INTEGRATED LANDSCAPE GENETICS OF THE LYME DISEASE VECTOR (*IXODES SCAPULARIS*) AND ITS HOST (*PEROMYSCUS LEUCOPUS*) IN SOUTHERN QUÉBEC

by

Sarah Sin Tong Leo

Department of Biology

McGill University, Montreal

April, 2016

A thesis submitted to McGill University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biology

For my Family

Thank you

SUPERVISORY COMMITTEE

Dr. Virginie Millien	Redpath Museum, McGill University
Dr. Andrew Gonzalez	Department of Biology, McGill University
Dr. Nicholas Ogden	Faculté de Médecine Vétérinaire, Université de Montréal
	Environmental and Zoonotic Infectious Diseases, Public Health Agency of Canada
Dr. Denis Réale	Département de Sciences Biologiques, Université du Québec à Montréal

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMOVA:	Analysis of Molecular Variance
ANOVA:	Analysis of Variance
bp:	base pairs
EID:	Emerging Infectious Disease
gpPCA:	between-group PCA
GPS:	Global Positioning System
IBD:	Isolation-by-Distance
IBR:	Isolation-by-Resistance
MCMC:	Markov chain Monte Carlo
MDNN:	Minimum Distance to the Nearest Neighbor
MTILG:	Multi-taxa Integrated Landscape Genetics
PCR:	Polymerase Chain Reaction
PCA:	Principal Component Analysis
PCoA:	Principal Coordinate Analysis
RDA:	Redundancy Analysis
SEM:	Structural Equation Model
sPCA:	Spatial Principal Components Analysis
UPGMA:	Unweighted Pair Group Method with Arithmetic Mean

ABSTRACT

Understanding how interactions between vectors, hosts, and pathogens, as well as their responses to climate and landscape conditions, influence disease emergence events is crucial for planning effective disease monitoring and control strategies. In southern Québec, the emergence of Lyme disease has been attributed to the northward range expansion and increase in abundance of both the disease vector - the black-legged ticks (*Ixodes scapularis* Say), and the reservoir host – the white-footed mice (*Peromyscus leucopus* Rafinesque). The goal of my thesis is to quantify interactions between populations of black-legged ticks and white-footed mice in southern Québec, as well as their individual responses to landscape heterogeneity using the multi-taxa integrated landscape genetics approach (MTILG). MTILG can provide valuable insights on biotic and abiotic variables that influence zoonotic disease emergence in a region by revealing correspondences in spatial genetic patterns that distinguish possible cases of species-coupled dispersal or shared spatial dependence, and identifying landscape variables and the extent to which they influence species movement.

When I investigated white-footed mouse population structure relative to the landscape and examined spatial morphological variation among populations, I found that white-footed mouse populations in southern Québec exhibited significant genetic differences among populations that could be explained by the presence of barriers to gene flow. There was also evidence of adaptations for dispersal and range expansion in the form of a significant and positive correlation between mouse foot length and latitude. I found evidence of weak genetic differentiation among populations of black-legged ticks, as well as multiple genetic signatures of frequent long-distance dispersal events, supporting the hypothesis that migratory passerine birds are the primary transport of black-legged ticks into southern Québec. Furthermore, I found that tick genetic diversity exhibited greater temporal variation than spatial variation, a result that I was able to utilize to determine stages of establishment for each tick population.

Finally, using structural equation modeling that incorporated insights from the previous chapters, along with climate and landscape variables, I was able to produce a model that successfully projected current Lyme disease risk observed in the field. The model revealed that Lyme disease infection risk (defined here as the prevalence of *Borrelia burgdorferi* at a site) was significantly dependent on climate-based disease vector occurrence probability. This model can

be a useful tool for projecting future Lyme disease infection risk in response to changing climate and landscape conditions, and can also be used to assess the potential effectiveness of different disease intervention strategies (e.g. land management policies vs. host culling).

The research in this thesis reveal the importance of an integrated approach for understanding some of the underlying variables influencing disease emergence events, and for more accurate and reliable predictions of disease risk that are essential for disease control and monitoring efforts. Insights obtained from my work will prove valuable in future efforts to control and monitor Lyme disease emergence in southern Québec, and at the same time serve as a useful source for raising both public and physician awareness of the emerging Lyme disease threat in Québec and the rest of Canada.

RÉSUMÉ

Comprendre comment les interactions entre les vecteurs, les hôtes et les agents pathogènes, ainsi que leurs réponses au changement climatique et la structure du paysage, influencent l'émergence des maladies zoonotiques infectieuses est essentielle pour la planification efficace des stratégies de surveillance et de contrôle de ces maladies. Dans le sud du Québec, l'émergence de la maladie de Lyme a été attribuée à l'expansion vers le nord et l'augmentation de l'abondance à la fois du vecteur de la maladie - la tique à pattes noires (*Ixodes scapularis* Say), et de l'hôte réservoir - la souris à pattes blanches (*Peromyscus leucopus* Rafinesque). Par conséquent, l'objectif de ma thèse est de quantifier les interactions entre les populations de tiques à pattes noires et de souris à pattes blanches dans le sud du Québec, ainsi que leurs réponses individuelles à l'hétérogénéité du paysage en utilisant une approche génétique intégrée multi-taxons du paysage (MTILG) (chapitre 2). Cette approche peut fournir des indications précieuses sur les variables biotiques et abiotiques qui influencent l'émergence des maladies zoonotiques dans une région, en révélant des correspondances spatiales entre les patrons de différentiation génétique qui distinguent les cas possibles de dispersion d'espèces-couplée ou partagée, et d'identifier la dépendance spatiale avec les variables du paysage et dans quelle mesure elles influencent le mouvement des espèces.

Lorsque j'ai étudié la structure de la population de la souris à pattes blanches par rapport au paysage et examiné la variation morphologique spatiale, j'ai trouvé que les populations de souris à pattes blanches dans le sud du Québec présentaient des différences génétiques importantes entre les populations qui pouvaient être expliquées par la présence de barrières au flux génique. J'ai aussi détecté une signature possible d'adaptations pour la dispersion et l'expansion de l'aire de distribution de la souris sous la forme d'une corrélation positive entre la longueur de pied et la latitude. La tique à pattes noires présente une faible différenciation génétique entre les populations, ainsi que la signature génétique de fréquentes dispersions à longue distance, qui appuie l'hypothèse selon laquelle les oiseaux passereaux migrateurs sont la source primaire des tiques à pattes noires dans le sud du Québec. En outre, j'ai trouvé que la diversité génétique de tiques présente une plus grande variation temporelle que spatiale, un résultat que j'ai été en mesure d'utiliser pour déterminer les stades d'établissement de chaque population de tiques.

Enfin, à l'aide d'un modèle d'équation structurelle et en utilisant les résultats des chapitres précédents sur l'influence du climat et des variables du paysage, j'ai été en mesure de produire un modèle qui a projeté avec succès le risque de la maladie de Lyme observé sur le terrain. Le modèle a révélé que bien que le risque d'infection par la maladie de Lyme (défini ici comme la prévalence de *Borrelia burgdorferi* à un site) était significativement dependant de la probabilité d'occurrence du vecteur de maladies en fonction du climat. Ce modèle peut être un outil utile pour la projection du risque futur d'infection par la maladie de Lyme en réponse à l'évolution du climat et des conditions du paysage, et peut également être utilisé pour évaluer l'efficacité potentielle de différentes stratégies d'intervention (par ex. les politiques de gestion des terres ou l'eradication des hôtes).

Mes travaux de recherche révèlent l'importance d'une approche intégrée pour la compréhension de certaines des variables sous-jacentes qui influencent l'émergence des maladies zoonotiques et obtenir des informations plus précises et fiables pour prédire le risque de maladie qui sont essentielles pour les efforts de contrôle et de surveillance de la maladie. Les resultats obtenus à partir de ce projet s'avéreront utiles pour les efforts futurs pour contrôler et surveiller l'émergence de la maladie de Lyme dans le sud du Québec, ainsi que de servir de ressource pour augmenter la sensibilisation du public et des médecins sur la menace de la maladie de Lyme en émergence au Québec et dans le reste du Canada.

ACKNOWLEDGEMENTS

There are many people I would like to thank for their generous support, guidance, assistance, and encouragement over the years of my degree. First on the list are my supervisors, Dr. Virginie Millien and Dr. Andrew Gonzalez. I am extremely grateful for the opportunity to undertake such an interesting project and during an exciting time of Lyme disease emergence in Canada. I would also like to thank the members of my supervisory committee, Dr. Nicholas Ogden (Université de Montréal, Public Health Agency of Canada), Dr. Denis Réale (Université du Québec à Montréal) for their guidance and assistance. I have learnt much from your advice and you have each contributed to making my thesis as well-rounded as possible. I would also like to convey my appreciation to Dr. Lauren Chapman (McGill University) and Dr. Andrew Hendry (McGill University) for sitting in my Ph.D. Qualifying Examination in 2013. And last but not least, I thank my external and internal thesis examiners for evaluating my thesis.

I am extremely thankful to my immediate family for their unwavering support, unconditional love and unlimited understanding through the years as I worked on this research project. Without their encouragement, I would not have been able to embark on this research project so effortlessly.

I am grateful to all members of the Millien and Gonzalez lab (both past and present) who have been most generous in sharing their experiences and wisdom: Ronan Ledevin, Juan Martinez, Adrien Andre, Cedric Boue, Jorge Gaitan, Julie Simon, Robby Marrotte, Ted Daigle, Anita Rogic, Rodrigo Lima, Emilie Roy-Dufresne, Alan Garcia-Elfring, Julia Nordlund, Shaun Turney, Bronwyn Rayfield, Katie Millette, Catalina Gomez, Alessandra Loria, Michael Pedruski, Patrick Thompson, Edward Wong, Matthew Mitchell, and Carly Ziter. Thank you for all the laughter, the games, the all-you-can-eat sushi outings, Friday afternoon drinks, and horrendously inappropriate lunch conversations. More importantly, thank you for always being available when I needed to rant and for being my sounding boards.

I must also express my sincerest gratitude to the many undergraduate volunteers and work study students that have been so very essential in helping me out in the field as well as with specimen processing: Samantha Morin, Noreen Mohsin, Krishanth Manokaran, Dahn Hahm, Joshua Baik, and Carol Zastavniouk. I hope all of you have great experiences as you wrap up

your degree programs in McGill and I wish you all the best and great success in all future endeavors.

I would also like to thank Dr. Robbin Lindsay and his team at the Zoonotics and Special Pathogens division of the National Microbiology Laboratory (Public Health Agency of Canada). I greatly appreciate your assistance in screening the specimens for *Borrelia* infections. The data you have provided are not only essential for the successful completion of this thesis; it has and will continue to be useful information for any future member of the Millien lab.

Last but not least, I would like to thank the landowners and organisations whom had kindly granted us access to their property in order to collect specimens.

This study was funded by a CGS-D NSERC Fellowship to me and a FQRNT Team Grant #147236 to Drs. Millien and Gonzalez. Dr. Gonzalez is further supported by the Canada Research Chair Program and an NSERC Discovery Grant. Dr. Millien is supported by an NSERC Discovery Grant. We also acknowledge the support of the Quebec Centre for Biodiversity Science.

PREFACE

All of the work presented henceforth were conducted the McGill University downtown campus (Montreal, QC). This research was approved by the McGill Animal Care Committee (AUP#5420) and the Quebec government (SEG permits 2011-05-15-014-00-S-F, 2012-07-16-1417-16-17-SF, 2013-07-04-1530-04-14-16-17-SF and 2014-05-02-1638-05-16-SF).

Chapter 2 has been published in Genome [Leo, S. S. T., Gonzalez, A., and Millien, M. Multi-taxa integrated landscape genetics for zoonotic infectious diseases: Deciphering variables influencing disease emergence]. I was the lead investigator, responsible for concept formation, data collection and analysis, as well as manuscript composition. Both Gonzalez A, and Millien V were also involved in concept formation and contributed to manuscript edits. They were also the supervisory authors on this project.

Chapters 3, 4 and 5 have all been formatted for publication. I was the lead investigator for the projects discussed in Chapters 3 to 5 where I was responsible for all major areas of concept formation, data collection and analysis, as well as the majority of manuscript composition. Lindsay R and his team at the National Microbiology lab contributed to data collection. Both Gonzalez A, and Millien V were the supervisory authors on these projects and were also involved in concept formation and contributed to manuscript edits.

This thesis is one of the first multi-taxa, integrated attempt to examine Lyme disease emergence patterns and rates in southern Québec based on molecular data collected from the disease vector (black-legged tick) and reservoir host (white-footed mouse). I identified mechanisms of species range expansion as well as uncover species adaptations/traits/roles that can potentially facilitate disease emergence rates. Furthermore, by quantifying disease prevalence relative to species interspecific interactions as well as their responses to the climate

and surrounding landscape, I was able to project disease infection risk in my study area. This thesis produced valuable insights and tools to help better understand disease emergence processes from an ecological perspective.

CHAPTER ONE

General Introduction

EMERGING INFECTIOUS DISEASES

An emerging infectious disease is defined as 'an infectious disease whose incidence increases following its first introduction into a new host population or whose incidence increases in an existing host population as a result of long-term changes in [its] underlying epidemiology' (Woolhouse and Dye 2001). Disease emergence events have often been attributed to several factors acting independently or in unison (Jones et al. 2008, Engering et al. 2013). For example, genetic drift or shifts resulting from selective pressures have resulted in microbial adaptations that have produced drug-resistant pathogen strain types (Schrag and Wiener 1995). Similarly, evolutionary processes and increase encounter rates between humans and wildlife have also allowed for cross species infections, producing new pathogenic species or strain types that are now capable of infecting humans (Daszak et al. 2000). Meanwhile, climate change and land use change have resulted in ecological changes across different regions of the world, promoting disease vector and pathogen invasion (Patz et al. 2008, Johnson et al. 2015). Finally, changes in human demography have further facilitated the rapid spread of diseases across the world, rendering disease outbreaks difficult to control without employing excessive measures like mandatory quarantines and mass health screenings. Struggles to contain and manage the sudden and rapid emergence of these diseases can be labour-intensive, time-consuming, and resource-draining (Daszak et al. 2000).

Emerging vector-borne zoonotic diseases

Zoonoses, or infectious diseases that can be transmitted between animals and humans, are of particular interest to health officials. According to the Centers for Disease Control and Prevention (CDC), over 70% of recently emerging infectious diseases has animal origins and roughly 60% of all human pathogens are zoonotic (Jones et al. 2008, Johnson et al. 2015). Well-known examples of zoonotic pandemics include the plague in the 14th century, as well as the more recent H5N1 and H1N1 avian influenza viruses. Furthermore, genetic evidence has suggested that many diseases now known to transmit freely human-to-human (such as HIV/AIDS, SARS, measles, tuberculosis, and smallpox) originated as adaptive strains of pathogens with animal origins (Li et al. 2005, Pearce-Duvet 2007, Li et al. 2007, Furuse et al. 2010, Sharp and Hahn 2011). Common wildlife species often identified as sources of emerging zoonotic diseases include bats, birds, primates, small rodents, and sometimes livestock and pets.

Zoonotic disease can come in many forms. Some of them involve direct transmission between wildlife and humans, whereby animals are reservoir hosts that maintain populations of disease agent within their bodies. In these cases, humans are often dead-end hosts for the pathogens in the sense that humans are essentially the end of the transmission line for the disease pathogen. Alternatively, zoonotic diseases may require a medium, also known as a vector, through which transmission occurs. Common vectors include hematophagous arthropod species like mosquitoes, fleas, and ticks. For vector-borne diseases, the presence and distribution of competent vectors are often employed as measures of disease risk. As a result, many control methods for these zoonotic disease transmission systems generally aim to control or reduce vector populations using combinations of chemical and biological controls (e.g. Vreysen et al. 2000).

In North America, a recent emerging zoonotic infectious disease that has been of increasing concern is the tick-borne spirochete disease – Lyme disease.

LYME DISEASE

Lyme disease, or Lyme borreliosis, is an emerging infectious zoonotic disease that is caused by spirochete bacteria from the *Borrelia burgdorferi* complex and has been reported across temperate regions of Europe, Asia and North America (Marques 2010). However, historical records suggest that the disease may have been present in Europe centuries prior to its discovery in 1975 in the towns of Lyme and Old Lyme, Connecticut, after which the disease was named (Benach and García-Moncó 2010). Today, Lyme disease is considered the most common vector-borne disease in North America (Ostfeld 2011).

Epidemiology

The Lyme disease transmission cycle involves three sets of organisms: the pathogen (*Borrelia burgdorferi* sensu lato), the vector (ticks from the genus *Ixodes*), and competent reservoir hosts (Ostfeld 2011). The pathogen is generally acquired by larval ticks during their first blood meal from infected hosts (usually small mammals) (Marques 2010) (Figure 1.1). After molting to the nymphal stage of the life cycle, infected ticks can transmit the pathogen to the next host that provides its second blood meal, creating new reservoirs that perpetuate the disease transmission cycle (Figure 1.1). The relative efficacy and success of disease transmission between ticks and hosts depends on host competency (defined as the host's ability to sustain or make pathogens available to a vector). In the final adult stage of the tick's life cycle, ticks will feed almost exclusively on larger mammals that are often poor reservoir hosts for the pathogen

(Lane et al. 1991, Jaenson and Tälleklint 1992) (Figure 1.1). Studies have shown that *Borrelia* spirochetes are rarely, if never, transmitted transovarially – from the adult female to its eggs (Magnarelli et al. 1987, Zhioua et al. 1994, Rollend et al. 2013), therefore the feeding behaviour of juvenile ticks is crucial for maintaining pathogen population in a community.

The risk of acquiring Lyme disease varies with distribution, density, and prevalence of infection in endemic tick populations (Marques 2010). Lyme disease is endemic in several areas across Europe and Asia. In North America, most reports of Lyme disease in human occur in Northeastern and Midwestern United States while incidences in the west and Southeastern USA remain low (Margos et al. 2011). Such regional variation in disease cases is attributed to host preference and human biting behaviour of the main vectors in each region, varying pathogenicity of different genetic strains of the bacteria, and other ecological factors (e.g. reservoir hosts competence or climate driven tick seasonality) (Killilea et al. 2008, Margos et al. 2011). Although both nymphal and adult ticks are capable of transmitting the pathogen, majority of Lyme disease cases in humans tend to be the result of bites from infected nymphs as they are often difficult to spot given their relatively small size (Marques 2010). Laboratory studies suggest that successful transmission of *Borrelia* from tick to human is a long process that requires at least 24 hours (Marques 2010). However, a recent literature review suggests that *Borrelia* infection can potentially occur in humans as early as 6 hours post-attachment although specific transmission times differ depending on the vector and pathogen species (Cook 2015).

The disease-causing organism – Borrelia burgdorferi *sensu lato*

The causative agents of Lyme disease in humans are bacteria from the genus *Borrelia*. *Borrelia* bacteria are members of the Spirochaetaceae family that have characteristic helically-shaped bodies with multiple endoflagella enclosed between the inner and outer cell membranes

(Aguero-Rosenfeld et al. 2005, Marques 2010). These bacteria can infect a wide variety of vertebrate animals including small mammals, lizards, and birds (Tilly et al. 2008). *Borrelia* spirochetes exhibit a great amount of taxonomic diversity, likely resulting from an explosive species radiation near the time of species origin (Brisson et al. 2012). It has been suggested that up to nine species of *Borrelia* spirochete, all of which belonged to the *Borrelia burgdorferi* sensu lato complex, can cause Lyme disease in humans (Schotthoefer and Frost 2015). However, the most commonly cited examples include three species: *B. burgdorferi* sensu stricto, *B. afzelii*, and *B. garinii* (Marques 2010, Brisson et al. 2012). These three *Borrelia* genospecies are antigenically distinct and have discrete geographical distributions (Marques 2010). While only *B. burgderferi* sensu stricto is present in North America, and *B. garinii* and *B. afzelii* common in Asia, all three genospecies can be found in Europe (Marques 2010).

Borrelia burgdorferi is the first spirochete bacteria to have had its complete genome sequenced (Aguero-Rosenfeld et al. 2005). Its genome is complex, consisting of segmentations and linear DNA molecules. (Brisson et al. 2012). Generally, *Borrelia* spirochetes have single linear chromosomes approximately 900,000 base pairs (bp) long and numerous linear and circular plasmids that range between 9,000 to 62,000 bp long (Casjens et al. 2010). Housekeeping genes are maintained on the chromosomes while plasmids encode for proteins that facilitate interactions between *Borrelia* and their arthropod vectors or vertebrate hosts (Casjens et al. 2010). Interestingly, despite its genomic complexity, *Borrelia* spirochetes chromosomes do not have genes that encode enzymes necessary for synthesizing amino acids, nucleotides, fatty acids, and enzyme cofactors (Brisson et al. 2012). Instead, the bacteria have over 50 genes that encode for transporters and/or binding proteins of carbohydrates, peptides, and amino acids, allowing the bacteria to import and scavenge what it needs from its vector and host (Brisson et

al. 2012). Researchers have also found that *Borrelia* spirochetes maintain a significant amount of chromosomal genes that encode for proteins important in chemotaxis and motility, allowing the pathogen to identify and approach optimal colonisation niches in their hosts (Brisson et al. 2012). *Borrelia* plasmids are the focus of much research due to their potential role in causing disease. Although no genes encoding for typical virulence factors have been found, *Borrelia* plasmids nonetheless carry important genes that facilitate bacterial infectivity, promote successful establishment in hosts, and aid in long term immune response evasion (Casjens et al. 2010). For example, outer surface proteins (e.g. *OspA* and *OspC*) play important roles in successful *Borrelia* transmission among hosts and vectors (Casjens et al. 2010, Brisson et al. 2012). Similarly, unusual genetic adaptations by *Borrelia* allow it to effectively evade host adaptive immune response in the mammalian host via antigenic variation of a surface lipoprotein (e.g. *vlsE* recombination) (Brisson et al. 2012).

Multilocus sequence analysis (MLSA) and multilocus sequence typing (MLST) studies on housekeeping genes of *Borrelia burgdorferi* sensu lato have revealed patterns of geographical structuring among populations of *Borrelia* spirochetes across the world (Heon et al. 2009, Margos et al. 2011, Margos et al. 2012). In Europe, *B. afzelii* and *B. garinii* exhibited population structures that imply restricted dispersal among different regions of Europe and Asia (Margos et al. 2011). However, the strength of these observed population structure differs depending on the host species with which the bacteria are associated. Unsurprisingly, European *Borrelia* populations associated with avian hosts tend to show more evidence of spatial mixing compared to those associated with small rodents (Margos et al. 2011). MLST studies on *B. burgdorferi* sensu stricto in North America have revealed that although the pathogen populations there originated from Europe (Margos et al. 2008), they are genetically distinct with limited gene flow

occurring between both continents (Margos et al. 2011). Additionally, divergence time estimates suggest that *B. burgdorferi* was present in North America several thousand years prior to European settlements (Heon et al. 2009). Further examination of North American *B. burgdorferi* spatial distributions of sequence types (STs) and inferred population boundaries revealed restricted pathogen gene flow among three separate regions: Western, Midwestern, and Northeastern USA (Heon et al. 2009, Margos et al. 2012). These three distinct sub-populations of the Lyme disease agent likely arose from a common ancestral population (Margos et al. 2012) that underwent multiple prehistoric population expansion events (Heon et al. 2009). Genetic evidence strongly suggests that *B. burgdorferi* populations in Northeastern USA originated from southern USA, followed by an east-to-west radiation into Midwestern USA (Heon et al. 2009). *B. burgdorferi* populations have since been persisting in separate refugia, only recently re-emerging in response to favourable population expansion conditions (i.e. increase in abundance of vectors and hosts).

The vector – Ixodes sp. ticks

Ticks are obligate hematophagous ectoparasites, feeding on animal blood and fluid (Allan 2001). Tick feeding activities can cause local irritation at the feeding sites which may become potential sites for secondary infections, and severe infestation on a single host individual can result in hypovolemia as well as tick paralysis (Allan 2001). More importantly, ticks are known to vector a wide variety of pathogens, making them important medical and veterinary arthropods (Allan 2001). All tick species today belong to one of three families: the Nuttalliellidae, the Argasidae (soft ticks), and the Ixodidae (hard ticks). Ixodidae forms the largest family of ticks, comprising over 650 species. Ixodid ticks are generally teardrop-shaped, with an anterior capitulum that is dorsally visible (Allan 2001). They have a hardened scutum that is sexually

dimorphic, covering the entire dorsal surface of the males and only the anterior portions in juvenile and female ticks (Allan 2001). All Ixodidae ticks have four life stages: egg, larva (three pairs of legs), nymph (4 pairs of legs and no genital pore), and adult (4 pairs of legs and a genital pore) (Allan 2001).

Lyme disease is transmitted primarily via bites of infected ticks from the largest genus of hard ticks in the world: *Ixodes.* The particular species of *Ixodes* tick vectoring Lyme disease differs from region to region. In Europe and Asia, the disease is transmitted by *I. ricinus* and *I. perculcatus*, respectively while *I. pacificus* and *I. scapularis* are the primary vectors in western and eastern North America, respectively (Andersen 1989, Lane et al. 1991, Eisen and Lane 2002, Xu et al. 2003). All four principal vectors of Lyme disease belong to the *Ixodes ricinus* species complex, however they are not sister species, suggesting that the ability to transmit *Borrelia* pathogens evolved independently multiple times (Xu et al. 2003). All *Ixodes* ticks are three-host ticks, with life cycles that span between one to two years (Allan 2001). Egg masses are usually laid in leaf litter. Larvae that emerge from these eggs will immediately seek to attach to their first host, usually small mammals, and will feed until fully engorged. Once blood meal is complete, the larvae drop to the ground and molt into nymphs. The nymphs in turn seek and attach themselves to a second host for their second round of feeding and will once again detach when replete and molt into adults on the ground. Unlike the immature ticks, adult ticks tend to seek out and attach to mid- or large-sized hosts. While female ticks require a blood meal for successful egg production, male ticks feed very little. Mating can occur on the host or on the ground prior to attachment (Allan 2001).

In eastern North America, *I. scapularis* is known to parasitize a broad range of host species (Allan 2001). Its distribution and abundance across eastern North America is influenced and

constrained by a combination of factors including habitat suitability, host abundance, climate, and host dispersal (Lane et al. 1991, Lindsay et al. 1998, Lindsay et al. 1999a, Lindsay et al. 1999b, Legér et al. 2013, Ostfeld 2011). For instance, *I. scapularis* ticks spend a significant part of their life cycle in the environment off the hosts, during which they are vulnerable to drowning, desiccation, predation, and freezing (Allan 2001, Legér et al. 2013). The presence of a protective layer of leaf litter can provide microhabitat conditions that offer the ticks some protection and contribute to off-host survival (Allan 2001, Lindsay et al. 1999b). In addition to habitat conditions, host availability also determines tick fitness and survival as blood meals from hosts are essential for the parasites' growth and development (Lane et al. 1991, Allan 2001, Legér et al. 2013). Finally, as arthropods, *I. scapularis* development rates are heavily dependent on climatic conditions. Insufficient degree day accumulations due to low ambient temperatures can hinder the completion of the tick's life-cycle (Lane et al. 1991), just as prolonged exposure to temperature extremes or drastic fluctuations in ambient temperature can severely reduce tick survival probability (e.g. Hermann and Gern 2013). Similarly, low ambient humidity (<80%) can greatly increase desiccation and reduce survival (Lane et al. 1991).

Studies on the phylogeography of *I. scapularis* in eastern North America using mitochondrial 16S ribosomal deoxyribonucleic acid (rDNA) revealed that *I. scapularis* populations there originated from a few migrants in southeastern USA that expanded rapidly into the northeastern states before subsequently colonising the Midwest in the last century (Humphrey et al. 2010). Qiu et al. (2002) found genetic signatures of recent colonization and rapid population expansion events in both Northeastern and Midwestern USA. The recent shared ancestry of *I. scapularis* populations in both regions is further evinced by the amount of phylogenetic similarity and lack of significant genetic structure among populations between both

regions. Indeed, estimates of divergence times found that *I. scapularis* populations in the Midwest are younger than those in the northeastern states, while populations in the southeastern states are the oldest (Humphrey et al. 2010).

The reservoir hosts and vector hosts

The Lyme disease transmission system comprises two groups of hosts, ones that are competent reservoirs for the disease pathogen, and others that are important for vector development, survival, and reproduction. Both sets of host species are not mutually exclusive. Species can simultaneously be competent hosts for both vectors and pathogens.

The local establishment of a disease transmission system requires maintaining a successful pathogen population, which in turn requires the presence of competent reservoir hosts that can efficiently receive and transmit the disease-causing pathogen. Zoonotic pathogens are usually host generalists, capable of infecting a broad spectrum of host species (Brunner at al. 2008). However, different host species vary extensively in their relative importance with regards to disease transmission and infection dynamics (Brunner at al. 2008). For example, variable host immune system responses can influence host infectivity/susceptibility, reservoir competence (ability of infected host to make pathogen available to a vector), and reservoir potential (ability of host to sustain transmission of pathogen to vector) (Mather et al. 1989, Brunner et al. 2008). This in turn determines host competency in disease transmission processes. This species-specific variation among host species has led to the hypothesis that disease incidence in a region is the function of host species community composition and biodiversity (Keesing et al. 2010). However, the exact mechanism and effects remain very much under debate (Randolph and Dobson 2012, Ostfeld 2013, Randolph 2013, Wood and Lafferty 2013, Ostfeld and Keesing 2013, Lafferty and Wood 2013).

Host species that serve as important reservoirs in Lyme disease transmission systems differ geographically. However for the most part, small mammals tend to be efficient reservoir hosts as large proportions of populations can often remain infected while exhibiting little to no disease symptoms (Lane et al. 1991, Tälleklint and Jaenson 1994, Gern et al. 1998, Marques 2010). Furthermore, small mammals tend to host large numbers of juvenile ticks and are often cited to be extremely efficient at transmitting *Borrelia* spirochetes to feeding ticks (Lane et al. 1991). Conversely, mid-size and large mammals usually serve as secondary reservoirs with low transmission efficacies or are altogether incompetent at disease transmission (Lane et al. 1991, Jaenson and Tälleklint 1992, Tälleklint and Jaenson 1994). Birds such as thrushes, wrens, and songbirds have also been proposed as competent reservoirs for *Borrelia burgdorferi* sensu lato in both Europe and North America but to a much lesser degree compared to small mammals (Lane et al. 1991, Gern et al. 1998, Comstedt et al. 2006, Marques 2010). However, it has also been suggested that the exact role avian hosts play in the Lyme disease transmission cycle in North America may actually be zooprophylactic, thus reducing *Borrelia* infection prevalence in ticks rather than amplifying it (Ogden et al. 2008).

In Europe, small reservoir hosts for *Borrelia garinii* and *B. afzelii* include field mice (*Apodemus sylvaticus, A. flavicollis*), voles (*Clethrionomys glareolus, Microtus agrestis*), dormice (genus *Glis*), and shrews (*Sorex araneus, S. minutus, Neomys fodiens*), while mid-sized secondary reservoirs consist of hedgehogs (genus *Erinaceus*), rats (genus *Rattus*), squirrels (genus *Sciurus*), and hares (*Lepus europaeus, L. timidus*) (Tälleklint and Jaenson 1994, Gern et al. 1998). Large mammals such as the roe and red deer are incompetent hosts (Jaenson and Tälleklint 1992). In eastern North America, the white-footed mouse (*Peromyscus leucopus*) is considered as the primary reservoir host of Lyme disease, although other small mammals like

chipmunks (*Tamias striatus*), squirrels (*Sciurus carolinensis*), and shrews (*Blarina brevicauda*) may serve as secondary reservoirs (Lane et al. 1991). Mid-size mammals like raccoons (*Procyon lotor*), opossums (*Didelphis virginiana*), and skunks (*Mephitis mephitis*) are known to carry *Borrelia burgdorferi* infection (Lane et al. 1991). Large ungulates like the white-tailed deer (*Odocoileus virginianus*) are poor reservoirs of the Lyme disease pathogen (Lane et al. 1991).

In addition to a thriving pathogen population, vector-borne zoonotic disease transmission system also requires the presence of well-established vector populations. In the Lyme disease transmission system, hosts are crucial components of the tick's life cycle. Tick development, survival, and reproductive success rates are heavily dependent on biological and behavioural differences among host species from which ticks obtain their blood meals (Allan 2001). Tick-host interactions are complex and dynamic, a perpetual arms race between host defences and tick counter-measures to said defense in the quest for successful blood feeding (Wikel 2013). Ticks are known to modulate host defenses via the secretion of various salivary peptides, lipids and proteins that counteract host itch or pain responses to cutaneous irritation, down-regulate hemostasis coagulation pathways, and evade or supress host immune responses (Wikel 2013). The ability of the tick vector to exploit one or several hosts will not only influence its own population dynamics and evolutionary trajectory, but also that of the pathogens it transmits (McCoy et al. 2013).

In Europe, the main vector of Lyme disease, *I. ricinus*, feeds indiscriminately on numerous species of reptiles, birds, and small to large mammals (Lane et al. 1991). For the most part, immature ticks tend to attach primarily to small rodents and insectivores while medium to large animals are preferred by adult ticks (Lane et al. 1991). Tick infestation in ground foraging birds can be five times that observed in other bird species, although the overall numbers are still

20-30 times lower than that observed in rodents (Comstedt et al. 2006). In western North America, immature *I. pacificus* frequently feed on western fence lizards (*Sceloporus accidentalis*) while adults are common on black-tailed deer (*Odocoileus hemionus columbianus*) (Lane et al. 1991). Rodents (*Peromyscus maniculatus* and *P. truei*), lagomorphs, and birds in the region generally yield much lower tick infestations despite their greater abundances (Lane et al. 1991). In eastern North America, *I scapularis* ticks have an extensive host range, with immature ticks parasitizing over 80 species of mammals and birds. The white-footed mouse (*P. leucopus*) is the host species most frequently parasitized by juvenile *I. scapularis* (Ostfeld 2011). Other small mammals often parasitized by *I. scapularis* include gray squirrels (*Sciurus carolinensis*), eastern chipmunks (*Tamias striatus*), and short-tailed shrews (*Blarina brevicauda*). Mid-sized, secondary mammalian hosts include racoons, opossums, and striped skunks. Adult *I. scapularis* are limited mostly to medium or large-sized mammals, with the white-tailed deer (*Odocoileus virginianus*) being often identified as the main and primary reproductive host (Lane et al. 1991). Tick infestations on white-tailed deer can be intense, with up to 500 individuals reportedly found on a single individual. High abundances of adult ticks in a region have often been associated with large deer populations (Ostfeld 2011).

Lyme disease pathogenesis, diagnosis, and treatment

Pathogenesis

Pathogenesis is the series of cellular events and pathologic mechanisms that result in the development of diseases in an organism. Types of pathogenesis can include microbial infections (invasion and multiplication of microbial organisms into host tissues), inflammation (protective response of immune system), disruption of cell function, and tissue breakdown (cellular damage and premature cell death). Most diseases are caused by a combination of these processes.

Pathogens usually employ a variety of techniques to cause pathology including rapid replication, physical damage to host cellular tissues, toxin production, and overstimulation of host immune responses. If pathogens successfully evade host immune defenses, they can prolong their presence within the host and increase the amount of damage inflicted.

The exact mechanism by which *Borrelia burgdorferi* sensu lato causes pathology in its human hosts is still a subject of much research attention. Attempts at sequencing the spirochete's complete genome have failed to identify genes that promote cellular biosynthetic reaction and none of the typical virulence factors (lipopolysaccharide, toxins, and secretion systems) have been found in either the bacteria's single chromosome, or the many associated plasmids (Tilly et al. 2008, Marques 2010). Instead, evidence is increasingly suggesting that Lyme disease is likely the result of human immune-pathological responses (McKisic et al. 2000, Wooten and Weis 2001, Tilly et al. 2008, Marques 2010, Norris et al. 2010). Therefore, it is the pathogen's prolonged ability to evade host immune responses, rather than classical virulence factors that is the mechanism by which the pathogen persists and causes disease in a susceptible human host (Tilly et al. 2008). *Borrelia* spirochetes have developed multiple genetic adaptations that increase their chances of successful infectivity and persistence within their mammalian hosts (Tsao 2008, Norris et al. 2010). These adaptations generally focus on changes in gene expression that promote successful transmission from tick vectors to host and evasion of host immune responses (Norris et al. 2010).

One method *Borrelia* spirochetes employ to evade host immune response during the transmission stage is to attach to tick salivary products that inhibit host immune responses. As *Borrelia* spirochetes are transmitted from their tick vectors to vertebrate hosts, the bacteria upregulates the production of the outer surface protein C (OspC) (Marques 2010, Norris et al.

2010). OspC has the ability to attach to tick salivary protein Salp15 which is itself known to inhibit CD4+ T cell activation in vertebrate hosts, a process that facilitates prolonged tick feeding periods (Tsao 2009). At initial disseminate stages in host dermal tissues, tick feeding activity and the presence of an invasive microbe will activate the alternative complement pathway of the host innate immune response, which aids in the opsonization and eradication of pathogens by macrophages (Tsao 2009). The alternative complement pathway is activated when the host C3b protein binds to the microbe cellular surface (Tsao 2009). However, C3b can attack both host and bacterial cells indiscriminately in the absence of regulator proteins (factor H or H-like proteins) produced by infected hosts (Tsao 2009). *Borrelia* pathogens take advantage of these host regulatory factors by producing complement regulator-acquiring surface proteins such as BbCRASP-1, BbCRASP-2, BbCRASP-3, OspE, and other Erp proteins that allows them to bind to regulatory proteins meant for host cells and thus permitting them to evade recognition and eradication by complement (Tsao 2009).

Upon successful infection and early evasion of host immune systems, *Borrelia* spirochetes enter host bloodstream (spirochetemia) where they may remain between 10 – 14 days as they disseminate to various host tissues (Tsao 2009). At this stage of infection, the spirochetes begin to down regulate OspC production and introduce new antigenic agents to cope with host adaptive immune responses. One of the mechanisms involves antigenic variation of an outer surface lipoprotein (vlsE) which comprises a huge repertoire of variants (Cabello et al. 2007, Tsao 2009, Marques 2010). In the late invasion stages, the spirochetes exit the bloodstream and invade host tissues with the aid of motility and chemotaxis proteins (Norris et al. 2010). At this stage, they continue to evade host immune systems by hiding in protective niches in host internal organs, central nervous system, synovial fluid, and collagenous tissue in skin, joints and tendons

(Wooten and Weis 2001, Cadavid et al. 2000, Tsao 2009). Specifically, *Borrelia* spirochetes produce proteins like DbpA, DpbB, BmpA and BmpB that facilitate attachment to host extracellular matrix (ECM) elements (e.g. proteoglycan decorin, glycosaminoglycans, and fibronectin) (Cabello et al. 2007, Tsao 2009). ECM components may inadvertently interfere with host immune responses by denying complements, antibodies and lymphocytes access to the pathogen (Cabello et al. 2007). Other ways *Borrelia* spirochetes evade host immune response in late dissemination stages includes hiding within host cells (Cadavid et al. 2000, Tsao 2009) or binding to invaginations of cell membranes, allowing them to avoid phagocytosis and other antibodies or antibiotics (Tsao 2009).

Diagnosis

Diagnosis of Lyme disease is dependent on compatible clinical symptoms, reasonable probability of exposure to ticks in an endemic Lyme region, and results from laboratory screens (Spach et al. 1993). There are several laboratory tests for *B. bugdorferi* infections and each approach has its own advantages and disadvantages (Spach et al. 1993). They include culturing the pathogen from patient skin samples, serological tests, and detecting pathogen genetic markers.

A test for *Borrelia burgdorferi* infection is to directly culture the pathogen from the patients' erythema migrans lesions (Spach et al. 1993, Marques 2010). However, culturing *B. burgdorferi* successfully requires a specially enriched bacteriologic medium that is not easily available, and results will usually not be available for weeks due to slow bacterial growth rates (Marques 2010). While culturing *B. burgdorferi* has a success rate of up to 90%, the test's lack of sensitivity for any samples other than that sampled from the defining symptom of Lyme disease (erythema migrans) renders the test redundant (Marques 2010). An alternate and most

commonly employed test for Lyme disease detects antibody production via serology. This method usually includes a two-tier approach, first with an enzyme-linked immunosorbant assay (ELISA), followed by a confirmation test via Western immune-blot that specifically targets the antibodies IgG and IgM (Spach et al. 1993, Marques 2010). This is because ELISA is prone to producing false-positives due to cross reactive antigens like flagellar and heat-shock proteins (Marques 2010). Unfortunately, serological tests for *Borrelia burgdorferi* infections are full of limitations. For one, the tests do not distinguish between active and inactive infection as patients can remain sero-positive for years post-infection (Marques 2010). Furthermore, serological tests operate on a time constraint as antibodies for *Borrelia burgdorferi* in patients are generally detectable only within 6 weeks of infection date, or after the disease itself has reached late dissemination stage (Spach et al. 1993). Additionally, antimicrobial treatments can potentially abort detectable antibody responses, generating false-negative results (Spach et al. 1993). With recent advancements in molecular techniques, laboratories have begun to screen for *Borrelia burgdorferi* presence by detecting genetic markers unique to the pathogen using polymerase chain reactions (PCR) (Marques 2010). Yet the sensitivity of this approach is heavily dependent on what type of Lyme disease the patients are suffering from and the source of DNA template. For example, PCR has an 85% detection rate for patient suffering from Lyme arthritis, but only if DNA amplification was performed on synovial fluid (Marques 2010). Similarly, PCR performed on cerebrospinal fluid (CSF) has a less than 40% detection rate for patients with neuroborreliosis (Marques 2010).

Treatment

If detected and identified early enough, Lyme disease can be successfully treated with antimicrobial agents (Spach et al. 1993). Delayed or inadequate treatments may lead to severe

chronic pathology that can be debilitating, however the existence of chronic Lyme disease and the acceptable treatment approach remains largely controversial (see Feder et al. 2007, Lantos 2015). Some have suggested that the persistence of inflammation with or without residual *Borrelia* infection may explain why some patients with Lyme borreliosis remain symptomatic despite treatment with adequate doses of antibiotics (Cadavid et al. 2000). Most Lyme disease patients with early localized infection usually respond well to 2-3 weeks of oral prescriptions like doxycycline, amoxicillin, cefuroxime, or azithromycin (Spach et al. 1993). For late Lyme infection stages, doctors generally prescribe intravenous drugs like ceftriaxone or penicillin for periods between 2 to 3 weeks (Spach et al. 1993). If patients exhibit no symptoms but report having been recently bitten by a tick in an endemic Lyme region, doctors may prescribe prophylactic antimicrobial agents but this practice is contentious (Spach et al. 1993).

LYME DISEASE EMERGENCE IN CANADA AND SOUTHERN QUÉBEC

Lyme disease has been garnering increasing amounts of interest in Canada as we witnessed a substantial rise in reported human cases across the country, with numbers increasing over ten-fold between 2005 and 2013 (Government of Canada 2015, Kulkarni et al. 2015). In southern Québec specifically, 6 cases of human Lyme disease were reported in 2004, while the first endemic (or locally acquired) case was reported in 2008. By 2014, this number had risen to over 120 (Institut National de Santé Publique du Québec 2014) (Figure 1.2). Due to its emerging status across the country, Lyme disease was declared a reportable disease by the Public Health Agency of Canada in 2009 (Ogden et al. 2009). However, recent studies have suggested that Lyme disease cases may be under-reported at rates from 40% (Henry et al. 2011) to 90% (Centers for Disease Control 2013), likely due to misdiagnosis resulting from the multi-systemic

nature of the disease, as well as lack of public and physician awareness. The risk of contracting Lyme disease in Canada is considered particularly high in regions where established populations of the vectors (black-legged ticks [*I. scapularis*] and western black-legged ticks [*I. pacificus*]) are present (Ogden et al. 2009). Currently, known and suspected endemic hotspots of Lyme disease extend from the province of British Columbia on the Pacific Coast and westward towards the Atlantic provinces of New Brunswick and Nova Scotia (Government of Canada 2015). At present, there is low risk of encountering infected ticks across most of Canada, although the number of risk areas is on the rise, particularly in eastern Canada (Ogden et al. 2009).

Unanimously, researchers and health officials agree that the emergence of Lyme disease in Canada resulted primarily from the range expansion of both the main disease vector and competent disease reservoir hosts into the region (Heon et al. 2009, Ostfeld and Brunner 2015). Currently, there are two driving forces hypothesized for Lyme disease emergence in Canada: [1] climate change and associated changes in species distribution (Legér et al. 2013, Simon et al. 2014), and [2] the effects of human land use on biodiversity composition in the region (Margos et al 2011, Simon et al. 2014). Studies have attempted to predict and project current and future disease risk across the country based on these predictions (e.g. Simon et al. 2014). Unfortunately, Lyme disease is a complex system made up of several layers of interactions within and among abiotic and biotic variables which makes it difficult to study (Ostfeld 2011).

