification of skunk species submitted for rabies testing in the desert southwest, Journal of Wildlife Diseases 40**:**371-376.

DRUMMOND, A. J., G. K. NICHOLLS, A. G. RODRIGO, AND W. SOLOMON. 2002. Estimating mutation parameters, population history, and genealogy simultaneously from temporally spaced sequence data, Genetics 161**:**1307-1320.

DRUMMOND, A. J., AND A. RAMBAUT. 2007. BEAST: Bayesian evolutionary analysis by sampling trees, BMC Evolutionary Biology 7**:**214.

DRUMMOND, A. J., A. RAMBAUT, B. SHAPIRO, AND O. G. PYBUS. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences, Molecular Biology and Evolution 22**:**1185-1192.

DUPANLOUP, I., S. SCHNEIDER, AND L. EXCOFFIER. 2002. A simulated annealing approach to define the genetic structure of populations, Molecular Ecology 11**:**2571-2582.

EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study., Molecular Ecology 14**:**2611-2620.

FAUNMAP WORKING GROUP. 1994. FAUNMAP: A database documenting late Quaternary distributions of mammal species in the United States. Vol. XXV. Illinois State Museum, Springfield, Illinois.

FORSTER, P., R. HARDING, A. TORRONI, AND H.-J. BANDELT. 1996. Origin and evolution of native American mtDNA variation: a reappraisal, American Journal of Human Genetics 59**:**935-945.

FU, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection, Genetics 147**:**915-925.

GEHRT, S. D., M. J. KINSEL, AND C. ANCHOR. 2010. Pathogen dynamics and morbidity of striped skunks in the absence of rabies, Journal of Wildlife Diseases 46**:**335-347.

GOUDET, J., N. PERRIN, AND P. WASER. 2002. Tests for sex-biased dispersal using bi-parentally inherited genetic markers, Molecular Ecology 11**:**1103-1114.

GRAHAM, R. W., ET AL. 1996. Spatial response of mammals to late Quaternary environmental fluctuations, Science 272**:**1601-1606.

HALL, E. R. 1981. The Mammals of North America, 2nd edn. John Wiley & Sons, New York.

HE, T., B.B. LAMONT, S.L. KRAUSS, N.J. ENRIGHT, AND B. MILLER. 2008. Covariation between intraspecific genetic diversity and species diversity within a plant functional group, Journal of Ecology 96:956-961.

HEWITT, G. 2000. The genetic legacy of the Quaternary ice ages, Nature 405**:**907-913.

HEWITT, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary, Philosophical Transactions of the Royal Society of London, B. Biological Sciences 359**:**183-195.

HEY, J., AND R. NIELSON. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics, Proceedings of the National Academy of Sciences of the United States of America 104**:**2785-2790.

HOBERG, E.P. AND D.R. BROOKS. 2008. A macroevolutionary mosaic: episodic host-switching, geographical colonization, and diversification in complex host-parasite systems, Journal of Biogeography 35:1533-1550.

HOFFMAN, E. A., AND M. S. BLOUIN. 2004. Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA, Evolution 58**:**145-159.

HOLMES, E.C. 2004. The phylogeography of human viruses, Molecular Ecology 13:745-756.

HOPPER, S.D. AND P. GIOIA. 2004. The Southwest Australia Floristic Region: evolution and convservation of a global hot spot of biodiversity, Annual Review of Ecology, Evolution, and Systematics 35:623-650

HOWES, B. J., B. LINDSAY, AND S. C. LOUGHEED. 2006. Range-wide phylogeography of a temperate lizard, the five-lined skink (*Eumeces fasciatus*), Molecular Phylogenetics and Evolution 40**:**183-194.

JENKINS, S. R., B. D. PERRY, AND W. G. WINKLER. 1988. Ecology and epidemiology of raccoon rabies, Reviews of Infectious Diseases 10**:**S620-S625.

KANDA, L. L., AND T. K. FULLER. 2004. Demographic responses of Virginia opossums to limitation at their northern boundary, Canadian Journal of Zoology 82**:**1126-1134.

KÚRTEN, B., AND E. ANDERSON. 1980. Pleistocene Mammals of North America. Columbia University Press, New York.

LARIVIÈRE, S., AND F. MESSIER. 1998. Spatial organization of a prairie striped skunk population during the waterfowl nesting season, Journal of Wildlife Management 62**:**199-204.

LARIVIÈRE, S., L. R. WALTON, AND F. MESSIER. 1999. Selection by striped skunks (*Mephitis mephitis*) of farmsteads and buildings as denning sites, American Midland Naturalist 142**:**96-101.

LEDJE, C., AND U. ARNASON. 1996. Phylogenetic analyses of complete cytochrome b genes of the order carnivora with particular emphasis on the caniformia, Journal of Molecular Evolution 42**:**135-144.

LESSA, E. P., J. A. COOK, AND J. L. PATTON. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary, Proceedings of the National Academy of Sciences of the United States of America 100**:**10331-10334.

LISTER, A. M. 2004. The impact of Quaternary ice ages on mammalian evolution, Philosophical Transactions of the Royal Society of London, B. Biological Sciences 359**:**221-241.

LUNT, D. H., K. M. IBRAHIM, AND G. M. HEWITT. 1998. MtDNA phylogeography and post-glacial patterns of subdivision in the meadow grasshopper, *Chorthippus parallelus*, Heredity 80.

MARMI, J., J. F. LÓPEZ-GIRÁLDEZ, AND X. DOMINGO-ROURA. 2004. Phylogeny, evolutionary history and taxonomy of the Mustelidae based on sequences of the cytochrome *b* gene and a complex repetitive flanking region, Zoological Scripta 33**:**481-499.

NEI, M., AND W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases, Proceedings of the National Academy of Sciences of the United States of America 76**:**5269-5273.

NGAMPRASERTWONG, T., I. J. MACKIE, P. A. RACEY, AND S. B. PIERTNEY. 2008. Spatial distribution and microsatellite DNA variation in Daubenton's bat within Scotland, Molecular Ecology 17**:**3243-3258.

NYLANDER, J.A.A. 2004. MrModeltest, version 2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

OLAH-HEMMINGS, V., ET AL. 2010. Phylogeography of declining relict and lowland leopard frogs in the desert southwest of North America, Journal of Zoology 280**:**343-354.

PETIT, R.J., ET AL. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity, Science 300:1563-1565.

POSADA, D., K. A. CRANDALL, AND A. R. TEMPLETON. 2000. GeoDis: A program for the Cladistic Nested Analysis of the Geographical Distribution of Genetic Haplotypes, Molecular Ecology 9**:**487-488.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure from multilocus genotype data, Genetics 155**:**945-959.