Role of climate change on Lyme disease emergence in southern Québec

Climate change has often been cited as the predominant cause of infectious disease emergence as it may influence multiple components of any zoonotic disease transmission system (Ostfeld and Brunner 2015). For instance, ambient temperature influences multiple vector, host and pathogen biological traits such as survival rates, phenology, fecundity, behaviour, and

vector/pathogen competency. Similarly, the combined effects of humidity and temperature can determine habitat suitability, microclimate conditions, and community species composition. However, it remains difficult to understand and predict the exact effects of climate conditions on disease emergence processes via the alteration of host, vector, and pathogen demographic processes and distributions.

In the context of the Lyme disease system in Canada, the role of climate change on the range expansion, establishment, and maintenance of endemic tick populations is undeniable (Legér et al. 2013). As exothermic arthropods, almost all aspects of tick biology and population establishment success are dependent on their surrounding environmental conditions. Specifically, warmer summer temperatures and steady winter temperatures are likely to promote tick developmental rates, survival, and fitness (Ogden et al. 2008, Wu et al. 2013). Similarly, greater relative humidity promotes greater activity and host seeking behaviour (Vail and Smith 2002). This in turn increases encounter rates between ticks and their hosts, promoting the transmission of the disease pathogen to vertebrate hosts (both humans and animals) and facilitating *B. burgdorferi* persistence in a community. Finally, since invertebrate immune responses are temperature-sensitive, changes in environmental temperatures could potentially influence tick immune responses which will in turn influence the vector's disease transmission efficacy (Ostfeld and Brunner 2015).

Unlike the tick vectors, hosts in the Lyme disease system are endothermic and thus the immediate effects of climate change will be buffered by physiological regulations. Nonetheless, climatic conditions can determine habitat suitability which in turn influences the distribution and demography of host species that are essential for the persistence of stable pathogen and vector populations (Legér et al. 2013, Ostfeld and Brunner 2015). For example, the distribution of

established populations of the white-footed mouse, the primary Lyme disease reservoir host in Eastern Canada, is influenced by winter length (Roy-Dufresne et al. 2013). Likewise, ground-foraging bird migratory range and timing, that are hypothesized to be the main mechanism of long-distance tick invasion into Canada, are significantly influenced by climate change (Ogden et al. 2008, Hulbert and Liang 2012). Finally, changes in the local species community composition in response to climate change will also influence disease emergence via dilution and/or amplification effects, though the specific mechanisms underlying the effect of biodiversity on disease risk remains context-specific (Miller and Huppert 2013).

Climate change also influences the rate at which the Lyme disease agent *B. burgdorferi* becomes established in Canada, simultaneously exerting evolutionary pressures that could determine pathogen virulence in human. We know that climate influences tick phenology, and studies have shown that this will in turn determine the evolution of *B. burgdorferi* genotypic strain types (Ostfeld and Brunner 2015). Specifically, it has been hypothesized that warmer climatic conditions that promote asynchronous tick life cycles are most likely to select for persistent *B. burgdorferi* strain types that are suspected to be particularly virulent/pathogenic in humans (Hamer et al. 2012). It has been postulated that it is this indirect selective pressure exerted by varying regional climate conditions that produced the differences in reported Lyme disease incidences and transmission risk across North America (Gatewood et al. 2009).

Role of land use change on Lyme disease emergence in southern Québec

In recent decades, with human population growth, resource use, and technological advances, anthropogenic activities have altered our surrounding landscape to a considerable extent and at a great rate. Forests have been cleared for logging, converted into agricultural land for food production, and to make way for urban and residential developments. Rivers have been

diverted for irrigation, generating electricity, and to create steady supplies of drinking water for consumption. These human-mediated, large-scaled changes on the landscape have significant effects on the local biodiversity, which has in turn facilitated disease emergence (Jousimo et al. 2014, McFarlane et al. 2013). Studies on the Lyme disease system in North America have shown that disease transmission risk in a region is correlated with landscape heterogeneity (Brownstein et al. 2005), habitat conditions (Ostfeld et al. 2006), and local biodiversity and species composition (Ostfeld and Keesing 2012). For instance, as habitat fragmentation decreases forest patch size, changes in local vertebrate species composition and abundances can alter disease risk (Allan et al. 2003, Brownstein et al. 2005). Since host species can exhibit varying levels of competency for transmitting or maintaining disease vector and pathogen populations (Ostfeld 2011), the composition and relative abundances of available competent and susceptible host species in a habitat patch will determine encounter rates among species that play important roles in a disease transmission cycle.

In southern Québec, forest patches have been increasing in number and decreasing in size (Bélanger and Grenier 2002, Ménard and Marceau 2007) as trees are cleared for agricultural purposes and urban development. This habitat fragmentation produces an environmental condition that is optimal for the colonisation and establishment of efficient Lyme disease reservoir hosts like the white-footed mouse and the eastern chipmunks (Brownstein et al. 2005). As the number of established local host populations and abundance increase, so does the probability of *Borrelia burgdorferi* pathogen gaining footholds within the region. We have also seen in recent years an increase in the white-tailed deer populations in southern Québec. White-tailed deer generally prefer edge habitats that have increased substantially in the past few decades due to the combined effects of habitat fragmentation and reforestation efforts, allowing

their abundance to reach levels previously unseen in the region (Legér et al. 2013). While white-tailed deer are not considered competent reservoir hosts for the Lyme disease pathogen, they are nonetheless important reproductive hosts for the disease vector and may also play a role in local vector dispersal (e.g. Pound et al. 2010, McShea 2012). We know that gene flow and dispersal of white-tailed deer and white-footed mouse can be significantly influenced by physical barriers like rivers and roads (Rogic et al. 2013, Marrotte et al. 2014, Robinson et al. 2012, Locher et al 2014). Therefore, land use changes that influence host abundance and dispersal behaviour may indirectly influence spatial patterns of disease emergence.

APPLYING INTEGRATED MOLECULAR APPROACH IN STUDIES ON DISEASE EMERGENCE DYNAMICS

Most studies on zoonotic disease ecology and emergence patterns tend to focus on a single representative species (the pathogen or the vector) and individual predictor variables (climate vs landscape). From there, attempts are made to predict current and future risk in order to inform disease management and control programs. However, disease emergence, transmission, and persistence dynamics can be simultaneously affected by multiple factors (Tabachnick 2010). As discussed above, changes in the climate and landscape can influence disease risk considerably. Furthermore, many zoonotic disease systems involve interactions among multiple species, each of which play independently crucial roles in the transmission and persistence of disease in a region (Alexander et al. 2012, Estrada-Peña et al. 2014). These species are also likely to react differently to changes in their environment which can in turn alter the facet of their individual role in disease transmission (Manel and Holderegger 2013, Johnson et al. 2015). To further

complicate disease models, climate, landscape, and biotic variables all operate at varying spatial and temporal scales.

It has become evident in recent years that zoonotic disease systems are more complex than expected, and that determining mechanisms of disease emergence based on constrained studies or broad definitions can result in erroneous conclusions regarding the dynamics of the disease system (Estrada-Peña et al. 2014, Rosenthal et al. 2015). Effective monitoring and control of zoonotic diseases therefore requires a thorough understanding on how host-vector-pathogen interactions define the role each species play in disease transmission, as well as how species responses to changing environmental conditions determine disease evolution, transmission, emergence, establishment, and persistence (Rosenthal et al. 2015). Comprehensive studies that integrate all relevant species and variables that influence species biology, behaviour, and demography will provide clearer insights that can inform disease control efforts (Johnson et al. 2015).

With recent advances in molecular techniques, it is now possible to conduct large-scale studies on zoonotic disease dynamics at the molecular level, allowing researchers to better study important evolutionary and ecological processes at historical and contemporary time scales (Gubler 2010, Biek and Real 2010). Molecular approaches have been employed in several studies on disease systems (Muellner et al. 2011). These studies ranged from identifying sources of a disease agent (e.g. Furuse et al. 2010, Sharp and Hahn 2011), tracing historical and contemporary spatial patterns of disease spread across a landscape (e.g. Lemey et al. 2010, Paquette et al. 2014, Biek et al. 2007), and uncovering pathogen evolutionary history and population dynamics (e.g. Margos et al. 2008, Black et al. 2009). Molecular approaches have also allowed for studies on evolutionary processes that influence disease virulence, as well as

interspecific interactions such as host-dependent dispersal that may affect disease emergence rate and spatial pattern (e.g. Biek et al. 2006).

The field of population genetics or genomics can allow us to investigate variables that can promote and facilitate disease emergence and rapid spread (Engering et al. 2013, Manel and Holderegger 2013, Estrada-Peña et al. 2014, Rosenthal et al. 2015). For example, population genetic analyses with appropriate genetic markers can reveal population evolutionary dynamics (McDonald and Linde 2002, King and Lively 2012, Rosenthal et al. 2015), likelihoods of successful population establishments (Lee 2002, Barrett and Schluter 2008, le Rouzic and Carlborg 2008, Hill et al. 2011, Rosenthal et al. 2015), and mechanisms underlying patterns of disease spread (Bialozyt et al. 2006, Excoffier et al 2009, Cristescu 2015). Application of landscape genetics or genomics, which detects genetic discontinuities and correlates such patterns with landscape features, can further reveal how geographical and environmental features structure genetic variation at both population and individual levels (e.g. Blanchong et al. 2008). For vector-transmitted diseases, landscape features that influence vector dispersal are the foremost determinants of rates of epizootic spread (McCallum 2008). Therefore, the ability to identify and estimate current and future rates and spatial patterns of disease emergence relative to landscape effects on critical disease vector or host species will be a significant contribution to disease monitoring efforts.

Using a combination of population and landscape genetics, we can begin to [1] uncover mechanisms underlying the range expansion of the disease vectors, pathogens, and host species, [2] identify established species populations that may translate to regions of greater disease ecological risk, and [3] categorize landscape features influencing the contemporary gene flow and movement of these critical species. These insights will be useful for extrapolating future

species range expansion and disease emergence patterns. For the Lyme disease system in southern Québec, spatio-temporal examination of population genetic patterns and turnover rate could help us identify the establishment stages of black-legged tick and white-footed mouse populations in southern Québec (Engering et al. 2013, Estrada-Peña et al. 2014). Meanwhile, landscape genetics analyses can help us understand how landscape features influence both current and future white-footed mouse and black-legged tick dispersal patterns across southern Québec. Finally, we can examine and quantify the role of host-dependent dispersal on black-legged tick movement across southern Québec. It is well established that ticks are poor dispersers (Falco and Fish 1991) and that any significant movement across a landscape will be dependent on the host (Madhav et al. 2004). Using population and landscape genetics, we could identify critical players of disease spread in the region (Ogden et al. 2015, Khatchikian et al. 2015) to help public health officials design efficient, targeted control approaches (Artois et al. 2011).

SCOPE OF THESIS

Traditionally, most studies on disease systems focus on a single species or variable. In this thesis, I focus on a multi-taxa approach using Lyme disease emergence in Québec as a case study. I employed an integrated multi-taxa approach by investigating the population and landscape genetics of two major players in the eastern North America Lyme disease system: the vector (the black-legged tick) and the reservoir host (white-footed mouse). There are several reasons why I chose to focus on these two species. First of all, as previously mentioned, both the ticks and mice are important players in the Lyme disease system and their presence is necessary for successful local establishment of the disease system. Secondly, both species are currently

undergoing range expansion into Canada which may have likely contributed to the rapid emergence of Lyme disease in the region. Finally, the current species range of ticks and mice in southern Québec are still somewhat limited, thus making it feasible to sample across their current known range in the region.

There are in total, five objectives to this thesis which are divided into four data chapters.

The first objective of this thesis is to present the conceptual framework and workflow of the multi-taxa approach I am using to studying the Lyme disease system (Chapter 2). In this chapter, I discussed and explored how multi-taxa integrated landscape genetics (MTILG), an approach that integrates concepts and methods from population genetics, landscape ecology, and spatial statistics, can reveal how interspecific interactions and landscape variables influence disease transmission system spread. I provided an example of the power of the power of MTILG by modeling the emergence of a hypothetical disease system. Simulated microsatellite data from these models were analysed via the MTILG approach to the different stages of analysis and the conclusions about disease spread that can be drawn from it. I was able to identify how landscape heterogeneity and geographical barriers indirectly influence disease emergence patterns by limiting species movement across a landscape. It also allowed me to conclude that the extent and rates of disease spread is dependent on dispersal capability of the host species responsible for dispersing vectors and pathogens across a landscape.

The second, third and fourth objectives involve field surveys and analyses on empirical data. I examined the spatial genetic diversity of both the white-footed mouse (Chapter 3) and black-legged tick (Chapter 4) across southern Québec. Both species were sampled across their current known overlapping ranges in the region and genotyped for sets of species-specific microsatellite loci. For the white-footed mouse, I further examined *B. burgdorferi* prevalence

data and compared it against mouse genetics and morphological traits to investigate possible associations between species traits and disease infection prevalence. My analyses revealed that white-footed mouse populations in Québec exhibit significant genetic differentiation, and that their gene flow is affected by linear barriers to migration. At the same time, I found evidence of morphological gradients (e.g. allometrically increasing hind-foot length) in the direction of species range expansion which has interesting implications for future rates of mouse range expansion. Finally, a comparison of mouse genetics, morphology, and *B. burgdorferi* infection prevalence revealed that while the white-footed mouse species traits do not appear to influence the species' susceptibility to *Borrelia* infection, white-footed mice play a critical role in Lyme disease emergence as hosts for both ticks and pathogen.

When I examined black-legged tick genetic variation in my study area (Chapter 4), I found evidence supporting the hypothesis that frequent long-distance dispersal events were the primary mechanism of tick invasion into southern Québec. Furthermore, such long distance events coupled with high rates of local extinction has produced rapid genetic turnover in tick populations over both space and time. This allowed me to accurately identify established tick populations, which in turn provides insights on hotspots of disease risk in the region.

Finally in the Chapter 5, I achieved the final objective of the thesis and integrated genetic information, field observations, climate-based species distribution information, and landscape effects on species abundances to create a model that projects Lyme disease infection risk in southern Québec. I first quantified the relative importance of white-footed mouse and black-legged tick on *B. burgdorferi* prevalence and subsequently examined the level of interactions and co-dependency between both ticks and mice with regards to movement across the landscape. I then determined the effect of various climatic conditions and landscape variables on both white-

footed mouse and black-legged tick distribution and abundance across the landscape using a combination of climate-niche modeling and redundancy analyses. From there, I identified and quantified relations among pathogen prevalence relative to the combined effects of climate and land-use changes on the species distribution of the disease vector and reservoir host using structural equation model (SEM) analysis. Armed with estimated parameter coefficients from my SEM analysis, I was able to project current *B. burgdorferi* prevalence across the study area. My results revealed that Lyme disease infection risk was greatest at the southwestern region of my study area, along the US-Canada border, and that infection risk declined in a north-easterly direction.

In summary, my thesis examined Lyme disease emergence patterns and rates in southern Québec, with particular emphasis on the disease vector (black-legged tick) and reservoir host (white-footed mouse). By examining each species independently, I was able to identify mechanisms of species range expansion as well as uncover species adaptations/traits/roles that can potentially facilitate disease emergence rates. When I examined both species simultaneously using the MTILG approach, I was able to identify how interspecific interactions, climate conditions, and surrounding landscape contribute to disease prevalence in a region. From there, I was able to project disease infection risk in my study area. The studies in this thesis exemplify the importance of applying an integrated, multi-taxa, molecular approach when studying the dynamics of an emerging disease system (e.g. West Nile Virus, Hantavirus, or other tick-borne diseases like Babesiosis). Doing so can produce valuable insights and tools that help us better understand disease emergence processes from an ecological perspective. Such information will in turn, be crucial in disease management and intervention policies.

REFERENCES

Aguero-Rosenfeld ME, Wang G, Schwartz I, Wormser GP. 2005. Diagnosis of Lyme Borreliosis. *Clinical Microbiology Reviews* **18**: 484 – 509.

Alexander KA, Lewis BL, Marathe M, Eubank S, Blackburn JK. 2012. Modeling of wildlife-associated zoonoses: Applications and Caveats. *Vector-borne and Zoonotic Diseases* **12**: 1005 – 1018.

Allan SA. 2001. Ticks (Class Arachnida: Order Acarina). *Parasitic Diseases of Wild Mammals, Second Edition*, pp. 72-106.

Allan BF, Keesing F, Ostfeld RS. 2003. Effect of forest fragmentation on Lyme disease risk. *Conservation Biology* **17**: 267 – 272.

Anderson JF. 1989. Epizootiology of *Borrelia* in *Ixodes* tick vectors and reservoir hosts. *Reviews of Infectious Diseases* **11**: S1451 – S1459.

Artois A, Blancou J, Dupeyroux O, Gilot-Fromont E. 2011. Sustainable control of zoonotic pathogens in wildlife: How to be fair to wild animals? *Scientific and Technical Review of the Office International des Epizooties* **30**: 733 – 743

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends in Ecology and Evolution* **23**: 38 – 44

Benach JL, García-Moncó JC. 2010. The worldwide saga of Lyme borreliosis. In *Borrelia: molecular biology, host interaction and pathogenesis* (eds. Samuels DS, Radolf JD). Caister Academic Press, Hethersett, Norwich, United Kingdom. p. 7 – 26.

Bélanger L, Grenier M. 2002. Agriculture intensification and forest fragmentation in the St. Lawrence Valley, Québec, Canada. *Landscape Ecology* **17**: 495 – 507.

Bialozyt R, Ziegenhagen B, Petit RJ. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**: 12 – 20

Biek R, Drummond AJ, Poss M. 2006. A virus reveals population structure and recent demographic history of its carnivore host. *Science* **311**: 538 - 541.

Biek R, Henderson JC, Waller LA, Rupprecht CE, Real LA. 2007. A high- resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proceedings of the National Academy of Science USA* **104**: 7993 - 7998.

Biek R, Real LA. 2010. The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* **19**: 3515 – 3531.

Black WC, Doty JB, Highes MT, Beaty BJ, Calisher CH. 2009. Temporal and geographic evidence for eolution of Sin Nombre virus using molecular analyses of viral RNA from Colorado, New Mexico, and Montana. *Virology Journal* **6**: 102

Blanchong JA, Samuel MD, Scribner KT, Weckworth BV, Langenberg JA, Filcek KB. 2008. Landscape genetics and the spatial distribution of chronic wasting disease. *Biology Letters* **4**: 130 – 133.

Brisson D, Drecktrah D, Eggers CH, Samuels DH. 2012. Genetics of *Borrelia burgdorferi.* *Annual Review of Genetics* **46**: 515 – 536.

Brownstein JS, Skelly DK, Holford TR, Fish D. 2005. Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. *Oecologia* **146**: 469 – 475.

Brunner JL, LoGiudice K, Ostfeld RS. 2008. Estimating reservoir competence of *Borrelia burgdorferi* hosts: Prevalence and infectivity, sensitivity, and specificity. *Journal of Medical Entomology* **45**: 139 – 147.

Cabello FC, Godfrey HP, Newman SA. 2007. Hidden in plain sight: *Borrelia burgdorferi* and the extracellular matrix. *Trends in Microbiology* **15**: 350 – 354.

Cadavid D, O'Neill T, Schaefer H, PAchner AR. 2000. Localization of *Borrelia burgdorferi* in the nervous system and other organs in a non-human primate model of Lyme disease. *Laboratory Investigation* **80**: 1043 – 1054.

Casjens SR, Eggers CH, Schwartz I. 2010. Borrelia genomics: chromosome, plasmids, bacteriophages and genetic variation. *In Borrelia: molecular biology, host interaction and pathogenesis* (eds. Samuels DS, Radolf JD). Caister Academic Press, Hethersett, Norwich, United Kingdom. pp. 27–53.

Centers for Disease Control. 2013. CDC provides estimate of Americans diagnosed with Lyme disease each year. http://www.cdc.gov/media/releases/2013/p0819-lyme-disease.html Assessed July 2015

Comstedt P, Bergström S, Olsen B, Garpmo U, Marjavaara L, Mejlon H, Barbour AG, Bunikis J. 1992. Migratory passerine birds as reservoirs of Lyme Borreliosis in Europe.*Emerging Infectious Diseases* **12**: 1087 – 1095.

Cook MJ. 2015. Lyme borreliosis: A review of data on transmission time after tick attachment. *International Journal of General Medicine* **8**: 1- 8.

Cristescu ME. 2015. Genetic reconstruction of invasion history. *Molecular Ecology* **24**: 2212 – 2225

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious disease of wildlife – threats to biodiversity and human health. *Science* **287**: 443 – 449.

Eisen L, Lane RS. 2002. Vectors of *Borrelia burgdorferi* sensu lato. *Lyme Borreliosis: Biology, Epidemiology and Control. Wallingford, Oxon, UK: CAB International*, 1-28.

Engering A, Hogerwerf L, Slingenbergh J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections* **2**: e5.

Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease risk: An ecological primer. *Trends in Parasitology* **30**: 205 – 214

Excoffier L, Foll M, etit RJ. 2009. Genetic consequences of range expansion. *Annual Review of Ecology, Evolution, and Systematics* **40**: 481 – 501.

Falco RC, Fish D. 1991. Horizontal movement of adult *Ixodes dammini* (Acari: Ixodidae) attracted to CO_2-baited traps. *Journal of Medical Entomology* **28**: 726 – 729

Feder HM, Johnson BJB, O'Connell S, Shapiro ED, Steere AC, Wormser GP, Ad Hoc International Lyme Disease Group. 2007. A critical appraisal of "Chronic Lyme disease". *The New England Journal of Medicine* **357**: 1422 – 1430.

Furuse Y, Suzuki A, Oshitani H. 2010. Origin of measles virus: Divergence from rinderpest virus between the 11^{th} and 12^{th} centuries. *Virology Journal* **7**: 52.

Gatewood AG, Lienman KA, Vourc'h G, Bunikis J, Hamer SA, Cortinas R, Melton F, Cislo P, Kitron U, Tsao J, Barboue AG, Fish D, Diuk-Wasser MA. 2009. Climate and tick seaonality are predictors of *Borrelia burgdorferi* genotype distribution. *Applied and Environmental Microbiology* **75**: 2476 – 2483.

Gern L, Estrada-Peña A, Frandsen F, Gray JS, Jeanson TGT, Jongejan F, Kahl O, Korenberg E, Mehl R, Nuttall PA. 1998. European reservoir hosts of *Borrelia burgdorferi* sensu lato. *Zentralblatt für Bakteriologie* **287**: 196 – 204.

Government of Canada. 2015. Surveillance of Lyme disease. http://www.healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/lyme/surveillance-eng.php. Assessed July 2015

Gubler DJ. 2010. The global threat of emergent/reemergent vector-borne diseases. In: *Vector Biology, Ecology and Control* (ed. Atkinson PW), pp. 39 – 62. Springer, Netherlands

Hamer SA, Hickling GJ, Sidge JL, Walker ED, Tsao JI. 2012. Synchronous phenology of juvenile *Ixodes scapularis*, vetertebrate host relationships, and associated patterns of *Borrelia burgdorferi* ribotypes in the mid-western United States. *Ticks and Tick-borne Diseases* **3**: 65 – 74.

Henry B, Roth D, Reilly R, MacDouugall L, Mak S, Li M, Muhamad M. 2011. How big is the Lyme problem? Using novel methods to estimate the true number of Lyme disease cases in British Columbia residents from 1998 to 2008. *Vector Borne Zoonotic Disease* **11**: 863 – 868.

Herrmann C, Gern L. 2013. Survival of *Ixodes ricinus* (Acari: Ixodidae) nymphs under cold conditions is negatively influenced by frequent temperature variations. *Ticks and Tick-borne Diseases* **4**: 445 – 451.

Hill JK, Griffiths, Thomas CD. 2011. Climate change and evolutionary adaptations at species' range margins. *Annual Review of Entomology* **56**: 143 – 159

Hoen AG, Margos G, Bent SJ, Diuk-Wasser MA, Barbour, Kurtenbach K, Fish D. 2009. Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. *Proceedings of the National Academy of Sciences USA* **106**: 15013 – 15018.

Hurlbert AH, Liang Z. 2012. Spatiotemporal variation in avian migration phenology: Citizen science reveals effects of climate change. *PLoS ONE* **7**: e31662.

Humphrey PT, Caporale DA, Brisson D. 2010. Uncoordinate phylogeography of *Borrelia burgdorferi* and its tick vector, *Ixodes scapularis*. *Evolution* **64**: 2653 – 2663.

Institut National de Santé Publique du Québec . 2014. Proposition d'un programme de surveillance intégré pour la maladie de Lyme et les autres maladies transmises par la tique *Ixodes scapularis* au Québec

Jaenson TG, Tälleklint L. 1992. Incompetence of roe deer as reservoirs of the Lyme borreliosis spirochete. *Journal of Medical Entomology* **29**: 813 – 817.

Johnson PTJ, de Roode JC, Fenton A. 2015. Why infectious disease research needs community ecology. *Science* **349**: 1259504-1 – 1259504-9.

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. *Nature* **451**: 990 – 993.

Jousimo J, Tack AJM, Ovaskainen O, Mononen T, Susi H, Tollenaere C, Laine A-L. 2014. Ecological and evolutional effects of fragmentation on infectious disease dynamics. *Science* **344**: 1289 – 1293.

Keesing F, Belden LK, Daszak P, Dobson A, Harvell CD, Holt RD, Hudson P, Jolles A, Jones KE, Mitchell CE, Myers SS, Bogich T, Ostfeld RS. 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**: 647 – 652.

Khatchikian CE, Prusinski MA, Stone M, Backenson PB, Wang I-N, Foley E, Seifert SN, Levy MZ, Brisson D. 2015. Recent and rapid population growth and range expansion of the Lyme disease tick vector, *Ixodes scapularis*, in North America. *Evolution* **69**:1678–1689.

Killilea ME, Swei A, Lane RS, Briggs CJ, Ostfeld RS. 2008. Spatial dynamics of Lyme disease: A review. *EcoHealth* **5**: 167 – 195.

King KC, Lively CM. 2012. Does genetic diversity limit disease spread in natural host populations? *Heredity* **109**: 199 – 203

Kulkarni MA, Berrang-Ford L, Buck PA, Drebot MA, Lindsay LR, Ogden NH. 2015. Major emerging vector-borne zoonotic diseases of public health importance in Canada. *Emerging MMicrobes and Infections* **4:** e33.

Lafferty KD, Wood CL. 2013. It's a myth that protection against disease is a strong and general service of biodiversity conservation: Response to Ostfeld and Keesing. *Trends in Ecology Evolution* **28**: 503 - 504.

Lane RS, Piesman J, Burgdorf W. 1991. Lyme borreliosis: Relation of its causative agent to its vectors and hosts in North America and Europe. *Annual Review of Entomology* **36**: 587 – 609.

Lantos PM. 2015. Chronic Lyme disease. *Infectious Disease Clinics of North America* **29**: 325 – 340.

le Rouzic A, Carlborg Ö. 2008. Evolutionary potential of hidden genetic variation. *Trends in Ecology and Evolution* **23**: 33 – 37.

Lee CE. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* **17**: 386 – 391

Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution* **27**:1877–1885.

Legér E, Vourc'h G, Vial L, Chevillon C, McCoy KD. 2013. Changing distributions of ticks: causes and consequences. *Experimental and Applied Acarology* **59**: 219 – 244.

Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S, Wang L-F. 2005. Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**: 676 – 679.

Li Y, Carroll DS, Gardner SN, Walsh MC, Vitalis EA, Damon IK. 2007. On the origin of smallpox: Correlating variola phylogenics with historical smallpox records. *Proceedings of the National Academy of Sciences* **104**: 15787 – 15792.

Lindsay LR, Barker IK, Surgeoner GA, McEwen SA, Gillespie TJ, Addison EM. 1998. Survival and development of the different life stages of *Ixodes scapularis* (Acari: Ixodidae) held within four habitats on Long Point, Ontario, Canada. *Journal of Medical Entomology* **35**: 189 – 199.

Lindsay LR, Mathison SW, Barker IK, McEwen SA, Surgeoner GA. 1999a. Abundance of *Ixodes scapularis* (Acari: Ixodidae) larvae and nymphs in relation to host density and habitat on Long Point, Ontario. *Journal of Medical Entomology* **36**: 243 – 254.

Lindsay LR, Mathison SW, Barker IK, McEwen SA, Gillespie TJ, Surgeoner GA. 1999b. Microclimate and habitat relation to *Ixodes scapularis* (Acari: Ixodidae) populations on Long Point, Ontario, Canada. *Journal of Medical Entomology* **36**: 255 – 262.

Locher A, Scribner KT, Moore JA, Murphy B, Kanefsky J. 2014. Influence of landscape features on spatil genetic structure of white-tailed deer in human-altered landscapes. *The Journal of Wildlife Management* **79**: 180 – 194.

Madhav NK, Brownstein JS, Tsao JI, Fish D. 2004. A dispersal model for the range expansion of black-legged tick (Acari: Ixodidae). *Journal of Medical Entomology* **41**: 842 – 852

Magnarelli LA, Anderson JF, Fish D. 1987. Transovarial transmission of *Borrelia burgdorferi in Ixodes dammini* (Acari: Ixodidae). *The Journal of Infectious Diseases* **156**: 234 – 236.

Manel S, Holderegger R. 2013. Ten years of landscape genetics. *Trends in Ecology and Evolution* **28**: 614 – 621.

Margos G, Gatewood AG, Ananensen DM, Hanicova K, Terekhova D, Vollmer SA, Cornet M, Piesman J, Donaghy M, Bormane A, Hurn MA, Feil EJ, Fish D, Casjens S, Wormser GP, Schwartz I, Kurtenbach K. 2008. MLST of housekeeping genes captures geographic population structure and suggest a European origin of *Borrelia burgdorferi*. *Proceedings of the National Academy of Science USA* **105**: 8730 – 8735.

Margos G, Tsao JI, Castillo-Ramírez S, Girard YA, Hamer SA, Hoen AG, Lane RS, Raper SL, Ogden NH. 2012. Two boundaries separate *Borrelia burgdorferi* populations in North America. *Applied and Environmental Microbiology* **78**: 6059 – 6067.

Margos G, Vollmer SA, Ogden NH, Fish D. 2011. Population genetics, taxonomy, phylogeny, and evolution of *Borrelia burgdorferi* sensu lato. *Infection, Genetics and Evolution* **11**: 1545 – 1563.

Marrotte RR, Gonzalez A, Millien V. 2014. Landscape resistance and habitat combine to provide an optimal model of genetic structure and connectivity at the range margin of a small mammal. *Molecular Ecology* **23**: 3983 – 3998.

Marques AR. 2010. Lyme disease: A review. *Current Allergy and Asthma Reports* 10: 13 – 20.

Mather TN, Wilson ML, Moore SI, Ribeiro JMC, Spielman A. 1989. Comparing the relative potential of rodents as reservoirs of the Lyme disease spirochete (*Borrelia burgdorferi*). *America Journal of Epidemiology* **130**: 143 – 150.

McCallum H. 2008. Chapter Five: Landscape structure, disturbance, and disease dynamics. In: *Infectious disease ecology: effects of ecosystems on disease and of disease on ecosystems* (eds. Ostfeld RS, Keesing F, Eviner VT), pp. 100 – 122. Princeton University Press, Princton, New Jersey

McCoy KD, Léger E, Dietrich M. 2013. Host specialisation in ticks and transmission of tick-borne diseases: a review. *Frontiers in cellular and Infection Microbiology* **3**: 57

McDonald BA, Linde C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**: 349 – 379

McFarlane RA, Sleigh AC, McMichael AJ. 2013. Land-use change and emerging infectious diseases on an island continent. *International Journal of Environmental Research and Public Health* **10**: 2699 – 2719.

McKisic MD, Redmond WL, Barthold SW. 2000. Cutting edge: T cell mediated pathology in Murine Lyme disease. *The journal of Immunology* **164**: 6096 – 6099.

McShea WJ. 2012. Ecology and management of white-tailed deer in a changing world. *Annals of the New York Academy of Sciences* **1249**: 45 – 46.

Ménard A, Marceau DJ. 2007. Simulating the inpact of forest management scenarios in an agricultural landscape of southern Québec , Canada, using a geographic cellular automata. *Landscape and Urban Planning* **79**: 253 – 265.

Miller E, Huppert A. 2013. The effects of host diversity on vector-borne disease: The conditions under which diversity will amplify or dilute the disease risk. *PLoS ONE* **8**: e80279

Mores SS, Mazet JAK, Woolhouse M, Parrish CR, Carroll D, Karesh WB, Zambrana-Torrelio C, Lipkin WI, Daszak P. 2012. Prediction and prevention of the next pandemic zoonosis. *Lancet* **380**: 956 – 965.

Muellner P, Zadoks RN, Perez AM, Spencer SEF, Schukken YH, French NP. 2011. The integration of molecular tools into veterinary and spatial epidemiology. *Spatial and Spatio-temporal Epidemiology* **2**: 159 – 171.

Myers P, Lundrigan BL, Hoffman SMG, Haraminac AP, Seto SH. 2009. Climate0-induced changes in the small mammal communities of the Northern Great Lakes Region. *Global Change Biology* **15**: 1434 – 1454

Norris SJ, Coburn J, Leong JM, Hu LT, Höök M. 2010. Pathobiology of Lyme disease *Borrelia. In Borrelia: molecular biology, host interaction and pathogenesis* (eds. Samuels DS, Radolf JD). Caister Academic Press, Hethersett, Norwich, United Kingdom. p. 299 – 331.

Ogden NH, Barker IK, Francis CA, Heagy A, Lindsay LR, Hobson KA. 2015. How far north are migrant birds transporting the tick *Ixodes scapularis* in Canada? Insights from stable hydrogen isotope analyses of feathers. *Ticks and Tick-borne Diseases* **6**: 715–720

Ogden NH, Lindsay LR, Hanicová K, Barker IK, Bigras-Poulin M, Charron DF, Heagy A, Francis CM, O'Callaghan CJ, Schwartz I, Thompson RA. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. *Applied and Environmental Microbiology* **74**: 1780 – 1790.

Ogden NH, Lindsay LR, Morshed M, Sockett PN, Artsob H. 2009. The emergence of Lyme disease in Canada. *Canadian Medical Association Journal* **180**: 1221 – 1224.

Ogden NH, St-Onge L, Barker IK, Brazeau S, Bigras-Poulin M, Charron DF, Francis CM, Heagy A, Lindsay LR, Maarouf A, Michel P, Milord F, O'Callaghan CJ, Trudel L, Thompson A. 2008. Risk map for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics* **7**: 24

Ostfeld RS. 2011. Lyme disease: The ecology of a complex system. Oxford University Press.

Ostfeld RS. 2013. A candide response to Panglossian accusations by Randolph and Dobson: Biodiversity buffers disease. *Parasitology* **140**: 1196 – 1198.

Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. 2006. Climate, Deer, Rodents, and Acorns as Determinants of Variation in Lyme-Disease Risk. *PLoS Biol* **4**: e145.

Ostfeld RS, Keesing F. 2013. Straw men don't get Lyme disease: response to Wood and Lafferty. *Trends in Ecology Evolution* **28**: 502 – 503.

Ostfeld RS, Brunner JL. 2015. Climate change and *Ixodes* tick-borne diseases of humans. *Philosophical Transactions B* **370**: 20140051

Patz JA, Olson SH, Uejio CK, Gibbs HK. 2008. Disease emergence from global climate and land use change. *Medical Clinics of North America* **92**:1473–1491.

Paquette SR, Talbot B, Garant D, Mainguy J, Pelletier F. 2014. Modelling the dispersal of the two main hosts of the raccoon rabies variant in heterogeneous environments with landscape genetics. *Evolutionary Applications* **7**: 734 – 749

Pearce-Duvet JMC. 2007. The origin of human pathogens: Evaluating the role of agriculture and domestic animals in the evolution of human diseases. *Biological Reviews* **81**: 369 – 382.

Pound JM, George JE, Kammlah DM, Lohmeyer KH, Davey RB. 2010. Evidence for role of ehite-tailed deer (Artiodactyla: Cervidae) in epizootiology of cattle ticks and southern cattle ticks (Acari: Ixodidae) in reinfestations along the Texas/Mexica border in south Texas: A review and update. *Journal of Economic Entomology* **103**: 211 – 218.

Qiu W, Dykhuizen DE, Acosta MS, Luft BJ. 2002. Geographic uniformity of the Lyme disease spirochete (*Borrelia burgdorferi*) and its shared history with tick vector (*Ixodes scapularis*) in the Northeastern United States. *Genetics* **160**: 83 – 849.

Randolph S. 2013. Commentary on 'A candide response to Panglossian accusations by Randolph and Dobson: Biodiversity buffers disease' by Dr. R. Ostfeld. *Parasitology* **140**: 1199 – 1200.

Randolph SE, Dobson ADM. 2012. Pangloss revisited: A critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology* **139**: 847 – 863.

Robinson SJ, Samuel MD, Lopez DL, Shelton P. 2012. The walk is never random: Subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. *Molecular Ecology* **21**: 4190 – 4205.

Rogic A, Tessier N, Legendre P, Lapointe F-J, Millien V. 2013. Gentic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Québec . *Ecology and Evolution* **3**: 2075 – 2088.

Rollend L, Fish D, Childs JE. 2013. Transovarial transmission of Borrelia spirochetes by Ixodes scapularis: a summary of the literature and recent observations. *Ticks and Tick-borne Diseases* **4**: 46 – 51.

Rosenthal SR, Ostfeld RS, McGarvey ST, Lurie MN, Smith KF. 2015. Redefining disease emergence to improve prioritization and macro-ecological analyses. *One Health* **1**: 17 – 23.

Roy-Dufresne E, Logan T, Simon JA, Chmura GL, Millien V. 2013. Poleward expansion of the white-footed mouse (*Peromyscus leucopus*) under climate change: Implications for the spread of Lyme disease. *PLoS ONE* **8**: e80724

Schotthoefer AM, Frost HM. 2015. Ecology and Epidemiology of Lyme Borreliosis. *Clinics in Laboratory Medicine* **35**: 723 – 743.

Schrag SJ, Wiener P. 1995. Emerging infectious disease: What are the relative roles of ecology and evolution? *Trends in Ecology and Evolution* **10**: 319 – 324.

Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harbor Perspectives in Medicine* **1**: a006841.

Simon JA, Marrotte RR, Desrosier N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of *Borrelia burgdorferi*, the agent of Lyme disease, at the northeastern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764.

Spach DH, Liles WC, Campbell GL, Quick RE, Anderson DE, Fritsche TR. 1993. Tick-borne diseases in the United States. *The New England Journal of Medicine* **329**: 936 – 947.

Tabachnick WJ. 2010. Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. *The Journal of Experimental Biology* **213**: 946 – 954.

Tälleklint L, Jaenson TG. 1994. Transmission of *Borrelia burgdorferi* sensu lato from mammal reservoirs to the primary vector it Lyme borreliosis, *Ixodes ricinus* (Acari: Ixodidae), in Sweden. *Journal of Medical Entomology* **31**: 880 – 886.

Tilly K, Rosa PA, Stewart PE. 2008. Biology of infection with Borrelia *burgdorferi. Infectious Disease Clinics of North America* **22**: 217 – 234.

Tsao JI. 2009. Reviewing molecular adaptations of Lyme borreliosis spirochetes in the context of reproductive fitness in natural transmission cycles. *Veterinary Research* **40**: 36.

Vail SG, Smith G. 2002. Vertical movement and posture of black-legged tick (Acari: Ixodidae) nymphs as a function of temperature of temperature and relative humidity in laboratory experiments. *Journal of Medical Entomology* **39**: 842 – 846

Vreysen MJB, Saleh KM, Ali MY, Abdulla AM, Zhu Z-R, Juma KG, Dyck A, Msangi AR, Mkonyi PA, Feldmann HU. 2000. *Glossina austeni* (Diptera: Glossinidae) eradicated on

the Island of Unguja, Zanzibar, using the sterile insect technique. *Journal of Economic Entomology* **93**: 123 – 135.

Wikel SK. 2013. Chapter 4 Tick-host interactions. *Biology of Ticks – Volume 2* (eds Sonenshine DE, Roe RM). Oxford University Press. pp 88 – 128

Wood CL, Lafferty KD. 2013. Biodiversity and disease: a synthesis of ecological perspectives on Lyme disease transmission. *Trends in Ecology Evolution* **28**: 239 – 247.

Woolhouse ME, Dye C. 2001. Population biology of emerging and re-emerging pathogens – preface. *Philosophical Transactions of the Royal Society B* **356**: 981 – 982.

Wooten RM, Weis JJ. 2001. Host-pathogen interactions promoting inflammatory Lyme arthritis: Use of mouse models for dissection or disease processes. *Current Opinion in Microbiology* **4**: 274 – 279.

Wu X, Duvvuri VR, Lou Y, Ogden NH, Pelcat Y, Wu J. 2013.Developing a temperature-driven map of the basic reproductive number of the emerging tick vector of Lyme disease *Ixodes scapularis* in Canada. *Journal of Theoretical Biology* **319**: 50 – 61

Xu G, Fang QQ, Keirans JE, Durden LA. 2003. Molecular phylogenetic analyses indicate that the *Ixodes ricinus* complex is a paraphyletic group. *The Journal of Parasitology* **89**: 452 – 457.

Zhioua E, Aeschlimann A, Gern L. 1994. Infection of field-collected *Ixodes ricinus* (Acari: Ixodidae) larvae with *Borrelia burgdorferi* in Switzerland. *Journal of Medical Entomology* **31**: 763 – 766.

FIGURES

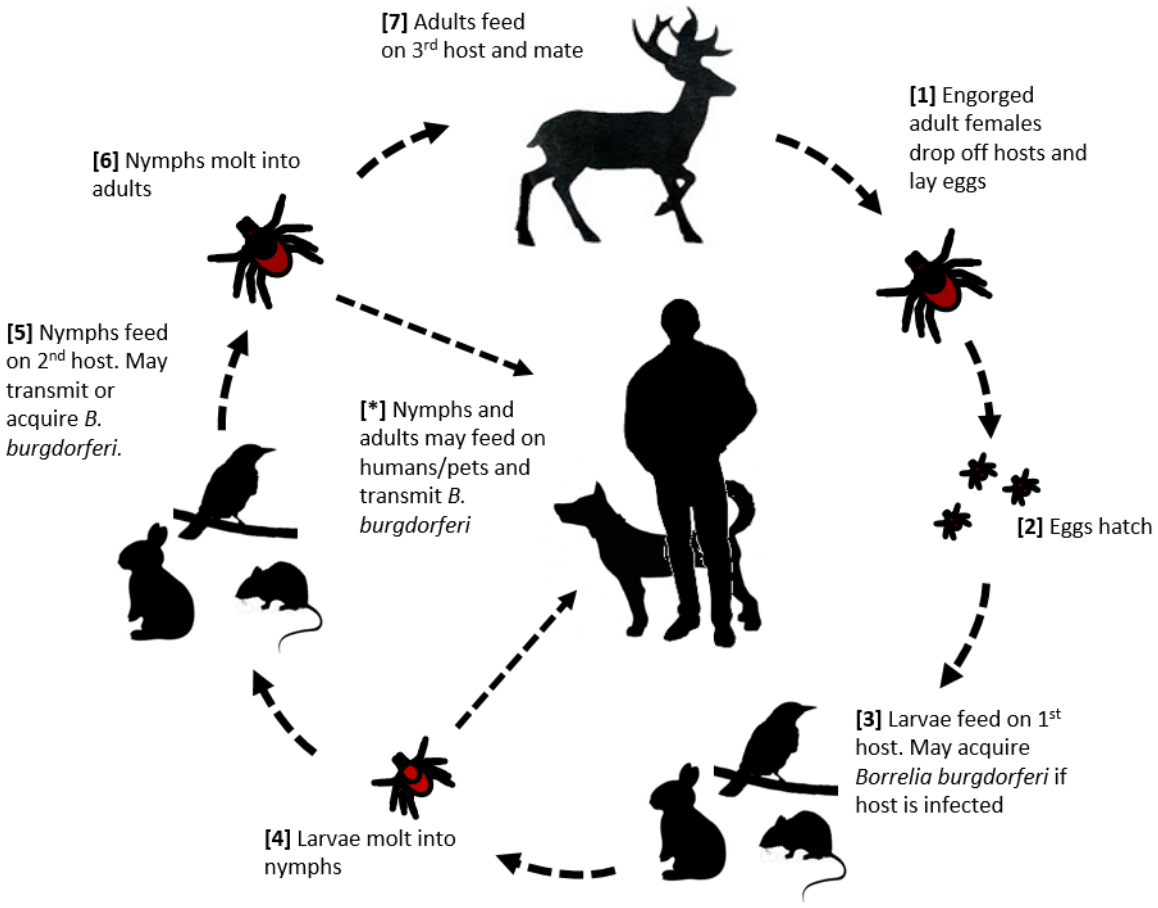


Figure 1.1. Typical life cycle of *Ixodes scapularis* and transmission cycle of Lyme disease in eastern North America.

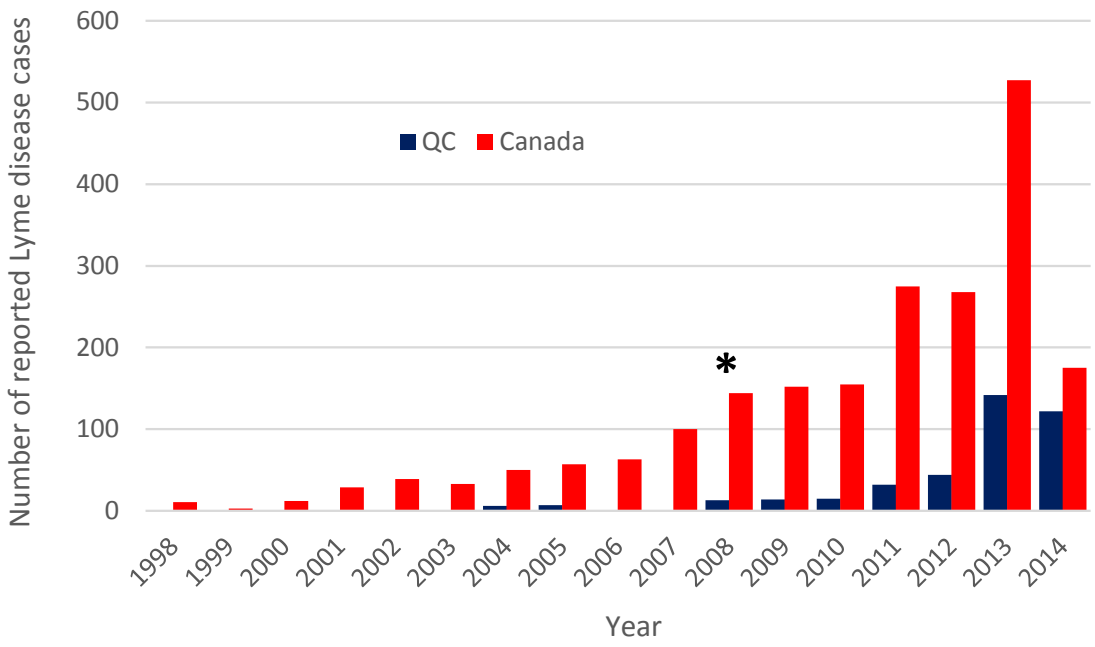


Figure 1.2. Annual number of cases of Lyme disease reported in humans in southern Québec (blue) and Canada (red) collated from annual reports from provincial health departments across the country. Numbers for Québec were obtained from the Institut National de Santé Publique du Québec. Please note that these numbers were reported as of July 2015 and that final numbers for 2014 may yet be incomplete. The '*' in the figure indicates when the first endemic case of Lyme disease in human was reported in Québec.

CHAPTER TWO

Multi-taxa integrated landscape genetics for zoonotic infectious diseases: Deciphering variables influencing disease emergence

* This chapter has been published:

Leo SST, Gonzalez A, Millien M. 2016. *Genome*, 10.1139/gen-2016-0039

Author contributions: *SSTL, AG, and VM conceived the ideas and designed the study. SSTL performed the simulations and analyses. SSTL, AG, and VM wrote the manuscript.*

ABSTRACT

Zoonotic disease transmission systems involve sets of species interacting with each other and their environment. This complexity impedes development of disease monitoring and control programs which require reliable identification of spatial and biotic variables and mechanisms facilitating disease emergence. To overcome this difficulty, I propose a framework that simultaneously examines all species involved in disease emergence by integrating concepts and methods from population genetics, landscape ecology, and spatial statistics. Multi-taxa integrated landscape genetics (MTILG), first described in a study on mountain pine beetle-fungi interaction, can reveal how interspecific interactions and landscape variables influence disease emergence patterns. I test the potential of the MTILG-based framework by modelling the emergence of a disease system across multiple species dispersal, interspecific interaction, and landscape scenarios. My simulations showed that both interspecific-dependent dispersal patterns and landscape characteristics significantly influenced disease spread. Using my framework, I was able to detect statistically similar inter-population genetic differences and highly correlated spatial genetic patterns that imply species-dependent dispersal. Additionally, species that were assigned coupled dispersal patterns were affected to the same degree by similar landscape variables. This study underlines the importance of an integrated approach to investigating emergence of disease systems. MTILG is a robust approach for such studies and can identify potential avenues for targeted disease management strategies.

INTRODUCTION

Emerging zoonotic infectious diseases can threaten biodiversity and impose significant pressures on public health systems at local, regional, and global scales (Taylor *et al.* 2001, Jones *et al.* 2008, Brito *et al.* 2012). Over 70% of recently emerging infectious diseases have animal origins and roughly 60% of all human pathogens are zoonotic (Jones et al. 2008, Johnson et al. 2015). For vector-borne diseases, the presence, distribution, and movement of competent vectors across a region are important variables determining disease risk and emergence patterns (Gubler 2010). We can therefore monitor zoonotic disease emergence by identifying landscape variables influencing movements of organisms, and hence disease spread, across a landscape (Gilchrist 2002, Roth 2011).

Spatial metacommunity dynamics are critical when species involved in a disease system are distributed across heterogeneous landscapes (Hess 1996, Arino *et al.* 2005, Alexander *et al.* 2012, Estrada-Peña *et al.* 2014). Changes in land use and habitat adjacency can influence species dispersal, population dynamics, and biotic interactions that in turn shape patterns of reservoir, vector, and/or pathogen distribution and disease occurrence (Grenfell & Harwood 1997, Suzán *et al.* 2015). For instance, in the Lyme disease system variability in habitat conditions that alter species composition, relative abundances, and interspecific interactions within an ecological community can influence vector abundance and pathogen persistence (LoGiudice *et al.* 2003, Biek & Real 2010, Ostfeld 2011, Ostfeld & Keesing 2012). Landscape structure influences disease spread as it affects both movement patterns and densities of hosts, disease vectors and pathogens (McCallum 2008). For vector-transmitted diseases, landscape features that influence vector dispersal are the foremost determinants of rates of epizootic spread. Physical barriers to

dispersal and habitat fragmentation can influence the distributions and affect rates of species movement and population connectivity, including on the epidemic front (McCallum 2008).

Additionally, most complex zoonotic disease transmission systems often include multiple sets of host, vector, and pathogen species, each with roles of varying importance in disease transmission, and it has become clear in recent years that disease emergence processes should be studied in a multi-species context (Estrada-Peña et al. 2014). The dispersal and evolution (i.e. host specificity, speciation) of pathogens and vectors in a vector-borne zoonotic diseases system are often dependent on host species (Blasco-Costa & Poulin 2013, Suzán et al. 2015, Mazé-Guilmo et al. 2016, Grégoir et al. 2015). Therefore, interspecies-dependent dispersal patterns should result in observable signs of shared spatial dependence (James et al. 2011, Blasco-Costa & Poulin 2013, Mazé-Guilmo et al. 2016) that can provide insights on key species facilitating the spread of pathogens across a region.

With recent advancements in molecular techniques and rapid improvements in bioinformatics, combinations of good genotypic and landscape datasets paired with appropriate statistical analyses can improve our understanding of biotic mechanisms and abiotic variables influencing disease emergence in a region (Hudson 2008, Archie et al. 2009, Ekblom & Galindo 2011). For example, we might be able to identify stages of disease emergence, detect species-specific genotypic or phenotypic adaptations and habitat characteristics that promote or inhibit species and disease establishment. In multi-taxa disease systems it is critical to quantify inter-specific interactions and estimate how landscape features impact these interactions and hence the emergence of the disease transmission system (Biek & Real 2010).

Here, I develop a framework using the multi-taxa integrated landscape genetics (MTILG) approach (James et al. 2011) to identify biotic and abiotic factors influencing disease emergence

(Figure 2.1). MTILG integrates concepts and methods from population genetics, community ecology, landscape ecology, and spatial statistics. It allows the simultaneous examination of how multiple species interact with each other and their surrounding environments. Different types of datasets, such as species-specific genetic or phenotypic information, site information, and landscape features can be used and paired with appropriate statistical analyses to identify and quantify biotic mechanisms and abiotic variables influencing disease emergence in a region (Hudson 2008, Archie *et al.* 2009, Ekblom & Galindo 2011) (Figures 2.1 & 2.2).

With this MTILG-based framework, we can potentially quantify species gene flow and movement across the landscape and estimate how landscape features impact spatial patterns of the disease transmission (Biek & Real 2010). Rates and amounts of gene flow among populations are often positively correlated with species' mobility and dependent on environmental constraints and dispersal barriers (Slatkin 1985). Landscape genetics can reveal how geographical and environmental features structure genetic variation at both population and individual levels (Manel & Holderegger 2013) by detecting genetic discontinuities and the correlation of these discontinuities with specific landscape features (Figure 2.1). Alternatively, interspecies-dependent dispersal patterns can result in correlated spatial genetic patterns (James et al. 2011, Blasco-Costa & Poulin 2013). Therefore, we would expect to observe high congruence in interspecific spatial genetic patterns if species within a disease system exhibit coupled-dispersal behaviour (e.g. pathogens and their vectors) or share similar spatial habitat extinction and colonization dynamics (e.g. in fragmented landscapes) (Figure 2.1). There are several approaches to studying concordance in spatial genetic patterns between species that range from simple Mantel comparisons of pairwise genetic distance among species, to complex graph theoretic approaches (e.g. Widmer et al. 2012) (Figure 2.1).

To test the effectiveness of my MTILG-based framework, I simulated the emergence of a disease system (consisting of two hosts species, one vector and one pathogen species) over four distinct landscape scenarios of increasing heterogeneous complexity and two host species dispersal scenarios (Figure 2.2) (Epperson *et al.* 2010). In order to simulate host dependent-dispersal patterns, vectors and pathogens dispersal were coupled with either of the host species. Upon completion of the disease emergence simulations, genetic data produced from the simulations were used in the MTILG framework to test the framework's effectiveness. I first examined the extent, rates, and time taken for disease emergence across simulation scenarios, testing for the effects and contributions of landscape characteristics, species dispersal rates, and species interactions on measures of disease emergence. I then analysed population genetic differentiation among host, vector, and pathogen populations, quantified interspecific interactions in the form of host-dependent dispersal, and identified the extent to which landscape features determine species gene flow and constrain disease emergence and establishment patterns.

MATERIAL AND METHODS

Model Overview - Purpose

I wanted to capture how landscape characteristics, host species dispersal rates, and interspecific interactions influence disease transmission system emergence within a region (Figure 2.2). Using computer simulations, I modelled the emergence of a hypothetical disease system over four landscape scenarios and two dispersal mechanisms (Epperson *et al.* 2010).

Model Overview - Variables and landscapes

Our simulated disease transmission system consists of four species – a pathogen (P), a disease vector (V), and two host species (H1 and H2) (Figure 2.2). The pathogen in the system has no free-living life stage and is transmissible only via an infected vector that picks up the pathogen from an infected host. The pathogen is not transmissible either horizontally (i.e. between members of the same host/vector species) or vertically (i.e. directly from the mother to its offspring). Both host species H1 and H2 are equally competent reservoirs of the disease-causing pathogen and are also suitable hosts for the parasitic vector (Figure 2.2). I assumed that the pathogen has negligible effects on hosts and vector behaviour, survival, or fitness. While the system contains two host species, the presence of either host within a habitat is sufficient to maintain the disease transmission system locally. I ran eight different simulation scenarios (Table 2.1). I modelled the spread of the disease system within 14 sites located across four distinct landscapes with increasing complexity, from a completely homogeneous landscape (e.g. undisturbed forests) to a heterogeneous landscape with a river (e.g. forest fragmentation with linear barriers to species movement) (Figure 2.2). I also simulated two distinct dispersal rates of the hosts driving the movement of the vector and the pathogen across these landscapes to examine the effects of species dispersal capabilities on disease emergence (Table 2.1).

Model Details - Initialization

I modelled the gene flow of all species in the disease transmission system using the program PEDAGOG. PEDAGOG is an evolutionary ecology software that simulates individual-level population dynamics while producing genotype, pedigree, and phenotype information that can be used in subsequent analyses (Coombs *et al.* 2010). I simulated population statistics and genetic data for all fourteen sites in the system over a period of 30 generations, by which time disease emergence events woud have reached equilibrium. Each simulation of unique landscape

and dispersal scenario was performed with 10 replicates each, which was sufficient to capture any inherent stochasticity due to demographic and genetic mechanisms and sampling. This resulted in a total of 80 replicate simulations. Fourteen sites were randomly placed on habitat patches across each landscape to serve as populations from which genetic data was sampled. All specimens were genotyped for ten microsatellite loci and sampled only once at initial capture. The number of alleles simulated per locus was random across loci and across simulation scenarios. For all subsequent genetic analyses, I used the simulated genetic data produced for the 30^{th} generation in my simulations (i.e. after disease emergence have reached equilibrium).

Model Details – Input

In the PEDAGOG simulations, I kept all initial species-specific genetic diversity, demographic rates, growth rates, mating designs, heritability, and mutation rates constant for all species, regardless of landscape or dispersal scenarios (see Table A1 in Appendix A).

Spatial variables were represented in raster format. Each simulated landscape measured 100 by 100 cells. In the homogeneous landscape, all cells were given the value 1. In a heterogeneous landscape, 18 habitat patches measuring 10 by 13 cells were randomly distributed across the landscape. In this scenario, all cells representing a habitat patch was assigned the value of 1 while any remaining cells in the landscape were assigned a value of 5 to represent greater resistance to species movement. In the presence of a geographical barrier in a landscape, the barrier extended across the full wideth of the landscape. The width of the barrier ranged from 8 to 15 cells, and all cells representing the barrier in the landscape were assigned a value 10 to indicated substantial barrier to species movement.

I simulated interspecific interactions in the form of host-dependent dispersal by coupling vector and pathogen dispersal to either host species H1 (short-distance disperser) or H2 (long-

distance disperser). I determined probabilities of movement from a source site and amongst sites for all species beforehand by incorporating least-cost distances calculated for each landscape scenario and species-specific dispersal capabilities as a simple negative exponential distribution function:

$$\rho_{ij} = \alpha e^{\left(\frac{-d_{ij}}{\beta}\right)} \quad \{\text{Eqn 1}\}$$

where ρ_{ij} is the probability of dispersal between sites *i* and *j*, α the probability of an individual remaining at the original site, d_{ij} the least-cost effect distance between sites *i* and *j* based on resistance values, and β the estimated mean dispersal capability of the species of interest.

I retrieved values of least-cost distances (d_{ij}) using the raster (Hijmans *et al.* 2015) and gdistance (van Etten 2015) packages in R. I first rasterized maps of our four landscape scenarios using the *raster* function and assigned resistance values of 1, 5, and 10 to three landscape variables respectively: habitat, non-habitat, and physical barrier. I then obtained transition values for each landscape using the *transition* function and from there calculated the least-cost distances (d_{ij}) among fourteen sample sites using the function *costDistance*. I set species dispersal capabilities - β (defined as the maximum resistance each species can overcome per dispersal event in the simulated landscapes) - at 80 and 160 for short-distance and long-distance dispersers, respectively. I set the probability of an individual remaining at the original site (α) at 0.85. Although the parameter values used in these simulations are not reflective of any real-world system, they are not unrealistic relative to the spatial and temporal scale of the model.

Analysis - Extents, rates, and time taken for disease system emergence

I determined differences in the spatial extent of disease emergence and time taken to reach maximum extent for all scenarios based on species-specific simulated data. The disease

transmission system was considered locally established at each site when populations of vector, pathogen, and at least one host species had established stable populations. Species populations were considered established when effective population sizes (N_e) were 20 or more for at least three consecutive generations. Rates of emergence were calculated as the final proportion of sites in which disease has established divided by the time taken from when the disease system began to colonize patches in the landscapes.

I first tested for the effect of simulation characteristics (landscape heterogeneity, presence of geographical barriers, and species dispersal rates) on the extent and rate of disease outbreak, along with time taken to reach equilibrium among the eight simulation scenarios via ANOVA. I then performed a series of linear regression analyses using simulation characteristics as explanatory variables and the extents, rates, and time taken for disease emergence to reach equilibrium as response variables. I subsequently performed stepwise model selection using the *stepAIC* function in R package MASS (Ripley *et al.* 2015) to identify which explanatory variables best explained disease emergence patterns. Finally, I assessed the relative importance and contribution of each predictor in the model with the *calc.relimp* and the *varpart* functions from the R package relaimpo and vegan, respectively.

Analysis - Genetic differentiation among species populations

I examined population structures using AMOVA in the software Arlequin ver. 3.5.1.3 (Excoffier *et al.* 2010).). Estimator options for AMOVA analyses in Arlequin are limited to R_{ST} (Slatkin 1995) or F_{ST} (Wright 1950). Since F_{ST} measures do not account for the stepwise mutation process inherent to microsatellite markers (Balloux & Lugon-Moulin 2002), I elected to employ R_{ST} as an estimator for measuring genetic differentiation. R_{ST} is a measure of genetic differentiation (Excoffier *et al.* 1992) and higher R_{ST} values generally indicate greater genetic

differentiation among populations. Similar R_{ST} values among associated species would be indicative of shared spatial dependence, coupled dispersal, or interspecific interactions. I was able to examine how interactions among landscape heterogeneity, geographical barriers, and species dispersal rates influence genetic differentiation among species populations by applying a stepwise model selection on a model that regressed observed R_{ST} values against landscape variables. Stepwise model selection was performed using the *stepAIC* function in R package MASS. The relative importance of each predictor in the model was measured using the *calc.relimp* function from the R package relaimpo (Groemping & Matthias 2013).

Our next step in the MTILG framework investigated interactions among species that exhibited similar genetic differentiation among populations by testing for concordance in their spatial genetic patterns.

Analysis - Quantifying interspecific interactions facilitating disease emergence

I quantified interspecific interactions in the form of host-dependent dispersal based on concordance of spatial genetic patterns among species by calculating ordination solutions (James *et al*. 2011) of allele frequency differences and pairwise genetic distance among populations of all species in each simulation across all replicates. I obtained ordination solutions of allele frequency differences among populations using the *rda* function in the vegan package in R (Oksanen *et al.* 2015). I also performed additional analyses based on pairwise genetic distances using Principal coordinate analyses (using the *pairwise.fst* and *capscale* functions in R packages adegenet (Jombart *et al.* 2014) and vegan, respectively). I compared PCA (obtained from raw allele frequencies) and PCoA ordination solutions (based on pairwise F_{ST} values) using a series of Procrustes rotation tests, a variant of co-inertia analysis to assess similarity between multivariate datasets. I found significant and high correlations between all PCA and PCoA

ordination solutions (average correlation coefficient = 0.8334, average p-value = 0.0115), suggesting that PCA ordinations likely do not violate evolutionary assumptions (e.g. mutation, selection, isolation, and mixing) and can be used in subsequent analyses.

I then performed a series of pairwise Procrustes rotation tests on ordination solutions of allele frequency differences among populations for all possible pairs of species in each replicate of disease emergence simulation scenario. This allowed us to evaluate the concordance in spatial genetic patterns of differentiation across species. In the event vector or pathogen species are dependent on a specific host species for dispersal, I expect the Procrustes Rotation test to register significantly similar patterns of allele frequency distributions between ordination solutions of these associated species. All Procrustes rotation tests were performed using the *protest* function from the R package vegan.

Analysis - Landscape features affecting spread

I converted spatial environmental data into site-specific measurements by calculating the connectivity of each site relative to all other sites based on the resistances values and least-cost effective distances estimated for each landscape variable. Connectivity was calculated using Eqn. 2:

$$S_i = \sum_{1}^{j} \frac{1}{1 + \beta d_{ij}}, \qquad \{\text{Eqn 2}\}$$

where S_i is the connectivity of site i to the rest of the metapopulation network (βd_{ij}), d_{ij} the least-cost effect distance between sites *i* and *j* based on resistance values, and β the estimated mean dispersal rate of the species of interest.

Using log-transformed allele frequencies as response matrices and measures of spatial connectivity as predictor variables in a series of RDA analyses, I identified ordination models that best described genetic similarities relative to landscape variables. I first standardized measures of spatial connectivity by standardizeding and detrending the values against euclidean distances. I subsequently selected optimal models via forward selection analyses using the *ordistep* function in the R package vegan.

RESULTS

Extent, rate, and time taken for disease system emergence

The spatial extent of disease emergence differed among simulation scenarios (Figure 2.3) and was dependent on both landscape characteristics and species dispersal rates (Table A2 in Appendix A). Interactions among landscape heterogeneity, geographical barriers, and species dispersal rates accounted for most of the variance observed in disease emergence extent (Table A2), with landscape heterogeneity and geographical barriers being the primary contributors to the variance (Table A2).

The amount of time taken for disease emergence to reach equilibrium also differed among simulation scenarios (Figure 2.3). Interactions between landscape heterogeneity and species dispersal rates best accounted for variance in time taken for emergence to reach equilibrium (Table A2). However, an assessment of the relative importance of the individual contributions of these variables, along with variation partitioning analyses revealed that landscape heterogeneity was the main driver of time taken for disease emergence to reach equilibrium, and that species dispersal capabilities rates contributed very little to none of the observed variance (Table A2).

Finally, ANOVA analyses indicated that rates of emergence also differed amongst simulation scenarios (Figure 2.3), and selection tests on multiple regression models revealed that interactions between landscape heterogeneity and geographical barriers best explained the observed variance in emergence rates. However, tests of relative importance and variance partitioning once again showed that landscape heterogeneity was the most important and primary contributor to the observed variance (Table A2).

Genetic differentiation among species populations

Stepwise model analyses of regression models (using R_{ST} values as response variables and landscape characteristics and species dispersal capabilities as explanatory variables) showed that population differentiation for all species was significantly influenced by, in order of importance, landscape heterogeneity, geographical barriers, and species dispersal rates. Populations distributed across homogeneous landscapes tended to exhibit lower R_{ST} values than those in heterogeneous landscapes (Table A3). Furthermore, R_{ST} values tended to be greater among populations of species with short distance dispersal than those capable of long distance dispersal (Figure 2.4, Table A3). These patterns matched the null expectations of isolation-by-distance, where populations experiencing greater amount of gene exchanges exhibit greater genetic similarities.

R_{ST} values also provided a proxy for the strength of interspecific interactions. I found that in six out of eight simulation scenarios, associated host, vector, and pathogen species generally shared similar patterns of genetic differentiation (Figure 2.4, Table A3). R_{ST} values were often similar among vector and pathogen species and the host species upon which they depended for dispersal (Figure 2.3) (i.e. coupled species in the simulation models). Conversely, R_{ST} values were generally significantly different among unassociated species (Figure 2.4).

Quantifying interspecific interactions facilitating disease emergence

The spatial genetic patterns of vector and pathogen species generally exhibited greater concordance with that of the associated host species upon which their dispersal depended (Table 2.2). I also found that statistically significant concordances were more frequently observed among associated species (~60% of 80 simulation replicates) than unassociated species (~35% of 80 simulation replicates) (Table 2.2). The unexpected statistically significant concordances among unassociated species may have occurred as a result of the landscape characteristics producing similar gene flow pressures across all species (i.e. shared spatial dependence) (James et al. 2011). A brief analysis of the differences in correlation coefficient values between associated and unassociated species revealed that the difference decreased with increasing landscape complexity. Additionally, I noted a positive relationship between the frequencies at which significant concordance between unassociated species was detected and landscape complexity. This underlines the importance of the ensuing step in the MTILG framework, in which I identified and quantified the relative roles of landscape variables on species gene flow via landscape genetics analyses. Landscape genetic analyses can differentiate whether any observed concordances in inter-specific spatial genetic resulted from genuine host-dependent dispersal or shared spatial dependence.

Landscape features affecting spread

Landscape heterogeneity and geographical barriers significantly affected species spatial genetic and dispersal patterns, which confirmed the model design. Results from selection tests revealed that the spatial genetic patterns of all species in the disease system can often be attributed to the same landscape variables, regardless of interspecific associations (Table A4). This is a strong indicator of the effects of shared spatial dependence, where the movements of all

species studied were affected by the same landscape variable. However, the extent to which these landscape variables explained species genetic patterns differed starkly depending on relations among species. Overall, I found that the proportion of genetic variation explained by a given landscape variable tended to be similar among host, pathogen, and vector species whose dispersal rates were coupled (Figure 2.5). This suggests that gene flow patterns of associated host, vector, and pathogen species were influenced to similar degrees by the same landscape characteristic (Figure 2.5), indicating potential coupled dispersal and intimate interactions among associated species in the simulations.

DISCUSSION

I chose to model disease emergence patterns over several landscape and dispersal scenarios to test the efficiency of my MTILG-based framework under these varying scenarios (Epperson *et al*. 2010). Despite the apparent simplicity of my computer simulations and disease system, I was able to reveal that interactions among landscape characteristics, interspecific interactions, and species dispersal rates can have different effects on disease emergence rates and extent (Figure 2.3). Furthermore, I was able to, using my framework, identify landscape variables that influence patterns of species (and therefore disease) spread, as well as identify potential coupled species dispersal patterns. However, the framework revealed a rich set of outcomes that depended on the landscape scenario. I now discuss these aspects and link to known patterns of emergence in zoonotic diseases.

Insights from MTILG on zoonotic infectious disease emergence

Most studies on emerging zoonotic disease to date tend to focus on a single, identified keystone species or a limited number of environmental variables. However, it is becoming evident in recent years that zoonotic disease systems are more complex than expected, and that determining mechanisms of disease emergence based on constrained studies often result in erroneous conclusions regarding the dynamics of the disease system (Estrada-Peña et al. 2014). For instance, while single-species landscape genetics study may provide important spatial information about population structure for a particular host, vector, or pathogen species, it cannot discern general underlying evolutionary and ecological factors that affect the disease system as a complex of species (Manel & Holderegger 2013). Interacting species in a complex disease system have distinct ecological niches that overlap only in certain portions of environmental and geographical space (Estrada-Peña et al. 2014). The landscape variable that influences one species at a particular spatial scale does not necessarily affect another at that scale. If disease risk is enhanced when multiple species are present, we need to simultaneously compare multiple host and vector species across spatial and temporal scales to identify cross-scale relations that facilitate disease emergence. These common variables can then be applied in models to accurately project realistic future geographic disease emergence patterns.

In addition to knowing how variables influence disease movement or establishment within the landscape, it is also imperative that we identify the species' roles and interspecific interactions in a disease system. This is especially true in the case of disease pathogens and vectors. If the dispersal of pathogens or vectors is dependent or coupled with host species movement, identification and quantification of the relative importance of these hosts as dispersers can help us identify likely spatial paths for disease emergence that can then be targeted for disease control. Targeted control strategies, such as vaccination, medication, host

population size manipulation, or environmental manipulation, will be more cost-efficient, and potentially decrease risks of detrimental impacts on ecosystems (Artois *et al.* 2011).

Challenges with the framework

The data requirement of MTILG may seem demanding in their extent and complexity. Nonetheless, recent advances in molecular techniques (e.g. high-throughput next-generation sequencing), satellite imaging, and the maintenance of online databases such as those maintained by the National Center for Biotechonology Information (NCBI) or the European Bioinformatics Institute (EBI) are all contributing to making MTILG-ready dataset more common. The nature of these datasets also promotes collaboration amongst experts across multiple disciplines, pooling and sharing resources and data, which reduces the amount of time required for data assembly and analyses.

The reliability and accuracy of results from the MTILG-based framework will depend on the temporal and spatial scales over which data are collected (Anderson *et al.* 2010). For instance, observed genetic patterns are the outcome of historical gene flow, and there is therefore a time lag between actual gene flow events and production of measurable genetic signals (Landguth *et al*. 2010). The types of molecular markers used will determine the temporal and spatial scales of data collection. Spatial scale is a crucial feature of landscape genetics and my simulations stress the importance of multiscale datasets. This is because the effects of landscape features vary greatly across scales (Cushman & Landguth 2010). This is in part functional because the relative impact of landscape features as barriers or as corridors depends strongly on a species' perceptions and movement rates. While analyses that partition populations according to assumed barriers of gene flow (e.g. hierarchical AMOVA) can potentially help identify the relative amount of variance explained by each landscape variable, erroneous classifications of

perceived barriers can result in incorrect conclusions. It is therefore critical to obtain or estimate appropriate and biologically realistic estimates for movement parameters (i.e. landscape resistance values and mean species dispersal distances) prior to analyses. Inaccurate values may result in over- and/or under-estimations of landscape effects on species gene flow, leading to unreliable projections and misinformed disease management decisions.

Additionally, my study suggests that the performance of the MTILG framework tends to improve with increasing landscape complexity. For instance, the framework was able to identify signals of coupled dispersal patterns (i.e. similar population structures, and concordant spatial genetic patterns) among hosts, vectors, and pathogen more frequently under heterogeneous landscape conditions. Conversely, application of the framework under homogenous landscape scenarios identified coupled-dispersal interactions at lower frequencies. I also found that the MTILG- framework was better at identifying landscape variables influencing species movement when the landscape was fragmented, although this is more likely a consequence of the inherent lack of variables to test in a homogeneous landscape. It is also interesting to note that the framework did not appear to perform differently under either of the scenarios of dispersal-driven disease emergence.

Applications of the framework

MTILG can be applied to a variety of emerging zoonotic disease systems. The emerging Lyme disease system in North America is a good example of a complex disease transmission system that lends itself to my MTILG-based framework . The disease system involves multiple host species, each of which exhibit varying efficacy in maintaining and supporting tick vectors and pathogen transmission (Brunner *et al.* 2008, Ostfeld 2011). In recent decades, climate and

land use changes have facilitated the rapid spread of Lyme borreliosis across North America (Allan *et al.* 2003, Brownstein *et al.* 2005, Simon *et al.* 2014). By simultaneously comparing and investigating interactions among ticks, their multiple hosts and the structure of the surrounding landscapes, researchers can begin to improve disease control strategies by identifying and targeting specific host species or landscape conditions that foster spread and mediate prevalence.

Another emerging zoonotic disease system that might benefit from the application of MTILG is rabies - a viral disease of the central nervous system in warm-blooded animals transmitted via scratches, bites, or contact of saliva with mucous membranes (Rupprecht *et al.* 2002). Wildlife hosts such as raccoons, skunks, and bats are competent hosts that can maintain rabies in a region, albeit with different antigenic variants or strains that may exhibit varying virulence (Smith *et al.*1984, Rupprecht *et al.* 2002). As climate and landscape changes alter wildlife species abundances, distribution, and behaviour, the distribution and prevalence of rabies may change as well. The framework can help to identify important variables and mechanisms that influence the spread of more virulent rabies strain types and allow development of targeted, more cost effective control programs (Rupprecht *et al.* 2004).

CONCLUSION

Zoonotic disease systems often involve complex networks of interactions among multiple species and their surrounding environment. To obtain a clearer understanding of the factors that influence disease emergence in a region, it is essential to simultaneously examine the species involved in the disease system and how their interactions and movements occur in the landscape. In this paper, I showed that the MTILG framework can reveal how biotic and abiotic variables influence zoonotic disease emergence in a region. Specifically, I was able to identify shared

spatial occurrences and the interspecies-dependent dispersal patterns between associated species. I was also able to demonstrate how interactions among landscape heterogeneity, geographical barriers, and species dispersal rates can influence the spatial extent and rate of disease transmission system spread through time. Modelling interactions among multiple species and how they interact with the landscape is essential for informed management decisions aimed at preventing disease emergence and spread. MTILG will become necessary and feasible as large datasets become available and can be tackled by the pipeline of analyses I described. New statistical and modeling tools facilitate regional models with greater forecasting accuracy (Sork & Waits 2010). Better forecasts will improve our ability to mitigate and even counter emergence of infectious diseases world wide.

REFERENCES

Alexander KA, Lewis BL, Marathe M, Eubank S, Blackburn JK. 2012. Modeling of wildlife-associated zoonoses: Applications and Caveats. *Vector-Borne and Zoonotic Diseases* **12**: 1005-1018. doi: 10.1089/vbz.2012.0987.

Allan BF, Keesing F, Ostfeld RS. 2003. Effect of forest fragmentation on Lyme disease risk. *Conservation Biology* **17**: 267-272. doi: 10.1046/j.1523-1739.2003.01260.x.

Anderson CD, Epperson BK, Fortin M-J, Holderegger R, James PMA, Rosenburg MS, Scribner KT, Spear S. 2010. Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology* **19**: 3565-3575. doi: 10.1111/j.1365-294X.2010.04757.x.

Archie EA, Luikart G, Ezenwa VO. 2009. Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology and Evolution* **24**: 21-30. doi: 10.1016/j.tree.2008.08.008.

Arino J, Davis JR, Hartley D, Jordan R, Miller JM, van den Driessche P. 2005. A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology* **22**: 129-142. doi: 10.1093/imammb/dqi003.

Artois A, Blancou J, Dupeyroux O, Gilot-Fromont E. 2011. Sustainable control of zoonotic pathogens in wildlife: How to be fair to wild animals? *Revue scientifique et technique* **30**: 733-743.

Balloux F, Lugon-Moulin N.2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**: 155 – 165.

Biek R, Real LA. 2010. The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* **19**: 3515-3531. doi: 10.1111/j.1365-294X.2010.04679.x.

Blasco-Costa I, Poulin R. 2013. Host traits explain the genetic structure of parasites: A meta-analysis. *Parasitology* **140**: 1316 – 1322. doi: 10.1017/S0031182013000784

Brito D, Moreira DO, Coutinho BR, Oprea M. 2012. Ill nature: Disease hotspots as threats to biodiversity. *Journal for Nature Conservation* **20**: 72-75. doi:10.1016/j.jnc.2011.10.003.

Brownstein JS, Holford TR, Fish D. 2005. Effect of climate change on Lyme disease risk in North America. *EcoHealth* **2**: 38-46. doi: 10.1007/s10393-004-0139-x.

Brunner JL, LoGiudice K, Ostfeld RS. 2008. Estimating reservoir competence of *Borrelia burgdorferi* hosts: Prevalence and infectivity, sensitivity, and specificity.*Journal of Medical Entomology* **45**: 139-147. doi: http://dx.doi.org/10.1093/jmedent/45.1.139.

Coombs JA, Letcher BH, Nislow KH. 2010. PEDAGOG: Software for simulation eco-evolutionary population dynamics. *Molecular Ecology Resources* **10**: 558-563. doi: 10.1111/j.1755-0998.2009.02803.x.

Cushman SA, Landguth EL. 2010. Scale dependent inferences in landscape genetics. *Landscape Ecology* **25**: 967-979. doi: 10.1007/s10980-010-9467-0.

Ekblom R, Galindo J. 2011. Applications of next-generation sequencing in molecular ecology of non-model organisms. *Heredity* **107**: 1-15. doi: 10.1038/hdy.2010.152.

Epperson BK, Mcrae BH, Scribner K, Cushman SA, Rosenberg MS, Fortin M-J, James PMA, Murphy M, Manel S, Legendre P, Dale MRT. 2010. Utility of computer simulations in landscape genetics. *Molecular Ecology* **19**: 3549 – 3564. doi: 10.1111/j.1365-294X.2010.04678.x

Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease rick: an ecological primer. *Trends in Parasitology* **30**: 205-214. doi: 10.1016/j.pt.2014.02.00.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows.*Moleculra Ecology Resources* **10**: 564-567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier L, Smouse P, Quattro J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

Gilchrist MJR. 2002. Chapter 5. Surveillance and Management of zoonotic disease outbreaks: Public health laboratory surveillance. *In* The emergence of zoonotic diseases. *Edited by* T Burroughs, S Knobler, J Lederberg. National Academies Press, Washington DC, pp. 79-85.

Grégoir AF, Hablützel PI, Vanhove MPM, Pariselle A, Bamps J, Volckaert FAM, Raeymaekers JAM. 2015 A link between hos dispersal and parasite diversity in two sympatric cichlids of Lake Tanganyika. *Freshwater Biology* **60**: 323 – 335. doi: 10.1111/fwb.12492

Grenfell B, Harwood J. 1997. (Meta)population dynamics of infectious diseases. *Trends in Ecology and Evolution* **12**: 395-399.

Groemping U, Matthias L. 2013. relaimpo: Relative importance of regresses in linear models. R package vers. 2.2-2

Gubler DJ. 2010. The global threat of emergent/reemergent vector-borne diseases. In Vector Biology, Ecology and Control. Edited by P.W. Atkinson. Springer, Netherlands, pp. 39-62.

Hess G. 1996. Disease in metapopulation models: Implications for conservation. *Ecology* **77**: 1617-1632. doi: 10.2307/2265556.

Hijmans RJ, van Etten J, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, Bevan AB, Racine E, Shortridge A. 2015. raster: Geographic Data Analysis and Modeling. R package vers 2.3-33

Hudson ME. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Moldecular Ecology Resources* **8**: 3-17. doi: 10.1111/j.1471-8286.2007.02019.x.

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH. 2011. Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PLoS ONE* **6** : e25359. doi: 10.1371/journal.pone.0025359.

Johnson PTJ, de Roode JC, Fenton A. 2015. Why infectious disease research needs community ecology. *Science* **349**: 1259504-1 – 1259504-9. doi: 10.1126/science.1259504.

Jombart T, Collins C, Solymos P, Ahmed I, Calboli F, Cori A. 2014. adegenet: an R package for the exploratory analysis of genetic and genomic data. R package vers 1.4-2

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. *Nature* **451**: 990-993. doi:10.1038/nature06536.

Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy M, Luikart G. 2010. Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**: 4179-4191. doi: 10.1111/j.1365-294X.2010.04808.x.

Letcher BH, Coombs JA, Nislow KH. 2011. Maintenance of phenotypic variation: Repeatability, heritability and size-dependent processes in a wild brook trout population. *Evolutionary Applications* **4**: 602-615. doi: 10.1111/j.1752-4571.2011.00184.x.

LoGiudice K, Ostfeld RS, Schmidt KA, Keesing F. 2003. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. *Proceedings of the National Academy of Sciences* **100**: 567-571. doi: 10.1073/pnas.0233733100.

Manel S, Holderegger R. 2013. Ten years of landscape genetics. *Trends in Ecology and Evolution* **10**: 614-621. doi: 10.1016/j.tree.2013.05.012.

Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G. 2016. Host dispersal as the driver of parasite genetic structure: A paradigm lost? *Ecology Letter* **19**: 336 – 347. doi: 10.1111/ele.12564

McCallum H. 2008. Chapter Five: Landscape structure, disturbance, and disease dynamics. In Infectious disease ecology: Effects of ecosystems on disease and of disease on ecosystems . Edited by R.S. Ostfeld, F. Keesing, V.T. Eviner. Princeton University Press, Princeton, New Jersey, pp. 100-122.

Nadin-Davis S, Huang W, Armstrong J, Casey GA, Bahloul C, Tordo N, Wandeler AI. 2001. Antigenic and genetic divergence of rabies virus from bat species indigenous to Canada. *Virus Research* **74**: 139-156.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2015. vegan: Community ecology package. R package vers. 2.2-1

Ostfeld RO. 2011. Lyme disease: The ecology of a complex system. Oxford University Press. Oxford, United Kingdom.

Ostfeld RS, Keesing F. 2012. Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics* **43**: 157-182. doi: 10.1146/annurev-ecolsys-102710-145022.

Pérez-Figueroa A, Wallen RL, Antao T, Coombs JA, Schwartz MK, White PJ, Luikart G. 2012. Conserving genomic variability in large mammals: Effect of population fluctuations and variance in male reproductive success on variability in Yellowstone bison. *Biological Conservation* **150**: 159-166. doi:10.1016/j.biocon.2012.02.022.

Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D. 2015 MASS: Support Functions and Datasets for Venables and Ripley's MASS. R package vers 7.3-40

Roth D. 2011. Surveillance for emerging infectious diseases: A Canadian perspective. National Collaborating Centre for Environmental Health, Vancouver, BC, Canada.

Rupprecht CE, Halon CA, Hemachudha T. 2002. Rabies re-examined. *The Lancet Infectious Diseases* **2**: 327-343.

Rupprecht CE, Halon CA, Slate D. 2004. Oral vaccination of wildlife against rabies: Opportunities and challenges in prevention and control. *Developmental Biology* **119**: 173-184.

Simon JA, Marrotte RR, Desrosiers N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of *Borrelia burgdorferi*, the agent of Lyme disease, at the northern limit of its distribution. *Evolutionary Applications* **7**: 750-764. doi: 10.1111/eva.12165.

Slatkin M.1985. Gene flow in natural populations. *Annual Review of Ecology, Evolution, and Systematics* **16**: 393-430.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457 – 462.

Smith JS, Sumner JW, Roumillat LF, Baer GM, Winkler WG. 1984. Antigentic characteristics of isolates associated with a new epizootic of raccoon rabies in the United States. *Journal of Infectious Diseases* **149**: S785-S798.

Sork VL, Waits L. 2010. Contributions of landscape genetics–approaches, insights, and future potential. *Molecular Ecology* **19**: 3489-3495. doi: 10.1111/j.1365-294X.2010.04786.x.

Suzán G, García-Peña GE, Castro-Arellano I, Rico O, Rubio AV, Tolsa MJ, Roche B, Hosseini PR, Rizzoli A, Murray KA, Zambrana-Torrelio C, Vittecoq M, Bailly X, Aguirre AA, Daszak P, Prieur-Richard A-H, Mills JN, Guegan J-F. 2015. Metacommunity and phylogenetic structure determine wildlife and zoonotic infectious disease patterns in time and space. *Ecology and Evolution* **5**: 865-873. doi: 10.1002/ece3.1404.

Taylor LH, Latham SM, Woolhouse ME. 2001. Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society* **356**: 983-989.

van Etten J. 2015. gdistance: Distances and routes on geographical grids. R package vers. 1.1-7

Whiteley AR, Coombs JA, Hudy M, Robinson Z, Nislow KH, Letcher BH. 2012. Sampling strategies for estimating brook trout effective population size. *Conservation Genetics* **13**: 625-637. doi: 10.1007/s10592-011-0313-y.

Widmer I, Grande FD, Excoffier L, Holdregger R, Keller C, Mikryukov V, Scheidegger C. 2012. European phylogeography of the epiphytic lichen fungus *Lobaria pulmonaria* and its green algal symbiont. *Molecular Ecology* **21**: 5827 – 5844. doi: 10.1111/mec.12051

Wright S. 1950. Genetical structure of populations. *Nature* **166**: 247 – 249.

TABLES

Table 2.1. Description of simulated scenarios A to H, with details on the different landscape characteristics (heterogeneity and barriers) and disease dispersal patterns employed in each scenario. The numbers in subscript following Ps and Vs indicate which host species their dispersal is dependent on. Short-distance and long-distance dispersal are driven by H1 and H2, respectively.

Scenarios	**Landscape heterogeneity**	**Physical Barrier**	**Dispersal scenario**	**Species Simulated**
A	Homogeneous	✖	Short-distance	H1, H2, V_1, P_1
B	Heterogeneous	✖	Short-distance	H1, H2, V_1, P_1
C	Homogeneous	✓	Short-distance	H1, H2, V_1, P_1
D	Heterogeneous	✓	Short-distance	H1, H2, V_1, P_1
E	Homogeneous	✖	Long-distance	H1, H2, V_2, P_2
F	Heterogeneous	✖	Long-distance	H1, H2, V_2, P_2
G	Homogeneous	✓	Long-distance	H1, H2, V_2, P_2
H	Heterogeneous	✓	Long-distance	H1, H2, V_2, P_2

Table 2.2. Summary of results from Procrustes rotation tests that determined strength and significance of correlation between interspecific spatial genetic patterns. Associated species refer to pathogen, vector, and host species that were assigned shared spatial dependence and dispersal rates in each scenario. Average correlation coefficients and standard deviations calculated from all replicates are provided. The frequencies of statistically significant correlations are also included in the table. Simulation scenarios are labelled as in Table 2.1.