RAYMOND, M., AND F. ROUSSET. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism, Journal of Heredity 86**:**248-249.

RICE, W. R. 1989. Analyzing tables of statistical tests, Evolution 43**:**223-225.

ROSATTE, R. C. 1987. Striped, Spotted, Hooded, and Hog-nosed skunk, Pp. 598-613 in Wild furbearer management and conservation in North America. Ontario Trappers Association, Toronto, Ontario.

ROZAS, J., J. C. S.-D. BARRIO, X. MESSEGUER, AND R. ROZAS. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods, Bioinformatics 19**:**2496-2497.

SANTUCCI, F., B. C. EMERSON, AND G. M. HEWITT. 1998. Mitochondrial DNA phylogeography of European hedgehogs, Molecular Ecology 7**:**1163-1172.

SARGEANT, A.B., R.J. GREENWOOD, J.L. PIEHL, AND W.B. BICKNELL. 1982. Recurrence, mortality, and dispersal of prairie striped skunks, *Mephitis mephitis*, and implications to rabies epizootiology, Canadian Field Naturalist 96:312-316.

SOLTIS, D. E., A. B. MORRIS, J. S. MCLACHLAN, P. S. MANOS, AND P. S. SOLTIS. 2006. Comparative phylogeography of unglaciated eastern North America, Molecular Ecology 15**:**4261-4293.

SWOFFORD, D. L. 2003. PAUP 4.0 beta version for the Macintosh PowerPC - Phylogenetic analysis using parsimony (and other methods). Sinauer Associates, New York, NY.

TABERLET, P., AND J. BOUVET. 1994. Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear, *Ursus arctos* in Europe, Proceedings of the Royal Society of London, B. Biological Sciences 255**:**195-200.

TEMPLETON, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history, Molecular Ecology 7**:**381-397.

TEMPLETON, A. R. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors, Molecular Ecology 13**:**789-809.

TEMPLETON, A. R. 2008. Nested clade analysis: An extensively validated method for strong phylogeographic inference, Molecular Ecology 17**:**1877-1880.

TEMPLETON, A. R., E. BOERWINKLE, AND C. F. SING. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in Drosophila, Genetics 117**:**343-351.

TEMPLETON, A. R., E. ROUTMAN, AND C. A. PHILLIPS. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*, Genetics 140**:**767-782.

VERTS, B. J. 1967. The biology of the striped skunk. University of Illinois Press, Urbana, Illinois.

WANG, X., AND Ó. CARRANZA-CASTAÑEDA. 2008. Earliest hog-nosed skunk, *Conepatus* (Mephitidae, Carnivora), from the early Pliocene of Guanajuato, Mexico and origin of South American skunks, Zoological Journal of the Linnean Society 154**:**386-407.

WANG, X., D. P. WHISTLER, AND G. T. TAKEUCHI. 2005. A new basal skunk *Martinogale* (Carnivora, Mephitinae) from late Miocene Dove Springs formation, California, and origin of New World mephitines, Journal of Vertebrate Paleontology 25**:**936-949.

WILLOUGHBY, R. E., ET AL. 2005. Survival after treatment of rabies with induction of coma, New England Journal of Medicine 352**:**2508-2514.

YANG, D.-S., AND G. J. KENAGY. 2009. Nuclear and mitochondrial DNA reveal contrasting evolutionary processes in populations of deer mice (*Peromyscus maniculatus*), Molecular Ecology 18**:**5115-5125.

Figures and Tables

Figure 3.1 Cytochrome *b* median-joining network based on 601 bp for 269 striped skunk specimens. Branch lengths are proportional to the number of substitutions, and circle sizes are proportional to the number of individuals represented. Ambiguous connections were resolved using the rooted maximum-likelihood phylogenetic tree. Clades are indicated with dotted lines, and the stock population is indicated with a dashed line. Results of NCA, including allopatric fragmentation (AF), restricted gene flow (RGF) and isolation by distance (IBD), are indicated where significant.