	Average correlation coefficients [SD]	**Frequency of statistically significant cases**
Scenario A		
Associated species	0.643 [0.179]	70.0%
Non-associated species	0.446 [0.156]	26.7%
Scenario B		
Associated species	0.791 [0.201]	73.3%
Non-associated species	0.656 [0.116]	60.0%
Scenario C		
Associated species	0.478 [0.210]	30.0%
Non-associated species	0.499 [0.147]	26.1%
Scenario D		
Associated species	0.716 [0.156]	50.0%
Non-associated species	0.624 [0.075]	50.0%
Scenario E		
Associated species	0.578 [0.151]	60.0%
Non-associated species	0.482 [0.141]	30.0%
Scenario F		
Associated species	0.688 [0.161]	76.7%
Non-associated species	0.618 [0.208]	56.7%
Scenario G		
Associated species	0.581 [0.137]	50.0%
Non-associated species	0.448 [0.157]	29.2%
Scenario H		
Associated species	0.647 [0.145]	56.7%
Non-associated species	0.599 [0.084]	50.0%

FIGURES

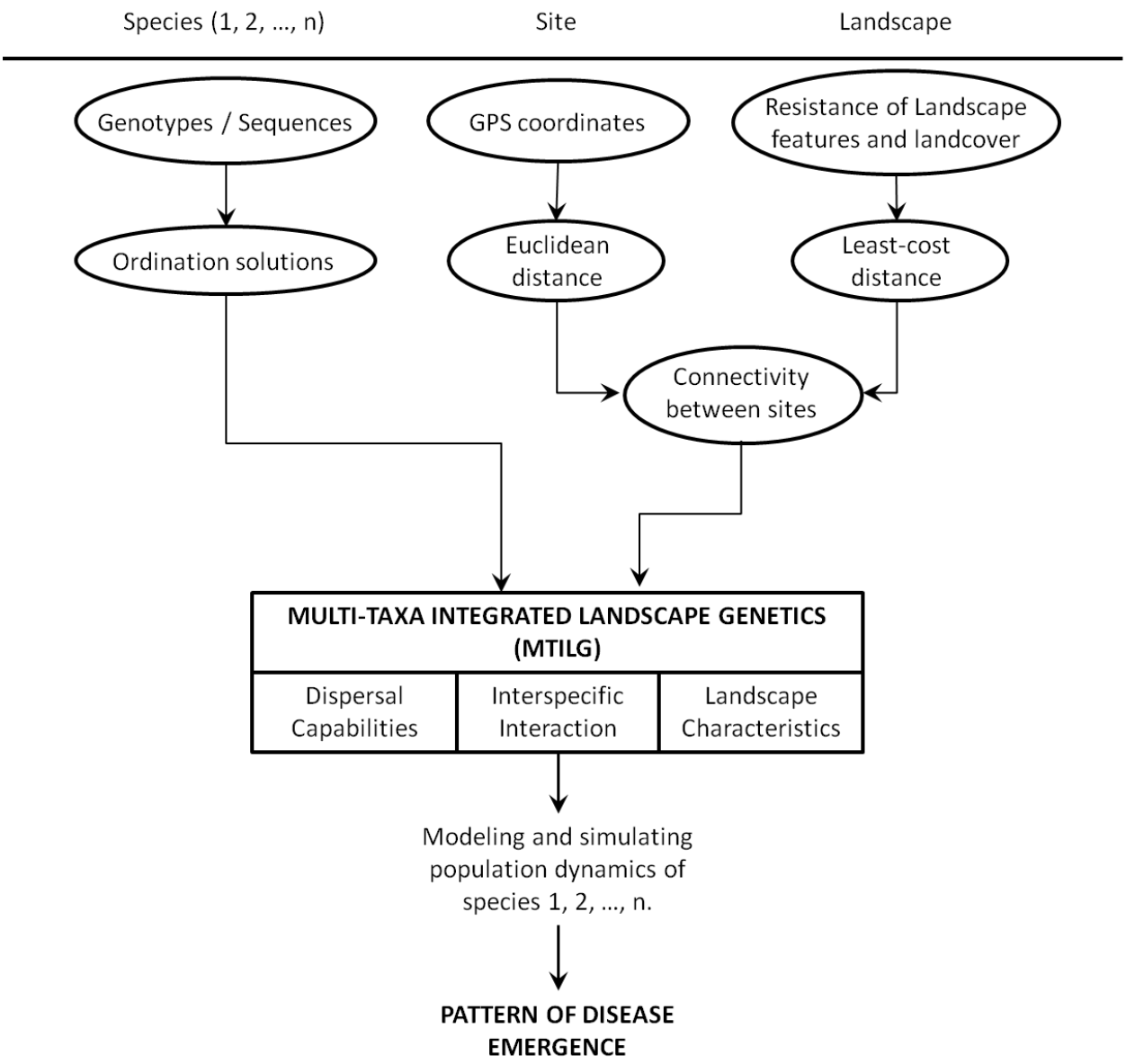


Figure 2.1. Visualisation of the multi-taxa integrated landscape genetic (MTILG) framework. I present examples of types of data that can be used (ovals) as well as how this data can be used for application of Multi-taxa Integrated Landscape Genetics (MTILG). Results and information from MTILG can then be incorporated into models and simulations that can help predict patterns (i.e. extent and rate) of disease emergence in a region.

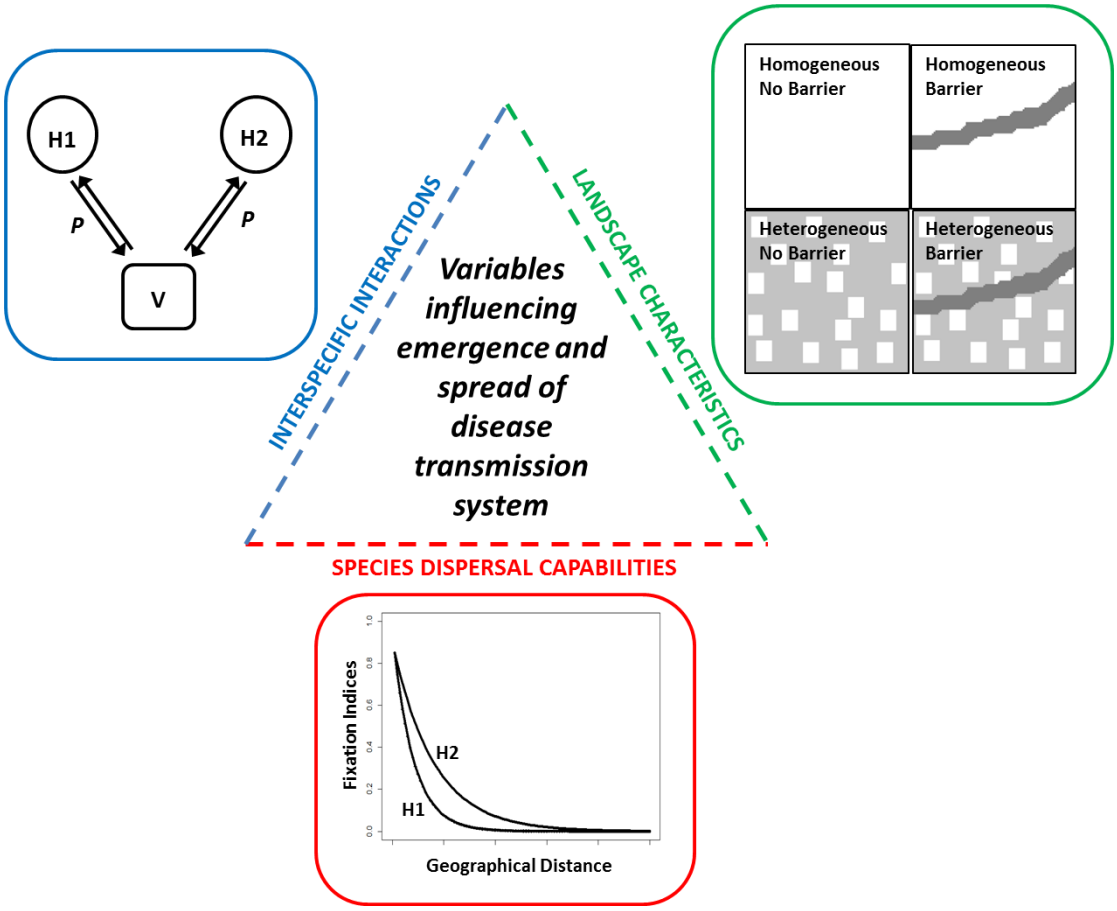


Figure 2.2. An integrative approach to examine how interspecific interactions, species dispersal rates, and landscape characteristics influence disease transmission system emergence in a region. H1: Short-distance dispersing host, H2: long-distance dispersing host, V: vector species, P: pathogen/disease-causing agent. Host dispersal probabilities were calculated using a negative exponential function and host movement across sites were calculated based on dispersal rates and resistance values assigned to landscape variables (Barrier=10, Suitable Habitat=1 Unsuitable habitat=5).

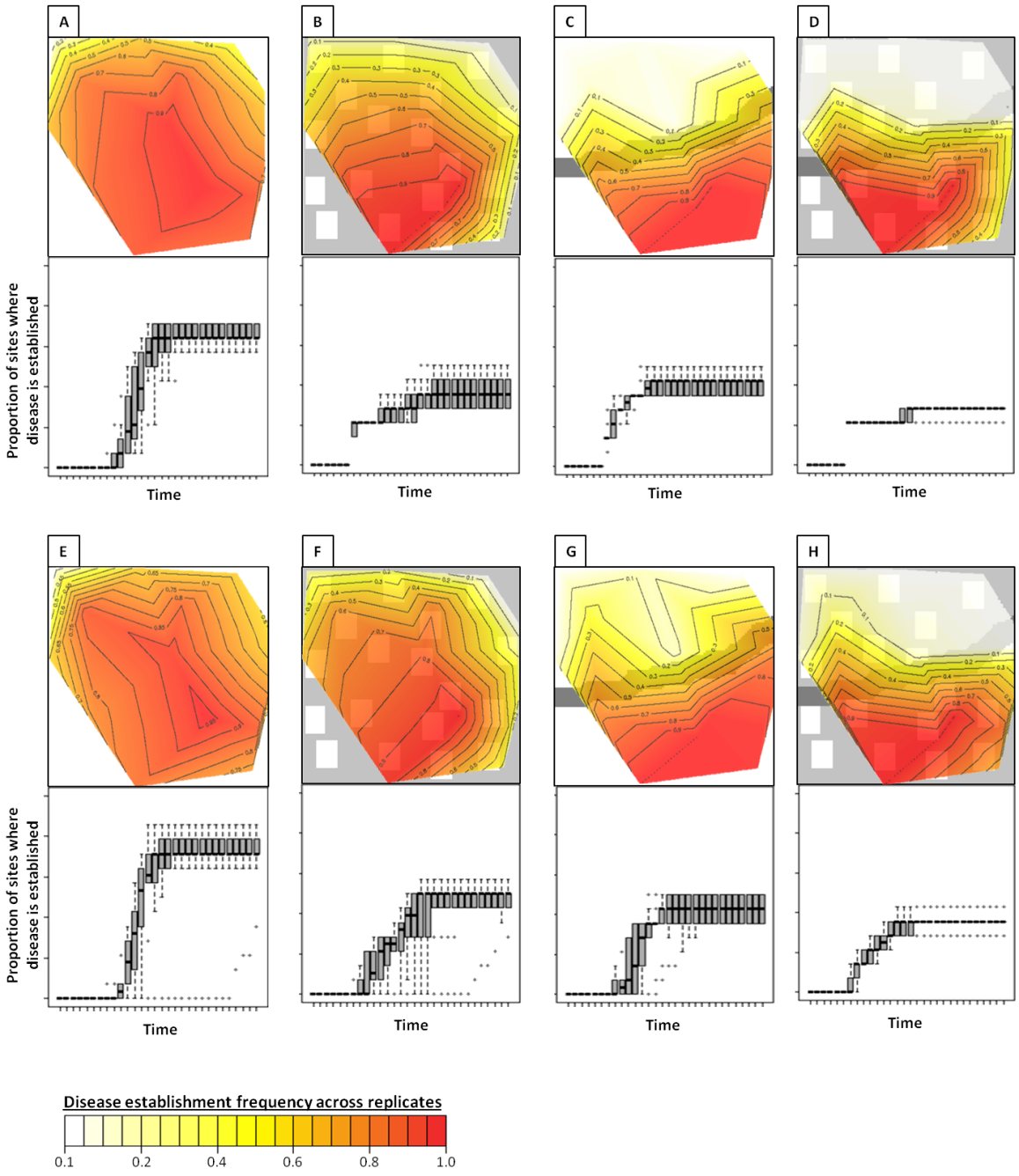


Figure 2.3. Extent and rate of disease system establishment across different landscapes and dispersal scenarios after 30 simulated generations. The contour lines and heat colours on the maps indicate frequencies with which the disease system had established successfully spatially over 10 simulation replicates. Warmer colours indicate greater frequencies of successful establishment. Plots below each map show proportion of sites in which the disease transmission system has established (y-axis) over time (x-axis) with variations across replicates accounted for. Simulation scenarios are labelled as in Table 2.1.

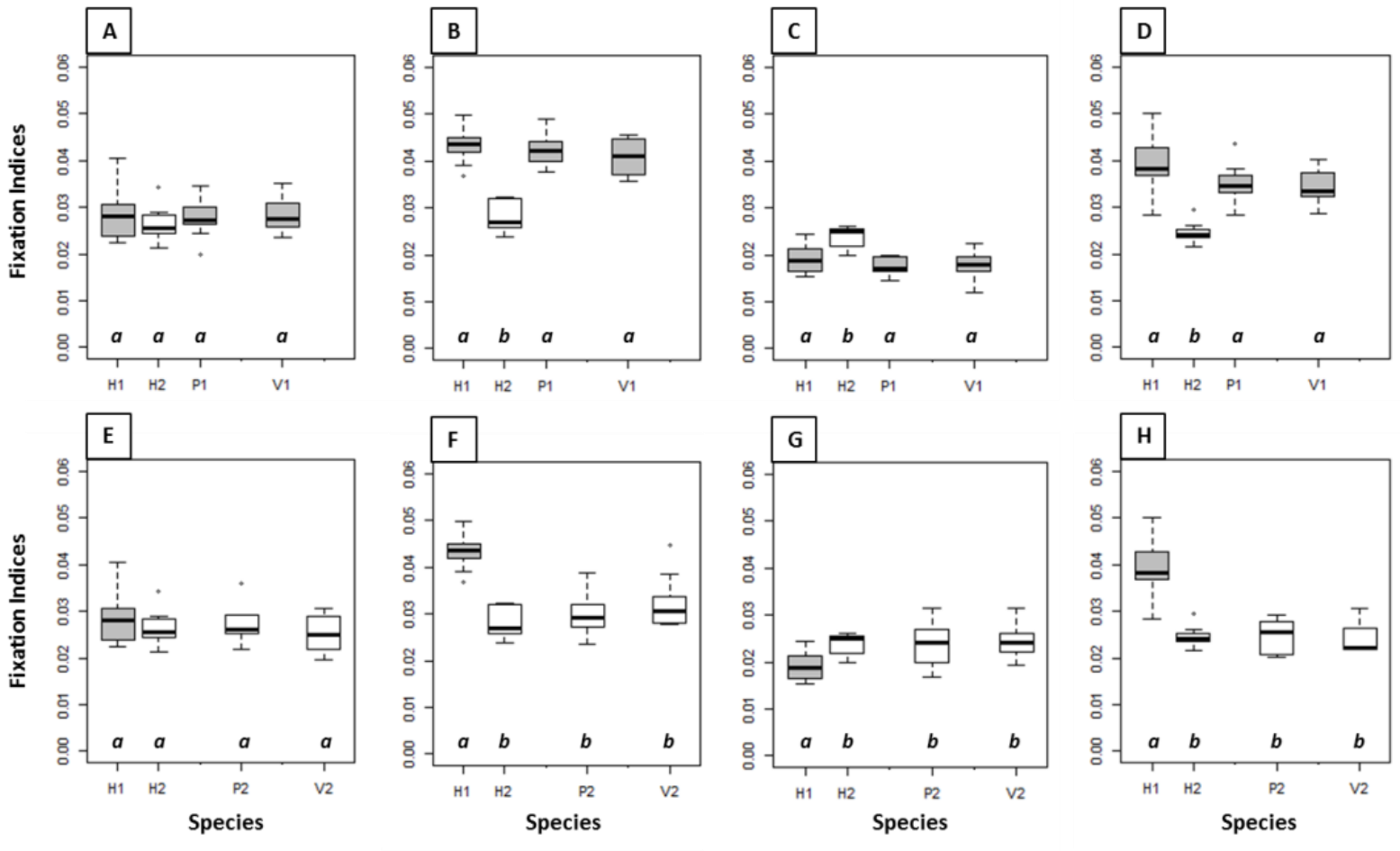


Figure 2.4. Comparisons of species-specific genetic differences between populations for each landscape scenario (columns) and dispersal mechanism (rows). Associated species share the same colours (e.g. pathogens and vectors dispersed by short distance dispersing host H1 are coloured grey). Significant differences in fixation indices values among species in each simulation are indicated with lower case ‘a’s and ‘b’s. Simulation scenarios are labelled as in Table 2.1.

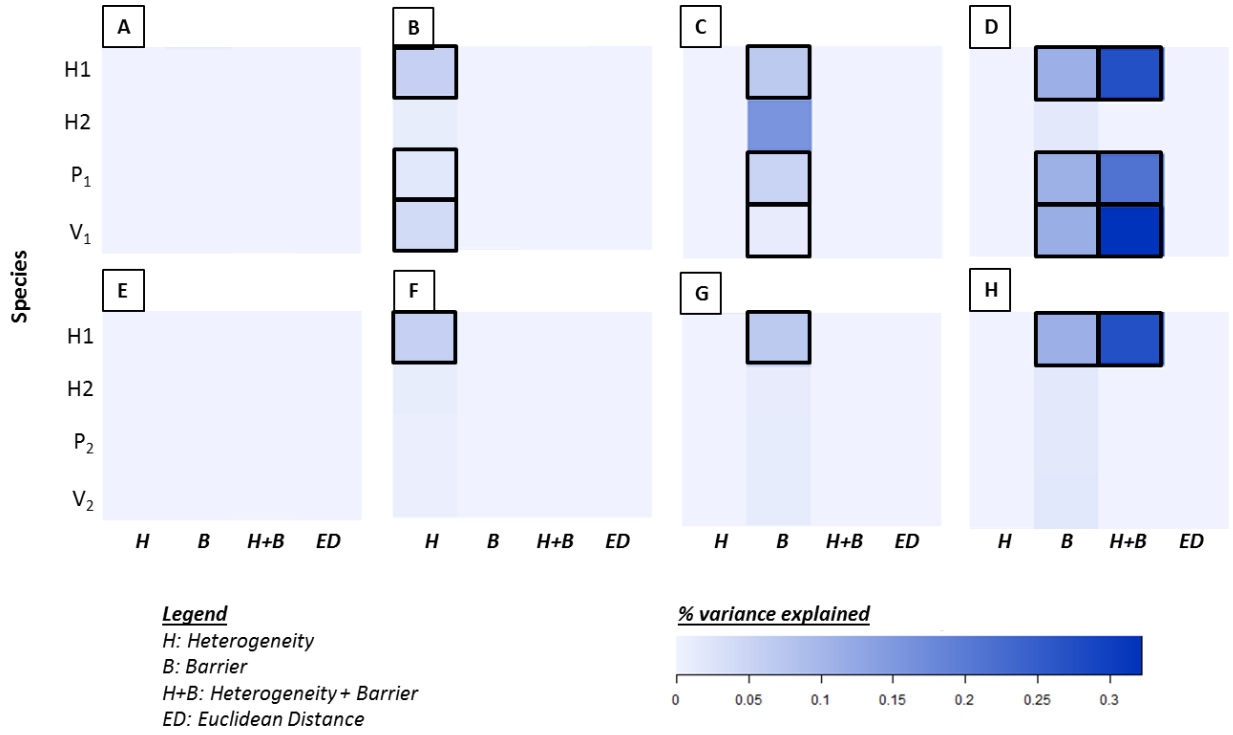


Figure 2.5. Species-specific landscape genetics results from redundancy analyses using allele frequencies as response matrices and landscape variables as predictor matrices. The relative amount of genetic variance accounted for by each landscape variable is given for each simulation scenario, with darker shading indicating greater percentage of variance explained. Boxes on the figures indicate predicted landscape variables that should exert observable effects on each species in each simulation scenario. Simulation scenarios are labelled as in Table 2.1.

CONNECTING STATEMENT

Multi-taxa integrated landscape genetics (MTILG) allows us to understand and identify ecological processes driving zoonotic disease emergence by simultaneously examining multiple key species in a disease system. In Chapter 2, I showed that the MTILG approach can reveal how interspecific interactions and landscape heterogeneity influence overall extent, rate, and pattern of disease spread across a landscape. In the subsequent chapters of this thesis, I begin to apply MTILG to the Lyme disease system in southern Québec. I first examine the host and vector species independently in Chapters 3 (white-footed mouse) and 4 (black-legged tick), before applying MTILG to both species in Chapter 5.

In Chapter 3, I focused first on white-footed mouse populations across southern Québec, examining their genetic diversity, population structures, morphological variation, tick burden, and pathogen prevalence. The goal is to obtain a better understanding of landscape effects on mouse movement, mouse morphological variation in the context of range expansion, and to identify the role the white-footed mice plays in the Lyme disease system.

CHAPTER THREE

Genetic and morphological variation in white-footed mouse populations at the leading edge of its expanding range: Implications for the emergence of Lyme disease

ABSTRACT

The northward range expansion of *Peromyscus leucopus* (white-footed mouse) into southern Québec is of considerable interest for public health as these rodents are excellent reservoir hosts of the Lyme disease pathogen *Borrelia burgdorferi*. A simultaneous examination of white-footed mouse genotypes, phenotypes, as well as pathogen prevalence and vector burdens, will help us better understand how the range expansion of this species may influence Lyme disease emergence at the northern edge of its range. White-footed mouse population genetics revealed that the effects of geographical barriers that limit gene flow (i.e. rivers) can be overcome, which suggests the potential for further range expansion. Comparisons of populations for mouse morphology and genetic diversity revealed some phenotypic variation among genetic clusters. While I found no significant differences in the measures of white-footed mouse genetic diversity along a latitudinal gradient, I observed a significant allometric increase in mouse hind-foot length as latitude increased. The latter may influence mouse dispersal rate and further range expansion and warrants further research. Finally, I found that the genotypic and phenotypic traits examined in this study were not associated with the mouse's susceptibility to infection or parasitism. However, I found indirect evidence suggesting that the white-footed mouse's role as a host of both the Lyme disease vector and pathogen is a major determinant of *B. burgdorferi* presence in a community. As such, future Lyme disease monitoring efforts in southern Québec

should account for white-footed mouse distribution in the region relative to landscape features that represent significant barriers to gene flow as well as potential adaptive traits that influence species' dispersal ability.

INTRODUCTION

Recent years have led to the rapid range expansion and emergence of several zoonotic diseases (Crowl et al. 2008, Engering et al. 2013). By examining how changes in species traits affect or reflect population fitness, survival, and dispersal ability, we can begin to predict population colonisation success which in turn influences patterns of disease emergence and spread (McCoy 2008, Tabachnick 2010, Hill et al. 2011). Similarly, surveillance of pathogen and parasite burdens on species undergoing range expansion can indirectly reveal rates and mechanisms of disease emergence in a region (Metz et al. 1983, Hadorn and Stärk 2008). By simultaneously examining phenotypic traits, spatial genetic structure, and pathogen and parasite burdens we can: (i) quantify species movement relative to landscape features (McCoy 2008, Manel and Holderegger 2013), (ii) identify phenotypic adaptations that may potentially increase rates of range expansion and disease emergence (Hill et al. 2011, King and Lively 2012), and (iii) characterise a species' role in disease transmission and emergence (Altizer and Pedersen 2008). When combined, this information can help predict future species distribution and identify possible avenues for controlling disease spread (McCoy 2008).

The white-footed mouse (*Peromyscus leucopus*) is one of the most abundant rodent species in North America. In recent years, the white-footed mouse range has expanded northwards at rates of approximately 10-15km/year (Myers et al. 2009, Roy-Dufresne et al. 2013, Simon et al. 2014). This is a public health concern in southern Québec as the white-footed mouse is widely recognised as the primary reservoir host of the Lyme disease pathogen *Borrelia burgdorferi* (Ostfeld 2011) and the Hantavirus (Plyusnin and Morzunov 2001). Indeed, in conjunction with the rapid range expansion of white-footed mouse in southern Québec, we have also observed a substantial increase in the abundance of the Lyme disease vector, the black-

legged tick (*Ixodes scapularis*) (Ogden et al. 2006), and incidences of human Lyme disease cases (Ogden et al. 2009, Institut National de Santé Publique Québec 2014).

Studies on the genetics of white-footed mouse populations in southern Québec have provided some insight on how the landscape influences mouse gene flow in the region. Specifically, Rogic et al. (2013) and Marrotte et al. (2014) found that while landscape heterogeneity such as agricultural land use posed little resistance to mouse movement, urban environments, rivers, and highways significantly determined mouse population genetic structure. While these results are useful for potential disease monitoring and control, both studies were performed in a small geographical area measuring approximately 633.5 km^2 located in the center of the current white-footed mouse distribution in Québec. Therefore, these results may not be applicable at broader spatial scales, as effects of landscape variables do not translate equally across geographical scales (Meentemeyer and Box 1987).

Another study on white-footed mouse in southern Québec had revealed surprisingly high concordance in mouse genetics and cranial morphology (Ledevin and Millien 2013). This concordance in inter-population genetic and morphological variation may be due to ongoing range expansion by founder effects from independent colonization of forest patches, or to local adaptation of white-footed mice to distinct habitat conditions (Simon et al. 2014). However, the cranial traits examined in Ledevin and Millien (2013) do not lend themselves well to direct tests of hypotheses about range expansion pattern.

Finally, while a study by Simon et al. (2014) revealed a strong association between the presence of *B. burgdorferi* and the occurrence of white-footed mice in a region, it is unknown if variation in mouse morphology or genetics might further influence *B. burgdorferi* prevalence and rates of black-legged tick parasitism locally. Ticks can develop host-associated genetic and

phenotypic differentiation (e.g. Dietrich et al. 2013) that arise either due to selection or drift from isolation. Similarly, *B. burgdorferi* strain types and virulence have been shown to undergo selection for particular tick phenology or host composition in a habitat (Ogden et al. 2008). It is likely therefore that local adaptation in the white-footed mouse may influence parasitism or pathogen infection success and rate. Correlations between white-footed mouse traits and measures of parasite and pathogen infection could reveal specialization or ongoing co-evolutionary processes that may modulate Lyme disease emergence in southern Québec.

While previous studies have provided valuable insights on the range expansion of white-footed mouse into southern Québec, none of them have integrated white-footed mouse genetics, morphology, *B. burgdorferi* prevalence, and measures of tick presence into a single comprehensive study. Here, I attempted to do just that to improve our ability to understand and predict Lyme disease risk and emergence patterns as driven by the white-footed mouse in southern Québec. My objectives were four-fold:

1) I evaluated how landscape heterogeneity affected neutral white-footed mouse gene flow across its current known distribution in Québec and its role on the ongoing range expansion northward into the rest of the province.
2) I compared observed patterns of variation between white-footed mouse genetics and external morphology to identify potential local adaptations within local populations.
3) I tested for spatial trends in measures of genetic diversity and morphological traits variation in white-footed mice along a latitudinal gradient.
4) Finally, I explored relationships among mouse morphology, genetic differentiation, *B. burgdorferi* infection rate, and black-legged tick burden. Significant correlations

among these variables are expected to influence the contribution of the white-footed mouse to Lyme disease emergence in Québec.

MATERIAL AND METHODS

Study sites and specimens

I sampled 485 individuals of white-footed mouse from 16 sites across an area of southern Québec measuring approximately 140 km by 110 km (15,400km^2) (Figure 3.1, Table B1 in Appendix B). Specimens were live-trapped with Sherman traps set out 5-10 meters apart in grids of 4x7 or 4x10. All captured specimens were euthanized the morning after capture and dissected on the same day. Harvested tissue samples (e.g. heart, liver, kidneys) were stored in 90% ethanol and stored in a -20°C freezer. DNA was extracted for all specimens from liver tissue using a standard 3-day phenol/chloroform extraction procedure as described by Sambrook et al. (1989).

Microsatellite genotyping

I genotyped the white-footed mouse specimens using eleven polymorphic microsatellite markers (Table B2). Polymerase chain reactions (PCR) were performed in 15 µL reactions containing 8.6 µL 10X PCR buffer, 0.2 mM of each dNTPs, 0.5 ρmol/µL of each primer, 0.05 U/µL of Taq polymerase and 2.5 µL DNA template. All PCR thermocycling conditions included an initial denaturation step at 95°C for 30 s, 30 cycles of 95°C for 15 s, loci-specific primer annealing temperature for 30 s, and 72°C for 45 s. PCR amplified products were visualized on an ABI3730 capillary sequencer (Applied Biosystems, Foster City, CA) at Genome Québec (Montreal, CA) and genotyped using the program GeneMarker (SoftGenetics LLC., State

College, PA) with GeneScan 500-LIZ (Applied Biosystems, Foster City, CA) as the size standard.

Analyses on genetic variation

I carried out individual-based clustering analyses in the software *structure* ver. 2.3 (Pritchard et al. 2000) using the 'no admixture' and 'independent allele frequency' models. I first analyzed each possible number of populations (k = 1 - 10) for ten iterations each before calculating the most likely number of genetic clusters (Evanno et al. 2005). A more thorough analysis that runs for 500,000 Markov chain Monte Carlo (MCMC) generations (burn-in = 10%) was then performed with the most likely 'k' value obtained.

I then assessed the presence and amount of genetic structuring among and within each genetic cluster identified by *structure* using Analyses of Molecular Variances (AMOVA), in Arlequin ver. 3.5 (Excoffier and Lischer 2010). AMOVA was performed based on the sum of squared differences (i.e. Slatkin's R-statistics (Slatkin 1995)) over 1000 permutations.

Landscape genetics analyses

I conducted isolation-by-distance (IBD) and isolation-by-resistance (IBR) analyses using Mantel Tests over 999 permutations. Pairwise genetic distances between sites were calculated using the *pairwise.fst* function in the R package *adegenet* (Jombart et al. 2014). For IBD analyses, I calculated pairwise geographical distances (as the crow flies) between sites using the *pointDistance* function from the R package *raster* (Hijmans et al. 2015). To obtain the resistance distance matrix for my IBR analysis, I first created a raster plot of the study area in which different landscape features were assigned pre-determined resistance values (Marrotte et al.

2014). I then calculated pairwise least-cost resistance distances among the sites using the *costDistance* function from the R package *gdistance* (van Etten 2015).

In addition to tests of dissimilarity among distance matrices, I also examined relations between mouse spatial genetic patterns and specific landscape features (Table B3) using redundancy analyses (RDA). RDA is a form of constrained ordination that investigates the extent to which a set of predictors explains variance in a set of observed variables. Analysis of Variance (ANOVA) was performed on the constrained ordination model to examine the effect of individual landscape variables on spatial genetic patterns. Using log-transformed allele frequencies as response matrices and measures of spatial connectivity as predictor variables in the RDA, I identified ordination models that best describe genetic similarities relative to landscape variables. Spatial environmental data (Table B3) were first converted into a site-specific index by calculating the connectivity of each site relative to all other sites in the metapopulation network using published resistances values (Marrotte et al. 2014) and least-cost effective distances for each landscape variable. Connectivity was calculated using a general ecological connectivity metric (James et al. 2011):

$$S_i = \sum_{1}^{j} \frac{1}{1 + \beta d_{ij}},$$

where S_i is the connectivity of site i to the rest of the meta-population network (βd_{ij}), d_{ij} the least-cost effect distance between sites *i* and *j* based on resistance values, and β the estimated mean dispersal rate of the species of interest.

Measurements of morphological traits

Standard external morphological traits measured on the mouse specimens were body mass (g), head-and-body length (mm), hind foot length (mm), tail length (mm), and ear length (mm). Missing data (e.g. broken tails) were estimated via linear regression using head-body length as a predictor variable. All raw measurements were log-transformed prior to analyses. I calculated two additional measures of white-footed mouse morphology: ratio of head-body length to tail length, and body condition. Body condition was estimated by residuals from a regression of log-transformed body mass (g) against log-transformed head-and-body length (mm). Juvenile and pregnant individuals were excluded from all morphological analyses.

Comparing morphological and genetic variation

Using a series of ANOVA analyses, I investigated variation in mouse morphology among genetic clusters and sub-clusters previously identified in *structure*. I then compared observed population-level morphological variation against population-level genetic variation to evaluate the congruence of pairwise morphological and genetic differences among my sampled populations. Inter-population pairwise genetic and morphological distances were calculated and compared using a Mantel test, and co-phenetic correlations between UPGMA trees produced from the distance matrices were quantified using the R package *dendextend* (Galili et al. 2015).

Latitudinal gradients in white-footed mouse morphology and genetics

I surveyed for gradients in white-footed mouse genetic diversity and morphological traits using a series of Pearson's product moment correlation coefficient tests. All correlation tests were performed using the *cor.test* function in R. Measures of genetic diversity in each population included the average number of alleles, allele richness, as well as observed and expected heterogeneity calculated from allele frequencies. Finally, I looked for changes in the effective population sizes of each white-footed mouse population sampled. Effective population sizes

were estimated from mouse genetic data using the software NeEstimator ver.2 (Do et al. 2014). Effective populations were estimated using the linkage disequilibrium, heterozygote excess, and molecular coancestry methods with a random mating model. Lowest allele frequency allowed for analyses was 0.05. Results from each model were averaged to estimate consensus mouse effective population sizes. For white-footed mouse external morphology, I tested for allometric changes by taking the residuals of the measurements (head-body length, tail length, ear length, hind-foot length, head-body: tail length ratio, and body condition) regressed against individual geometric size and comparing them against latitudinal data from collection records. I also examined changes in the variance of the measurements of white-footed mouse external morphology from south to north to indirectly identify the presence of directional selection on specific mouse phenotype.

Estimating Borrelia burgdorferi *presence at each site and prevalence in white-footed mice*

Only white-footed mouse specimens captured from 2011 onwards (N=389) were tested for *B. burgdorferi* infection at the National Microbiology Laboratory of the Public Health Agency (Winnipeg, MB, Canada) using DNA extracted from heart tissues. Mice specimens were considered infected if they tested positive for two *B. burgdorferi* genetic markers: the 23S ribosomal RNA (rRNA), and the *ospA* gene (Ogden et al. 2011). *Borrelia burgdorferi* prevalence in each mouse population was calculated as the proportion of mice that tested positive over the total number of mice tested from each population. I also tested black-legged ticks and other small mammal species sampled from the same sites for *B. burgdorferi* infection. Positive screening results from any of the species tested were taken as indicative of the pathogen presence at the site.

Measures of tick burden on white-footed mouse

All mouse specimens were thoroughly examined for ectoparasites and number of black-legged ticks collected from each specimen was recorded. For each population, measures of black-legged tick burden on white-footed mice were represented as tick prevalence (proportion of mice parasitized out of all examined), tick intensity (mean number of ticks found on parasitized mice only), and tick abundance (mean number of ticks found on all mice examined).

Relations between B. burgdorferi *prevalence, tick burden, and mouse genetics and morphology*

I performed a series of partial RDA analyses to hierarchically investigate relations among *B. burgdorferi* presence at a site, *B. burgdorferi* prevalence in white-footed mouse, tick burden on host species, and mouse genetics and morphology. Since my variables may be confounded by site proximity, I specified site coordinates (latitude and longitude) as conditional matrices in the analyses. All partial RDAs were performed using the *rda* function from R package *vegan* (Oksanen et al. 2015).

In the first and second analyses, I independently determined how white-footed mouse traits (genetics and morphology) influenced black-legged tick burden and *B. burgdorferi* prevalence at the population level. Population-level white-footed mouse genetic variation was estimated by extracting scores of the first axis obtained from a spatial principal components analysis (sPCA) on log-transformed allele frequencies, using the *spca* function from the R package *adegenet* (Jombart et al. 2014). I chose to employ sPCA as it allowed me to investigate allele frequency distribution while accounting for spatial correlation among sites. Meanwhile, population-level measures of white-footed morphological variation were determined based on scores extracted from the first axis of a between-group (i.e. population) PCA analysis (gpPCA) on log-

transformed morphological data. gpPCA was performed using the *dudi.pca* and *bca* functions from the R package *ade4* (Dray et al. 2015).

In the third RDA, I examined how the combination of white-footed mouse species traits, and vector burden determined observed *B. burgdorferi* prevalence among white-footed mouse population.

RESULTS

Genetic variation among white-footed mouse populations

Preliminary AMOVA analyses on white-footed mouse genetics revealed no significant temporal variation among mice collected between 2007 and 2014 (R_{CT} = 0.000, p-value = 0.954, % variance = 0.000).

Structure analysis sorted the white-footed mouse individuals into two distinct genetic clusters (Ln likelihood = -24735.06, ΔK = 267.636). The first genetic cluster (Cluster 1) consisted primarily of individuals from populations on the northern shore of the St. Lawrence River, as well as individuals from Longueuil Parc (LP) (Figure 3.2A). The other genetic cluster (Cluster 2) consisted solely of individuals collected from populations on the southern shore of the St. Lawrence River (Figure 3.2A). Variance partitioning analysis in locus-by-locus AMOVA revealed that these genetic clusters accounted for 10.42% of the overall observed genetic variation (R_{CT} = 0.104, p-value = 0.001). However, there appeared to be further structuring within genetic clusters (R_{SC} = 0.116, p-value < 0.001, % variance = 10.42%). Individuals from populations assigned to Cluster 1 in the initial analysis could be further divided into two sub-clusters (Ln likelihood = -5189.48, ΔK = 42.004) (Figure 3.2B). Individuals from populations

assigned to Cluster 2 formed three distinct genetic sub-clusters (Ln likelihood = -5189.48, ΔK = 42.004) (Figure 3.2C). These sub-clusters explained 4.00 % and 2.98 % of observed genetic variation within Clusters 1 (R_{CT} = 0.040, p-value = 0.021) and 2 (R_{CT} = 0.030, p-value = 0.000), respectively.

Landscape genetics on white-footed mouse populations

There was a positive but not statistically significant correlation between genetic distance and geographical distance between mouse populations (r = 0.248, p = 0.076, R^2_{adj} = 0.0536), but a significant positive correlation between genetic distance and landscape connectivity based on resistance values (r = 0.302, p-value = 0.049, R^2_{adj} = 0.0597). RDA revealed that all measures of site connectivity based on resistance values (see Marrotte et al. 2014) assigned to different landscape variables examined significantly explained approximately 13.2% (F = 1.457, p = 0.0225) of genetic variation observed amongst white-footed mouse populations (Table 3.1). However, only water bodies significantly accounted for variation in allele frequencies (F = 1.947, p = 0.018, R^2_{adj} = 0.062) (Table 3.1).

Morphological vs. genetic variation in white-footed mouse

Overall, white-footed mouse from Clusters 1 and 2 differed in their morphology (p < 0.001). Specifically, ear length was longer in mice from Cluster 2 (F = 13.14, p < 0.001), while body condition was generally poorer in mice assigned to Cluster 1 (F = 4.97, p = 0.027) (Table 3.2). Within Cluster 1, white-footed mice assigned to different sub-clusters also exhibited significant morphological differences (p = 0.001). Individuals in sub-cluster 1 (red in Figure 3.2B) had longer tails, shorter ears, lower head-body: tail length ratios, and better body conditions (Table 3.2). There was no significant morphological difference among individuals from different sub-clusters in Cluster 2 (Table 3.2). Direct comparisons revealed no significant

correlation between mouse genetics and morphology (Mantel r = 0.006, p = 0.516). Similarly, UPGMA trees estimated from both datasets were not significantly correlated with each other (co-phenetic correlation = - 0.099, p = 0.690).

Latitudinal gradients in white-footed mouse morphology and genetics

Correlation tests revealed no significant changes in measures of overall genetic diversity among white-footed mouse populations on a south-north gradient (Figure B1 in Appendix B). Estimates of effective population sizes calculated from genetic data revealed an overall decreasing trend on a south-north gradient, although the final correlations were not statistically significant (Figure 3.4). I identified a statistically significant allometric increase in white-footed mouse hind-foot length (Figure 3.3A). This increase in hind-foot length was accompanied by a significant decrease in foot length variance among mouse populations on the southern shore of the St. Lawrence River (Figure 3.3B). I also found a significant increase in the variance of head-body: tail length ratio among mouse populations on the northern shore of the river. No other statistically significant latitude-associated gradients were identified in the remaining morphological traits (Figures B2 and B3).

Borrelia burgdorferi *prevalence and black-legged tick burden on white-footed mice*

Borrelia burgdorferi was detected in 4 out of 15 white-footed mouse populations (26.7%) tested, with prevalence in white-footed mice ranging from 0.056 to 0.133 at each site (Table B4). Eleven out of 389 white-footed mice tested positive for *B. burgdorferi* (*B. burgdorferi* prevalence = 0.0282). Of the 16 sites from which small mammals, black-legged ticks, and white-footed mice were screened, *B. burgdorferi* was present in 8 sites.

Black-legged ticks were found in 11 out of 15 sites (73.3%) for which white-footed mice were examined for ectoparasites (Table B4). Overall, 28% of white-footed mice in this study were found to carry black-legged ticks. An average of 0.835 ticks were found on all white-footed mice examined (ranging from 0 to 33 ticks per individual), while the mean number of ticks reported on infected mice only was 2.981. Site-specific black-legged tick prevalence, intensity, and abundance on white footed mice are summarised in Table B4.

Relations among B. burgdorferi *prevalence, tick burdens, and mouse genetics and morphology*

The first and second partial RDAs revealed that neither white-footed mouse genetic nor morphological variation explained observed tick burden or *B. burgdorferi* prevalence on white-footed mice in this study.

However, using RDA, I found that *B. burgdorferi* prevalence in white-footed mice was significantly correlated to the proportion of white-footed mice found carrying ticks (i.e. tick prevalence in mice) ($F = 35.0175$, $p = 0.002$) and the mean number of ticks found on each mouse individual (i.e. tick abundance on mice) ($F = 5.6985$, $p = 0.037$).

DISCUSSION

I investigated population-level relations among white-footed mouse genetics, morphology, tick burden, and *B. burgdorferi* prevalence to improve our understanding of the species' rapid range expansion in southern Québec. Significant correlations among these variables reveal factors mediating the rate and pattern of Lyme disease emergence in the region.

Genetic variation of white-footed mouse in southern Québec

As white-footed mice have limited dispersal ability (Goundie and Vessey 1986), I expected to find significant spatial genetic structure among mouse populations. I found two distinct genetic clusters that were for the most part associated with northern and southern shores of the St Lawrence River. This pattern is similar to that found in a phylogeographic study performed on white-footed mouse mitochondrial DNA (Fiset et al. 2015) and strongly supports the hypothesis that current white-footed mouse populations in southern Québec originated from two refugia lineages during the Last Glacial Maximum. Furthermore, the Saint Lawrence River serves as a barrier which kept both lineages from panmixia. Isolation-by-resistance and landscape genetics analyses revealed that landscape heterogeneity significantly influenced observed spatial genetic patterns of white-footed mouse populations in the region. Specifically, water bodies accounted for the observed spatial genetic patterns among white-footed mouse populations in southern Québec (Rogic et al. 2013, Marrotte et al. 2014). Unlike results from studies by Rogic et al. (2013) and Marrotte et al. (2014) however, I found that roads were not effective barriers against mouse gene flow at the spatial scale of my study (Meentemeyer and Box 1987).

An interesting finding from this study is that the barrier effect of the St. Lawrence River on mouse gene flow was not absolute, as is evident by the unexpected genetic similarity between white-footed mice collected from Longueuil Parc, on the south shore, and those sampled on the northern shore of the St. Lawrence River. Genetic similarity was also observed among white-footed mice sampled from both shores of the much smaller Richelieu River. Several studies have shown that white-footed mice are proficient swimmers and may swim for distances up to 40 meters (Sheppe 1965). However, it remains highly unlikely that a small species like the white-footed mouse could cross the St. Lawrence River, which has an average stream channel width of

over 2700 meters (Downing et al. 2012). Population connectivity facilitated by human-mediated movement (e.g. major tunnels and bridges across the St. Lawrence River) is a more likely explanation for the unexpected genetic similarity in some populations on either side of the river. Such human-mediated movement of organisms within or beyond their past or present range and dispersal potential can break natural distribution boundaries. This has important implications on the potential for continued range expansion of the white-footed mouse in the southern Québec. However, the rate and final extent of white-footed mouse colonisation of novel habitats in the region remains limited by local habitat suitability and environmental conditions (Roy-Dufresne et al. 2013). Given the important role white-footed mice play as reservoir hosts for the Lyme disease pathogen, the rate and pattern of ongoing northward range expansion of this species has important implications for Lyme disease emergence in southern Québec (Ostfeld 2011).

In addition to results from the population and landscape genetics analyses on white-footed mouse populations, this study did not reveal any significant changes in genetic diversity or effective population size in the direction of range expansion. Generally, in cases of species range expansion, one expect to observe patterns of decreasing genetic diversity and population size with increasing distance from the species distribution core (Excoffier et al. 2009). The lack of such a pattern in the mouse populations was unexpected, but could be explained by several factors. First of all, my current study area was small relative to the range of the white-footed mouse across North America. Once again spatial scale may be a factor influencing the results of this study and I note that any future attempts to study changes in genetic diversity resulting from founder effect or genetic drift among mouse populations in this region will have to be conducted over a greater spatial scale. Second, it has been shown that populations at the leading-edge of species range have the potential to rapidly restore their genetic diversity through the combined

effects of long-distance gene flow and inbreeding avoidance (e.g. Hampe et al. 2013). We know that the white-footed mouse tends to avoid breeding with closely related individuals as doing so can result in substantial inbreeding depression (i.e. small litter size and lower offspring weight) (Keane 1990a, 1990b, Jimenez et al. 1994). Furthermore, the results from this study suggest that the white-footed mouse may have the potential to travel great distances if human-mediated dispersal is at play.

Morphological variation of white-footed mouse in southern Québec

I found evidence of minor morphological differentiation among white-footed mice associated with different genetic clusters, as well as variation in certain white-footed mouse morphological traits along a latitudinal gradient, which may be the results of directional selection. Although some of the results in this study were inconclusive due to the large amount of variance in the correlations, I nonetheless consider that morphological variation in white-footed mice warrants more research attention, particularly in the context of disease emergence. Species experiencing range expansion often exhibit changes in life history traits that increase colonization success and dispersal ability (Hill et al. 2011). In the case of the white-footed mouse, increased hind-foot length due to rapid adaptive change or spatial selection can potentially improve dispersal ability (e.g. Philips et al. 2006, Sparrow 2015), which can in turn accelerate species range expansion and influence rate of disease spread. Similarly, changes in body conditions among populations may not only indicate a change in dispersing individual's potential to establish and persist in novel habitats, but also provide insight regarding the individual's susceptibility to pathogen infection or large parasite burden (e.g. Schmidt et al. 1999).

When I compared variation in mouse genetics and morphology, I expected to observe inter-population morphological variation that corresponded with that observed in the mouse genetics (Haldane 1930). However, I found no such correspondence at the population-level. This result runs counter to what was previously described by Ledevin and Millien (2013), where strong concordance was observed between white-foot mouse skull morphological differentiation and genetic structure among populations. While Ledevin and Millien (2013) focused on subtle variation in white-footed mouse skull morphology, I examined only broad external morphological traits for which subtle signals of phenotypic variation may not be easily detectable. Nonetheless, the poor concordance between white-footed mouse genetic and morphological variation in this study may also be explained by metapopulation dynamics. Generally, one expects to observe correlations between species gene flow patterns and population adaptation to local conditions, as the former plays an important role introducing or reducing genetic variation in a population (Haldane 1930, Gillespie 1975, Takahata 1983). However, the combined effects of migration and environmental heterogeneity in a metapopulation context can produce mosaics of varying degrees of local adaptation across a region (Kaltz and Shykoff 1998). Strong drift-migration dynamics as well as local habitat quality can also interfere with local selection processes (Kaltz and Shykoff 1998).

Role of white-footed mouse on Lyme disease prevalence

The Lyme disease transmission system comprises two distinct types of hosts: competent reservoirs that facilitate disease pathogen persistence and transmission (Brunner et al. 2008), and host species that are important for vector development, survival, and reproduction. In eastern North America, the white-footed mouse is widely recognised as the most competent reservoir host for Lyme disease and it is also the host species most frequently parasitized by juvenile

black-legged ticks (Ostfeld 2011). Through its role as a host for both disease vector and disease-causing agent, the white-footed mouse plays an important role in determining Lyme disease presence in a region. I found that as tick burden on white-footed mouse increased, so did *B. burgdorferi* prevalence in said mouse populations. This is unsurprising, since black-legged ticks are primary vectors of the *Borrelia* pathogen, and increased encounter rate between hosts and infected vectors should lead to an increase in pathogen transmission probability (Ostfeld 2011).

White-footed mouse species traits (genotypes and morphology) examined in this study appeared to have little effect on the species' ability to host both the Lyme disease vector and disease-causing agent. *Borrelia burgdorferi* and black-legged tick infections are reported to have little effect on the fitness, immunological, and behavioural responses of white-footed mice (Ostfeld et al. 1996, Schwanz et al. 2011), suggesting that the pathogen and ticks exert little selective pressures on the white-footed mouse. However, a more targeted approach that sequences or genotypes candidate genes commonly associated with white-footed mouse immunological responses or morphological species traits that promote or inhibit species susceptibility to infections or parasitism might reveal otherwise. Additionally, since this study took place at the northern-most limit of the white-footed mouse species range, disease dynamics may differ from that in established areas. Studies that compare white-footed mouse populations from both the core region in the United-States and the edges of the range in southern Québec will produce a more comprehensive picture of pathogen-vector-host interactions that will help inform disease emergence surveillance strategies.

CONCLUSION

In this study, I found that white-footed mouse populations in southern Québec exhibit significant genetic differentiation and gene flow affected by physical barriers. However, it appeared that these barriers may potentially be circumvented by human activities. This implies that the combination of human activities and the white-footed mouse's natural ability to navigate a heterogeneous landscape can facilitate the continued and rapid northward range expansion of the species. Additionally, the observable trend in a functional trait associated with species dispersal capability (i.e. hind-foot length) may have implications for future white-footed mouse dispersal rate and merits further research. Finally, I found that white-footed mouse genetic and morphological traits examined in this study do not exhibit concordant variation, nor do they appear to be associated with rates of tick parasitism and pathogen infection in the species (although I also note that the markers employed in this study are assumed to be neutral and therefore not under selection). Rather, *B. burgdorferi* presence in this region of active white-footed mouse movement appeared to be indirectly determined by the rodent's role as hosts for both disease agent and vector. Lyme disease monitoring strategies and projection models should therefore account for white-footed mouse distribution and predictive modelling of its future range, keeping in mind that changes in functional traits may influence a species' ability to navigate a landscape.

REFERENCES

Altizer S, Pedersen AB. 2008. Host–pathogen evolution, biodiversity and disease risks for natural populations. *Conservation biology: Evolution in action*: 259-277.

Brunner JL, LoGiudice K, Ostfeld RS. 2008. Estimating reservoir competence of *Borrelia burgdorferi* hosts: Prevalence and infectivity, sensitivity, and specificity. *Journal of Medical Entomology* **45**: 139 – 147.

Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment* **6**: 238 – 246.

Dietrich M, Beati L, Elguero E, Boulinier T, McCoy KD. 2013. Body size and shape evolution in host races of the tick *Ixodes uriae*. *Biological Journal of the Linnean Society* **108**: 323 – 334.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* **14**: 209 – 214.

Downing JA, Cole JJ, Duarte CM, Middelburg JJ, Melack JM, Prairie YT, Kortelainen P, Striegl RG, McDowell WH, Tranvik LJ. 2012. Global abundance and size distribution of streams and rivers. *Inland Waters* **2**: 229 – 236.

Dray S, Dufour A-B, Thioulouse J. 2015. ade4: analysis of ecological data – exploratory and Euclidean methods in environmental sciences. R package vers. 1.7-2.

Engering A, Hogerwerf L, Slingenbergh J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections* **2**: e5.

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, Foll M, Peti RJ. 2009. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics* **40**: 481 – 501. doi: 10.1146/annurev.ecolsys.39.110707.173414

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

Fiset J, Tessier N, Millien V, Lapointe F-J. 2015. Phylogeographic structure of the white-footed mouse and the deer mouse, two Lyme disease reservoir hosts in Québec. *PLoS ONE* **10**: e0144112.

Galili T, Simpson G, Jefferis G, Gallotta M, Renaudie J, Hornik K, Ligges U, Spiess A-N, Horvath S, Langfelder P, Van Der Loo M, de Vries A, Benjamini Y. 2015. dendextend: Extending R's Dendrogram Functionality. R package vers 0.18.3.

Gillespie JH. 1975. The role of migration in the genetic structure of populations in temporarily and spatially varying environments. I. Conditions for polymorphism. *The American Naturalist* **109**: 127 – 136.

Goundie TR, Vessey SH. 1986. Survival and dispersal of young white-footed mice born in nest boxes. *Journal of Mammalogy* **67**: 53 – 60.

Hadorn DC, Stärk KDC. 2008. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Veterinary Research* **39**: 57 – 68.

Haldane JBS. 1930. A mathematical theory of natural and artificial selection. VI. Isolation. *Proceedings of the Cambridge* Philosophical *Society* **26**: 220–230.

Hampe A, Pemonge M-H, Petit RJ. 2013. Efficient mitigation of founder effects during the establishment of a leading-edge oak population. *Proceedings of the Royal Society B* **280**: 20131070. doi: 10.1098/rspb.2013.1070

Hijmans RJ, Cheng J, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, Bevan A, Racine EB, Shortridge A. 2015. raster: Geographic data analysis and modeling. R package vers. 2.4-15.

Hill JK, Griffiths HM, Thomas CD. 2011: Climate change and evolutionary adaptations at species' range margins. *Annual Review of Entomology* **56**: 143 – 159.

Institut National de Santé Publique du Québec. 2014. Proposition d'un programme de surveillance intégré pour la maladie de Lyme et les autres maladies transmises par la tique *Ixodes scapularis* au Québec.

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH. 2011. Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PloS ONE* **6**: e25359.

Jombart T, Collins C, Solymos P, Ahmed I, Calboli F, Cori A. 2014. adegenet: an R package for the exploratory analysis of genetic and genomic data. R package vers 1.4-2

Jimenez JA, Hughes KA, Alaks G, GrahamL , Lacy RC. 1994. An experimental study of inbreeding depression in a natural habitat. *Science* **266**: 271-273. doi: 10.1126/science.7939661

Kaltz O, Shykoff J. 1998. Local adaptation in host-parasite systems. *Heredity* **81**: 361 – 370.

Keane B. 1990a. The effect of relatedness on reproductive success and mate choice in the white-footed mouse, *Peromyscus leucopus*. *Animal Behaviour* **39**: 264–273.

Keane B. 1990b. Dispersal and inbreeding avoidance in the white-footed mouse, *Peromyscus leucopus*. *Animal Behaviour* **40**: 143–152.

King KC, Lively CM. 2012. Does genetic diversity limit disease spread in natural host populations. *Heredity* **109**: 199 – 203. (doi:10.1038/hdy.2012.33)

Ledevin R, Millien V. 2013. Congruent morphological and genetic differentiatioin as a signature of range expansion in a fragmented landscape. *Ecology and Evolution* **3**: 4172 – 4182.

Manel S, Holderegger R. 2013. Ten years of landscape genetics. *Trends in Ecology and Evolution* **28**: 614 – 621.

Marrotte R, Gonzalez A, Millien V. 2014. Landscape heterogeneity and habitat quality combine to provide and optimal model of connectivity and estimate gene flow at the range margin of a small mammal. *Molecular Ecology* **23**: 3983 – 3998.

McCoy KD. 2008. The population genetic structure of vectors and our understanding of disease epidemiology. *Parasite* **15**: 444 – 448.

Meentemeyer V, Box EO. 1987. Scale effects in landscape studies. In *Landscape heterogeneity and disturbance*. Springer New York. pp. 15 – 34.

Metz JAJ, Wedel M, Angulo AF. 1983. Discovering an epidemic before it has reached a certain level of prevalence. *Biometrics* **39**: 765 – 770.

Myers P, Lundrigan BL, Hoffman SMG, Haraminac AP, Seto SH. 2009. Climate-induced changes in the small mammal communities of the Northern Great Lakes Region. G*lobal Change Biology* **15**: 1434 – 1454.

Ogden NH, Bigras-Poulin M, Hanincová K, Maarouf A, O'Callaghan CJ, Kurtenbach K. 2008. Projected effects of climate change on tick phenology and fitness of pathogens transmitted by the North American tick *Ixodes scapularis*. *Journal of Theoretical Biology* **254**: 621 – 632.

Ogden NH, Lindsay LR, Morshed M, Sockett PN, Artsob H. 2009. The emergence of Lyme disease in Canada. *Canadian Medical Association Journal* **180**: 1221 – 1224.

Ogden NH, Margos G, Aanensen DM, Drebot MA, Feil EJ, Hanincova K, Schwartz I, Tyler S, Lindsay LR. 2011. Investigation of genotypes of *Borrelia burgdorferi* in *Ixodes scapularis* ticks collected during surveillance in Canada. *Applied and Environmental Microbiology* **77**: 3244 – 3254.

Ogden NH, Trudel L, Artsob H, Barker IK, Beauchamp G, Charron DF, Drebot MA, Galloway TD, O'Handley R, Thompson RA, Lindsay LR. 2006. *Ixodes scapularis* ticks collected by passive surveillance in Canada: Analysis of geographic distribution and infection with Lyme borreliosis agent *Borrelia burgdorferi*. *Journal of Medical Entomology* **43**: 600 – 609.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2015. vegan: Community ecology package. R package vers. 2.2-1

Ostfeld RS. 2011.Lyme disease: The ecology of a complex system. Oxford University Press, Oxford, United Kingdom..

Ostfeld RS, Miller MC, Hazler KR. 1996. Causes and consequences of tick (*Ixodes scapularis*) burdens on white-footed mice (*Peromyscus leucopus*). *Journal of Mammalogy* **77**: 266 – 273.

Philips BL, Brown GP, Webb JK, Shine R. 2006. Invasion and the evolution of speed in toads. *Nature* **439**: 803.

Plyunin A, Morzunov SP. 2001. Virus evolution and genetic diversity of hantaviruses and their rodent hosts. In *Hantaviruses*. Springer Berlin Heidelberg, pp. 47 – 75.

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

Rogic A, Tessier N, Legendre P, Lapointe FJ, Millien V. 2013. Genetic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Québec. *Ecology and Evolution* **3**: 4172 – 4182.

Roy-Dufresne E, Logan T, Simon JA, Chmura GL, Millien V. 2013. Poleward expansion of the white-footed mouse (*Peromyscus leucopus*) under climate change: implications for the spread of Lyme disease. *PLoS ONE* **8**: e80724.

Sambrook J, Fritsch EF, Maniatis T. 1989. Chapter 6 – Preparation and analysis of eukaryotic genomic DNA. In: *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schmidt KA, Ostfeld RS, Schauber EM. 1999. Infestiation of *Peromyscus leucopus* and *Tamias straitus* by *Ixodes scapularis* (Acari: Ixodidae) in relation to the abundance of host and parasites. *Journal of Medical Entomology* **36**: 749 – 757.

Schwanz LE, Voordouw MJ, Brisson D, Ostfeld RS. 2011. *Borrelia brugdorferi* has minimal impact on the Lyme disease reservoir host *Peromyscus leucopus*. *Vector-borne and Zoonotic Diseases* **11**: 117 – 124.

Sheppe E. 1965. Dispersal by swimming in *Peromyscus leucopus*. *Journal of Mammalogy* **46**: 336 – 337.

Simon JA, Marrotte RR, Desrosier N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton P, Lindsay R, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Milllien V. 2014. Climate change and habitat fragmentation drive the occurrence of *B. burgdorferi*, the agent of Lyme disease, at the northern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457 – 462

Sparrow L. 2015. The effect of limb length on the cost of locomotion and gait in the Longshanks mouse. M.Sc. Thesis. University of Calgary.

Tabachnick WJ. 2010. Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. *The Journal of Experimental Biology* **213**: 946 – 954.

Takahata N. 1983. Gene identity and genetic differentiation of populations in the finite island model. *Genetics* **104**: 497 – 512.

van Etten J. 2015. gdistance: Distances and Routes on Geographical Grids. R package vers 1.1-9

TABLES

Table 3.1. Summary of results from performing ANOVA on redundancy analyses (RDA) investigating the influence of spatial heterogeneity on white-footed mouse genetic variation. Connectivity among sites was estimated from resistance values (Marrotte et al. 2014) assigned to each landscape variable. Statistically significant results ($p < 0.05$) are in bold.

Analysis	Landscape variable	F	p	R^2	R^2_{adj}	Resistance value assigned
RDA	***All***	***1.457***	***0.0225***	***0.421***	***0.132***	
	Roads	1.497	0.0751	0.087	0.021	2.12
	Water bodies	**1.947**	**0.0180**	**0.125**	**0.062**	**3.01**
	Agriculture	1.468	0.0831	0.072	0.006	0.49
	Urban / Developed	1.052	0.3924	0.139	0.078	4.00
	Forest	1.321	0.1502	0.091	0.026	0.00

Table 3.2. Summary of white-footed mouse morphological variation relative to genetic clusters. Genetic clusters were assigned to individuals based on results from *structure.* Additional analyses on sub-clusters within clusters were also performed. N: number of individuals. Statistically significant results ($p < 0.05$) are in bold.

	Global Structure Cluster 1 vs 2 (N = 280)		Cluster 1 Sub-clusters 1 vs 2 (N = 60)		Cluster 2 Sub-clusters 1 vs 2 vs 3 (N = 220)	
	F	p	F	p	F	p
HB	2.559	0.111	0.156	0.694	0.084	0.773
Tail	0.260	0.610	**7.634**	**0.008**	0.420	0.518
Foot	0.034	0.853	1.764	0.190	2.037	0.155
Ear	**13.14**	**0.000**	**6.348**	**0.015**	0.242	0.632
HB:Tail	0.593	0.442	**6.400**	**0.014**	0.353	0.553
Body Condition	**4.975**	**0.027**	**5.023**	**0.029**	2.364	0.126

FIGURES

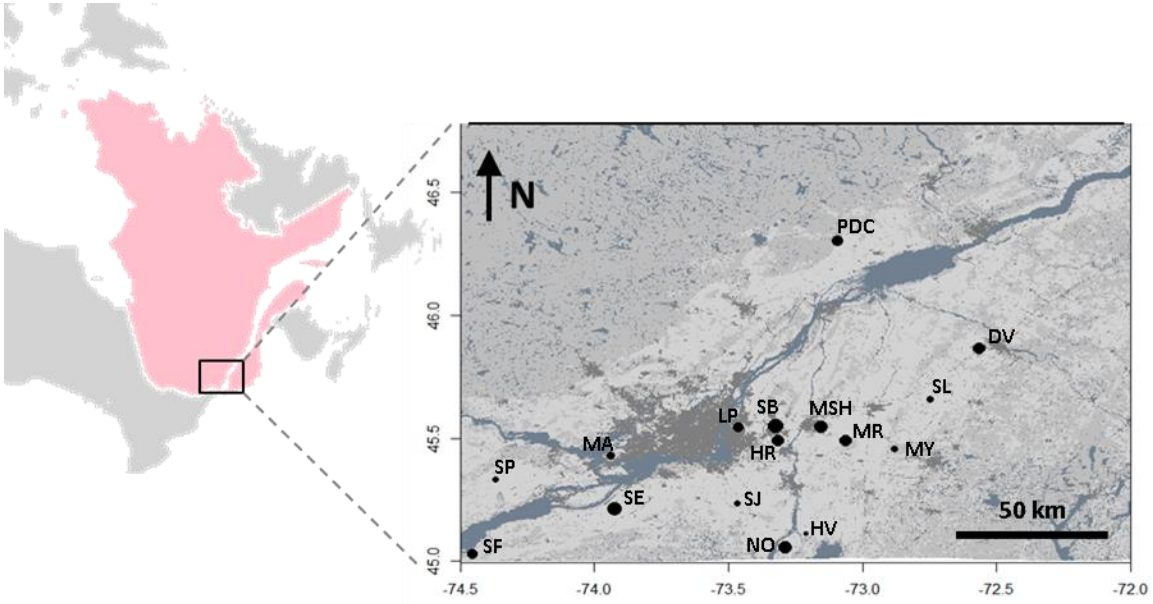


Figure 3.1. Study area and sixteen white-footed mouse sampling sites in southern Quebec, Canada. Details of each sampling locality are provided in Table B1.

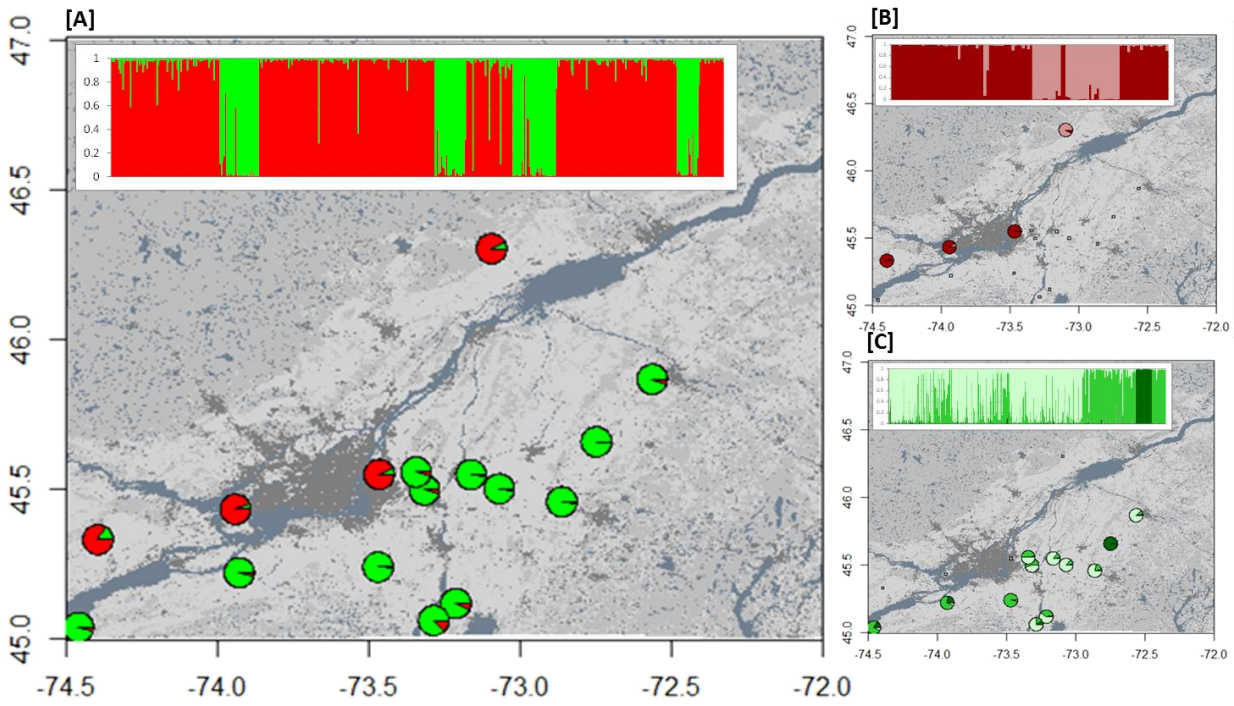


Figure 3.2. Results from *Structure*'s clustering analysis on [A] all 16 white-footed mouse populations; [B] 4 populations that were assigned to Cluster 1 (red in A); and [C] 12 populations that were assigned to Cluster 2 (green in A). Pie charts on each map show proportion of membership of each white-footed mouse population in each cluster. Bar-plots indicating probability of individual assignment are provided as inserts in each map.

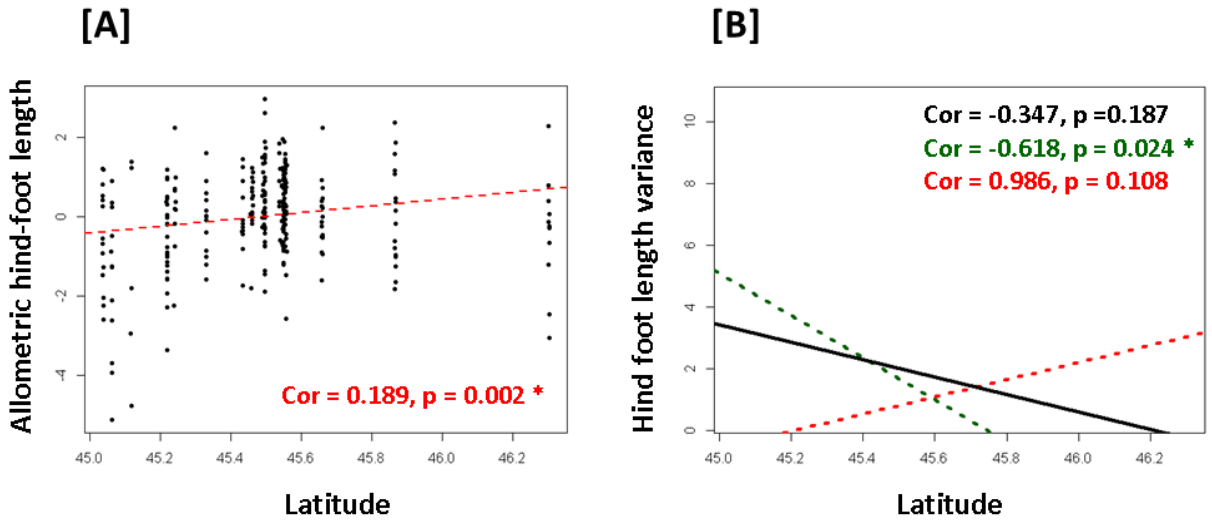
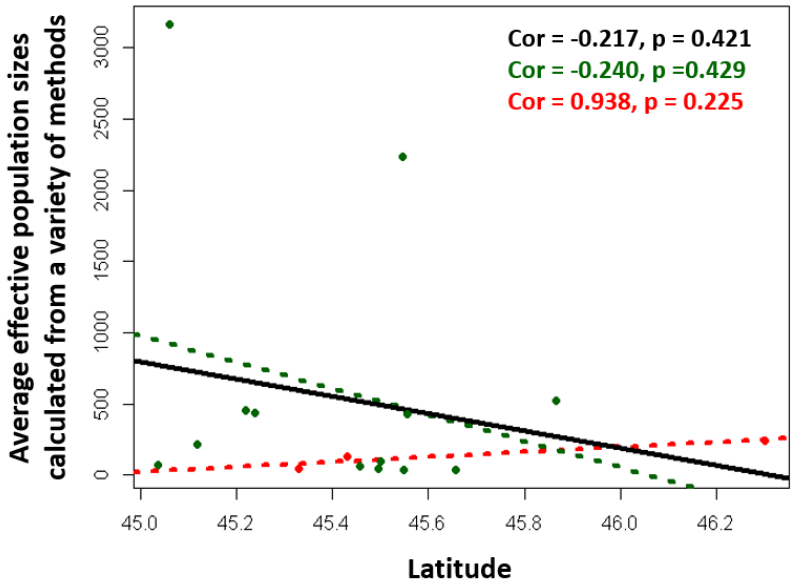


Figure 3.3. Comparisons of *Peromyscus leucopus* [A] allometric hind-foot length and [B] amount of variance observed in hind-foot length measurements against latitude. The coloured lines in plot [B] represents results for all populations (black), populations on the northern shores of the St. Lawrence River (red), and populations on the southern shore (green).



Colour legend
Black = All populations
Green = Populations on the south shore
Red = Populations on the north shore

Figure 3.4. Effective population sizes estimated from genetic data for each *Peromyscus leucopus* population over a latitudinal gradient.

CONNECTING STATEMENT

In Chapter 3, I examined the spatial genetic diversity of the white-footed mouse across my study area. I found significant genetic differentiation among white-footed mouse populations in Québec, and that gene flow is affected by linear barriers to migration. I also found increasing hind-foot length in the direction of the species range expansion, which has potential consequences for mouse dispersal ability. Finally, a comparison of mouse traits, tick burden, and *B. burgdorferi* infection prevalence revealed that the white-footed mouse's role as host for both ticks and pathogen play a critical role in Lyme disease emergence.

Armed with a better understanding regarding how white-footed mouse movement, species traits and morphology potentially contributes to the pattern of Lyme disease emergence in Québec, my focus in Chapter 4 now turns to the black-legged tick. In the following chapter, I study the spatio-temporal genetic diversity of tick populations in the region. Information from this study has allowed me to empirically identify the primary mode of tick movement into southern Québec (i.e. long vs. short distance dispersal). At the same time, information on genetic turnover rates within a tick population allowed me to identify stages of tick population establishment. This information will be useful for projecting regions of greatest Lyme disease infection risk in Chapter 5.

CHAPTER FOUR

Genetic signature of range expansion in a disease vector and the pattern of Lyme disease emergence.

ABSTRACT

Monitoring and predicting the spread of an emerging infectious disease requires that I understand the mechanisms of range expansion amongst its vectors. Here, I examined spatial and temporal variation of genetic structure among populations of the Lyme disease vector, the black-legged tick, in southern Quebec, where this tick species is currently expanding and Lyme disease is emerging. I found multiple genetic signatures of frequent long-distance dispersal events, supporting the hypothesis that migratory birds are the primary transport of black-legged ticks into southern Quebec. I also found that temporal genetic variation among the tick samples greatly exceeds that of spatial variation, indicating that while cohorts of black-legged ticks can rapidly invade large areas across southern Quebec, they are also undergoing frequent turnover. Finally, I was able to determine stages of establishment and identify established versus adventitious tick populations. My results allow us to better understand the seasonality of Lyme disease infection risk and help public health officials to predict the future spread and risk of emerging zoonotic disease systems in general.

INTRODUCTION

Species range expansion is occurring globally across a wide range of taxa and is driven by climate and land use changes (Chen et al. 2011, Dawe et al. 2014). Range expansion is causing the rapid emergence of a number of zoonotic diseases (Crowl et al. 2008, Engering et al. 2013) because new pathogen, arthropod vector, or vertebrate host species are being introduced into naïve regions (e.g. Legér et al. 2013, Fuller et al. 2012). Species range expansion produces observable genetic signals, such as lower genetic diversity and greater genetic structuring among populations at range edges (Excoffier et al. 2009, Cristescu 2015). We can thus apply concepts from population genetics to identify and quantify the degree of establishment of invasive pathogen, vector, or host populations that facilitate disease emergence and rapid spread (Engering et al. 2013, Estrada-Pena et al. 2014).

The number of reported Lyme disease cases in Canada has been increasing rapidly in recent years (Ogden et al. 2009), a phenomenon largely attributed to the range expansion of the Lyme disease vector, the black-legged tick, *Ixodes scapularis*, into the country (Ogden et al. 2008). The number of established black-legged tick populations in Canada has increased substantially from one at Long Point, Ontario in the 1990s (Ogden et al. 2009), to numerous populations now found across five Canadian provinces (Ogden et al. 2014). Despite the threat black-legged ticks present for Lyme disease exposure risk, attempts have been seldom made to examine the genetic diversity of black-legged ticks in Canada. Research performed to date focused on active or passive surveillance, and predictive modelling of future tick distributions (e.g. Ogden et al 2006b, Leighton et al. 2012). To my knowledge, only a few studies have examined black-legged tick genetic diversity in Canada using the 16S ribosomal RNA mitochondrial marker (Krakowetz et al. 2011, 2014) or the mitochondrial cytochrome C cox1

marker (Mechai et al. 2013). Although these studies assess tick genetic diversity, population structure, and potential source populations, these genetic markers do no reflect contemporary gene flow among tick populations.

Here, I applied population genetics using a set of fast-evolving, selectively neutral microsatellite markers to investigate the range expansion of black-legged ticks into southern Quebec. Black-legged ticks are unlikely to travel great geographical distances on their own (Falco and Fish 1991), because their movement across a landscape is dependent on their hosts (Ogden et al. 2006a). Migratory birds may contribute to the long-distance movement of ticks from USA into Canada during their annual migrations (Ogden et al. 2008, Ogden et al. 2015), while terrestrial hosts play additional roles in the expansion of tick populations at the edge of their current distribution range (Khatchikian et al. 2015). Theoretical population genetics predicts that long-distance and short-distance dispersal produce distinct and observable patterns of allele frequency distributions across the landscape (patchy allele distributions versus a gradient of allele frequency in the direction of spread, respectively) (Excoffier et al 2009, Bialozyt et al. 2006). The frequency of dispersal may be also apparent in the pattern of genetic diversity observed across the landscape (Bialozyt et al. 2006). If the introduction of ticks into Quebec was driven primarily by hosts that exhibit frequent long-distance dispersal behaviour (i.e. migratory birds), I expect tick populations in southern Quebec to present patchy allele distributions, low population structuring, and consistently high genetic diversity across the region (Bialozyt et al. 2006).

Temporal variation in allele frequency is also the net result of colonisation and extinction in a population. While abundance in the population may remain stable in time, genetic diversity may vary and populations at earlier stages of establishment exhibit greater genetic turnover rate compared to well-established populations (Figure 4.1). As new individuals are introduced into a population, I expect to observe influxes of new genetic material. However, the relative amount of novel genetic material relative to that already present locally will differ depending on the establishment stage (Figure 4.1). Such differences in temporal genetic differentiation within each population can provide insights on population establishment stages. One would expect to observe great genetic turnover rates within populations undergoing recolonisation-extinction cycles (Figure 4.1A), moderate turnover rates in populations at the colonisation stage (Figure 4.1B), and low genetic turnovers in well-established populations (Figure 4.1C). Based on these hypotheses, I attempted to quantify rates of genetic turnovers in tick populations for each of my sampling sites to estimate the degree to which these populations are established across the landscape. This is something that to my knowledge has never been attempted on black-legged ticks. In order to test the validity of my hypothesised relation between the amount of genetic turnover and tick population establishment stage, I compared my estimations against results from a study by Ogden et al. (2008b) which estimated the certainty for the occurrence of established tick populations based on surveillance survey in the region.

MATERIAL AND METHODS

Study sites and data collection

I sampled and analysed data from 613 black-legged ticks of all three life stages (larva, nymph, and adult) from ten sites across a study area in southern Quebec spanning approximately

2080 km^2 between 2011 and 2014 (Figure C1). Ticks were collected by dragging through vegetation and from small mammals captured using Sherman traps (Table C1). The sampling for this study occurred primarily in the summer months of July and August. As a result, over 85% of my specimens were larvae or nymphs (Table C1: larvae = 21% of specimens, nymphs = 65% of specimens, adults = 14% of specimens). Given that nymphs and adults are frequently sampled near the end of each sampling season, I am relatively confident there is a high probability that individuals sampled in the same year belong to the same cohort. In the event multiple larvae were sampled from a single drag or host individual, I used only one specimen in my genetic analyses to remove the potential confounding effect of kinship or cryptic population structuring. A set of ten microsatellite loci developed by Fagerberg et al. (2001) for black-legged ticks were used to genotype all individuals (Table C2). A subset of randomly selected specimens (10 per sampling site) was genotyped twice to ensure consistent alleles calling.

Genetic analyses

I performed Analyses of Molecular Variances (AMOVA) to examine the genetic differentiation amongst individuals sorted by sampling localities using an analog of Slatkin's R-statistics (Slatkin 1995) in Arlequin ver. 3.5 (Excoffier and Lischer 2010). I also conducted isolation-by-distance analyses via Mantel Tests over 999 permutations to investigate overall genetic variation relative to geographical distance. Pairwise Euclidean geographical distances and genetic distances between sites were calculated using the *pointDistance* function from the R package vegan (Oksanen et al. 2015) and the *pairwise.fst* function in the R package adegenet (Jombart et al. 2014), respectively. I further correlated measures of genetic diversity (i.e. expected heterozygosity, observed heterozygosity, and number of allele) against a latitudinal gradient based on sampling locality GPS coordinates. Measures of genetic diversity were

estimated using the Excel Microsatllite Toolkit (Park 2011). I identified spatial patterns of allele frequency distributions among tick populations using both individual-based clustering analyses and spatial principal components analyses (sPCA), a multivariate method to identify both local and global spatial genetic patterns. The former analysis was performed using *STRUCTURE* ver. 2.3 using the no admixture and independent allele frequency models (Pritchard et al. 2000). The latter was performed using the *spca* function from the R package adegenet, using the Delaunay triangulation approach to estimate a connection network among tick individuals.

I then sorted the specimens according to sampling sites and sampling years and removed samples for which I had only one year of sampling or fewer than 12 individuals in a given year to examine temporal genetic variation (Table C1). I performed principal components analyses (PCA) at individual and population levels using raw genotype data and log-transformed allele frequency respectively using the *dudi.pca* function from the R package ade4 (Dray et al. 2015). Next, I performed hierarchical AMOVA in Arlequin to quantify genetic variation among samples to test the effect of sampling site nested within sampling time, as well as the effect of sampling time nested within sampling site. This allowed us to estimate the respective effects of time and geographic location on genetic variation. I also performed individual within-sample (either within year or within site) AMOVA analyses (Table C1).

I then examined genetic differentiation across years for specimens sampled within each site using pairwise genetic distances calculated in Arlequin from the AMOVA analyses.

RESULTS

Spatial genetic patterns in black-legged tick populations

There was no significant correlation between genetic distance and geographical distance (Mantel's r = 0.0350, p = 0.422), and measures of genetic diversity exhibited no significant evidence of clines along latitudinal gradient. Expected heterozygosity, observed heterozygosity, and allele numbers in all sites remained relatively constant along a latitudinal gradient ($p > 0.05$).

I detected patchy allele frequency distributions among my sampling sites, with additional cryptic structures among individuals within each site (Figure 4.2). There was a significant but low amount of genetic structuring among ticks sampled from different sites across the study area (R_{ST} = 0.03746, p = 0.000, % variance = 3.75%). However I also obtained a significantly positive R_{IS} (R_{IS} = 0.39775, p = 0.000, % variance = 38.29%) value, which suggest the presence of cryptic subpopulation structures. I identified three distinct genetic clusters among the 613 black-legged ticks included in the analysis (K = 3, mean LnP(k) = -19692.96, ΔK = 177.989), but did not detect any clear pattern in the geographic distribution of these clusters (Figure 4.2.). Specifically, I found that more than one genetic cluster may be found in 12 out of 13 sampling sites. Finally, sPCA on individual data revealed significant global (p = 0.001) and local (p = 0.017) genetic structures among black-legged tick samples, as well as a patchy allele distribution patterns (Figure C2).

Spatio-temporal genetic variation in black-legged tick populations

A preliminary analysis comparing pairwise differences among ticks collected from different sites at different years revealed that genetic variation observed among tick individuals was better explained by collection year than sampling locality (Figure C3). This suggests that there were greater temporal than spatial genetic variation in my tick specimens.

PCA analyses at the individual and sample levels both revealed observable grouping according to collection years (Figure C4). However the grouping was weak in the individual-level analysis, with the first two PCA axes accounting only for 15.16% of total variation (Figure C4A). Conversely, PCA analysis based on allele frequencies calculated for each sampling site revealed much stronger grouping by sampling year (1^{st} PCA axis = 25.528%, 2^{nd} PCA axis = 21.780%) (Figure C4B), where samples from sites sampled in 2012 and 2014 grouped together, away from samples from sites sampled in 2011 and 2013 (Figure C4B).

Hierarchical AMOVA performed with sampling time nested within sites, revealed no significant genetic structuring among black-legged tick sampling sites ($p = 0.600$) (Table 4.1). Conversely, AMOVA analysis with sampling sites nested within collection year, revealed genetic structures, and sampling time accounted for 6.23% of the genetic variation observed among black-legged tick specimens (Table 4.1). Site-specific analyses (i.e. testing genetic differentiation among sampling years within sites) revealed consistently significant low to high amounts of genetic structure (R_{ST} values: 0.0519 – 0.209; average R_{ST} values = 0.1120) among the black-legged tick specimens (Table 4.1). Sites for which specimens were sampled two years apart (i.e. in 2011 and 2013) had generally lower R_{ST} values (0.0519 – 0.1138) compared to sites that had ticks sampled in both odd and even years (R_{ST} values = 0.1328 – 0.2029).

Status of black-legged tick establishment

Pairwise genetic distances of samples of individuals sampled at different years within each sampling site ranged extensively from 0.0553 to 0.251. However, for the most part, these pairwise genetic distances tended to be lower when there had been even number of years between sampling time (i.e. 2011 vs. 2013), as opposed to odd number of years between sampling time (i.e. 2011 vs. 2014, 2012 vs. 2013, 2013 vs. 2014) (Table 4.2). Pairwise genetic

distance between sampling years were also lower for sampling sites located closer to the Canadian-US border compared to sampling sites located further north (Table 4.2). Specifically, I note that populations exhibiting lower amounts of temporal genetic differentiation were consistently located in regions where the certainty of the presence of established tick populations was comparatively higher, as determined by Ogden et al. (2008b - Figure 4.3). I therefore propose that black-legged tick populations located in the most southern part of the study area are more established than the populations located further north, which are still experiencing substantial recolonisation-extinction cycles (Figure 4.3).

DISCUSSION

We have witnessed in recent years the rapid emergence of Lyme disease in southern Quebec, along with the range expansion of the disease vector into the region. To obtain a better understanding of the processes and mechanisms underlying this phenomenon, I examined both the spatial and temporal genetic variation among black-legged ticks using a set of species-specific microsatellite markers.

Employing microsatellites in tick population genetic studies

The employment of microsatellite loci on any species carries its own challenges and technical difficulties (null alleles or cryptic population structures) and this applies to ticks too. Yet, we continue to see many published studies on tick population genetics that have used microsatellite markers and most have managed to provide significant insights on tick evolutionary history, ecology, gene flow, and population structure (e.g. McCoy et al. 2002, McCoy et al. 2003, Ullman et al. 2003, Kempf et al. 2009, Kempf et al. 2011, Leo et al. 2014 - see review by Araya-Anchetta et al. 2015 for full list). Aware of such potential problems, I was

extremely careful in my sampling to avoid cryptic population structures in my analyses. Specifically, my specimens include nymphs almost exclusively. If multiple larvae were sampled from a single drag or host individual, I used only one specimen to avoid potentially sampling C3siblings. Technical issues were countered by genotyping a subset of specimens twice to ensure consistent alleles were being called.

Additionally, there has previously been concern in the literature regarding the low amounts of polymorphisms exhibited by the microsatellite loci designed by Fagerberg et al. (2001) – the same ones I employed in this study. However, a closer examination of the work by Fagerberg et al. (2001) and the subsequent paper that examined the same markers (Ullman et al. 2002), revealed that the markers were tested only on ticks originating from Bridgeport, Connecticut. Specifically, Fagerberg et al. (2001) assessed their markers using F1 and F2 progenies descended from a single female collected from the site. Meanwhile, Ullman et al. (2002) tested the same markers from F1 and inter-crossed F2 progenies descended from 4 female individuals, also collected from the same site. It is not surprising that one would observe low heterogeneity given such limited sampling of lineages. Here, I used large sample sizes collected from multiple distinct populations scattered across a landscape. Therefore, I expected greater levels of polymorphism in my data relative to that observed in previous studies that had used the same markers.

Mechanism of black-legged tick range expansion into southern Quebec

My first objective was to detect the genetic signature of black-legged tick range expansion into southern Quebec. Depending on the frequency and type of dispersal mechanism, I expect distinct genetic signals (Bialozyt et al. 2006). In the case of pure diffusion or short-distance dispersal, I expect allele distribution to occur in clines with steady loss of diversity

along the colonisation corridor (Bialozyt et al. 2006). However, long-distance dispersal events will produce patchy allele distributions due to "embolism effects", whereby the migration of genetic material along a colonisation corridor is interrupted or blocked due to the establishment of certain genotypes ahead of the main colonisation front (Bialozyt et al. 2006). This blocking effect tends to increase in intensity as the frequency of long distance dispersal increases (Bialozyt et al. 2006). Yet, past a certain frequency threshold of long-distance dispersal events, any patterns resulting from blocking can be masked by homogenising effects, whereby we find consistently great amounts of genetic diversity across sites and lower genetic differentiation among sites (Bialozyt et al. 2006).

I found genetic patterns supporting the hypothesis that ticks are being introduced into the region via long-distance dispersal events occurring at a high frequency. Specifically, my analyses revealed a relatively patchy distribution of allele frequency across the study area. I also uncovered patterns that correspond with signatures of high frequencies of dispersal events such as no significant isolation-by-distance, consistently high amounts of genetic diversity in ticks across my sampling sites, and low amounts of genetic differentiation among sites. The primary candidate hosts responsible for facilitating frequent long distance dispersal events in ticks are migratory passerine birds that perform annual long-distance migrations between their over-wintering and breeding sites (Ogden et al. 2008a).

Frequent long distance dispersal events driving vector introduction into a region will affect the rates of vector and disease spread in the region. Frequent long distance dispersal events can greatly reduce the time taken for successful tick population colonisation and establishment (Bialozyt et al. 2006), given suitable habitat and host availability. It has previously been estimated that tick range expansion into Canada may occur at rates of up to 46 km per year

(Leighton et al. 2012, but see Simon et al. 2014). However this value is subject to factors that locally influence tick development rates, survival, behaviour and fitness such as temperature, rainfall, and host availability (Simon et al. 2014, Brunner et al. 2012). Furthermore, specific rates of *B. burgdorferi* emergence will be dependent in part, on the relative susceptibility of these avian hosts to tick parasitism and *B. burgdorferi* infection (Newman et al. 2015, Schneider et al 2015).

My conclusion differs from that of another study performed in New York in which mitochondrial markers suggested that tick dispersal patterns were driven not by long-distance migration, but by local dispersal events, possibly over multiple tick generations (Khatchikian et al. 2015). There are several possible explanation for the differences in the results, the first being my choice of genetic markers (mitochondrial sequences vs. microsatellites), and the second being the locations of the study relative to tick's range edge (i.e. established tick range vs. leading edge of range expansion). Future research should attempt to quantify the roles of local dispersal and long distance dispersal once tick populations have established in the region (Khatchikian et al. 2015). This can be done by directly comparing the population structure and spatial genetic patterns of the ticks against those observed in other candidate host species (See Chapter 2: Leo et al. 2016). Similarities in such structures or patterns can be indicative of interspecific dependency for dispersal (Chapter 2: Leo et al. 2016). Additionally, while I was able to identified long-distance dispersal events as the primary mechanism of tick northwards dispersal into Canada from the United States, it would be interesting to examine if this same process is responsible for the north-to-south dispersal of ticks identified by van Zee et al (2015) within the United States.