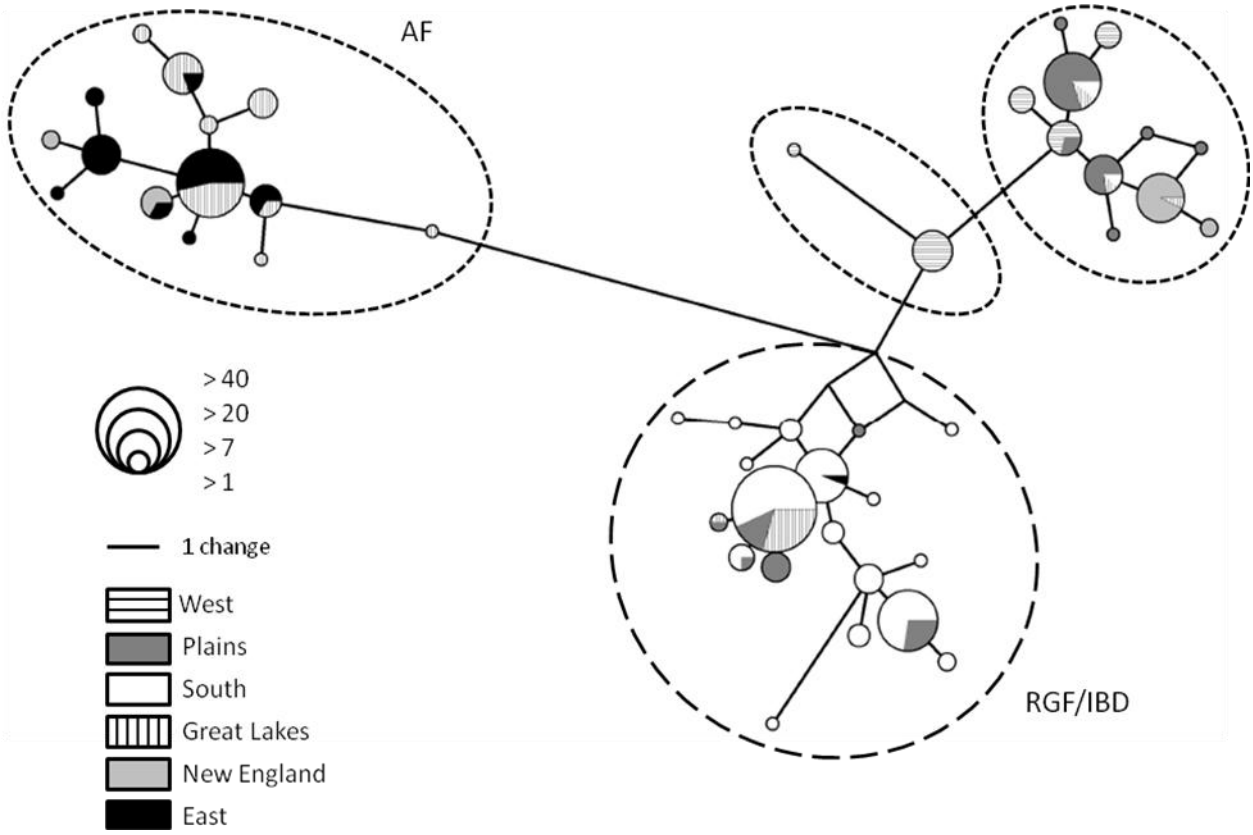


Figure 3.2 Geographic distribution of cytochrome *b* mtDNA striped skunk clades throughout the United States. Pie charts indicate the proportional representation of clades and stock population in each state. The hypothesized Pleistocene and Holocene dispersal patterns for striped skunk clades and stock population are indicated by different line-dash types.

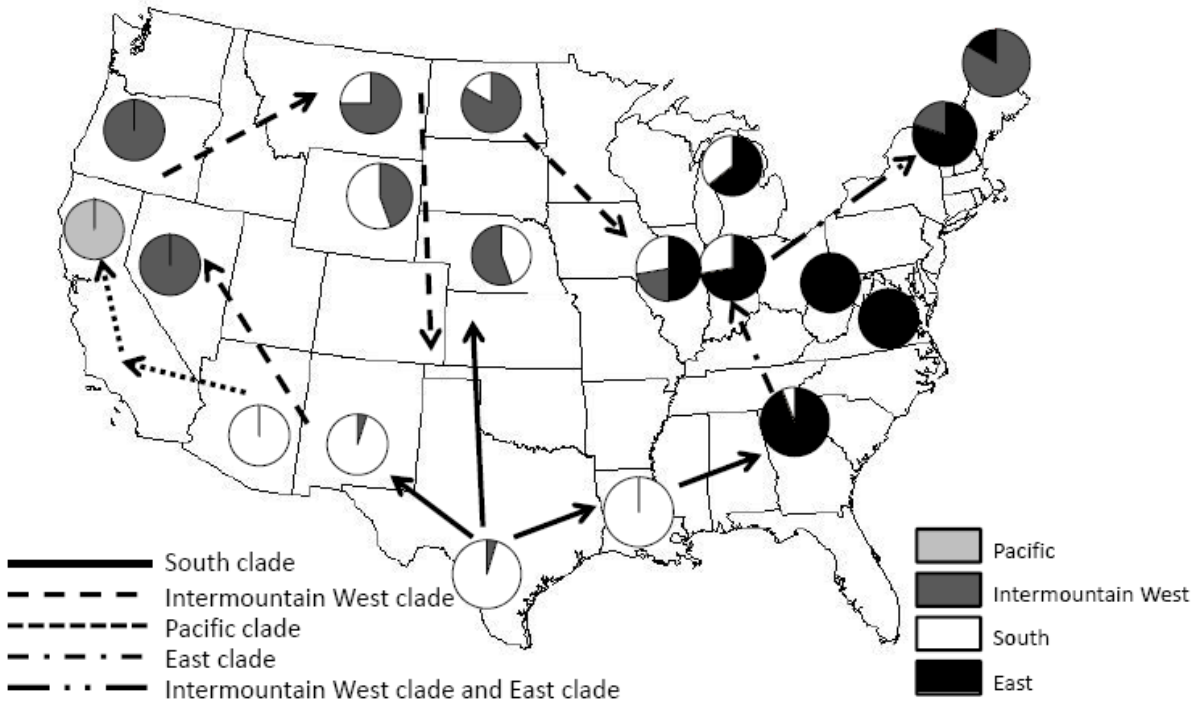


Figure 3.3 Maximum-likelihood phylogram of cytochrome *b* haplotypes constructed under a GTR+I+G model of evolution based on 601 bp for 269 striped skunk specimens. Adjacent numbers indicate bootstrap values for supported nodes ≥ 50%. We used eastern spotted skunk (*S. putorius*) as the outgroup. Shaded bars at right of tree indicate clade designation (South stock population: white, Intermountain West clade: dark gray, East clade: black, Pacific clade: light gray).