Establishment stages of black-legged tick populations in southern Quebec

Temporal genetic structure in black-legged ticks examined in this study greatly exceeded any observed spatial structure, with individuals sampled two years apart being more genetically similar to each other than specimens that were sampled either one or three years apart. This two-year temporal pattern in tick genetic variation is consistent with the species' life history. Sample collection in this study was performed predominantly in the summer months, when larvae and nymphs were at their peak activity levels. Due to the two-year life cycle of black-legged ticks (Ostfeld 2011), the juvenile ticks sampled in a single year are the offspring of a generation that were themselves in their juvenile stages two years before. It stands to reason that offspring would be more likely to exhibit genetic similarity with their parental generation than with an alternate generation that may not be closely related. It has been suggested that tick phenology changes with climatic conditions (Gatewood et al 2009, Ostfeld and Brunner 2015). Changes in black-legged tick phenology that synchronises larval and nymphal tick activity periods can exert selection pressures that promote changes in pathogen virulence within a transmission system (Gatewood et al. 2009, Ogden et al. 2008c). As global temperatures increase and tick populations establish in the region, it will be interesting to continue examining the temporal genetic variation of tick populations to see if the two-year pattern is maintained over time.

The substantial temporal genetic turnover patterns and the relatively low amounts of spatial genetic structure in the black-legged tick strongly suggest that while ticks can be introduced over a wide geographical area, populations are constantly replaced due to the combined effects of local extinction and constant introduction of new individuals (Dinsdale et al. 2012). My analyses suggested that temporal genetic variation among individuals within each site appeared to increase towards the north, suggesting that black-legged ticks sampled from different

sites in this study exhibit varying degrees of population establishment. I propose that sites in this study exhibiting low (genetic distances < 0.10) to moderate (genetic distances between 0.10 and 0.15) amounts of temporal variation, while located in regions that Ogden et al. (2008b) identified to be at high risk of established tick populations, may be considered more established (Figure 4.3). Meanwhile, the sites with great amounts of temporal variation (genetic distances > 0.20) and located in regions of 'moderate' risk (as defined by Ogden et al. 2008b) can be considered less established (Figure 4.3).

The status of black-legged tick spread has important implications for Lyme disease risk across a landscape. Both adventitious and established ticks are capable of transmitting Lyme disease to humans (Ogden et al. 2006b). However, the timing of disease risk differs: risk of infection by adventitious ticks is higher during the early spring and autumn activity periods of adult ticks, whereas risk is greater during spring-early summer peak activity period of nymphs in areas where tick populations are established (Ogden et al. 2006b). Furthermore, regions with establish black-legged tick populations are likely to have higher tick abundances, allowing for increased human-tick encounter rates, further increasing risks of tick bites and disease transmission (Ostfeld 2011). Finally, a recent study has estimated a temporal gap between tick and *B. burgdorferi* establishment into Canada, where tick population establishment is generally followed by *B. burgdorferi* establishment approximately three to five years later (Ogden et al. 2013). Early detection of established tick populations in a region can foretell the possibility of future emergence of local Lyme disease hotspots.

CONCLUSION

This is the first study to investigate contemporary population genetic variation of black-legged ticks at the regional level in Canada since records of tick presence first began. The salient point of this study is the application of concepts from population genetics to assess mechanism of species range expansion. I was able to quantify empirical patterns of genetic diversity and structure among populations of an important invasive disease vector species and identified specific range expansion processes and mechanisms (i.e. frequent long-distance dispersal events) that would create such patterns. I also further identified stages of population establishment at the leading edge of a disease vector undergoing range expansion, eliminating the need for labour-intensive and time-consuming surveillance efforts in the field. Furthermore, the identification of established vector populations will allow public health officers to identify when disease risks in a region will be greatest (e.g. adventitious vs. established ticks) and disseminate crucial information to medical professionals and the general public for diagnostic and preventative efforts respectively.

REFERENCES

Araya-Anchetta,A, Busch JD, Scoles GA, Wagner DM. 2015. Thirty years of tick population genetics: A comprehensive review. *Infection, Genetics and Evolution* **29**: 164 – 179.

Bialozyt R, Ziegenhagen B, Petit RJ. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**: 12 – 20.

Brunner JL, Killilea M, Ostfeld RS. 2012. Overwintering survival of nymphal *Ixodes scapularis* (Acari: Ixodidae) under natural conditions. *Journal of Medical Entomology* **49**: 981 – 987.

Chen I-C, Hill JK, Ohlimüller R, Roy DB, Thomas CD. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science* **333**: 1024 – 1026. doi: 10.1126/science.1206432

Cristescu ME. 2015. Genetic reconstruction of invasion history. *Molecular Ecology* **24**: 2212 – 2225. doi: 10.1111/mec.13117

Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment* **6**: 238 – 246. doi: 10.1890/070151

Dawe KL, Bayne EM, Boutin S. 2014. Influence of climate and human land use on the distribution of white-tailed deer (*Odocoileus virginianus*) in the western boreal forest. *Canadian Journal of Zoology* **92**: 353 – 363. doi: 10.1139/cjz-2013-0262

Dinsdale A, Schellhorn NA, de Barro P, Buckley YM, Riginos C. 2012. Rapid genetic turnover in populations of the insect pest *Besmisia tabaci* Middle East: Asia Minor 1 in an agricultural landscape. *Bulletin of Entomological Research* **102**: 539 – 549. doi: 10.1017/S0007485312000077

Dray S, Dufour A-B, Thioulouse J. 2015. ade4: Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences. R package 1.7-2

Engering A, Hogerwerf , Slingenbergh J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections* **2**: e5. doi: 10.1038/emi.2013.5

Estrada-Pena A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease risks: An ecological primer. *Trends in Parasitology* **30**: 205 – 214. doi: 10.1016/j.pt.2014.02.003

Excoffier L, Foll M, Peti RJ. 2009. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics* **40**: 481 – 501. doi: 10.1146/annurev.ecolsys.39.110707.173414

Excoffier L, Lischer HEL. 2010 Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**: 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Fagerberg AJ, Fulton RE, Black WC IV. 2001. Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. *Insect Molecular Biology* **10**: 225 – 236.

Falco RC, Fish D. 1991 Horizontal movement of adult *Ixodes dammini* (Acari: Ixodidae) attracted to CO_2-baited traps. *Journal of Medical Entomology* **28**: 726 – 729.

Fuller T, Bensch S, Müller I, Novembre J, Pérez-Tris J, Ricklefs RE, Smith TB, Waldenström J. 2012. The ecology of emerging infectious disease in migratory birds: An assessment of the role of climate change and priorities for future research. *EcoHealth* **9**: 80 – 88. doi: 10.1007/s10393-012-0750-1

Gatewood AG, Liebman KA, Vourc'h G, Bunikis J, Hamer SA, Cortinas R, Melton F, Cislo P, Kitron U, Tsao J, Barbour AG, Fish D, Diuk-Wasser MA. 2009. Climate and tick seasonality are predictors of *Borrelia burgdorferi* genotype distribution. *Applied and Environmental Microbiology* **75**: 2476 – 2483. doi: 10.1128/AEM.02633-08

Jombart T, Collins C, Kamvar ZN, Lustrik R, Solymos P, Ahmed I, Calboli F, Cori A, Schliep K. 2014. adegenet: an R package for the exploratory analysis of genetic and genomic data. R package vers 1.4-2

Kempf F, Boulinier T, de Meeûs T, Arnathau C, McCoy KD. 2009. Recent evolution of host-associated divergence in the seabird tick *Ixodes uriae*. *Molecular Ecology* **18** : 4450 – 4462.

Kempf F, de Meeus T, Vaumourin E, Noel V, Taragel'ova V, Plantard O, Heylen DJA, Eraud C, Chevillon C, McCoy KD. 2011. Host races in *Ixodes ricinus*, the European vector of Lyme borreliosis. *Infection, Genetics and Evolution* **11**: 2043 - 2048

Khatchikian CE, Prusinski MA, Stone M, Backenson PB, Wang IN, Foley E, Seifert SN, Levy MZ, Brisson D. 2015. Recent and rapid population growth and range expansion of the Lyme disease tick vector, *Ixodes scapularis*, in North America. *Evolution* **69**: 1678 – 1689. doi: 10.1111/evo.12690

Krakowetz CN, Lindsay LR, Chilton NB. 2011. Genetic diversity in *Ixodes scapularis* (Acari: Ixodidae) from six established populations in Canada. *Ticks and Tick-borne Diseases* **2**: 143 – 150. doi: 10.1016/j.ttbdis.2011.05.00

Krakowetz CN, Lindsay LR, Chilton NB. 2014. Genetic variation in the mitochondrial 16S ribosomal RNA gene of *Ixodes scapularis* (Acari: Ixodidae). *Parasites and Vectors* **7**: 530 – 544. doi: 10.1186/s13071-014-0530-6

Legér E, Vourc'h G, Vial L, Chevillon C, McCoy KD. 2013. Changing distributions of ticks: causes and consequences. *Experimental and Applied Acarology* **59**: 219 – 244. doi: 10.1007/s10493-012-9615-0

Leighton PA, Koffi JK, Pelcat Y, Lindsay LR, Ogden NH. 2012. Predicting the speed of tick invasion: An empirical model of range expansion for the Lyme disease vector *Ixodes scapularis* in Canada. *Journal of Applied* Ecology **49**: 457 – 464. doi: 10.1111/j.1365-2664.2012.02112.x

Leo SST, Gonzalez A, Millien V. 2016. Multi-taxa integrated landscape genetics for zoonotic infectious diseases: Deciphering variables influencing disease emergence. *Genome* (doi: 10.1139/gen-2016-0039)

Leo SST, Samuel WM, Pybus MJ, Sperling FAH. 2014. Origin of Dermacentor albipictus (Acari: Ixodidae) on elk in the Yukon, Canada. *Journal of Wildlife Diseases* **50**: 544- 551. (doi: 10.7589/2013-03-078)

Madhav NK, Brownstein JS, Tsao JI, Fish D. 2004. A dispersal model for the range expansion of black-legged tick (Acari: Ixodidae). *Journal of Medical Entomology* **41**: 842 – 852.

McCoy KD, Tirard C. 2002. Reproductive strategies of the seabird tick *Ixodes uriae* (Acari: Ixodidae). *Journal of Parasitology* **88**: 813 – 816.

McCoy KD, Tirard C, Michalakis Y. 2003. Spatial genetic structure of the ectoparasite *Ixodes uriae* within breeding cliffs of its colonial seabird host. *Heredity* **91**: 422 – 429.

Mechai S, Feil EJ, Gariepy TD, Gregory TR, Lindsay LR, Millien V, Ogden NH. 2013 Investigation of the population structure of the tick vector of Lyme disease *Ixodes scapularis* (Acari: Ixodidae) in Canada using mitochondrial Cytochrome C Oxidase subunit I gene sequences. *Journal of Medical Entomology* **50**: 560 – 570**.**

Newman EA, Eisen L, Eisen RJ, Fedorova N, Hasty JM, Vaughn C, Lane RS. 2015. *Borrelia burgdorferi* sensu lato spirochetes in wild birds in northwestern California: Associations with ecological factors, bird behavior and tick infestation. *PLoS ONE* **10**: e0118146. doi: 10.1371/journal.pone.0118146

Ogden NH, Barker IK, Francis CM, Heagy A, Lindsay LR, Hobson KA. 2015. How far north are migrant birds transporting the tick *Ixodes scapularis* in Canada? Insights from stable hydrogen isotope analyses of feathers. *Ticks and Tick-borne Diseases* **6**: 715 – 720. doi: 10.1016/j.ttbdis.2015.06.004

Ogden NH, Bigras-Poulin M, Hanincová K, Maarouf A, O'Callaghan CJ, Kurtenbach K. 2008c. Projected effects of climate change on tick phenology and fitness of pathogens transmitted by North American tick *Ixodes scapularis*. *Journal of Theoretical Biology* **254**: 621 – 632.

Ogden NH, Koffi KJ, Pelcat Y, Lindsay LR. 2014. Environmental risk from Lyme disease in central and eastern Canada: A summary of recent surveillance information. *Canada Communicable Disease Report* **40**: 74 – 82. ISSN: 1481-8531

Ogden NH, Lindsay LR, Morshed M, Sockett PN, Artsob H. 2009. The emergence of Lyme disease in Canada. *Canadian Medical Association Journal* **180**: 1221 – 1224. doi: 10.1503/cmaj.080148

Ogden NH, Lindsay LR, Hanincová K, Barker IK, Bigras-Poulin M, Charron DF, Heagy A, Francis CM, O'Callaghan CJ, Schwartz I, Thompson RA. 2008a Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. *Applied and Environmental Microbiology* **74**: 1780 – 1790. doi: 10.1128/AEM.01982-07

Ogden NH, Lindsay LR, Leighton PA. 2013. Predicting the rate of invasion of the agent of Lyme disease *Borrelia burgdorferi. Journal of Applied Ecology* **50**: 510 – 518. doi: 10.1111/1365-2664.12050

Ogden NH, Maarouf A, Barker IK, Bigras-Poulin M, Lindsay LR, Morshed MG, O'callaghan CJ, Ramay F, Waltner-Toews D, Charron DF. 2006a. Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. *International Journal for Parasitology* **36**: 63 – 70.

Ogden NH, St-Onge L, Barker IK, Brazeau S, Bigras-Poulin M, Charron DF, Francis CM, Heagy A, Lindsay LR, Maarouf A, Michel P, Milord F, O'Callaghan CJ, Trudel L, Thompson RA. 2008b. Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics* **7**: 24 – 38. doi: 10.1186/1476-072X-7-24.

Ogden NH, Trudel L, Artsob H, Barker IK, Beauchamp G, Charron DF, Drebot MA, Galloway TD, O'Handley R, Thompson RA, Lindsay LR. 2006b. *Ixodes scapularis* ticks collected by passive surveillance in Canada: Analysis of geographical distribution and infection with Lyme borreliosis agent *Borrelia burgdorferi*. *Journal of Medical Entomology* **43**: 600 – 609.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2015. vegan: Community ecology package. R package vers. 2.2-1

Ostfeld RS. 2011. Lyme disease: The ecology of a complex system. Oxford University Press, Oxford, United Kingdom.

Ostfeld RS, Brunner JL. 2015. Climate change and *Ixodes* tick-borne diseases of humans. *Philosophical Transactions of the Royal Society B* **370**: 20140051.

Park SDE. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. PhD Thesis, University of Dublin, Dublin, Ireland.

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

Schneider SC, Parker CM, Miller JR, Fredericks LP, Allan BF. 2015. Assessing the contribution of songbirds to the movement of ticks and *Borrelia burgdorferi* in the Midwestern United States during Fall migration. *Ecohealth* **12**: 164 – 173. doi: 10.1007/s10393-014-0982-3

Simon JA, Marrotte RR, Desrosiers N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe FJ, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of *Borrelia burgdorferi*, the agent of Lyme disease, at the northeastern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764. doi: 10.1111/eva.12165

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457-462.

Ullman A, Piesman J, Dolan MC, Black WC IV. 2003. A preliminary linkage map of the tick, Ixodes scapularis. In *Ticks and Tick-Borne Pathogens* (pp. 107-126). Springer Netherlands.

van Zee J, Piesman JF, Hojgaard A, Black WC IV. 2015. Nuclear markers reveal predominantly North to South gene flow in *Ixodes scapularis*, the tick vector of the Lyme disease spirochete. *PLoS ONE* **10**: e0139630. doi:10.1371/journal.pone.0139630

TABLES

Table 4.1. Hierarchical AMOVA table and corresponding values for R_{ST}. R_{ST} values are statistically significant when $p < 0.05$. Amount of observed variance explained at each hierarchical level is provided.

Hierarchical Structure	$\boldsymbol{R_{ST}}$	p	% variance
Location > Time [β]	-0.009	0.600	0.00%
Time > Location [γ]	0.0623	0.000	6.23%
Among sites within sampling years			
2011 (Nsites = 7)	0.0651	0.000	6.51%
2013 (Nsites = 8)	0.1072	0.000	10.72%
2014 (Nsites = 3)	0.0168	0.509	1.66%
Among sampling years within sites			
BMF (2011, 2013)	0.0519	0.000	5.19%
DV (2012, 2013, 2014)	0.2029	0.000	20.29%
HR (2011, 2013)	0.1138	0.000	11.38%
HV (2011, 2013)	0.0602	0.000	6.02%
MSB (2011, 2013, 2014)	0.1656	0.000	16.56%
NO (2011, 2013)	0.0921	0.000	9.21%
SF (2011, 2013, 2014)	0.1328	0.000	13.28%
SV (2011, 2013)	0.0763	0.000	7.63%

[β] Sampling time is nested within site

[γ] Sampling site is nested within sampling year

Table 4.2. Pairwise genetic differentiation (F_{ST} values) of *Ixodes scapularis* specimens across time for each sampling site. Analyses were performed on individuals collected 1, 2, and 3 years apart between 2011 and 2014. Significant F_{ST} values are indicated with an '*'. Sites for which I had only one year of sampling or fewer than 12 individuals in a given year were excluded from the analysis (See Table C1).

Locality	**2011 vs. 2013 (2 years)**	**2011 vs. 2014 (3 years)**	**2012 vs. 2013 (1 year)**	**2012 vs. 2014 (2 years)**	**2013 vs. 2014 (1 year)**
BMF	0.055 *	*NA*	*NA*	*NA*	*NA*
DV	*NA*	*NA*	0.251 *	0.210 *	0.167 *
HR	0.120 *	*NA*	*NA*	*NA*	*NA*
HV	0.064 *	*NA*	*NA*	*NA*	*NA*
LP	*NA*	*NA*	*NA*	*NA*	*NA*
LS	*NA*	*NA*	*NA*	*NA*	*NA*
MSB	0.134 *	0.215 *	*NA*	*NA*	0.156 *
NO	0.095 *	*NA*	*NA*	*NA*	*NA*
SdV	*NA*	*NA*	*NA*	*NA*	*NA*
SE	*NA*	*NA*	*NA*	*NA*	*NA*
SF	0.095 *	0.163 *	*NA*	*NA*	0.133 *
SJ	*NA*	*NA*	*NA*	*NA*	*NA*
SV	0.079 *	*NA*	*NA*	*NA*	*NA*

FIGURES

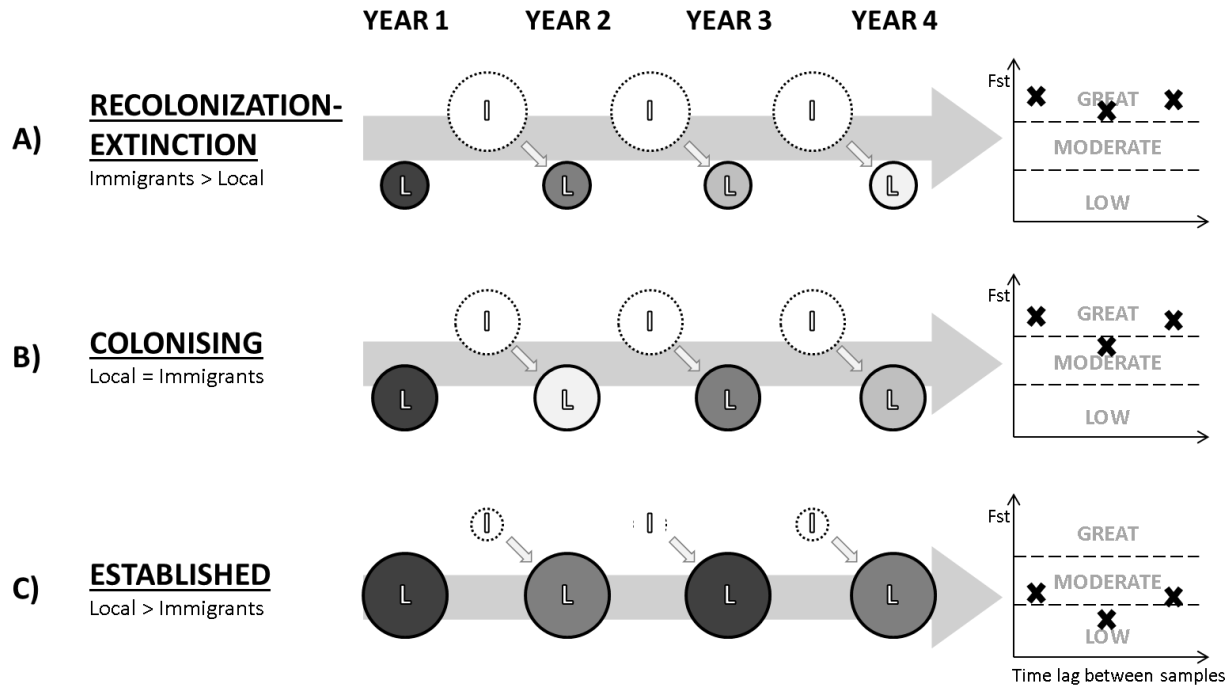


Figure 4.1. A hypothesis for the genetic turnover and differentiation within sites over time based on stages of population establishment (row). The size of the circles indicate relative abundance of local "L" versus immigrant "I" individuals within the site at each point in time (column). The graphs on the right summarise the relative patterns of genetic differentiation I expect to observe for each establishment stage. In this hypothesis, the two-year life cycle of the ticks is taken into account. The different shades of grey in the local ("L") circles represent hypothesised relatedness of individuals with those from previous years.

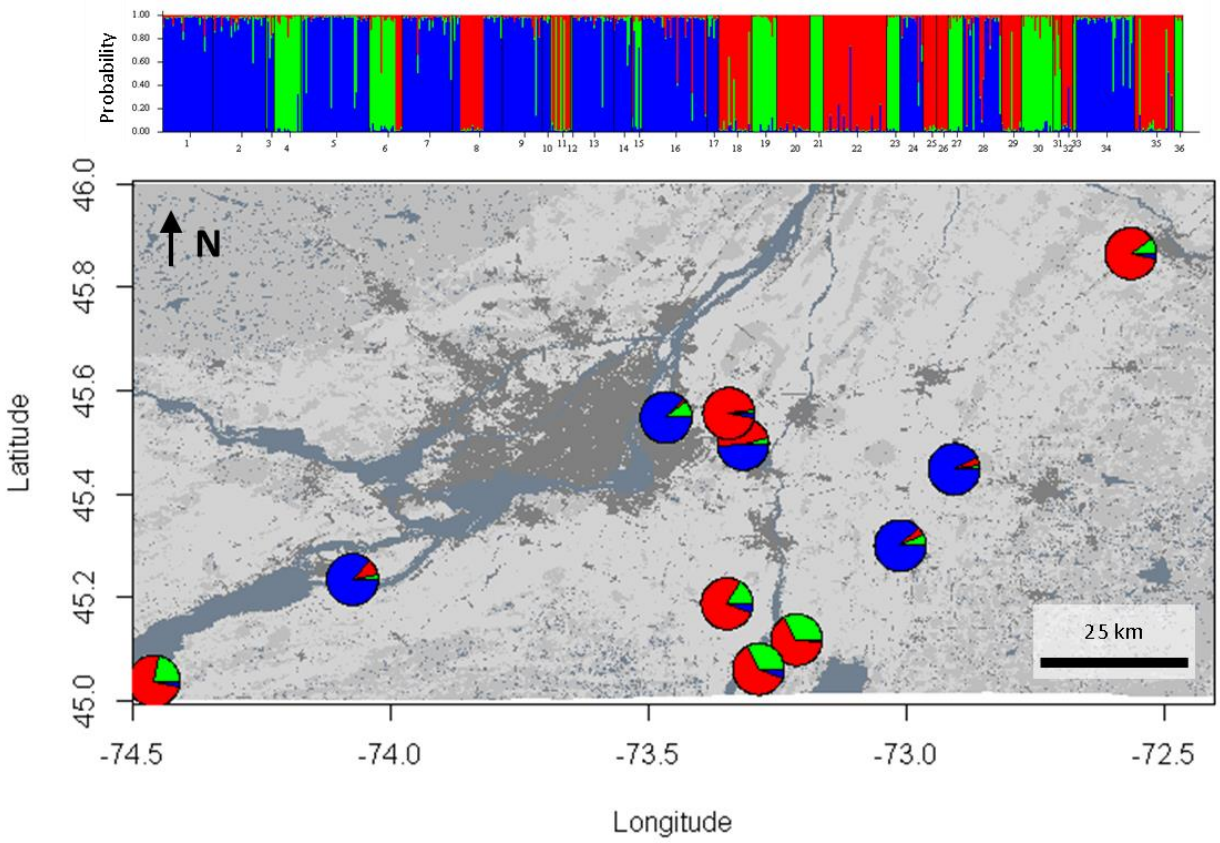


Figure 4.2. Results from structure's clustering analysis on all 613 black-legged tick individuals (K = 3, mean LnP(k) = -19692.96, ΔK = 177.989). All analyses were performed using no admixture and independent allele frequency models. Ten iterations for each number of populations (k) equaling 1 through 10 was initially analyzed for 100,000 Markov Chain Monte Carlo (MCMC) generations with an initial burn-in of 10,000 generations. The most likely number of populations was subsequently calculated from likelihood outputs produced by structure using the approach by Evanno et al. (2005). Based on this result, a more thorough analysis with 'k' defined was completed with MCMC running for 500,000 generations with an initial burn-in of 50,000. Pie charts on the map show proportion of membership of each pre-defined black-legged population in each cluster. The bar-plot indicates the probability of individual assignment to these clusters.

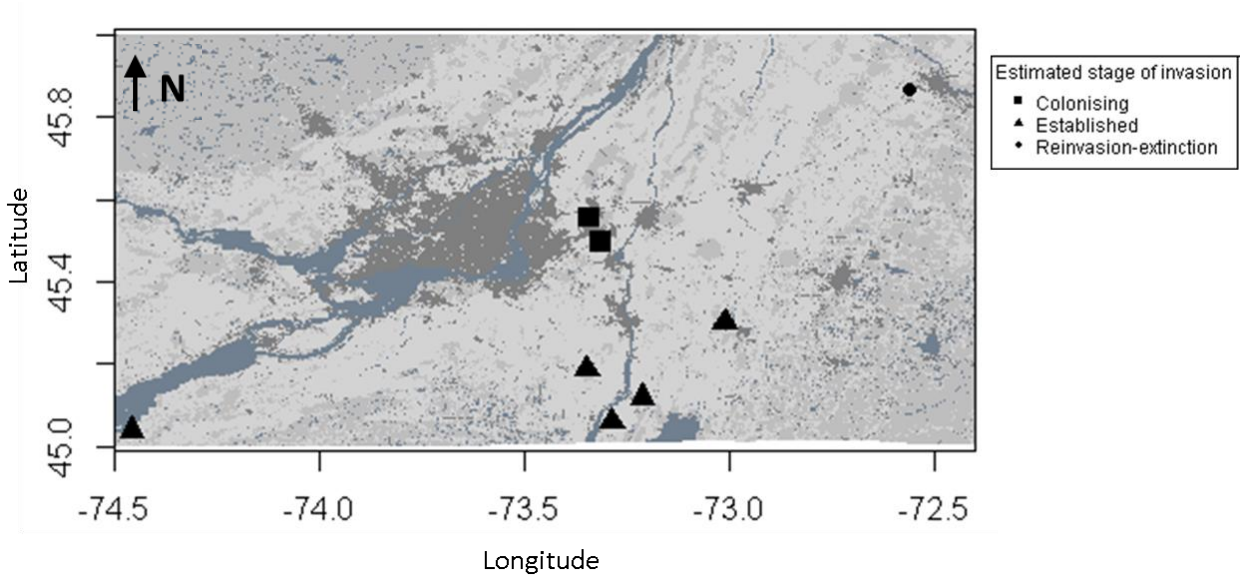


Figure 4.3. Estimated stages of tick population establishment (established [triangles], colonizing [squares], and recolonization-extinction [circles] respectively) based on genetic turnover as hypothesized in Figure 4.1.

CONNECTING STATEMENT

In Chapters 3 and 4, I examined the genetic diversity of both the white-footed mouse and black-legged tick, respectively using microsatellite markers. I have a deeper understanding of species-specific movement across the landscape. Specifically, white-footed mouse dispersal was heavily dependent on landscape structure (Chapter 3) while black-legged tick movement appeared less constrained (Chapter 4). Furthermore, I found that the white-footed mouse, as a host of both pathogen and ticks, is an important determinant of Lyme disease presence in an area (Chapter 3). Meanwhile, large annual influxes of black-legged ticks into southern Québec can potentially promote Lyme disease establishment (Chapter 4).

In Chapter 5, I create an integrative model that projects Lyme disease infection risk in southern Québec based on what I have learnt from previous chapters regarding the two main components of the Lyme disease transmission system – the white-footed mouse reservoir host (Chapter 3), and the black-legged tick vector (Chapter 4). I focus on three variables: [i] interspecific interactions between tick and mice (as determined using MTILG described in Chapter 2), [ii] landscape and habitat effects on species abundances, and [iii] climate effects on species distributions. This approach led to a projection of disease risk across the region. This prediction will be essential for disease control and monitoring efforts of Lyme disease into the future.

CHAPTER FIVE

Lyme disease risk from vector and reservoir host species distributions are mediated by climate and landscape heterogeneity

ABSTRACT

Recent climate and land use changes have facilitated the emergence of Lyme disease in Canada by promoting the range expansion of the primary disease reservoir host (white-footed mouse - *Peromyscus leucopus*) and vector (black-legged tick - *Ixodes scapularis*). As both species are essential for local transmission and persistence of the disease-causing pathogen (*Borrelia burgdorferi*), I created an Structural Equation Model to project Lyme disease infection risk in southern Québec based on host and vector distributions with respect to [i] host-dependent dispersal estimated from species population genetics, [ii] the effect of landscape and habitat characteristics on species abundance, and [iii] the effect of climate on species distribution. I found that white-footed mouse density tended to increase in smaller forest patches located close to urban environments while none of the landscape variables examined in this study explained observed tick abundance. I also found that warmer winter temperatures tend to promote tick occurrence probability. Similarly, regional winter length and mean temperatures can influence white-footed mouse distribution. Finally, my projections revealed that Lyme disease ecological risk is greatest at the US-Canada border, gradually decreasing northwards and eastwards. Such an integrated modeling approach promises a more accurate and reliable prediction of disease risk that will be essential for disease control and monitoring efforts.

INTRODUCTION

The emergence of zoonotic disease is an ongoing global phenomenon (Jones et al. 2008). Predicting disease emergence is difficult because of the complexity of the interactions between underlying variables (Daszak et al. 2000, Wilcox and Gubler 2005, Patz et al. 2008, Tabachnick 2010). Many factors must be simultaneously accounted for to obtain accurate estimates of disease risk. For example, changes in climate influence the distribution, phenology, and survival of important biological components of a disease system (Martinez and Marino 2011, Ogden 2014, Ostfeld and Brunner 2015, Bozick and Real 2015). Similarly, changes in the landscape and habitat characteristics can exert selective pressures and change species fitness. Many other factors can amplify or dilute disease risk including changing species' dispersal patterns (Wilcox and Gubler 2005), population demography (Grenfell and Harwood 1997, Taylor et al. 2011, Suzán et al. 2015) and community composition (Ostfeld and Keesing 2012).

Many zoonotic disease systems involve interactions among multiple species, each of which can play distinct role in the transmission and persistence of disease in a region (Patz et al. 2008, Engering et al. 2013, Estrada-Pena et al. 2014). To further complicate disease models, species are also likely to respond differently to human and natural forms of changes in their environment, which often occur at different spatial and temporal scales. Therefore, models of disease risk and spread require an integrative approach that takes into account how critical disease vectors and hosts respond to changes in climate, landscape, and biotic conditions (Arino et al. 2005, Estrada-Pena et al. 2014, Chapter 2: Leo et al. 2016). Here, I modeled changing Lyme disease infection risk in southern Québec, Canada by integrating population genetics, projected effects of climate variables on species distribution, and landscape effects on the

abundance of the key disease vectors and reservoir hosts. From there, I determined how each variable contribute to disease risk in terms of pathogen prevalence across the region.

Over the past decade, Lyme disease cases in Québec had increased from 6 in 2004 to over 120 in 2014 (Institut National de Santé Publique du Québec 2014). This rapid emergence was attributed to the range expansion of both the disease vector (*Ixodes scapularis*– black-legged ticks) and the primary disease reservoir host (*Peromyscus leucopus* – white-footed mouse) (Simon et al. 2014). The combined effects of climate change at the continental scale and land use change at the regional scale have produced conditions that promoted the successful establishment and increase in abundances of white-footed mouse and black-legged tick populations, which in turn promoted the introduction and persistence of the disease-causing bacterium *Borrelia burgdorferi* in the region. For example, white-footed mouse abundance across the region has increased in response to warmer and shorter winters and habitat fragmentation (Anderson et al. 2003, Roy-Dufresne et al. 2013). Meanwhile, warming temperatures and ambient relative humidity have increased tick survival, development, and reproductive rates, as well as altered tick behavior and questing activity (Lindsay et al. 1999a, Vail and Smith 2002). Changes in small mammal host species community composition in response to climate and land use change have also promoted the emergence of the pathogen in the region (Ostfeld 2011, Ostfeld and Keesing 2012). Lastly, changes in tick phenology in response to climate change can exert selective pressures that maintain or increase pathogen virulence in humans (Haven et al. 2012).

Here I present an integrated approach to projecting disease risk across a region by integrating different types of variables, each operating on different temporal and spatial scales into a single comprehensive model. I examined how interspecific interactions (in the form of coupled dispersal patterns), climate conditions, and landscape and habitat characteristics

influence black-legged tick and white-footed mouse distribution across southern Québec, and how that in turn affect pathogen prevalence in the region (Figure 5.1). The model I produced incorporates a combination of genetic information, statistical analyses, and climate niche models. The purpose of this study is three-fold:

1. To simultaneously determine how two different yet important species of a disease system respond to variables operating across varying spatial and temporal scales.
2. To quantify the indirect contribution of different variables underlying observed disease emergence patterns (e.g. climate vs. landscape) based on their effects on key species of the disease system.
3. To identify areas of greatest disease risk to humans by comparing maps of projected disease ecological risk with maps of human population densities.

This information obtained by this approach will be critical for informing the public and physicians about the necessity of preventive measures and diagnostic approaches, respectively. This information can also help inform future policies (e.g. land use planning) that can influence ongoing disease emergence patterns.

MATERIAL AND METHODS

Field data

Small mammals were live trapped from 40 sites across (southern Quebec, Canada between July and September from 2011 to 2014) with Sherman traps set 5-10 meters apart in grids of 4x7 or 4x10. All small mammals were euthanized upon capture and dissected on the same day. Tissues collected (i.e. heart, liver, kidneys) were stored in 90% ethanol and in -20°C freezers.

Ticks were collected either by cloth dragging in the vegetation or removed from captured host animals. Ticks were also stored in 90% ethanol and in -20°C freezers. I made sure to store individual nymphs and adults in individual containers. Larvae belonging to the same cluster were stored within a single tube. I identified all small mammal and tick specimens to species either based on morphology or genetic marker amplification (i.e. *Peromyscus* species – Rogic et al. 2013). I also tested them for infection by the Lyme disease agent - *Borrelia burgdorferi* at the National Microbiology Laboratory of the Public Health Agency (Winnipeg, MB, Canada) following the protocol described in Ogden et al. (2011). *Borrelia burgdorferi* prevalence for each site was calculated as the proportion of individuals that tested positive over the total number of host and tick specimens tested. Positive screening results were taken as indicative of pathogen presence at the site.

Variables influencing Borrelia burgdorferi *prevalence*

I investigated relations between *B. burgdorferi* prevalence at a site and a series of black-legged tick, white-footed mouse, and small mammal related variables (Table 5.1) (see Appendix D details on variables examined). Since all explanatory variables exhibited collinearity with each other (based on a test using the function *vif* from the R package *car* (Fox et al. 2015)), I performed a hierarchical partitioning analysis using the *hier.part* package in R (Walsh and MacNally 2013). This allowed us to estimate the individual contribution of each explanatory variable to *B. burgdorferi* prevalence. I evaluated the statistical significance of the independent contribution by each variable by running a randomization test with 1000 permutations using the *rand.hp* function in the *hier.part* package in R.

Population genetics and multi-taxa integrated landscape genetics on mouse and tick

DNA was extracted from all specimens using a standard 3-day phenol/chloroform extraction procedure as described by Sambrook et al. (1989). I genotyped white-footed mouse and black-legged tick specimens for eleven and ten polymorphic microsatellite loci respectively (Table D1).

I first examined the population structures of tick and mice across the study area using individual-based clustering analyses in the software *structure* ver. 2.3 (Pritchard et al. 2000), as well as an Analysis of Molecular Variance (AMOVA) in the population genetics software Arlequin ver. 3.5.1.3 (Excoffier et al. 2010).

Because Lyme disease risk is enhanced by *B. burgdorferi* prevalence at sites where both white-footed mice and black-legged ticks are present, I investigated whether interspecific interactions between these species in the form of host-dependent dispersal may contribute to disease emergence patterns. I applied the multi-taxa integrated landscape genetics (MTILG) (James et al. 2011) approach (Chapter 2: Leo et al. 2016) on populations of white-footed mouse and black-legged ticks across the study area (see Appendix D for details). In order to make comparisons between species, all analyses in this section were performed on the 9 sites from which both white-footed mice and black-legged ticks were sampled.

Landscape and habitat effects on mouse density and tick abundance

I performed redundancy analysis (RDA) to examine the influence of landscape and habitat characteristics on white-footed mouse density and black-legged tick abundance in each of the 40 sites (Figure 5.1). Habitat characteristics included estimates of the proportion of forest, agricultural areas, and urban development within a 5km radius of each patch, the area and perimeter of the patches, and the minimum distance to the nearest neighbor (MDNN) (Table 5.2). For the white-footed mouse, I also included additional site connectivity variables with

respect to the distribution of roads, agricultural areas, and urban development, as these landscape variables significantly influenced mouse gene flow across the landscape (Table D4). I first identified predictor variables that significantly accounted for variance in mouse densities and tick abundances by performing an ANOVA on the RDA model before quantifying the relative effects of these variables on mouse density and tick abundance using a linear model. Coefficients from the linear model were then used to estimate mouse density and tick abundance, which were subsequently regressed against observed field data using Spearman's rank correlation tests.

Climatic variables on mouse and tick occurrence probabilities

I used climate niche models to predict vector and host occurrence probabilities at the sites based on a combination of climatic variables associated with winter conditions (Figure 5.1). Specifically, I focused on variables reflecting winter conditions including snow depth, winter precipitation, minimum and maximum winter temperatures, and winter length (Table D2). These variables were chosen for their potential to determine both mouse and tick species distributions. For example, a warmer climate with milder winters and earlier spring snowmelt may alter white-footed mouse breeding behaviour and success (Pratt and Barrett 2012, Kaseloo et al. 2014). Similarly, winter temperatures, humidity, and snow cover may substantially influence tick overwintering mortality rates (Brunner et al. 2012, Hayes et al. 2015). The application of winter temperature variables in tick climate niche models is controversial as similar rates of tick mortality between the winter and summer seasons suggest that cold-related overwintering mortality does not restrict tick populations (see Ostfeld and Brunner 2015). Instead, Ogden et al. (2014) suggested that tick population survival is primarily dependent on cumulative annual degree days > 0°C in the non-winter seasons. Nonetheless, I wanted to employ similar variables between species to identify tick and mouse relative responses to similar climate variables.

Furthermore, one can postulate that winter temperatures and annual cumulative degree days are correlated; therefore winter temperatures and winter lengths may still be employed as valid factors influencing vector distribution.

I performed the climate niche models using the R package *BIOMOD2* (Thuiller et al. 2014) (see Appendix D for details on climate niche models). Consensus species occurrence probabilities estimated by these models were extracted for the 40 sites in the study and the fit between predicted occurrence probability models and field observations was tested using Spearman's rank correlation tests.

Climate and landscape effect on Lyme disease risk through disease host and vector species

I used a structural equation model (SEM) to determine and quantify the relative effects of climate and landscape and habitat variables on the distribution of the disease vector and reservoir host and their relative effects on *B. burgdorferi* prevalence at each site (Figure 5.1) (see Appendix D for details on SEM analysis). Coefficients obtained from the structural equation model, along with occurrence probabilities of white-footed mouse and black-legged ticks extracted from their respective climate niche models, and species abundances estimated from landscape and habitat variables, were then used to produce a linear equation model to estimate *B. burgdorferi* prevalence.

Projection of Borrelia burgdorferi *prevalence across the study area*

I projected Lyme disease infection risk (in the form of *B. burgdorferi* prevalence) in my study area using the linear equation described above ((Figure 5.1). First, I extracted climate-based white-footed mouse and black-legged tick occurrence probability from the output of my climate-based species distribution model (Figure 5.1). I also calculated mouse density from

landscape and habitat characteristics for all 9043 habitat patches in the study area based on results from the analyses above (Figure 5.1). Since both mouse hosts and tick vectors are essential for the transmission and persistence of *B. burgdorferi* populations, the infection risk was considered null in the absence of both species. I subsequently compared projected Lyme disease infection risk against human population density. Information on population density within each census dissemination block in Québec was extracted from the 2011 census data from Statistics Canada (Geographic Attribute File, 2011 Census. Statistics Canada Catalogue no. 92-151-X, and Dissemination Block Boundary File, 2011 Census. Statistics Canada Catalogue no. 92-163-X).

RESULTS

Variables influencing Borrelia burgdorferi *prevalence*

Tests screening for *B. burgdorferi* infection in small mammals and black-legged ticks uncovered pathogen presence at 14 out of 40 study sites, with pathogen prevalence ranging from 0.000 to 0.2641 (mean = 0.0378, standard deviation = 0.067). When I investigated relations between *B. burgdorferi* prevalence at a site relative to a set of tick and small mammal variables using hierarchical partitioning analyses, I found that *B. burgdorferi* prevalence at a site is mostly driven by tick abundance as well as tick burden on white-footed mice, and to a lesser extent by tick burden on all small mammal hosts (Table 5.1).

White-footed mouse and black-legged tick population genetics

Individual-based clustering analyses revealed that mice populations can be sorted into two genetic clusters that correspond for the most part with either shore of the St. Lawrence River.

Conversely, tick populations formed three genetic clusters that exhibit no clear spatial trend (Figure D1). AMOVA via Arlequin further revealed that both species have very different population structures (Table D3). White-footed mouse populations exhibit much greater genetic structuring (R_{ST} = 0.153, $p < 0.001$) than black-legged ticks (R_{ST} = 0.043, $p < 0.001$).

When I compared the pairwise genetic distances among populations in both white-footed mice and black-legged ticks as part of the MTILG analyses, I found no significant correlations (Mantel cor = -0.0589, p = 0.575). Similarly, Procrustes rotation tests on PCA and PCoA ordination solutions calculated from allele frequency distributions and inter-population pairwise genetic distances respectively, revealed no concordant spatial genetic pattern between both species (PCA: cor = 0.4125, p = 0.517; PCoA: cor = 0.3599, p = 0.659). Subsequently, landscape genetic analyses revealed that the landscape influenced white-footed mouse and black-legged tick gene flow differently. I found that mouse gene flow across the landscape is significantly influenced by site connectivity relative to the presence and distribution of roads, agriculture areas, and urban development across the landscape (Table D4). Conversely, gene flow among black-legged tick populations was not affected by any of the landscape variables examined (Table D4).

Results from the MTILG analyses suggest that black-legged tick and white-footed mouse in the study area do not exhibit concordant spatial genetic patterns or shared spatial dependence. This implies that local black-legged tick dispersal across my study area is not dependent on white-footed mouse movement. Therefore, subsequent species distribution models for both species were performed independently.

Landscape and habitat variables on species density and abundance

None of the landscape or habitat characteristics I examined significantly accounted for patterns of tick presence or abundance at my study sites, suggesting that the landscape variables examined in this study have little to no effect on tick distribution at the spatial scale of this study (Table 5.2). As no landscape/habitat variables were found to contribute significantly to black-legged tick abundance at the study sites, I did not attempt to predict tick abundance from landscape variables.

Conversely I found that the area of a forest patch and the proportion of area dedicated to urban development within a 5km radius of a patch significantly affected mouse density (Table 5.2). Mouse density tends to increase with decreasing patch area (estimated coefficient = -3.47×10^{10}, $p = 0.034$) and increasing amount of urban development around the patch (estimated coefficient = 0.013, $p = 0.037$). When I applied coefficients from the linear model to estimate mouse density as a function of patch area and amount of urban development surrounding the patch (Figure 5.2A), my predicted density was significantly correlated with observed white-footed mouse density (rho = 0.504, $p = 0.001$).

Climatic variables on tick and mouse occurrence probabilities

Specific variable importance scores varied among climate niche models for the two species. The white-footed mouse distribution appeared to be primarily determined by mean winter length and mean maximum winter temperatures (Table D2), while black-legged tick distribution was primarily influenced by both minimum and maximum winter temperatures (Table D2). Climate-based species distribution models predicted that white-footed mouse occurrence probability ranged from 0.46 to 0.98 (mean = 0.966, standard deviation = 0.056) across the study area (Figure 5.2B). Meanwhile, the predicted probability of presence for black

legged tick across the study area ranged from 0.08 to 0.98 (mean = 0.911, standard error = 0.008, standard deviation = 0.177) (Figure 5.2C).

I assessed the performance of my climate-based species distribution models by comparing site-specific projected species occurrence probabilities against observed white-footed mouse density (rho = 0.417, p-value = 0.007) and black-legged tick abundance (rho = 0.634, p-value < 0.001). Therefore my climate-based species distribution models were accurate and as such, the projections of species occurrence probabilities from these models were extracted where necessary and used to project Lyme disease infection risk.

Modeling B. burgdorferi *prevalence based on climate and landscape effects on disease host and vector*

My structural equation model was not rejected (p-value = 0.126; Table D5) and revealed that *B. burgdorferi* prevalence at a site was significantly explained only by climate-based black-legged tick occurrence probability (estimated coefficient = 0.39, p = 0.03) (Table 5.3). However, since the model performed acceptably well, explaining 36.2% of variance observed in *B. burgdorferi* prevalence with a p-value > 0.05, I chose to include all three variables in my linear equation for projecting *B. burgdorferi* prevalence in the region.

Estimates of *B. burgdorferi* prevalence based on SEM coefficients were significantly correlated with observed pathogen prevalence (rho = 0.48, p-value = 0.002) in the study area. These results suggest that the parameter coefficients estimated from my SEM model are reliable and may be used to project and extrapolate Lyme disease risk across the study area consisting of over 9043 forest patches.

Projecting Lyme disease infection risk across the study area

My projections of current *B. burgdorferi* prevalence ranged from 0.000 to 0.735 (mean = 0.544, standard deviation = 0.154) across my study area (Figure 5.3A). When I plotted the results of my projections across the study area, I found that estimated measures of *B. burgdorferi* prevalence appeared to be greatest in the southwestern region of the study area, decreasing northwards and eastwards (Figure 5.3A). Ten out of fourteen (71.4%) study sites where *B. burgdorferi* was detected were located in regions where predicted pathogen prevalence ll in the upper 4thquartile of projected risk (Figure 5.3A). A correlation test comparing projected regions of disease risk against human population density based on the 2011 census data from Statistics Canada revealed that disease ecological risk was projected to be higher in regions of high human population density (cor = 0.189, $p < 0.001$) (Figures 5.3A and 5.3B).

DISCUSSION

Zoonotic disease systems are complex and difficult to predict because they are driven by a large number of biotic and abiotic variables (Alexander et al. 2012). These variables may include biological, genetic, ecological, environmental, and socio-economic factors (Wilcox and Gubler 2005, Estrada-Pena et al 2014). However, integrated models can provide insights into the ecological dynamics of a disease system (Plowright et al. 2008), which will help inform public health interventions and guide public health policies (Leach and Scoones 2013). By simplifying the different aspects of a zoonotic disease system into constrained components that are easier to understand, models can answer specific questions about disease system dynamics, emergence, and spatial patterns of risk (Plowright et al. 2008). My model for Lyme disease in this study can be broken down into three components: [1] interspecific interactions (Ostfeld and Keesing 2012, Simon et al. 2014, Turney et al. 2014), [2] landscape effects on species abundances (Brownstein

et al. 2005, Legér et al. 2013), and [3] climate-based species occurrence probabilities (Lane et al. 1991, Legér et al. 2013, Ostfeld 2011).

The integrated model projecting Lyme disease ecological risk across the study area in southern Quebec focuses particularly on how the disease vector and the primary reservoir host respond to climate and landscape factors. This is because both the white-footed mouse and the black-legged tick are essential components of the Lyme disease transmission system (Ostfeld 2011). This is evident from my initial analyses on variables accounting for *B. burgdorferi* prevalence in the study sites (Table 5.1), where I found that pathogen prevalence was dependent on interactions amongst ticks and mice, confirming the conclusions in Simon et al. (2014). However, while the white-footed mouse may play an important role in supporting both vector and pathogen populations within a site, it did not appear to actively contribute to tick movement in this study, nor did the mice experience shared spatial dependence with the ticks as MTILG analyses revealed no significant concordance between either species in the study area.

In addition to host-vector-pathogen interactions, the model also incorporated landscape-heterogeneity. When I examined the effects of landscape and habitat variables on mouse density and tick abundance, I found that white-footed mouse density was influenced by habitat patch size and the relative amount of urban development around the patch. White-footed mouse density tended to increase in smaller forest patches located close to urban environments. Studies have shown that white-footed mouse populations tend to do well in fragmented habitats (Anderson et al. 2003, Brownstein et al. 2005). Meanwhile, increased urban development around their habitat may have contributed to increased mouse density by increasing the amount of resources available to the mouse, or by limiting mouse dispersal out of a habitat patch due to increased resistance to species' movement (Marrotte et al. 2014). None of the landscape variables

examined in this study explained observed tick abundance. It could be that the landscape variables considered in this study were not direct or important contributors to tick survival, or that the variables were measured at the wrong scale. Future studies should incorporate other habitat characteristics important for ticks (e.g. forest types, soil conditions, and biota composition).

Climate also influences multiple components of a zoonotic disease transmission system (Ogden 2014, Ostfeld and Brunner 2015). The effect of climate on the exothermic tick vector is well-established (Legér et al. 2013). Warmer temperatures, greater relative humidity, and steady winter temperatures are most likely to promote tick developmental rates and survival (Wu et al. 2013, Vail and Smith 2002). Similarly, regional winter length and mean temperatures can influence white-footed mouse distribution directly (Roy-Dufresne et al. 2013), or indirectly by altering habitat suitability (Legér et al. 2013, Ogden 2014, Ostfeld and Brunner 2015). My climate-niche models revealed that projected white-footed mouse and black-legged tick occurrence probability across the study area differs substantially. While the white-footed mouse had an evenly high occurrence probability across the study area, tick occurrence probability appeared more constrained.

Using the model and parameters I estimated, I was able to create a map projecting current *B. burgdorferi* prevalence. This map can be representative of disease ecological risk across my study area, with estimated risk being greatest at the US-Canada border, gradually decreasing northwards and eastwards. Disease risk was also higher close to large urban areas (Figure 5.3). This information can aid public awareness and provide an early warning for disease diagnosis by identifying potential regions at greater risk of infection, given exposure and encounter probabilities (e.g. Finch et al. 2014). The model can be useful for other purposes. For example,

we can infer future Lyme disease risk in the region using future climate and landscape scenarios (e.g. Radeloff et al. 2012, Huard et al. 2014). Furthermore, this study also revealed that white-footed mouse density, which directly influences in *B. burgdorferi* prevalence, was determined by landscape and habitat variables. Therefore, land use management policies in southern Québec (e.g. Patz et al. 2004) that result in further forest fragmentation and urban development can potentially increase Lyme disease ecological risk by facilitating greater reservoir host densities. Application of the model can further simulate potential outcomes of ecological perturbations in the region and help policy makers identify the least damaging alternatives. Finally, as climate change continues to modulate disease emergence patterns in conjunction with ecological mechanisms, we can expect white-footed mouse and black legged tick occurrence probabilities to continue to change. Specifically, we expect both species to continue their northwards range expansion over the next 50 years, resulting in species distributions and disease risk that will exceed that predicted in this study (Roy-Dufresne et al. 2013, Simon et al. 2014). Furthermore, changes in climatic conditions can exert substantial evolutionary pressures that promote the increased transmission of more persistent and virulent *B. burgdorferi* strains (Levi et al. 2015).

The model produced in this study was created based on empirical data and I was able to project estimates of *B. burgdorferi* prevalence across my study area. However, since the Lyme disease system is currently emerging across this study area, observed prevalence of the pathogen is expected to be very low. As a result, I cannot confidently claim that, despite the extensive sampling effort, zero observed pathogen prevalence in this study equates to true absences. Nonetheless, the workflow I employed may prove useful for the prediction of infection risk in other disease systems such as Babesiosis, West Nile Virus, or Hantavirus. For a truly comprehensive analysis of changing disease infection risk, disease models should incorporate

patterns of human use of the landscape (Fenichel et al. 2011). This is particularly important, considering how projected regions of greatest disease ecological risk in this study coincided with regions of highest human population density (based on Statistics Canada 2011 Census) in this study. Disease infection dynamics are also dependent on human behaviour, which varies depending on lifestyle, socio-economic conditions, gender, age, psychology, and level of risk awareness or education. Identifying behaviours or lifestyles that increase the chance of contact between humans and a disease system will inform projections of disease infection probabilities (Finch et al. 2014).

CONCLUSION

In this study, I have presented, using the emerging Lyme disease system in Québec as a case study, the importance of a multi-species, integrated approach to disease modeling which enabled me to produce a simple model that successfully and accurately predicted disease infection ecological risk in my study area. By focusing specifically on the disease vector and hosts, their interactions with each other, along with their individual responses to climate and landscape variables, I was able to identify correlations among these variables and pathogen prevalence in the field. I was then able to employ these insights to further predict and project disease ecological risk across the region (Simon et al. 2014). Maps of current pathogen prevalence in this region would assist public health in designing and implementing their programs at a regional scale. Finally, if used in conjunction with appropriate data (e.g. future climate conditions, future land use changes), the model can also potentially project future ecological risk of disease in the region.

REFERENCES

Alexander KA, Lewis BL, Marathe M, Eubank S, Blackburn JK. 2012. Modeling of wildlife-associated zoonoses: Applications and caveats. *Vector Borne Zoonotic Disease* **12**: 1005–1018.

Anderson CS, Cady AB, Meikle DB.2003. Effects of vegetation structure and edge habitat on the density and distribution of white-footed mice (*Peromyscus leucopus*) in small and large forest patches. *Canadian Journal of Zoology* **81**: 897 – 904.

Arino J, Davis JR, Hartley D, Jordan R, Miller JM, van den Driessche P. 2005.A multi-species epidemic model with spatial dynamics. *Mathematical Medicine and Biology* **22**: 129 – 142

Balloux F, Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**: 155 – 165.

Bozick BA, Real LA. 2015. Integrating parasites and pathogens into the study of geographical range limits. *The Quarterly Review of Biology* **90**: 361 – 380.

Brownstein JS, Skelly DK, Holford TR, Fish D. 2005. Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. Oecologia **146**: 469 – 475.

Brunner JL, Killilea M, Ostfeld RS.2012. Overwintering survival of nymphal *Ixodes scapularis* (Acari: Ixodidae) under natural conditions. *Journal of Medical Entomology* **49**: 981 – 987.

Buskirk JV, Ostfeld RS. 1998. Habitat heterogeneity, dispersal, and local risk of exposure to Lyme disease. *Ecological Applications* **8**: 365 – 378.

Chirhart SE, Honeycutt RL, Greenbaum IF. 2005. Microsatellite variation and evolution in the *Peromyscusmaniculatus* species group. *Molecular Phylogenetics and Evolution* **34**: 408 – 415.

Clifford, C. M., Anastos, G., Elbl, A. 1961. The larval Ixodid ticks of the eastern United States (Acarina - Ixodidae). *Miscellaneous Publications of the Entomological Society of America* **2**: 213 – 237

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious disease of wildlife – threats to biodiversity and human health. *Science* **287**: 443 – 449.

Durden LA, Keirans JE. 1996. Nymphs of the genus *Ixodes* (Acari: Ixodidae) of the United States: Taxonomy, identification key, distribution, hosts, and medical/veterinary importance. Entomological Society of America, USA

Engering A, Hogerwerf L, Slingenbergh J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections* **2**: e5

Estrada-Pena A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease risk: an ecological primer. *Trends in Parasitology* **30**: 205 – 214

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

Fagerberg AJ, Fulton RE, Black WC. 2001. Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. *Insect Molecular Biology* **10**: 225 – 236

Falco RC, Fish D. 1991 Horizontal movement of adult *Ixodes dammini* (Acari: Ixodidae) attracted to CO2-baited traps. *Journal of Medical Entomology* **28**: 726 – 729.

Fenichel EP, Castillo-Chavez C, Ceddia MG, Chowell G, Parra PAG, Hickling GJ, Holloway G, Hora R, Morin B, Perrings C, Springborn M, Velazquez L, Villalobos C. 2011. Adaptive human behaviour in epidemiological models. *Proceedings of the National Academy of Science USA* **108**: 6306 – 6311.doi: 10.1073/pnas.1011250108

Finch C, Al-Damluji MS, Krause PJ, Niccolai L, Steeves T, O'Keefe CF, Diuk-Wasser M. 2014.Integrated assessment of behavioural and environmental risk factor for Lyme disease infection on Block Island, Rhode Island. *PLoS ONE* **9**: e84758. doi:10.1371/journal.pone.0084758

Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, Ellison S, Firth D, Friendly M, Gorjanc G, Graves S, Heiberger R, Laboissiere R, Monette G, Murdoch D, Nilsson H, Ogle D, Ripley B, Venables W, Zeileis A. 2015. car: Companion to applied regression. R package 2.0-25

Grenfell B, Harwood J. 1997. (Meta)population dynamics of infectious diseases. *Trends in Ecology and Evolution* **12**: 395 – 399

Government of Canada. 2015. Surveillance of Lyme disease. http://www.healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/lyme/surveillance-eng.php. Assessed July 2015

Haven J, Magori K, Park AW. 2012. Ecological and inhost factors promoting distinct parasite life-history strategies in Lyme borreliosis. *Epidemics* **4**: 152 – 157.

Hayes LE, Scott JA, Stafford III KC. 2015. Influences of weather on *Ixodes scapularis* nymphal densities at long-term study sites in Connecticut. *Tick and Tick-borne Diseases* **6**: 258 – 266.

Hoen AG, Margos G, Bent SJ, Diuk-Wasser MA, Barbour, Kurtenbach K, Fish D. 2009. Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. *Proceedings of the National Academy of Sciences USA* **106**: 15013 – 15018.

Huard D, Chaumont D, Logan T, Sottile M-F, Brown RD, St-Denis BG, Grenier P, Braun M. 2014: A decade of climate scenarios: The Ouranos consortium modus operandi. *Bulletin of the American Meteorological Society* **95**: 1213–1225. doi: http://dx.doi.org/10.1175/BAMS-D-12-00163.1

Institut National de Santé Publique du Quebec. 2014. Proposition d'un programme de surveillance intégré pour la maladie de Lyme et les autres maladies transmises par la tique Ixodes scapularis au Québec.

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH. 2011. Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PLoS ONE* **6**: e25359

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. *Nature* **451**: 990 – 993.

Kaseloo PA, Crowell MG, Heideman PD. 2014. Heritable variation in reaction norms of metabolism and activity across temperatures in a wild-derived population of white-footed mice (*Peromyscus leucopus*).*Journal of Comparative Physiology B* **184**: 525 – 534.

Kulkarni MA, Berrang-Ford L, Buck PA, Drebot MA, Lindsay LR, Ogden NH. 2015. Major emerging vector-borne zoonotic diseases of public health importance in Canada. *Emerging Microbes and Infections* **4**: e33.

Lane RS, Piesman J, Burgdorf W. 1991. Lyme borreliosis: Relation of its causative agent to its vectors and hosts in North America and Europe. *Annual Review of Entomology* **36**: 587 – 609.

Leach M, Scoones I. 2013. The social and political lives of zoonotic disease models: Narratives, science and policy. *Social Science and Medicine* **88**: 10–17

Legér E, Vourc'h G, Vial L, Chevillon C, McCoy KD. 2013. Changing distributions of ticks: causes and consequences. Experimental and Applied Acarology **59**: 219 – 244.

Leo SST, Gonzalez A, Millien V. 2016. Multi-taxa integrated landscape genetics for zoonotic infectious diseases: Deciphering variables influencing disease emergence. Genome. doi: 10.1139/gen-2016-0039

Levi T, Keesing F, Oggenfuss K, Ostfeld RS. 2015. Accelerated phenology of blacklegged ticks under climate warming. *Philosophical Transactions B* **370**: 20130556. doi: 10.1098/rstb.2013.0556

Madhav NK, Brownstein JS, Tsao JI, Fish D. 2004. A dispersal model for the range expansion of black-legged ticks (Acari: Ixodidae). *Entomological Society of America* **41**: 842 – 852.

Margos G, Vollmer SA, Ogden NH, Fish D. 2011.Population genetics, taxonomy, phylogeny, and evolution of *Borrelia burgdorferi* sensu lato. Infection, Genetics and Evolution **11**: 1545 – 1563.

Marrotte RR, Gonzalez A, Millien V. 2014. Landscape resistance and habitat combine to provide an optimal model of genetic structure and connectivity at the range margin of a small mammal. *Molecular Ecology* **23**: 3983 – 3998

Martinez J, Merino S. 2011. Host-parasite interactions under extreme climatic conditions. *Current Zoology* **57**: 390 – 405.

Ogden NH. 2014. Lyme Disease and Climate Change. In: *Climate Change and Global Health*. (Ed. Butler C). CABI, Boston, Massachusetts. pp 328.

Ogden NH, Margos G, Aanensen DM, Drebot MA, Feil EJ, Hanincova K, Schwartz I, Tyler S, Lindsay LR. 2011. Investigation of genotypes of *Borrelia burgdorferi* in *Ixodes scapularis* ticks collected during surveillance in Canada. *Applied and Environmental Microbiology* **77**: 3244 – 3254

Ogden NH, Radojevic M, Wu X, Duvvuri VR, Leighton P, Wu J. 2014. Estimated effects of projected climate change on the basic reproductive number of the Lyme disease vector *Ixodes scapularis*. *Environmental Health Perspectives* **122**: 631 – 638. doi: 10.1289/ehp.1307799

Ostfeld RS. 2011. Lyme disease: The ecology of a complex system. Oxford University Press Inc., Oxford, United Kingdom.

Ostfeld RS, Brunner JL. 2015. Climate change and *Ixodes* tick-borne diseases of humans. *Philosophical Transactions B* **370**: 20140051

Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. 2006.Climate, Deer, Rodents, and Acorns as Determinants of Variation in Lyme-Disease Risk. *PLoS Biol* **4**: e145.

Ostfeld RS, Keesing F. 2012.Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics* **43**: 157-182

Patz JA, Daszak P, Tabor GM, Aguirre AA, Pearl M, Epstein J, Wolfe ND, Kilpatrick AM, Foufopoulos J, Molyneux D, Bradley DJ, Working Group on Land Use Change Disease Emergence. 2004. Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environmental Health Perspectives* **112**: 1092 – 1098. doi: 10.1289/ehp.6877

Patz JA, Olson SH, Uejio CK, Gibbs HK. 2008. Disease emergence from global climate and land use change. *Medical Clinics of North America* **92**:1473–1491

Plowright RK, Sokolow SH, Gorman ME, Daszak P, Foley JE. 2008. Causal inference in disease ecology: Investigating ecological drivers of disease emergence. *Frontiers in Ecology and the Environment* **6**: 420–429. doi: dx.doi.org/10.1890/070086

Pratt NL, Barrett GW. 2012. Timing of breeding in *Ochrotomys nuttalli* and *Peromyscus leucopus* is related to a latitudinal isotherm. *Landscape Ecology* **27**: 599 – 610.

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

Radeloff VC, Nelson E, Plantinga AJ, Lewis DJ, Helmers D, Lawler JJ, Withey JC, Beaudry F, Martinuzzi S, Butsic V, Lonsdorf E, White D, Polasky S. 2012. Economic-based projections of future land use in the conterminous United States under alternative policy scenarios. *Ecological Applications* **22**:1036–1049. http://dx.doi.org/10.1890/11-0306.1

Rogic A, Tessier N, Legendre P, Lapointe FJ, Millien V. 2013. Genetic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Quebec. *Ecology and Evolution* **3**: 4172 – 4182.

Roy-Dufresne E, Logan T, Simon JA, Chmura GL, Millien V. 2013. Poleward expansion of the white-footed mouse (*Peromyscus leucopus*) under climate change: Implications for the spread of Lyme disease. *PLoS ONE* **8**: e80724

Sambrook J, Fritsch EF, Maniatis T. 1989. Chapter 6 – Preparation and analysis of eukaryotic genomic DNA. In: *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schmidt CA. 1999. Variation and congruence of microsatellite markers for *Peromyscus leucopus*. *Journal of Mammalogy* **2**: 522 – 529

Simon JA, Marrotte RR, Desrosier N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of *Borrelia burgdorferi*, the agent of Lyme disease, at the northeastern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457 – 462

Suzán G, García-Peña GE, Castro-Arellano I, Rico O, Rubio AV, Tolsa MJ, Roche , Hosseini PR, Rizzoli A, Murray KA, Zambrana-Torrelio C, Vittecoq M, Bailly X, Aguirre AA, Daszak P, Prieur-Richard A-H, Mills JN, Guegan J-F. 2015. Metacommunity and

phylogenetic structure determine wildlife and zoonotic infectious disease patterns in time and space. *Ecology and Evolution* **5**: 865 – 873.

Tabachnick WJ. 2010. Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. *The Journal of Experimental Biology* **213**: 946 – 954.

Taylor LH, Latham SM, Woolhouse ME.2001. Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B* **356**: 983 – 989

Thuiller W, Georges D, Engler R. 2014. biomod2: Ensemble platform for species distribution modeling. R package 3.1-64

Turney S, Gonzalez A, Millien V. 2014. The negative relationship between mammal host diversity and Lyme disease incidence strengthens through time. *Ecology* **95**: 3244-3250.

Urban JM, Cryan JR. 2012. Two ancient bacterial endosymbionts have coevolved with the plant hoppers (Insecta: Hemiptera: Fulgoroidea). *BMC Evolutionary Biology* **12**: 87.

Vail SG, Smith G. 2002. Vertical movement and posture of black-legged tick (Acari: Ixodidae) nymphs as a function of temperature of temperature and relative humidity in laboratory experiments. *Journal of Medical Entomology* **39**: 842 – 846

Walsh C, Mac Nally R. 2013. hier.part: Hierarchical partitioning. R package 1.0-4

Wilcox BA, Gubler DJ. 2005. Disease ecology and the global emergence of zoonotic pathogens. *Environmental Health and Preventive Medicine* **10**: 263 – 272

Wright S. 1950. Genetical structure of populations. *Nature* **166**: 247 – 249

Wu X, Duvvuri VR, Lou Y, Ogden NH, Pelcat Y, Wu J. 2013.Developing a temperature-driven map of the basic reproductive number of the emerging tick vector of Lyme disease *Ixodes scapularis* in Canada. Journal of Theoretical Biology **319**: 50 – 61.

TABLES

Table 5.1. Summary of results from hierarchical partitioning analyses examining the independent effects of each explanatory variable on *Borrelia burgdorferi* prevalence across 40 sites in southern Québec, Canada. 'I' is the percentage of explained variance accounted for by each variable and its significance is estimated by calculating Z-scores over 1000 randomisations. Correlation tests were subsequently performed for variables with statistically significant 'I' to examine their relations with *B. burgdorferi* prevalence in the study area. Statistically significant values are in bold and the top four contributors are further underlined.

Variables examined	***Borrelia burgdorferi* prevalence**		
	I (%)	**Z-scores**	**Correlation**
Number of questing ticks	**12.848**	**3.51**	**0.7711**
Number of feeding ticks	5.026	0.57	N.A
Total number of ticks	**15.327**	**4.21**	**0.7499**
Tick prevalence on small mammals	**8.320**	**1.72**	**0.6416**
Tick abundance on small mammals	**8.611**	**1.84**	**0.5134**
Tick prevalence on white-footed mice	**12.822**	**3.55**	**0.7034**
Tick abundance on white-footed mice	**11.079**	**2.88**	**0.6354**
Density of small mammals	4.968	0.48	N.A
Density of white-footed mice	7.242	1.42	N.A
Effective species richness	5.955	0.86	N.A
Species richness	3.165	-0.10	N.A
Proportion of WFM to small mammals	4.666	0.39	N.A

Table 5.2. Summary of results from redundancy analyses (RDA) performed to identify landscape and habitat characteristics that significantly account for white-footed mouse density and black-legged tick abundance across 40 sites in southern Québec, Canada. Statistically significant values are in bold.

Variables examined	**White-footed mouse density** [**F=2.6096, p=0.02875**]		**Black-legged tick abundance** [F=1.0332, p=0.34]	
	F	**p**	**F**	**p**
Area (meters-squared)	**7.4107**	**0.009**	1.3412	0.224
Perimeter (meters)	0.1539	0.673	0.0070	0.896
MDNN	3.2742	0.077	3.5500	0.078
% Agriculture	1.1091	0.283	0.3448	0.554
% Urban	**8.7389**	**0.009**	0.9266	0.343
% Forest	2.7050	0.118	0.0295	0.840
Connectivity (Road)	0.0848	0.778		
Connectivity (Agriculture)	0.0037	0.951		
Connectivity (Urban)	0.0056	0.948		

Table 5.3. Estimated coefficients obtained from a Structural Equation Model analysis quantifying the relative contributions of climate and landscape to *Borrelia burgdorferi* prevalence across the study area in southern Québec, Canada (see Figure 1). The effects of climate and landscape were represented as predicted species occurrence probabilities (based on climate-based species distribution models) and estimated species densities (from specific landscape variables). Statistically significant values are in bold.

Regressions *B. burgdorferi* prevalence ~	**Estimate**	**Std. Error**	**Z-value**	**P-value**
Mouse occurrence probability (climate)	0.309	0.221	1.395	0.163
Tick occurrence probability (climate)	**0.391**	**0.180**	**2.167**	**0.030**
Estimated mouse density (landscape)	0.067	0.258	0.259	0.795

FIGURES

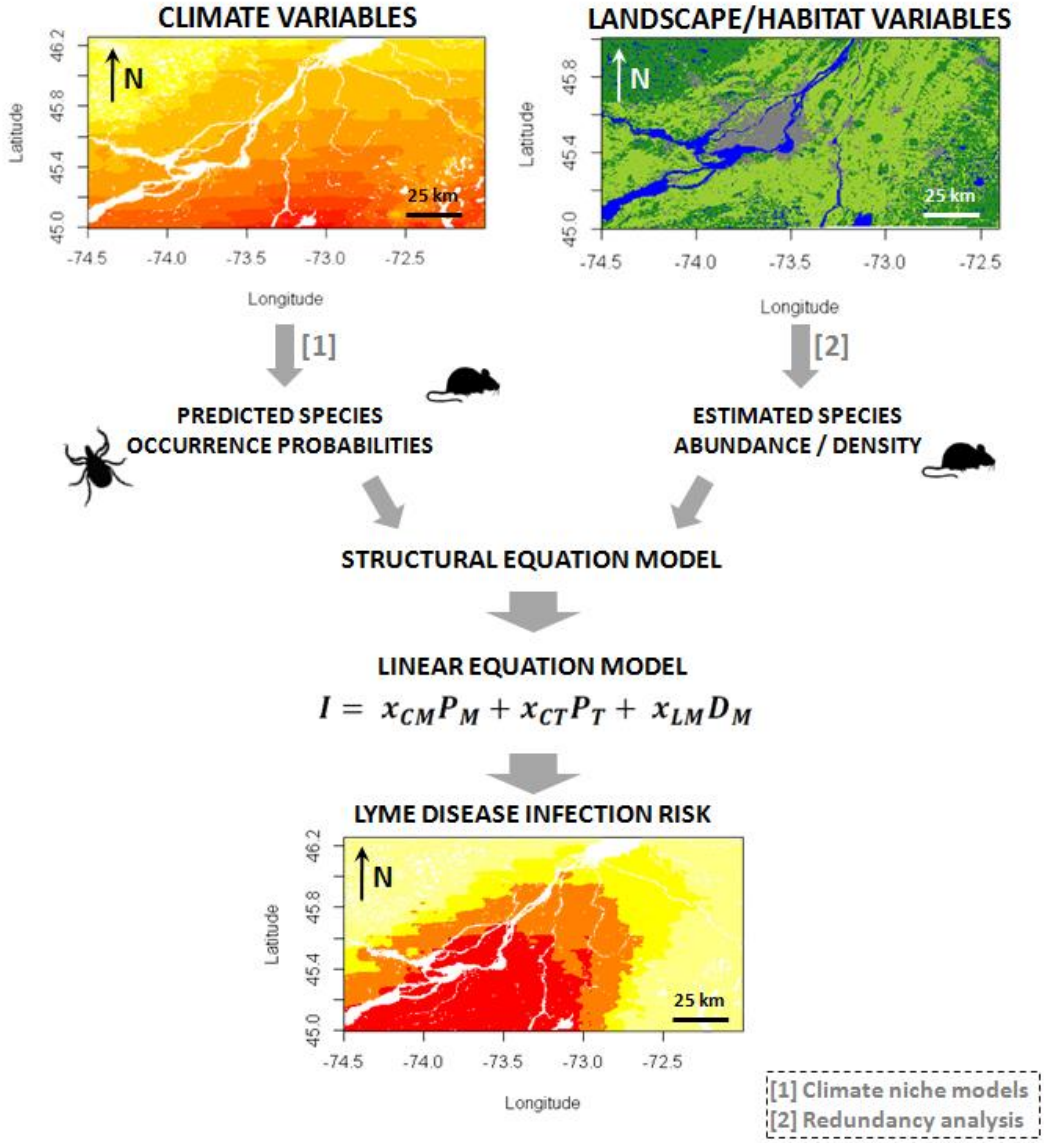


Figure 5.1. The workflow employed in this study. Focusing on the study area, I first predicted species occurrence probabilities using climate niche models and estimated species abundance or density using landscape and habitat variables. Information from these analyses (i.e. species occurrence probabilities and estimated density) were subsequently entered into a structural equation model to assess the indirect contributions of climate and landscape to *Borrelia burgdorferi* prevalence among 40 study sites across southern Québec Canada through their influences on white-footed mouse and black-legged tick occurrence probabilities and estimated mouse density. The resulting estimated coefficients were used to extrapolate Lyme disease infection risk in the form of *B. burgdorferi* prevalence in all 9043 forest patches across the study area.

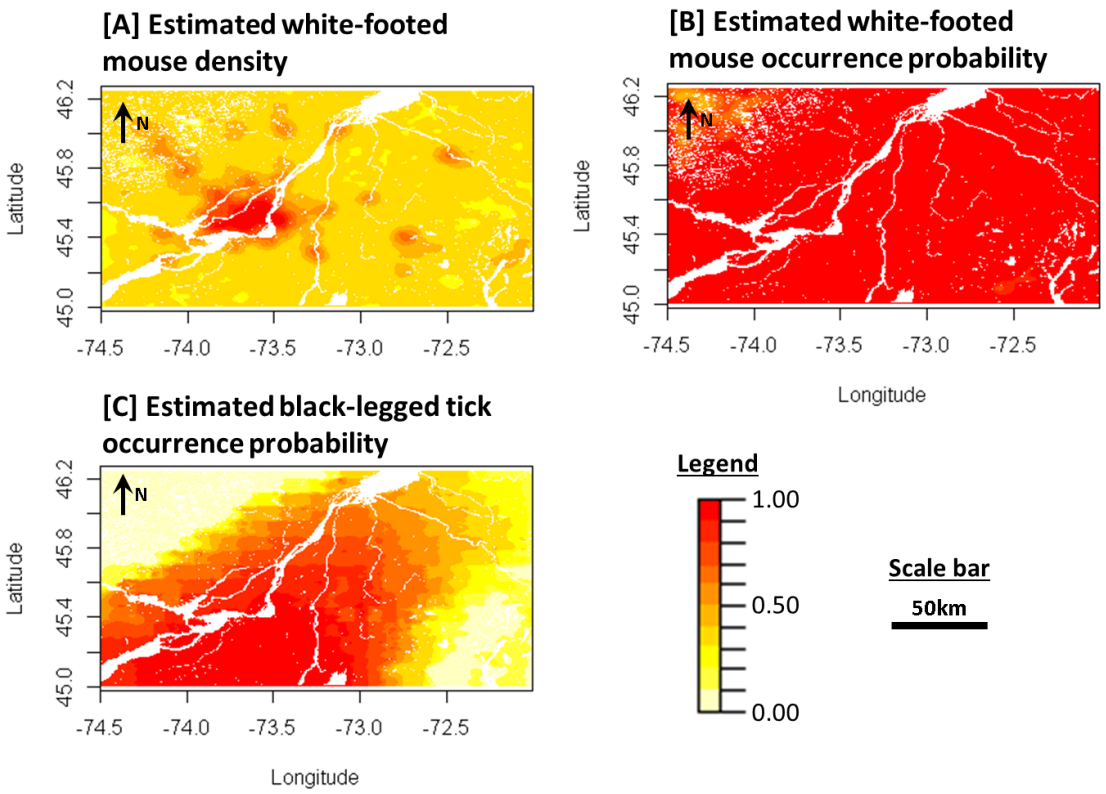


Figure 5.2. Plots of [A] white-footed mouse density across the study area estimated from landscape variables, [B] white-footed mouse, and [C] black-legged tick occurrence probabilities estimated from climate based species distribution models. The scale bar represents a distance of approximately 50km on each map.

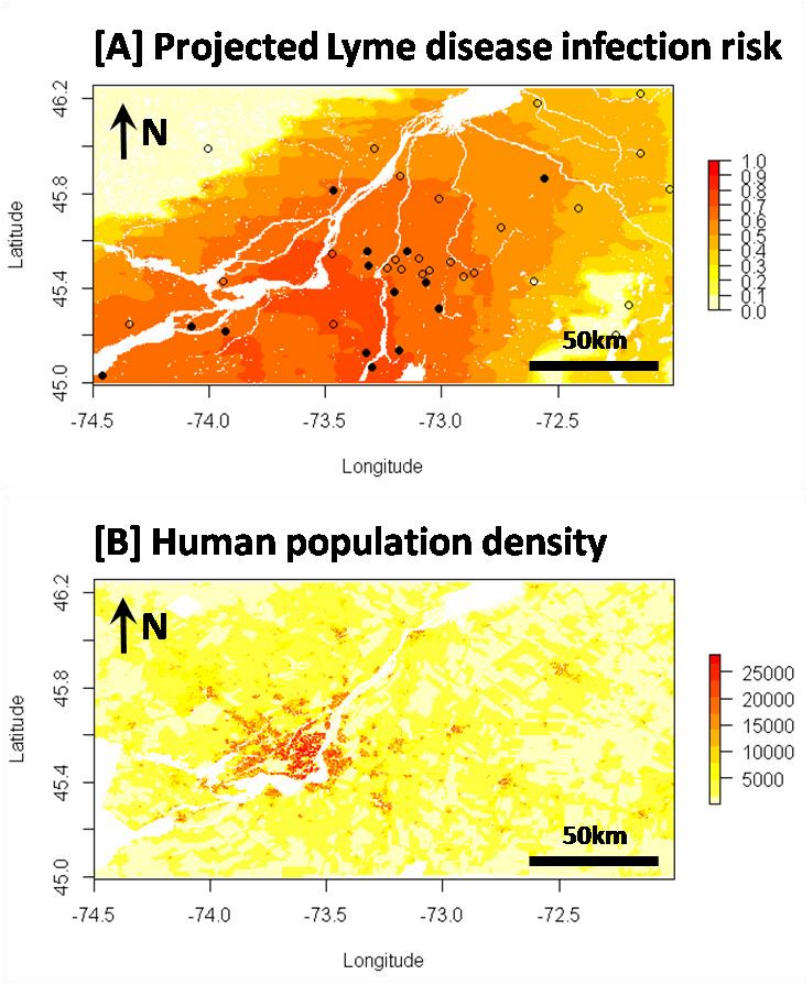


Figure 5.3. Plots of [A] projections of Lyme disease infection risk (in the form of estimated *Borrelia burgdorferi* prevalence) calculated from my linear equation model, and [B] human population density in the study area based on the 2011 census data from Statistics Canada. Population density was calculated for each census dissemination block by dividing the total population by land area (in square kilometers). The scale bar represents a distance of approximately 50km on each map.

CHAPTER SIX

General Conclusions

THESIS SUMMARY

In the past seven decades, we have witnessed over 300 cases of emerging infectious disease (EIDs) across the world. EIDs can result from the evolution of novel species or genetic strains of known pathogens (Schrag and Wiener 1995, Daszak et al. 2000), as an unintentional outcome of changing ecological factors (Patz et al. 2008), or simply by overcoming geographical resistance and spreading into populations in which individuals had no prior exposure to the pathogen (Taylor et al 2001). Approximately 70% of identified emerging infectious diseases have zoonotic origins (Jones et al. 2008). In many cases, these zoonotic disease transmission systems involve multiple species, each with their own sets of complex interactions with each other, and their environment (Estrada-Peña et al. 2014, Engering et al. 2013). Studies that identify and quantify such interactions among disease pathogens, vectors, hosts, and their environment will provide insights on mechanisms and processes that facilitate disease emergence. Such information is crucial for informed management decisions prioritized for disease mitigation, prevention, or surveillance (Rosenthal et al. 2013).

In this thesis, I focused on an multi-taxa, integrated, approach using Lyme disease emergence in Québec as a case study. I did this by investigating the mechanisms, processes, and variables influencing the range expansion of two major members of the disease system (the black-legged tick and the white-footed mouse) in a region of highly dynamic disease emergence. I first presented an integrated molecular approach for comparing the population structures of the primary disease vector and reservoir host species in Chapter 2 (Sork and Waits 2010). The approach – Multi-taxa integrated landscape genetics (MTILG) – integrates concepts and tools

from population genetics, landscape ecology, and spatial statistics (James et al. 2011). Using simulated genetic data, I showed that MTILG can ascertain combinations of biotic and abiotic variables that influence zoonotic disease emergence within a region. I was able to identify shared spatial occurrences and interspecies-dependent dispersal patterns as well as quantify how interactions among landscape heterogeneity, geographical barriers, and species dispersal rates influence the spatial extent and rate of disease transmission system spread through time.

I then examined the population genetics of the disease reservoir host (the white-footed mouse, *Peromyscus leucopus*) and the disease vector (the black-legged tick, *Ixodes scapularis*) individually in Chapters 3 and 4, respectively. Using sets of species-specific microsatellite loci, I investigated white-footed mouse and black-legged population structures in southern Québec, quantified species interactions with the landscape, uncovered the mechanism of species movement across the region, and determined stages of population establishment for the black-legged tick.

I found that white-footed mouse populations in southern Québec exhibit significant genetic structuring and differentiation (Chapter 3). Additionally, gene flow among mouse populations was affected by landscape variables such as linear barriers to migration (as seen in Rogic et al. 2013, Marrotte et al. 2014). However, the effects of these barriers can be potentially circumvented, most likely by human activities. This suggests that human activities and the white-footed mouse's natural ability to navigate landscape heterogeneity will likely continue to facilitate the northward range expansion of the species and any pathogens it hosts. I also found evidence of gradients in morphological variation (e.g. hind-foot length) in the direction of the species range expansion, which may be indicative of the potential for greater white-footed mouse dispersal rates in populations located at the edge of the species' range (e.g. Philips et al. 2006,

Sparrow 2015). Such trends warrant further research. I propose that future Lyme disease monitoring and control strategies in southern Québec should continue to focus on the surveillance of white-footed mouse distribution range as well as incorporate predictive models of the species' potential distribution range based on its ability to navigate a heterogeneous landscape.

In Chapter 4, I investigated the genetic variation present within populations of black-legged ticks in southern Québec. By quantifying both spatial and temporal patterns of genetic diversity and structure among populations, I was able to detect several genetic patterns that were indicative of high frequencies of long-distance dispersal events (Bialozyt et al. 2006, Excoffier et al 2009, Cristescu 2015). This lends support to the hypothesis by Ogden et al. (2008, 2015) that migratory birds play important roles in the south-north movement of ticks into Canada. Additional analyses of temporal genetic variation within black-legged tick populations in this study further allowed me to determine stages of establishment for each tick population. Such information is important for public health officers to identify regions at greatest disease risk, because it is crucial for diagnostic and preventative efforts. This molecular approach can not only replace current labour-intensive and time-consuming surveillance efforts in the field, it will likely be more accurate (Gubler 2010, Biek and Real 2010, Muellner et al. 2011).

Finally in Chapter 5, I integrated species genetic information, along with climate and landscape effects on host and vector distribution and abundance to determine how these variables influence *Borrelia burgdorferi* prevalence in my study sites. From there I produced a model to project pathogen prevalence across a region, in the process empirically uncovering several critical aspects of modeling disease risk. First, disease models should use a multispecies approach to obtain more accurate risk projections (Estrada-Peña et al. 2014). Secondly, disease

emergence models should avoid focusing only on a single predictor variable, as doing so risks omitting the effects of variables that operate on other spatial scales (Raghaven et al. 2013, Riley et al. 2015). Finally, it is important to note that the relative contributions by climate, landscape, and species variables changes depending on whether models focus on disease presence versus prevalence as evidence in this study and that by Simon et al. 2014.

FUTURE WORK

While I have examined interactions among two species recognised as major players in the North American Lyme disease transmission system, there is a plethora of other potential species involved in the system that, while commonly overlooked, still contribute substantially to the disease system (Ostfeld 2011). Some of these species include alternate tick and pathogen hosts (i.e. the eastern chipmunk and the short-tailed shrew), critical reproductive hosts for the disease vector (i.e. the white-tailed deer), as well as migratory birds, which I have identified as the primary mode of tick transportation into southern Québec (Ostfeld 2011, Ogden et al. 2008, Ogden et al. 2015). More importantly, future integrated studies on the Lyme disease system should also incorporate genetic information from the disease pathogen itself – *Borrelia burgdorferi*. The potential presence of coupled interactions and co-evolution among strains of *B. burgdorferi* and genetic clusters of vector or hosts can possibly influence disease emergence rates, spread patterns, virulence risk, or transmission efficacy (Margos et al. 2011, Hamer et al. 2012, Ostfeld and Brunner 2015). Due to logistical limitations I was unable to include data from alternate host species and the *Borrelia burgdorferi* pathogen itself in this thesis project. Nonetheless, as I have emphasised repeatedly throughout this thesis, it is important to employ multi-species, integrated approaches to disease modeling, because they incorporate key species

and their individual, population, and community responses to climate and landscape variables (Johnson et al. 2015). Disease risk models that further incorporate the human factors will certainly result in even more powerful and reliable prediction of disease risk.

This thesis is the first step in the application of an integrated approach to studying disease emergence patterns and mechanisms. The methods employed in this study had successfully provided critical insights regarding underlying processes and variables that influence Lyme disease emergence in my study area. It stands to reason that the approach may be applied to other disease systems across the world (e.g. Lyme disease in Europe, the Zika pandemic in the Americas). By applying a standardized approach for the study of multiple disease systems, we can begin to identify and understand, based on insights obtained from observed similarities or differences among different systems, overarching factors influencing different types of disease emergence events. This knowledge will contribute substantially to our ability to mitigate, prevent, monitor, and predict different types of infectious diseases as they emerge worldwide (Rosenthal et al. 2015).

REFERENCES

Bialozyt R, Ziegenhagen B, Petit RJ. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**: 12 – 20.

Biek R, Real LA. 2010 The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* **19**: 3515 – 3531.

Cristescu ME. 2015. Genetic reconstruction of invasion history. *Molecular Ecology* **24**: 2212 – 2225.

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious disease of wildlife – threats to biodiversity and human health. *Science* **287**: 443 – 449.

Engering A, Hogerwerf L, Slingenbergh J. 2013. Pathogen-host-environment interplay and disease emergence. *Emerging Microbes and Infections* **2**: e5.

Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014. Effects of environmental change on zoonotic disease risks: An ecological primer. *Trends in Parasitology* **30**: 205 – 214

Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansion. *Annual Review of Ecology, Evolution, and Systematics* **40**: 481 – 501.

Gubler DJ. 2010. The global threat of emergent/reemergent vector-borne diseases. In: Vector Biology, Ecology and Control (ed. Atkinson PW). Springer, Netherlands. pp. 39 – 62.

Hamer SA, Hickling GJ, Sidge JL, Walker ED, Tsao JI. 2012. Synchronous phenology of juvenile *Ixodes scapularis*, vetertebrate host relationships, and associated patterns of *Borrelia burgdorferi* ribotypes in the mid-western United States. *Ticks and Tick-borne Diseases* **3**: 65 – 74.

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH. 2011. Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PLoS ONE* **6**: e25359.

Johnson PTJ, de Roode JC, Fenton A. 2015. Why infectious disease research needs community ecology. *Science* **349**: 1259504-1 – 1259504-9.

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. *Nature* **451**: 990 – 993.