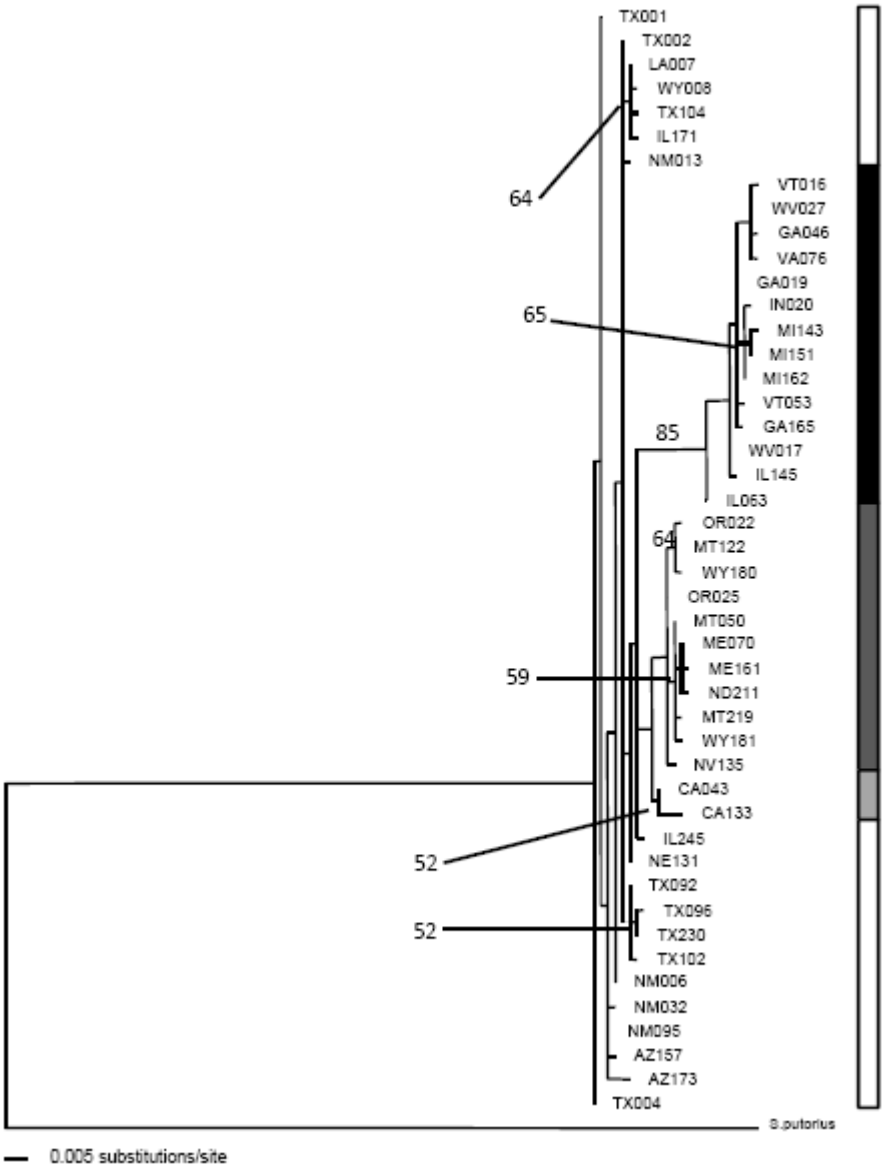


Figure 3.4 Linear regression of average q-value against latitude for the South and Plains populations (black circles), and for the Great Lakes/East population (open squares). Linear regression was significant for the South and Plains populations (r^2 = 0.80, P = 0.02), but not for the Great Lakes/East population (r^2 = 0.56, P = 0.09), indicating a moderate amount of admixture in the Great Plains and minimal admixture east of the Mississippi River.

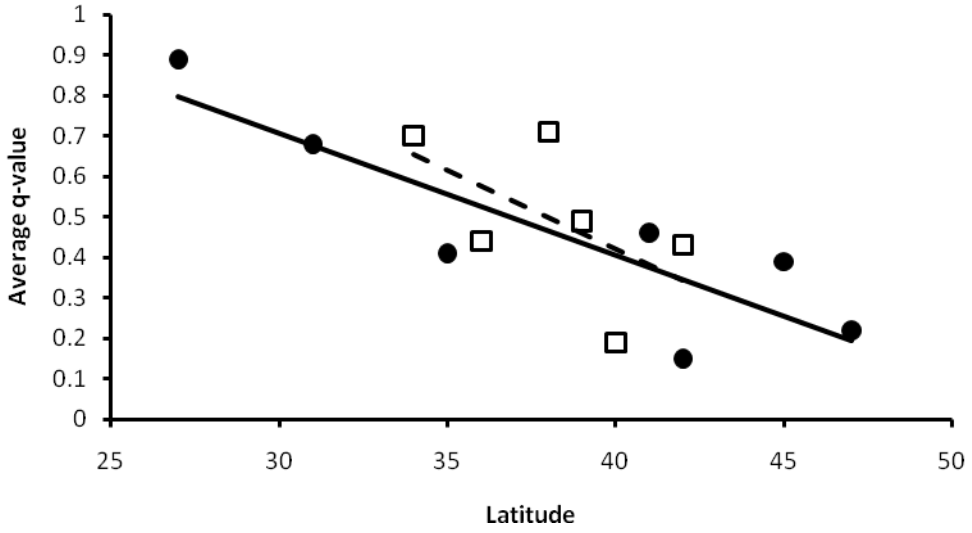


Figure 3.5 Demographic fluctuations based on a Bayesian skyline plot derived from 601 bp of the cytochrome *b* gene for striped skunks. The x-axis represents time in the past calculated in units of mutations per site; the y-axis depicts the population size as $N_e\theta$. The black line is the median population size estimate, and the shaded lines represent the upper and lower 95% highest posterior density intervals. For reference, we have highlighted the coldest period around the Last Glacial Maximum (20,000 years ago).

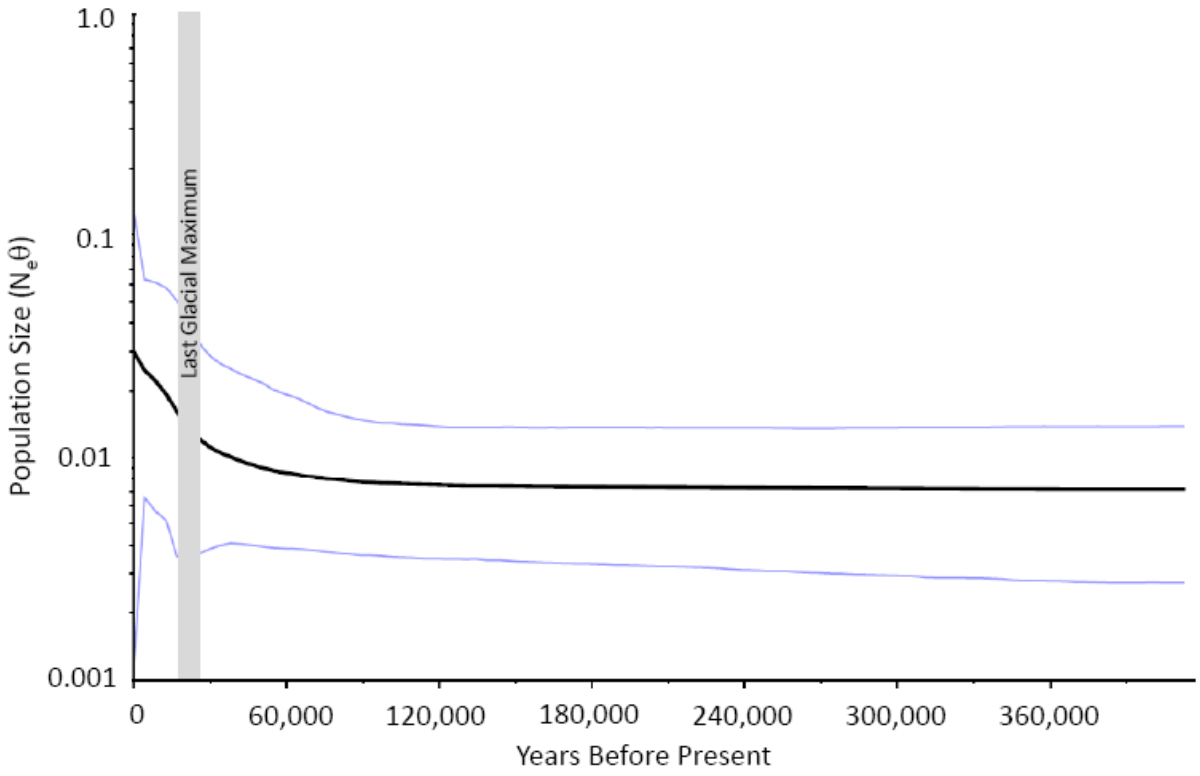


Figure 3.6 Timeline of Pleistocene glacial cycles. Glacial periods are indicated in bold. Glacials or interglacials during which stock-clade and clade-clade splits occurred are indicated by arrows.

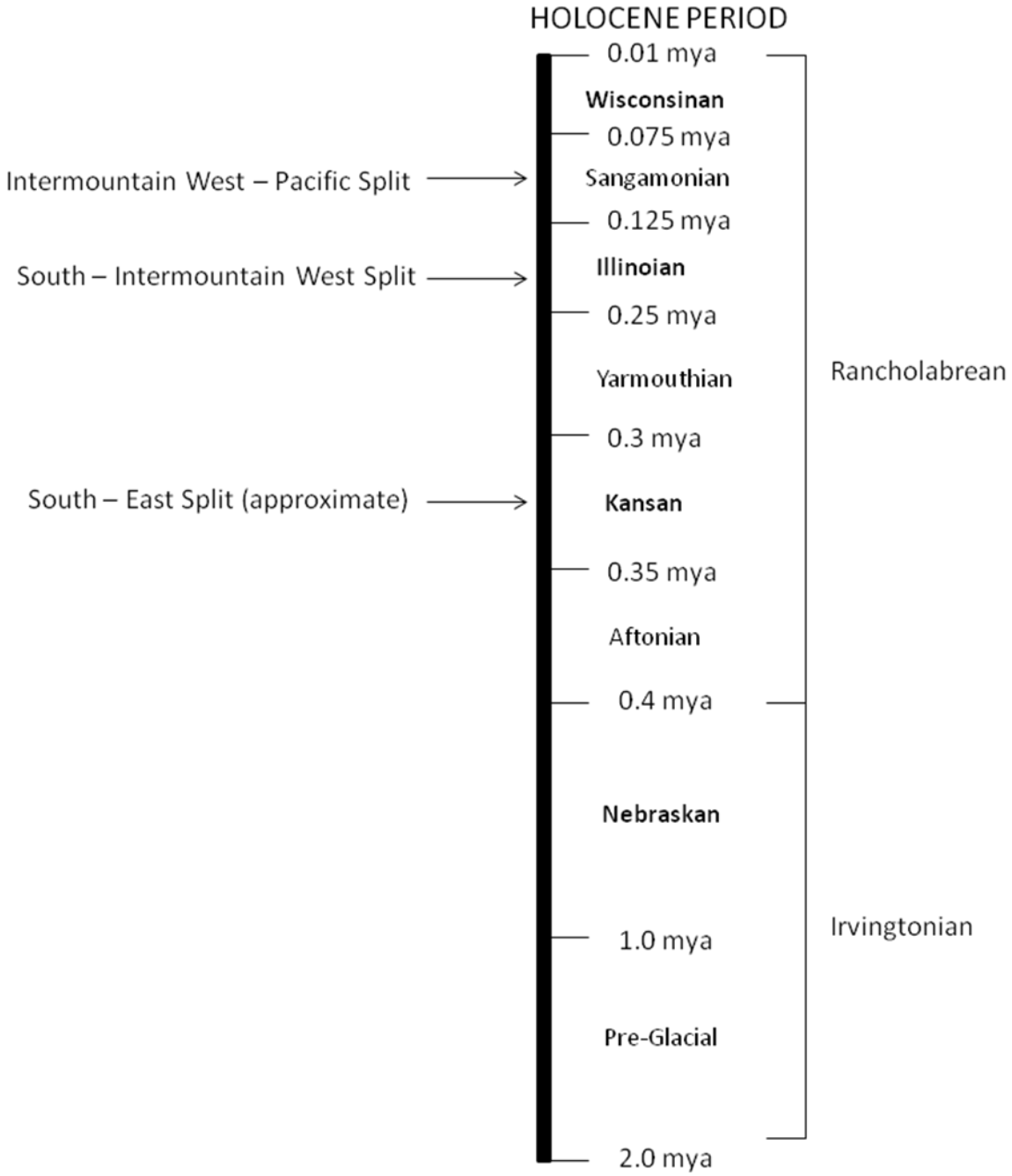


Table 3.1 SAMOVA groupings (A-E) based on the cytochrome *b* dataset for K = 2, 3, 4, and 5 groups.

Population	K=2	K=3	K=4	K=5
West	A	A	A	A
Plains	A	A	A	A
South	A	C	C	C
Great Lakes	B	B	B	B
New England	A	A	D	D
East	B	B	B	E
Φ_{CT}	0.42	0.44	0.45	0.47
P	0.08	0.04	0.004	0.07

Table 3.2 Pairwise Φ_{ST} estimates based on cytochrome *b* sequence data from the 6 geographic sampling units (below the diagonal), and pairwise F_{ST} estimates based on microsatellite markers for the 5 contemporary striped skunk populations (above the diagonal). Asterisks indicate statistical significance ($\alpha = 0.05$) based on sequential Bonferroni correction. All 6 geographic sampling units and all 5 modern populations show significant differentiation from one another, which is evidence of population structuring throughout North America.

Population	West	Plains	South	Great Lakes	New England	East
West	-	0.04*	0.03*	0.07*†	0.13*	-
Plains	0.18*	-	0.02*	0.03*†	0.08*	-
South	0.64*	0.36*	-	0.03*†	0.09*	-
Great Lakes	0.46*	0.41*	0.53*	-	0.08*†	-
New England	0.32*	0.26*	0.64*	0.33*	-	-
East	0.82*	0.73*	0.82*	0.21*	0.67*	-

† indicates Great Lakes/East population combined.

Table 3.3 Estimates for stock-clade and clade-clade divergence times using 3 different programs based on cytochrome *b* sequences. The 95% credibility intervals for the posterior distributions of the estimated parameters in IMa2 and BEAST are indicated in parentheses.

Splitting Event	IMa2: Isolation with migration (x 1000 years)	BEAST: T_{MRCA} (x 1000 years)	NETWORK: Rho estimate	NETWORK: Rho estimate in years (x 1000 years)
South – East	209 (123, 2144)	358 (145, 602)	10.7 ± 2.52	287 ± 68
South – Intermountain West	149 (87, 228)	130 (60, 213)	6.11 ± 1.82	164 ± 49
Intermountain West – Pacific	132 (74, 207)	97 (17, 215)	3.87 ± 1.64	104 ± 44

Table 3.4 Haplotype and nucleotide diversities for cytochrome *b* sequences of 601 bp and for concatenated sequences of 982 bp in length for each of the 6 striped skunk geographic sampling units, and for all specimens.

	Cytochrome *b*				Concatenated			
Population	*n*	Haplotypes	Haplotype diversity	Nucleotide diversity	*n*	Haplotypes	Haplotype diversity	Nucleotide diversity
West	24	5	0.757 ± 0.056	0.004 ± 0.003	24	9	0.873 ± 0.042	0.006 ± 0.003
Plains	49	13	0.746 ± 0.056	0.009 ± 0.005	49	28	0.965 ± 0.011	0.009 ± 0.005
South	85	17	0.845 ± 0.034	0.005 ± 0.003	84	34	0.906 ± 0.022	0.006 ± 0.003
Great Lakes	50	14	0.842 ± 0.023	0.014 ± 0.007	49	24	0.931 ± 0.023	0.014 ± 0.007
New England	21	4	0.865 ± 0.027	0.011 ± 0.006	21	4	0.590 ± 0.105	0.010 ± 0.005
East	40	9	0.590 ± 0.033	0.004 ± 0.002	40	18	0.873 ± 0.039	0.006 ± 0.003
All Specimens	69	46	0.938 ± 0.006	0.015 ± 0.008	267	105	0.970 ± 0.005	0.014 ± 0.007

Table 3.5 Population genetic analyses of each of the 5 striped skunk populations, and for pooled data. Effective population size for the West and New England populations was smaller than for the other 3 populations, but all 5 populations had similar levels of heterozygosity and allelic richness. Two of the populations showed the presence of a population bottleneck, and the pooled data showed a marginally significant bottleneck signature. All 8 loci were out of Hardy-Weinberg equilibrium when all samples were pooled.

	AR	H_o	H_e	θ	Bottleneck	HWE
West	9.55 ± 0.46	0.659 ± 0.03	0.816 ± 0.02	2.45	P = 0.02	---
Plains	11.70 ± 0.40	0.791 ± 0.02	0.881 ± 0.01	3.18	P = 0.008	---
South	12.59 ± 0.34	0.807 ± 0.01	0.883 ± 0.02	3.41	P = 0.25	---
New England	8.38 ± 0.43	0.660 ± 0.03	0.745 ± 0.04	1.75	P = 0.25	---
Great Lakes/East	11.61 ± 0.34	0.763 ± 0.02	0.879 ± 0.01	3.19	P = 0.31	---
Pooled Data	12.88 ± 0.43	0.764 ± 0.01	0.896 ± 0.01	---	P = 0.05	---
22-70	23.00	0.756	0.933	---	---	P << 0.001
22-67	20.00	0.789	0.888	---	---	P << 0.001
22-14	18.00	0.759	0.910	---	---	P << 0.001
42-26	17.00	0.736	0.916	---	---	P << 0.001
42-15	18.00	0.699	0.881	---	---	P << 0.001
42-25	22.00	0.843	0.889	---	---	P = 0.02
22-19	17.00	0.789	0.886	---	---	P << 0.001
42-73	15.00	0.742	0.855	---	---	P << 0.001

CHAPTER 4 - Conclusions

Over the last 30 years, the rate of emergence of novel zoonotic diseases has risen at an increasingly rapid pace (Greger 2007). Disease emergence is dependent on several factors that all interact to form a cohesive disease pattern. These factors include the epidemiological properties and molecular evolution of the pathogen; host ecology and phylogeographic history; and the interactions of hosts, pathogens, and the landscape on which they persist. Thus, it is necessary to investigate multiple parts of a disease system in order to fully understand the emergence and spread of that disease. In our study of rabies virus in striped skunk populations in the central Great Plains, we combined host and pathogen genetic data and spatial data to characterize the landscape epidemiology of an important zoonotic disease. By combining genetic information and spatial data, I believe that we have enhanced the breadth of knowledge of rabies disease emergence and helped to form a more complete picture of the interactions among different influencing factors in disease systems. The genetic data that we collected for both striped skunks and rabies provided insight into host population genetics and pathogen properties throughout the central Great Plains where two different strains of striped skunk rabies coexist. The results of our genetic analyses also provided the basis for spatial modeling of rabies in the central Great Plains; our spatial analyses revealed that pathogen properties influence disease emergence and epidemiological properties on a landscape scale. The integration of pathogen epidemiology, host population genetics, and landscape features has led to several important conclusions regarding striped skunk rabies emergence and distribution across the landscape in the central Great Plains.

Through the integration of pathogen and host genetic analyses and spatial analysis in Chapter 2, we concluded that 1) North Central skunk rabies strain and South Central skunk

rabies strain exhibit different intrinsic viral properties, 2) striped skunks living in the central Great Plains constitute a single, large admixed population, and 3) while striped skunks are not confined by landscape features like rivers, these same features differentially influence the emergence and spread of the two rabies strains. South Central rabies exhibited more purifying selection in the N gene, less purifying selection in the G gene, and a higher infectivity than North Central rabies. These data indicated that the South Central rabies strain has more epizootic qualities than the North Central strain, which appears to be a slower burning, enzootic virus. Taken together, South Central rabies appears to be more transmissible than the North Central strain, but these same traits also make South Central rabies more susceptible to landscape barriers. Conversely, striped skunks displayed high levels of genetic admixture and mobility in our study area.

Spatial features have proven to be barriers to dispersal for many hosts and parasites. However, the high dispersal capacity of striped skunks combined with their ability to swim means that for healthy skunks, rivers should be minimal barriers to dispersal. Additionally, habitat alteration in the form of modern roads has led to more bridges, making many rivers easily traversable. Thus, rivers in the central Great Plains are not barriers to striped skunk dispersal. Yet, the different epidemiological characteristics exhibited by North Central and South Central rabies mean that incubation period, virulence, and transmission efficiency potentially interact differently with rivers. Incubation period is most likely responsible for the differential emergence patterns that we observed (Smith, Lucey, Waller, Childs, & Real 2002); a long incubation period allows an infected skunk to cross rivers because it is asymptomatic longer, while a short incubation period means that symptoms emerge much faster and result in an inability of the infected skunk to cross water or travel long distances.

These results have implications for rabies emergence in the central Great Plains. Striped skunks are ubiquitous in the region, and have the ability to disperse relatively long distances for a small-bodied carnivore. Depending on which rabies strain a particular skunk is infected with, that skunk may be able to travel further and to cross rivers before becoming symptomatic. Rabies can remain latent for a variable time, from a few days to a few months (Carey & McLean 1983), potentially allowing infected skunks to contact susceptible individuals in areas that are relatively distant from the original location of infection, increasing the potential for epidemic outbreaks in novel biota and locations.

We gained further insight into the movement and distribution of striped skunks by investigating the phylogeography of striped skunks throughout North America in chapter 3. Through phylogeographic analysis, we concluded that 1) the highly mobile nature of striped skunks allowed for multiple dispersal events during the Pleistocene, which gave rise to three distinct clades out of a stock population of striped skunks and 2) once Holocene warming permitted large tracts of habitable areas to become available, high levels of admixture among previously isolated lineages of striped skunks was responsible for the contemporary continental scale population structure of striped skunks. The Pleistocene was a time of fluctuating climate, alternately permitting and restricting striped skunk dispersal, which allowed individual lineages to develop on either side of major barriers (e.g. the Rocky Mountains, the Sierra Nevada Mountains, and the Mississippi River), but the Holocene marked the beginning of a warming trend, during which time increasing temperatures and improved habitat enabled the expansion of striped skunks throughout North America. Different amounts of admixing during the Holocene and Recent eras among clades in different regions of the United States has led to different levels of secondary contact among clades, and resulted in contemporary populations that are still

largely structured according to biogeographic barriers, which is evidenced by the fact that rivers in the Great Plains are not barriers to dispersal of striped skunks, but the Mississippi River is.

Based on the conclusions from chapters 2 and 3, we infer that a holistic approach is essential to understanding the ecology underlying emerging diseases. Emerging diseases are tied very closely to the landscape on which they emerge, whether directly or through their host connections. As the landscape changes, so too does the range of the host and its pathogens and parasites. We know from studying the biogeography of striped skunks through the Pleistocene, Holocene, and Recent eras that striped skunks have responded favorably to past climate warming; they expanded throughout all of the United States into their current distribution. Thus, they should also respond favorably to current warming trends by expanding their range further north into Canada. Climatological data have indicated that the greatest warming will occur at high latitudes, and tundra will likely be replaced by taiga in the near future (Waltari, Hoberg, Lessa, & Cook, 2007); as ecosystems continue to shift north as a result of global climate change, striped skunks should expand further north into Canada as well, and because striped skunks are major disease reservoirs, we would also expect to see a corresponding range shift in their pathogens. As these pathogens and parasites expand, we should see diseases emerging in areas where they have not been previously observed at northern latitudes. For instance, persistent epizootics of red fox rabies have occurred in southern Ontario since the 1950's (Nadin-Davis, Muldoon, & Wandeler 2006). A number of canid species (i.e. grey fox, red fox, gray wolf, coyote, and striped skunk and raccoon to a lesser degree) reside in this region, all of which can serve as rabies reservoirs, and as rabies epizootics continue, the potential for emergence of a novel terrestrial rabies strain in the region increases. Another example is the emergence of bluetongue virus (BTV) among ruminants in northern Europe. BTV gradually expanded

northward from northern Africa and the south Mediterranean (Tomley & Shirley 2009) as midges, which are responsible for transmitting the virus, shifted their range to the north in response to increasing northern temperatures, leading to BTV emergence at increasingly northern latitudes, and most recently, northern Europe. As certain ruminant species are very susceptible to Bluetongue disease, and mortality rates can be high (Tomley & Shirley 2009), continued epizootics of Bluetongue disease in northern Europe could have adverse effects on the livestock industry.

Comparing the phylogeographies of disease reservoir species and their pathogens also provides insight into how future climate change is likely to affect disease emergence. Because many emerging infectious diseases are caused by RNA viruses, which by necessity require a host in order to reproduce, the expansion and emergence of these diseases should mirror the range expansion of their hosts. Several comparative studies of host and parasite phylogeography have shown that comparative phylogeography is a useful tool to determine how parasites disperse in relation to their hosts (Brooks & Hoberg, 2007; Hoberg & Brooks, 2008; Holmes, 2004). Investigating the phylogeography of both host and pathogen can aid in determining the likelihood for pathogen host-shifts and the potential for pandemics in novel environments, as in the case of *Staphylococcus aureus* jumping from humans to poultry (Lowder, et al., 2009), or even the co-evolution of hosts and parasites, as in lice in primates (Reed, Light, Allen, & Kirchman, 2007) and seabird ticks (Kempf, Boulinier, de Meeûs, Arnathau, & Mccoy, 2009). Finally, by investigating phylogeographies of host and pathogen, we can better predict where a disease may have originated as well as where it might emerge in the future, as many studies of hantaviruses and their rodent hosts in North and South America and Europe have shown (Dragoo, et al., 2006; Kang, et al., 2009; Medina, et al., 2009). These studies, among others,

indicate that as habitat changes, hosts shift their ranges accordingly, and their pathogens shift with them. These range shifts open up entirely new environments with susceptible hosts, providing pathogens with ample opportunities to switch host, thereby establishing new reservoirs and new possibilities for zoonosis. Thus, understanding the connections among epidemiological characteristics of pathogens, host ecology and phylogeography, and landscape features is critical to accurately predict disease emergence patterns. Indeed, the increased emergence of several zoonotic diseases, including severe acute respiratory syndrome (SARS), Avian Influenza, and West Nile Virus have been well documented in recent years (Greger, 2007; Lee & Krilov, 2005; Reisen & Brault, 2007), reinforcing the need to simultaneously investigate all of the factors that influence disease emergence. The comparative host-pathogen system presented in this dissertation may be useful as a model on which to base future studies of emerging infectious diseases.

References

Brooks, D. R., & Hoberg, E. P. (2007). How will global climate change affect parasite-host assemblages? *Trends in Parasitology, 23*, 571-574.

Carey, A. B., & McLean, R. G. (1983). The ecology of rabies – evidence of co-adaptation. *Journal of Applied Ecology, 20*, 777-800.

Dragoo, J. W., Lackey, J. A., Moore, K. E., Lessa, E. P., Cook, J. A., & Yates, T. L. (2006). Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a predictive framework for research on hantaviruses. *Journal of General Virology, 87*, 1997-2003.

Greger, M. (2007). The Human/Animal Interface: Emergence and Resurgence of Zoonotic Infectious Diseases. *Critical Reviews in Microbiology, 33*, 243-299.

Hoberg, E. P., & Brooks, D. R. (2008). A macroevolutionary mosaic: episodic host-switching, geographical colonization, and diversification in complex host-parasite systems. *Journal of Biogeography, 35*, 1533-1550.

Holmes, E. C. (2004). The phylogeography of human viruses. *Molecular Ecology, 13*, 745-756.

Kang, H. J., Arai, S., Hope, A. G., Song, J.-W., Cook, J. A., & Yanagihara, R. (2009). Genetic diversity and phylogeography of Seewis virus in the Eurasian common shrew in Finland and Hungary. *Virology Journal, 6*, 1-6.

Kempf, F., Boulinier, T., de Meeûs, T., Arnathau, C., & Mccoy, K. D. (2009). Recent evolution of host-associated divergence in the seabird tick, *Ixodes uriae*. *Molecular Ecology, 18*, 4450-4462.

Lee, P. J., & Krilov, L. R. (2005). When animal viruses attack: SARS and Avian Influenza. *Pediatric Annals, 34*(1), 42-52.

Lowder, B. V., Guinana, C. M., Ben Zakour, N. L., Weinert, L. A., Conway-Morris, A., Cartwright, R. A., et al. (2009). Recent human-to-poultry host jump, adaptation, and

pandemic spread of *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 19545-19550.

Medina, R. A., Torres-Perez, F., Galeno, H., Navarrete, M., Vial, P. A., Palma, R. E., et al. (2009). Ecology, genetic diversity, and phylgeographic structure of Andes Virus in humans and rodents in Chile. *Journal of Virology, 83*, 2446-2459.

Nadin-Davis, S. A., Muldoon, F., & Wandeler, A. I. (2006). Persistence of genetic variants of the arctic fox strain of rabies virus in southern Ontario. *Canadian Journal of Veterinary Research, 70,* 11-19.

Reed, D. L., Light, J. E., Allen, J. M., & Kirchman, J. J. (2007). Pair of lice lost or parasites regained: the evolutionary history of anthropoid primate lice. *BMC Biology, 5*, 1-11.

Reisen, W., & Brault, A. C. (2007). West Nile virus in North America: perspectives on epidemiology and intervention. [Review]. *Pest Management Science, 63*(7), 641-646.

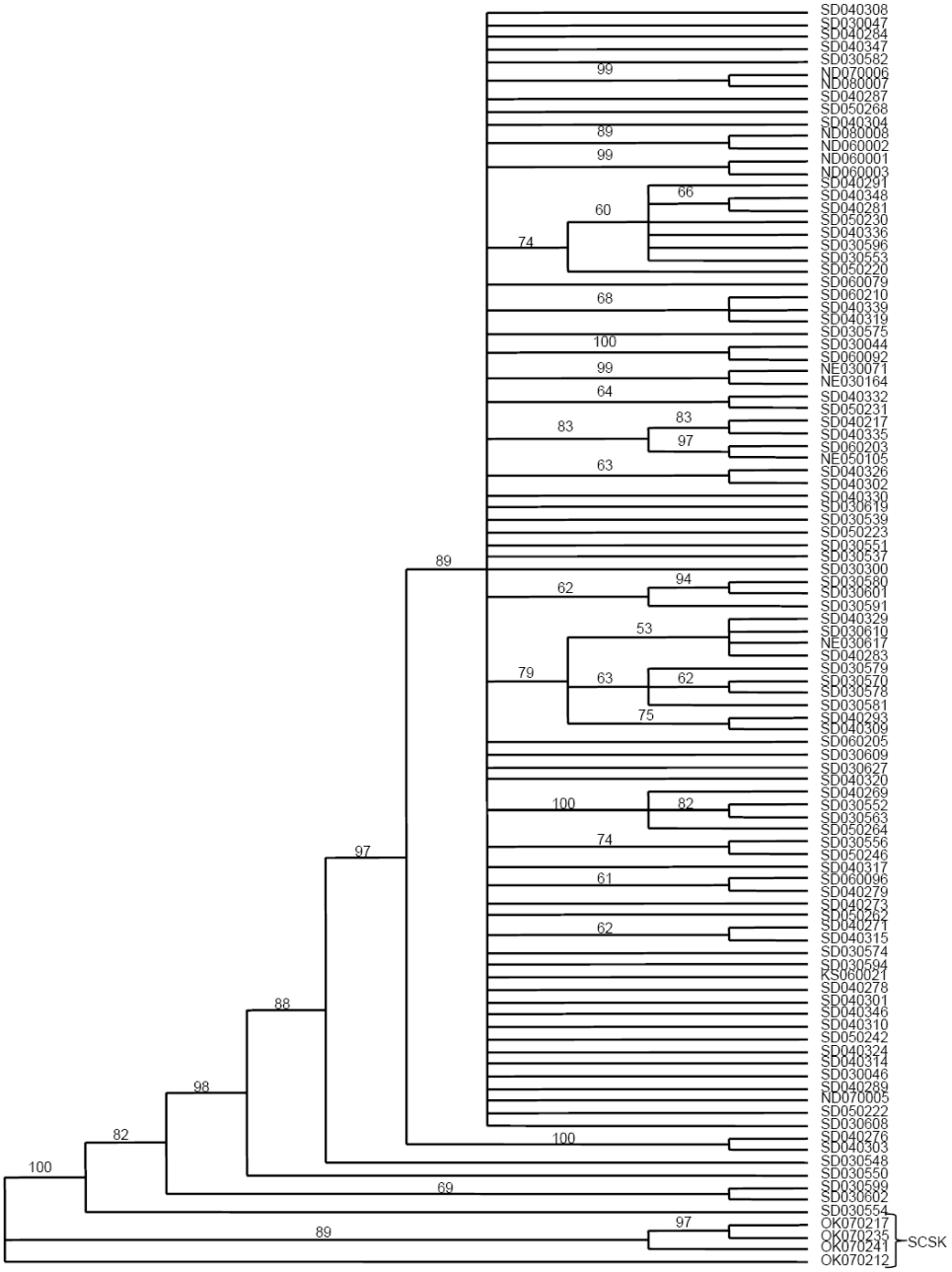
Smith, D. L., Lucey, B., Waller, L. A., Childs, J. E., & Real, L. A. (2002). Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 3668-3672.

Tomley, F. A., & Shirley, M. W. (2009). Livestock infectious diseases and zoonoses. *Philosophical Transactions of the Royal Society of London, B. Biological Sciences, 364*, 2637-2642.

Waltari, E., Hoberg, E. P., Lessa, E. P., & Cook, J. A. (2007). Eastward Ho: phylogeographical perspectives on colonization of hosts and parasites across the Beringian nexus. *Journal of Biogeography, 34*, 561-574.

Appendix A - Electronic Supplementary Material to Chapter 2

Figure 4.1 Bootstrap consensus cladogram of all 98 NCSK N gene samples implemented using the neighbor-joining method in PAUP* and rooted with 4 SCSK N gene samples. Numbers on the branches indicate bootstrap support for the branch.



Appendix B - Electronic Supplementary Material to Chapter 3

Figure 4.2 Concatenated median-joining network based on 982 bp for 304 striped skunk specimens. Geographic sampling units are indicated with unique colors. Branch lengths are proportional to the number of substitutions, and circle size is proportional to the number of individuals represented. Clades are indicated by dotted lines, and the stock population is indicated by a dashed line.