Margos G, Vollmer SA, Ogden NH, Fish D. 2011. Population genetics, taxonomy, phylogeny, and evolution of *Borrelia burgdorferi* sensu lato. *Infection, Genetics and Evolution* **11**: 1545 – 1563.

Marrotte RR, Gonzalez A, Millien V. 2014. Landscape resistance and habitat combine to provide an optimal model of genetic structure and connectivity at the range margin of a small mammal. *Molecular Ecology* **23**: 3983 – 3998.

Muellner P, Zadoks RN, Perez AM, Spencer SEF, Schukken YH, French NP. 2011. The integration of molecular tools into veterinary and spatial epidemiology. *Spatial and Spatio-temporal Epidemiology* **2**: 159 – 171.

Ogden NH, Barker IK, Francis CA, Heagy A, Lindsay LR, Hobson KA. 2015. How far north are migrant birds transporting the tick *Ixodes scapularis* in Canada? Insights from stable hydrogen isotope analyses of feathers. *Ticks and Tick Borne Diseases* **6**: 715 – 20.

Ogden NH, Lindsay LR, Hanicová K, Barker IK, Bigras-Poulin M, Charron DF, Heagy A, Francis CM, O'Callaghan CJ, Schwartz I, Thompson RA. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and *Borrelia burgdorferi* and

Anaplasma phagocytophilum in Canada. *Applied and Environmental Microbiology* **74**: 1780 – 1790.

Ostfeld RS. 2011. Lyme disease: The ecology of a complex system. Oxford University Press, Oxford, United Kingdom.

Ostfeld RS, Brunner JL. 2015. Climate change and *Ixodes* tick-borne diseases of humans. *Philosophical Transactions B* **370**: 20140051.

Patz JA, Olson SH, Uejio CK, Gibbs HK. 2008. Disease emergence from global climate and land use change. *Medical Clinics of North America* **92**:1473–1491.

Philips BL, Brown GP, Webb JK, Shine R. 2006. Invasion and the evolution of speed in toads. *Nature* **439**: 803.

Raghavan RK, Brenner KM, Harrington JA, Higgins JJ, Harkin KR. 2013. Spatial scale effects in environmental risk-factor modelling for diseases. *Geospatial Health* **7**:169-82.

Riley S, Eames K, Isham V, Mollison D, Trapman P. 2015. Five challenges for spatial epidemic models. *Epidemics* **10**: 68 – 71.

Rogic A, Tessier N, Legendre P, Lapointe F-J, Millien V. 2013. Genetic structure of the white-footed mouse in the context of the emergence of Lyme disease in southern Québec. *Ecology and Evolution* **3**: 2075 – 2088.

Rosenthal SR, Ostfeld RS, McGarvey ST, Lurie MN, Smith KF. 2015. Redefining disease emergence to improve prioritization and macro-ecological analyses. *One Health* **1**: 17 – 23.

Schrag SJ, Wiener P. 1995. Emerging infectious disease: What are the relative roles of ecology and evolution? *Trends in Ecology and Evolution* **10**: 319 – 324.

Simon JA, Marrotte RR, Desrosier N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of Borrelia burgdorferi, the agent of Lyme disease, at the northeastern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764.

Sork VL, Waits L. 2010. Contributions of landscape genetics–approaches, insights, and future potential. *Molecular Ecology* **19**: 3489-3495.

Sparrow L. 2015. The effect of limb length on the cost of locomotion and gait in the Longshanks mouse. M.Sc. Thesis. University of Calgary, Canada.

Taylor LH, Latham SM, Woolhouse ME. 2001. Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **356**: 983 – 989.

APPENDICES

APPENDIX A

Table A1. Summary of parameters used in PEDAGOG simulations. All parameters were kept constant across populations and genetic loci across all simulations.

Parameter Areas	**Specific parameters**	**Details**
Genetic Diversity		Total number of loci = 10 Locus repeat units = 2 Number of alleles per loci = random Allele frequencies = Normal distribution Loci linkage rates = 0
Demographics: Population sizes	Cohort size	Abundance = Geometrically determined Initial cohort size = 1 Lambda = 1.2
	Bottleneck	No bottleneck events specified
	Sex proportion	Distribution = Normal Mean = 0.5 Standard deviation = 0.0
	Density dependence	Carrying capacity = 250 (constant) Ages to include in population total = All
Demographics: Demographic rates	Age-based survival, maturation, capture and movement probabilities *Kept equal for both males and females*	Number of age classes = 5 Maturation probability: Age class 0: Mean = 0.0, standard deviation=0.0 Age class 1 - 4: Mean = 1.0, standard deviation=0.0 Mature survival probability: Age class 0 - 4: Mean = 0.75, standard deviation=0.0 Immature survival probability: Age class 0 - 4: Mean = 0.75, standard deviation=0.0 Capture probability: Age class 0 - 4: Mean = 0.5, standard deviation=0.0 Movement probability: Age class 0 - 4: Mean = 1.0, standard deviation=0.0
Demographics: Distance	Pairwise distance among populations	Calculated from population spatial coordinates
Demographics: Movment	Movement probability for emmigrators	Depends on species and landscape scenario being simulated
Demographics: Sampling	Number of samples between reproductive events	N = 1
	Sample times	Sample timing distribution: Normal, mean = 0.5, standard deviation = 0.0
	Number of generations	N = 30
	Initialize simulations using population pools	Initial population pool distribution: Normal, mean = 10000, standard deviation = 0.0
	Burn-in	Number of generations = 5
Mating design	*None specified*	-
Mutation/Error: Genetic error and mutation	Mutation rates	Mutation rate: Normal distribution, mean = 1.0, standard deviation = 1.0 Mutation source: 50% IAM, 50% SMM
	Error rates	Micsall rate: Normal distribution, mean = 1.0, standard deviation = 1.0 Miscall source: 50% adjacent bin, 50% false allele
Mutation/Error: Demographicl error	Age-based cohort misclassification rate	Error probability distribution (Age class 0 - 4): Normal, mean = 0.5, standard deviation=0.0
Individual growth	*None specified*	-
Heritability/Selection	*None specified*	-

Table A2. ANOVA tests on the effects of landscape heterogeneity, geographical barriers, and species dispersal rates on disease emergence extent, time taken for disease emergence to reach equilibrium, and rates of disease emergence across simulation scenarios. Selection tests were then performed to identify the best models describing variances between simulation scenarios. Results from variance partitioning (var.part) and the absolute relative importance (rel.impt) of each predictor variable on all three measures of disease emergence are presented.

	Extent of disease spread		**Time taken to reach equilibrium**		**Emergence rate**	
Differences among simulation scenarios						
ANOVA	F	p-value	F	p-value	F	p-value
	41.53	< 0.001	6.636	< 0.001	11.19	< 0.001
Selection tests						
	Model	R^2_{adj}	Model	R^2_{adj}	Model	R^2_{adj}
	~ Heterogeneity + Barrier + Dispersal	0.714	~ Heterogeneity + Dispersal	0.342	~ Heterogeneity + Barrier	0.396
Relative importance (rela.impt) and variance partitioning (var.part)						
	rela.impt	var.part	rela.impt	var.part	rela.impt	var.part
Heterogeneity	48.50%	34.33%	94.91%	33.03%	82.74%	33.17%
Barrier	44.77%	31.59%	-	-	17.25%	0.059%
Dispersal	0.067%	0.036%	0.051%	0.007%	-	-

Table A3. Average statistically significant global R_{ST} values for each species in each simulation along with standard deviations in square brackets. The proportions of replicates in which R_{ST} values are statistically significant are provided. Lowercase alphabets indicate species in each scenario that shared similar R_{ST} values. Simulation scenarios are labelled as in Table 2.1.

	# significant R_{ST}	Average R_{ST} values [SD]	
Scenario A			
H1	100%	0.029 [0.005]	a_
H2	100%	0.026 [0.004]	a_
P_1	100%	0.027 [0.004]	a_
V_1	100%	0.028 [0.003]	a_
Scenario B			
H1	100%	0.043 [0.004]	a_
H2	100%	0.028 [0.003]	_b
P_1	100%	0.042 [0.003]	a_
V_1	100%	0.041 [0.004]	a_
Scenario C			
H1	100%	0.019 [0.003]	a_
H2	100%	0.024 [0.002]	_b
P_1	100%	0.018 [0.002]	a_
V_1	100%	0.018 [0.003]	a_
Scenario D			
H1	100%	0.039 [0.006]	a_
H2	100%	0.024 [0.002]	_b
P_1	100%	0.035 [0.004]	a_
V_1	100%	0.034 [0.004]	a_
Scenario E			
H1	100%	0.029 [0.005]	a_
H2	100%	0.026 [0.004]	a_
P_2	100%	0.027 [0.004]	a_
V_2	100%	0.025 [0.004]	a_
Scenario F			
H1	100%	0.043 [0.004]	a_
H2	100%	0.028 [0.003]	_b
P_2	100%	0.030 [0.004]	_b
V_2	100%	0.032 [0.006]	_b
Scenario G			
H1	100%	0.019 [0.003]	a_
H2	100%	0.024 [0.002]	_b
P_2	100%	0.024 [0.005]	_b
V_2	100%	0.024 [0.003]	_b
Scenario H			
H1	100%	0.039 [0.006]	a_
H2	100%	0.024 [0.002]	_b
P_2	100%	0.025 [0.004]	_b
V_2	100%	0.025 [0.005]	_b

Table A4. Summary of model selection on constrained ordination models that described the influences of landscape variables (barrier, heterogeneity, others) on species spatial genetic patterns. The frequency at which a variable is selected for each species is shown, along with the percentage of genetic variance explained by said variable. Simulation scenarios are labelled as in Table 2.1. "-" indicate that the variable(s) in question did not significantly explained observed genertic patterns for a species in a particular simulation scenario.

	Barrier		**Heterogeneity**		**Barrier + Heterogeneity**		**Euclidean Distance**	
	Frequency	**% Var**	**Frequency**	**% Var**	**Frequency**	**% Var**	**Frequency**	**% Var**
Scenario A								
H1	-	-	-	-	-	-	-	-
H2	-	-	-	-	-	-	-	-
P_1	-	-	-	-	-	-	-	-
V_1	-	-	-	-	-	-	-	-
Scenario B								
H1	-	-	-	-	-	-	-	-
H2	-	-	0.2	0.0098	-	-	-	-
P_1	-	-	0.1	0.0191	-	-	-	-
V_1	-	-	0.3	0.0399	-	-	-	-
Scenario C								
H1	0.9	0.0686	-	-	-	-	-	-
H2	1	0.1580	-	-	-	-	-	-
P_1	0.7	0.0517	-	-	-	-	-	-
V_1	1	0.0120	-	-	-	-	-	-
Scenario D								
H1	0.8	0.1118	-	-	0.2	0.2728	-	-
H2	0.9	0.0164	-	-	-	-	-	-
P_1	0.6	0.1104	-	-	0.1	0.2147	-	-
V_1	0.3	0.1156	-	-	0.4	0.3224	-	-
Scenario E								
H1	-	-	-	-	-	-	-	-
H2	-	-	-	-	-	-	-	-
P_2	-	-	-	-	-	-	-	-
V_2	-	-	-	-	-	-	-	-
Scenario F								
H1	-	-	0.4	0.0580	-	-	-	-
H2	-	-	0.2	0.0098	-	-	-	-
P_2	-	-	0.1	0.0067	-	-	-	-
V_2	-	-	0.2	0.0069	-	-	-	-
Scenario G								
H1	0.9	0.0686	-	-	-	-	-	-
H2	1	0.0120	-	-	-	-	-	-
P_2	0.9	0.0113	-	-	-	-	-	-
V_2	1	0.0116	-	-	-	-	-	-
Scenario H								
H1	0.8	0.1118	-	-	0.2	0.2728	-	-
H2	0.9	0.0164	-	-	-	-	-	-
P_2	0.9	0.0174	-	-	-	-	-	-
V_2	0.2	0.0184	-	-	-	-	-	-

APPENDIX B

TABLES

Table B1. Total number of white-footed mouse specimens sampled for each study site. Global Positioning System (GPS) coordinates for each site are also presented.

Code	Longitude	Latitude	Sampling years							Total
			2007	2008	2009	2011	2012	2013	2014	
DV	-72.56	45.86	-	-	-	5	12	12	7	36
HR	-73.31	45.49	-	-	-	31	-	5	-	36
HV	-73.21	45.11	-	-	-	5	-	9	-	14
LP	-73.46	45.54	-	-	-	-	-	-	31	31
MA	-73.94	45.43	-	-	-	-	-	-	24	24
MR	-73.06	45.49	-	-	17	18	-	-	-	35
SB	-73.34	45.55	-	-	32	-	-	-	13	45
MSH	-73.16	45.54	14	8	16	-	-	-	-	38
MY	-72.86	45.45	-	-	9	12	-	-	-	21
NO	-73.28	45.06	-	-	-	7	11	20	-	38
PDC	-73.09	46.30	-	-	-	5	8	21	-	34
SE	-73.92	45.21	-	-	-	6	-	13	24	43
SF	-74.45	45.03	-	-	-	7	-	4	19	30
SJ	-73.46	45.23	-	-	-	2	-	17	-	19
SL	-72.74	45.65	-	-	-	-	-	-	22	22
SP	-74.39	45.32	-	-	-	-	-	-	19	19

Table B2. List of microsatellite primers used to genotype *Peromyscus leucopus* specimens in this study. Information on repeat motifs, primer sequences, observed size ranges and literature references is provided.

Loci Name	Repeat Motif	Forward / Reverse Primer Sequences (3' – 5')	Size Ranges	References
Pml01	CA	*F:* CATTCAAGACCTGGCTTTTT *R:* TGGGTTTCATCAGTGCTTCT	149 – 205	Chirhart et al. 2005
Pml03	CA	*F:* GCCATTAGTCTATGTGACAG *R:* GCGATGTACCCAGAAAT	231 – 259	Chirhart et al. 2005
Pml04	CA	*F:* CATAAGGTGGCTCGGAATCA *R:* CAGGAAGGGGAAATGACCAT	186 – 244	Chirhart et al. 2005
Pml05	CA	*F:* CTGAGCCAAAAGTGGTCCTT *R:* TGAAGACAGCCCCTCTCTG	182 – 248	Chirhart et al. 2005
Pml06	CA	*F:* CAGGGCTGTAGAGGGAGAAC *R:* ACTGGAGCAGAGGCATTTG	142 – 178	Chirhart et al. 2005
Pml09	CA	*F:* GAATCCATACACCCATGC *R:* TTGCTTTTCGTCAAGTTTT	186 – 226	Chirhart et al. 2005
Pml11	CA	*F:* ACCCCCGAGTGCTGAGATT *R:* TTTGCTGCTTTCCCCAGAGA	218 – 264	Chirhart et al. 2005
Pml12	CA	*F:* GCAGCCTGTATTCTCTCACA *R:* GCCAACCATTTCTTCAAGTG	141 – 167	Chirhart et al. 2005
PLGT58	GT	*F:* GATCTTGTGAACACGCTTCT *R:* TTGATGGCTCTGGAGAGGCT	132 – 184	Schmidt 1999
PLGT66	GT	*F:* CTCTGTCTGCCACACATGCT *R:* GTGCCATCACAGATGTGACA	129 – 165	Schmidt 1999
GATA70	GATA	*F:* CTTGGTATGCATCGCCATCT *R:* TAATCTCTGTAGCTTCATGT	252 – 324	Schmidt 1999

Table B3. List of connectivity-based predictor variables used in constrained ordinations.

Data	Source	Spatial resolution
Roads	GeoBase - National Road Network	1 : 10,000
Water bodies	GeoBase - Land Cover, circa 2000	1 : 50,000
Agriculture	GeoBase - Land Cover, circa 2000	1 : 50,000
Urban / Developed	GeoBase - Land Cover, circa 2000	1 : 50,000

Table B4. Summary of *Borrelia burgdorferi* and *Ixodes scapularis* (black-legged ticks) presences in each site included in the study. The presence/absence of *B. burgdorferi* in each site was determine from pathogen screening results on white-footed mouse, black-legged ticks, and other small mammals sampled from the site. *B. burgdorferi* prevalence and measures of tick burden (prevalence, intensity, and abundance) on white-footed mouse specimens are also presented. No *B. burgdorferi* or *I. scapularis* data was available for white-footed mice sampled from MSH.

Code	*B. burgdorferi* presence/absence	*B. burgdorferi* prevalence in White-footed mouse	Tick prevalence	Tick intensity	Tick abundance
DV	Present	0.055556	0.555556	4.15	2.305556
HR	Present	0.111111	0.555556	5.15	2.861111
HV	Present	0	0.357143	1.8	0.642857
LP	Absent	0	0.290323	1.555556	0.451613
MA	Absent	0	0.041667	1	0.041667
MR	Absent	0	0	0	0
SB	Present	0.076923	0.769231	2.4	1.846154
MSH	NA	NA	NA	NA	NA
MY	Absent	0	0.166667	1	0.166667
NO	Present	0	0.473684	1.555556	0.736842
PDC	Absent	0	0	0	0
SE	Present	0	0.116279	1	0.116279
SF	Present	0.133333	0.533333	2.875	1.533333
SJ	Absent	0	0.157895	3.333333	0.526316
SL	Absent	0	0	0	0
SP	Absent	0	0	0	0

FIGURES

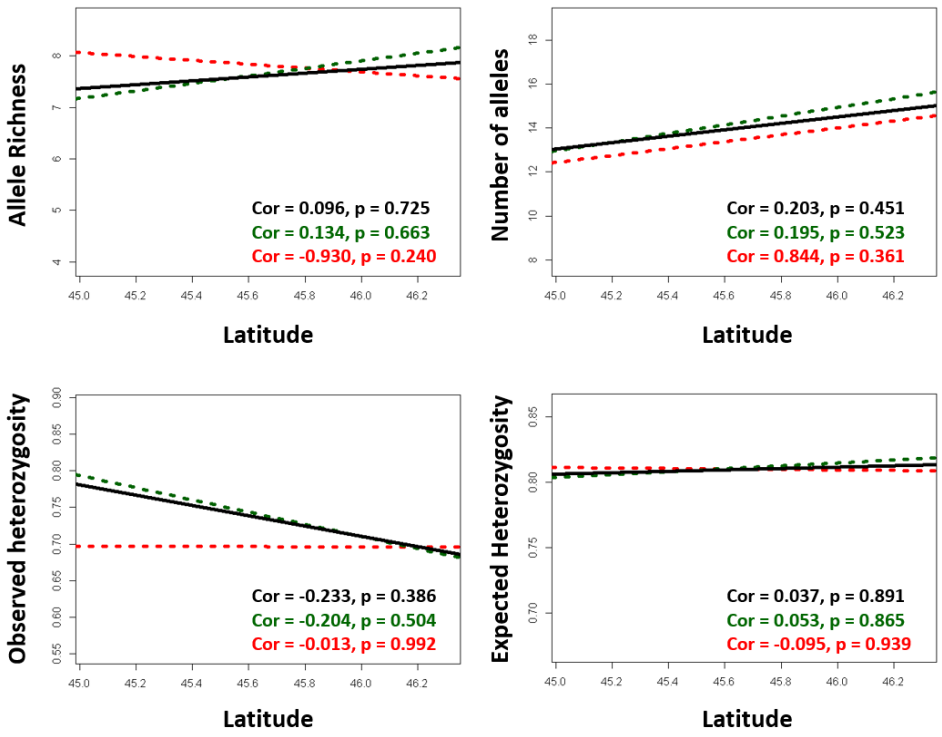


Figure B1. Comparisons of measures of genetic diversity within *Peromyscus leucopus* populations against latitude.

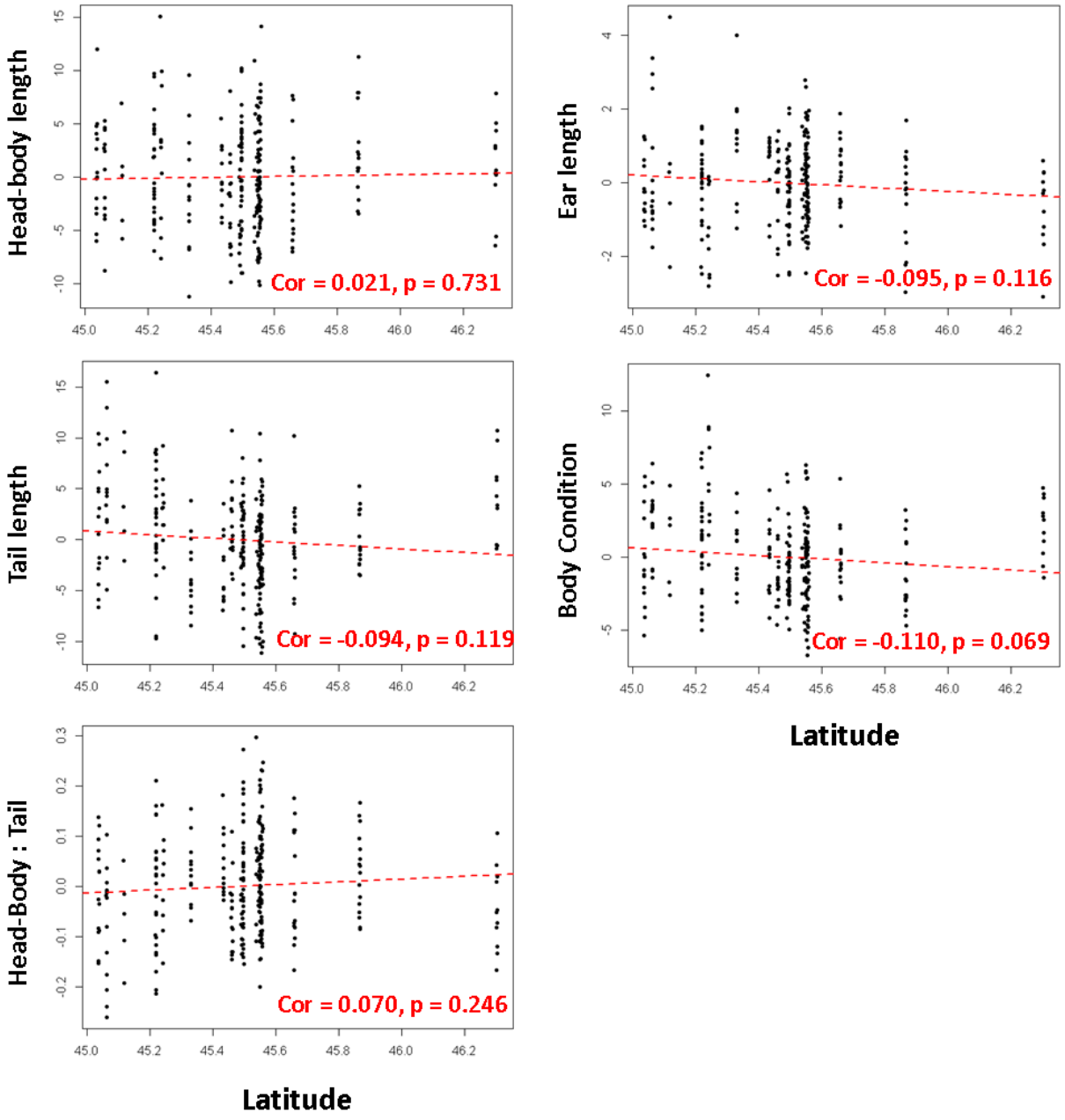


Figure B2. Comparisons of allometric measurements of *Peromyscus leucopus* external morphology (head-body length, tail length, head-body: tail ratio, ear length, body condition) against latitude.

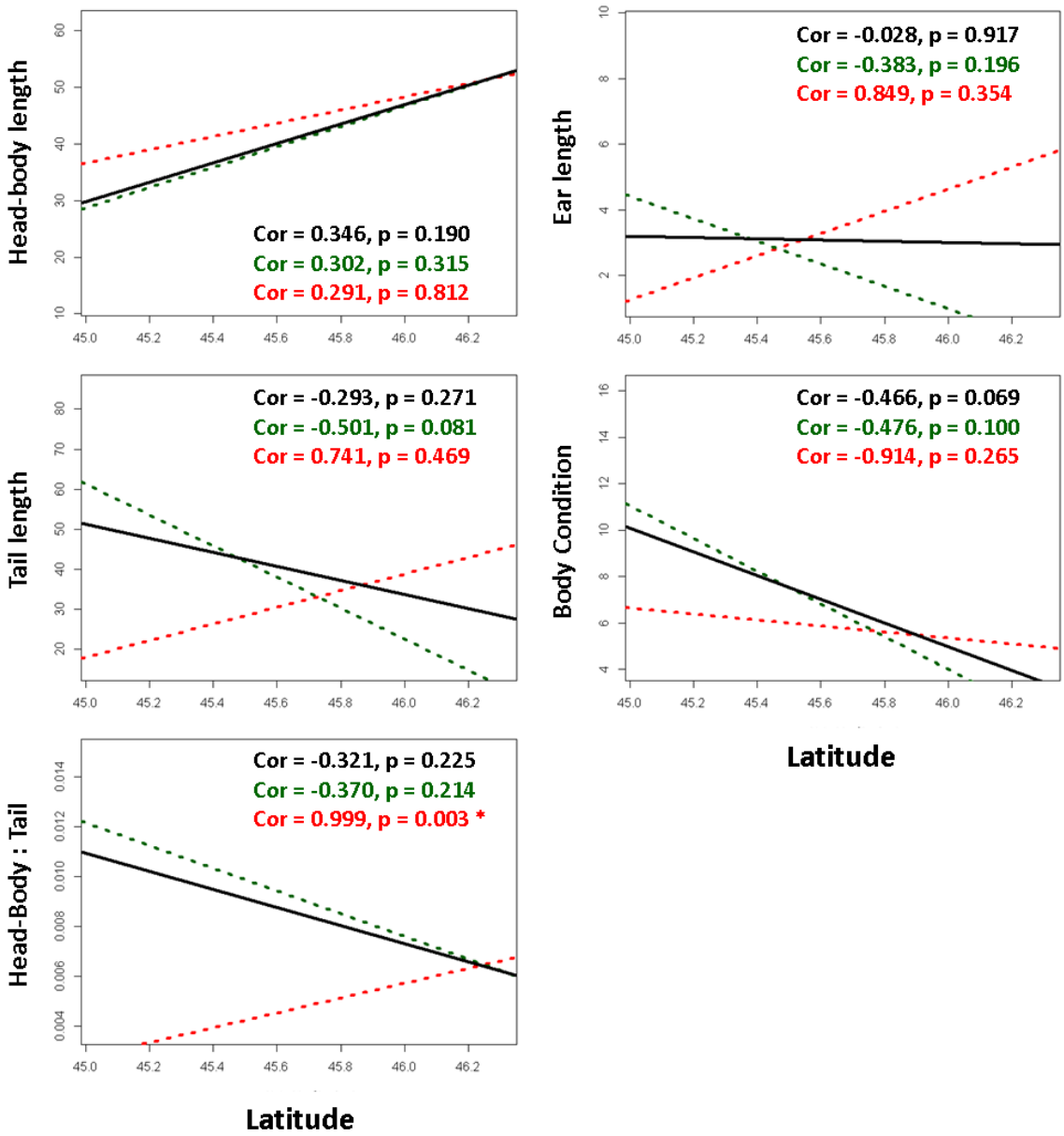


Figure B3. Amount of variance observed within *Peromyscus leucopus* populations for each external morphological trait (head-body length, tail length, head-body: tail ratio, ear length, body condition) plotted against latitude.

APPENDIX C

TABLES

Table C1. Number of *Ixodes scapularis* specimens collected from 13 sites between 2011 and 2014.

Locality	Latitude	Longitude	N_{Total}	N_{2011}	N_{2012}	N_{2013}	N_{2014}
BMF	45.29	-73.01	62	30	-	32	-
DV	45.86	-72.56	82	5	16	42	20
HR	45.49	-73.31	42	24	5	12	1
HV	45.11	-73.21	60	30	-	30	-
LP	45.54	-73.46	17	11	-	-	-
LS	45.44	-72.90	25	25	-	-	-
MSB	45.55	-73.34	81	39	7	20	15
NO	45.06	-73.28	74	20	8	38	8
SdV	45.23	-74.07	14	14	-	-	-
SE	45.21	-73.92	24	8	-	7	9
SF	45.03	-74.45	54	23	-	12	19
SJ	45.23	-73.46	14	5	-	7	2
SV	45.18	-73.34	64	35	-	24	5

Table C2. List of microsatellite primers used to genotype *Ixodes scapularis*. Polymerase chain reactions (PCR) were performed in 15 µL reactions containing 8.6 µL 10X PCR buffer, 0.2 mM of each dNTPs, 0.5 ρmol/µL of each primer, 0.05 U/µL of Taq polymerase and 2.5 µL DNA template. All PCR thermocycling conditions included an initial denaturation step at 95°C for 30 s, 30 cycles of 95°C for 15 s, loci-specific primer annealing temperature for 30 s, and 72°C for 45 s. PCR amplified products were visualised on an ABI3730 capillary sequencer (Applied Biosystems, Foster City, CA) at Genome Québec (Montreal, CA) and genotyped using the program GeneMarker (SoftGenetics LLC., State College, PA) with GeneScan 500-LIZ (Applied Biosystems, Foster City, CA) as the size standard.

Loci Name	**Repeat Motif**	**Forward / Reverse Primer Sequences (3' – 5')**	**Size Range**
IsAC8	AC	F: GAGCTACCCCTTTCATCGTCTTCG R: TCTTCCCGCTGCTGTCTCGTATTC	104 – 142
IsCTGY17	CTGY	F: TTCTTGTTTTATTGGTTGGGTG R: AATGCAGGGTAAGTTGAGATTG	122 – 154
IsAC22	AC	F: TATTGTAAGGCCAGTCGCCGCTGC R: CAGCTGGGCCCCTCCTTTTAATCC	149 – 187
IsAG4	AG	F: ATAAGCAATTCATACGAGATAGT R: AAGAAAATAAAGCGAACAAG	150 – 242
IsGATA4	GATA	F: CAGACAATGTCATTCAATCGCA R: CGCACAATGCAAAACAAATCTA	187 - 267
IsGATA3	GATA	F: AGTCCCCTCAGAGCGATTTTCA R: GGCCGCCAGTTTGATGGATA	113 – 165
IsAG25	AG	F: AAATGTCCGAACAGCCTTAT R: GCCCTTGAGTCTACCCACTA	138 – 202
IsAC4	AC	F: AAGCGTATCCGATTTGCCCTTCAT R: GGGTCCCAACGATTGCTAAACCAG	130 – 196
IsCAG12	CAG	F: GAAGAACACCGAGCGAACCAGAAC R: GTGAGCTGAGGGTTGCTGTTGATG	126 – 156
IsAC20	AC	F: AGAAACACGGAAGGAGAAAGGAGA R: AACTGTGCCAGATGGGAAAGAAGA	117 – 147

FIGURES

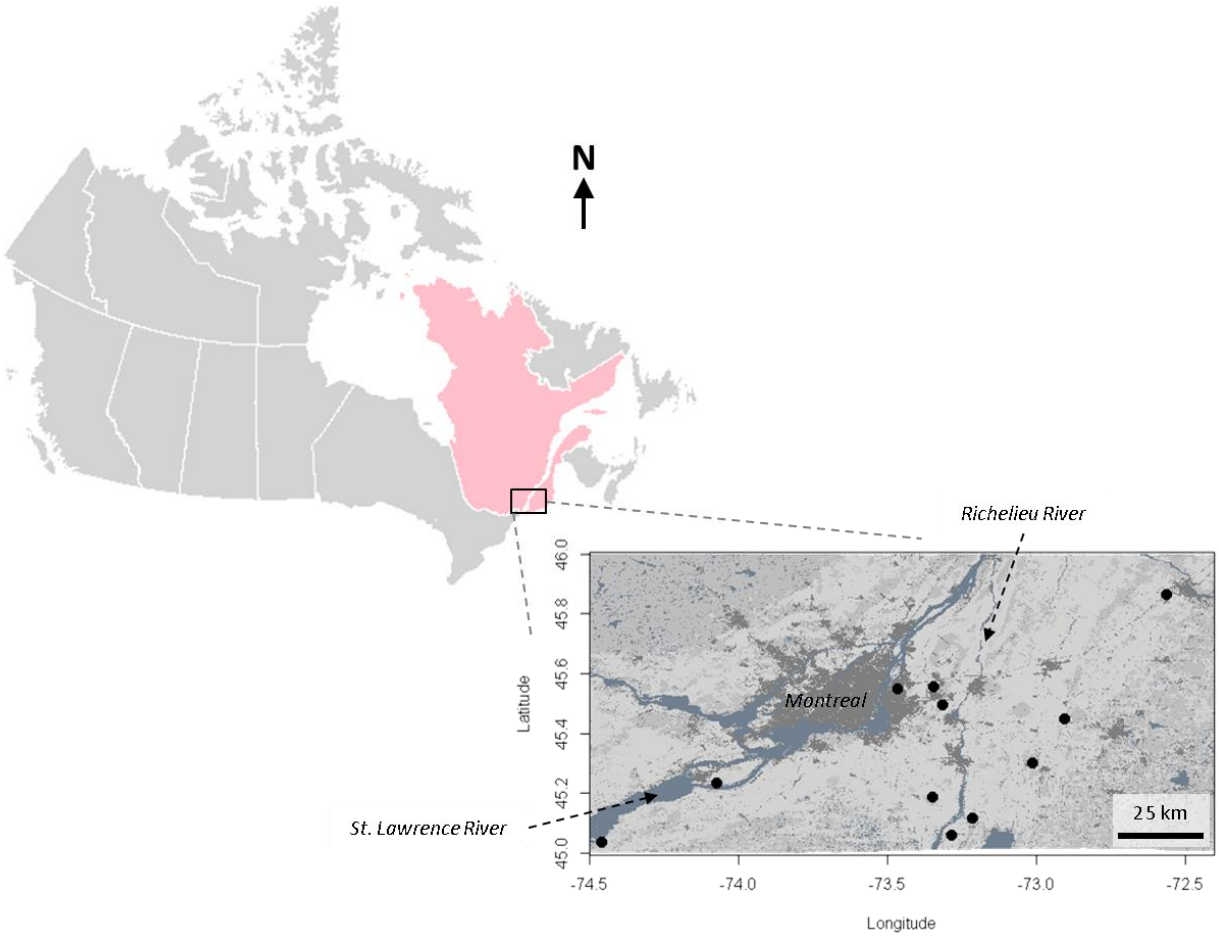


Figure C1. Location of sampling sites in southern Québec from which *Ixodes scapularis* examined in this study were collected.

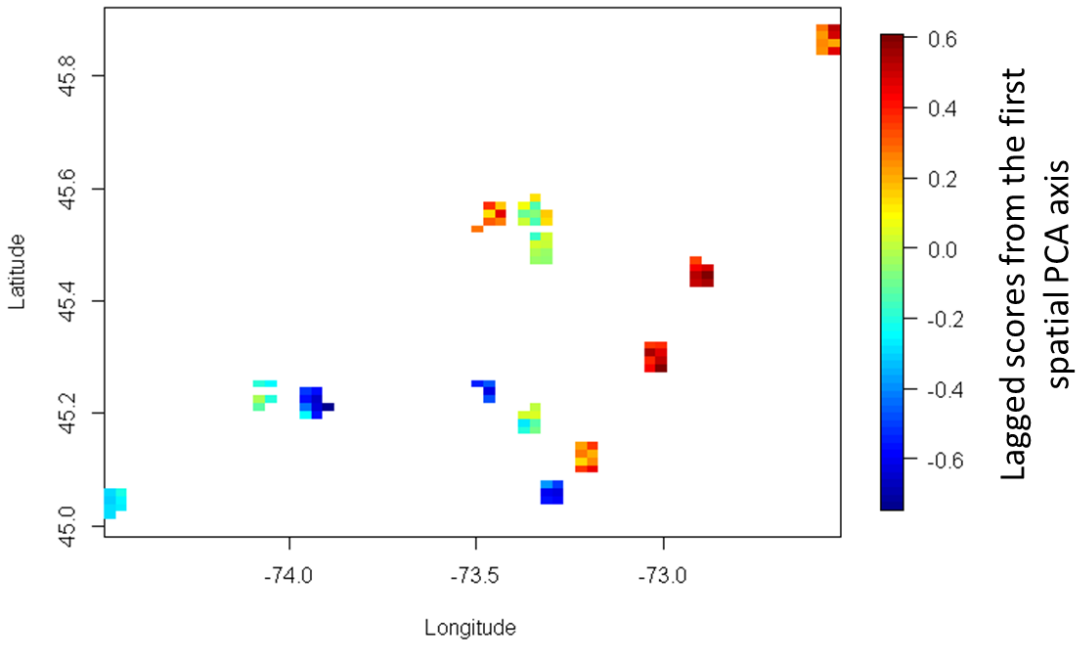


Figure C2. Quilt plot of the first axis of spatial principal components analysis (sPCA) scores for all 613 *Ixodes scaluparis* examined in this study. sPCA scores are indicated by colour and each dot represents an individual tick. Scores are plotted spatially according to latitude and longitude for each individual tick. Tests revealed significant global and local structure among ticks examined ($p < 0.05$). sPCA analysis, I used the Delaunay triangulation as the connection network. I then assessed global and local structures using the *global.rtest* and *local.rtest* functions respectively from the adegenet package over 999 permutations.

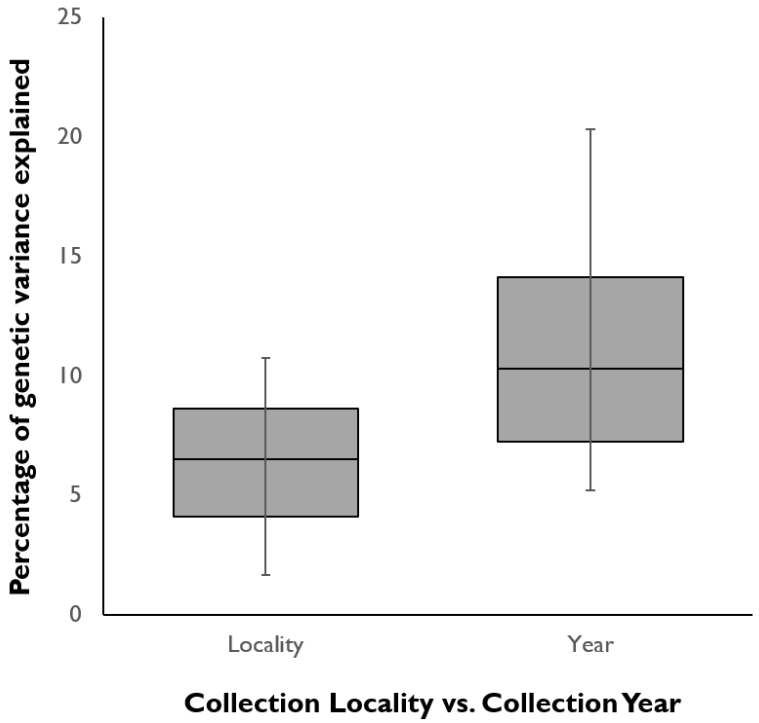


Figure C3. Results from a preliminary analysis comparing pairwise differences among ticks collected from different sites at different years. I found that genetic variation observed among tick individuals were better explained by collection year than sampling locality.

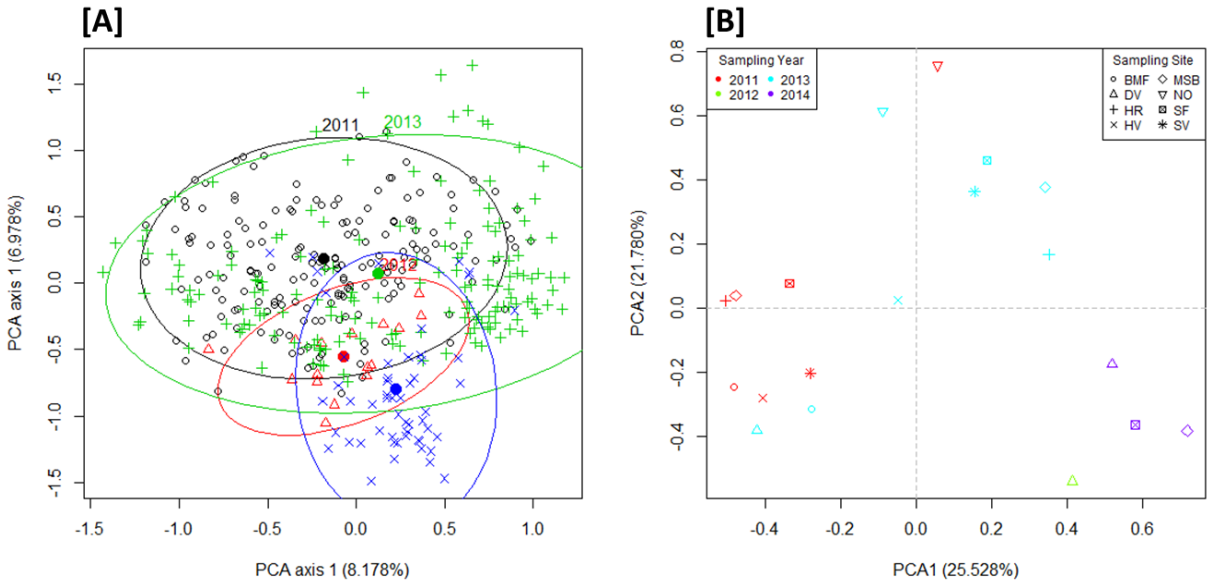


Figure C4. Results of principal components analyses (PCA) analyses on *Ixodes scapularis* at the (A) individual and (B) population levels.

APPENDIX D

Material and Methods

Variables influencing Borrelia burgdorferi *prevalence*

Tick-related variables examined include the number of questing ticks collected from vegetation at the site, the number of ticks found feeding on captured small mammal hosts, total number of ticks (or tick abundance – estimated as the sum of ticks collected by dragging and from trapped animals) at the site, the prevalence of black-legged tick on small mammals and white-footed mice (i.e. proportion of hosts examined that were parasitised), and the abundance of black-legged ticks on small mammals and white-footed mice (i.e. mean number of ticks found on all hosts examined). I decided not to break down my estimates of tick abundance and prevalence by instar as overall tick abundance and number of t ick infected with *Borrelia burgdorferi* was low. Nonetheless, I performed the analysis using the number of questing and feeding ticks as variables to account for possible bias in tick sampling (i.e. whereby observed tick collection in this resulting as a function of the number of mammals captured rather than true tick abundance at a given site).

I also considered the density of small mammals and white-footed mice at each site, observed and effective species richness (the latter calculated as an exponent of Shannon diversity index), as well as the proportion of small mammals captured that were identified to be white-footed mice. In order to standardise density measures to account for varying sampling efforts through the years, I calculated annual small mammal and white-footed mice densities by dividing the number of specimens by the number the number of trap-nights (number of traps x number of trapping nights) and the total area in km^2 covered by the sampling grids. I then used the average of the yearly specimen densities estimated for 2011 to 2014.

Multi-taxa Integrated Landscape Genetics on ticks and mice

I first performed a Mantel test comparing both tick and mouse pairwise population genetic distances in the form of pairwise F_{ST} values using the *mantel.randtest* function from the R package *ade4* (Dray et al. 2015). Pairwise F_{ST} values were calculated using the *pairwise.fst* function from the *adegenet* R package (Jombart et al. 2014). I also compared ordination solutions of allele frequency differences (via principle components analysis) and pairwise F_{ST} values among populations of tick and mice (via principle coordinates analysis) using a series of Procrustes rotation tests performed using the *protest* function from the R package *vegan* (Oksanen et al. 2015). I obtained ordination solutions of allele frequency differences and pairwise F_{ST} values using the *rda* and *capscale* functions from the R packages *vegan*.

Using log-transformed allele frequencies as response matrices and measures of spatial connectivity as predictor variables in a series of redundancy analyses (RDA) analyses, I also identified landscape variables that best describe observed tick and mice spatial genetic patterns. For each of the 40 sites, I calculated site connectivity relative to the distributions of specific landscape variables (roads, water, agriculture, urban, forest) from least-cost effective distances estimated from previously described resistance values associated with each landscape variable (Marotte et al. 2014). Connectivity among sites was calculated as described in Leo et al. 2016.

Climatic niche models on mouse and tick in southern Québec

Following Roy-Dufresne et al. (2013) and Simon et al. (2014), species presence data for the white-footed mouse were compiled using a combination of collection records for southern Quebec between 1966 – 2011, and records obtained from the Arctos Collection Management Information System (http://arctos.database.museum, using records from the University of Alaska Museum of the North, the Museum of Southwestern Biology and the Museum of Vertebrate

Zoology), the Banque de données sur les micrommamifères et les chiroptères du Québec, and the mammals collection database from the Field Museum of Natural History. A total of 404 records were obtained for the white-footed mouse. For the black-legged tick, presence data was downloaded from collection records uploaded onto the VectorMap website (http://www.vectormap.org/ - a product of the Walter Reed Biosystematics Unit based in the Smithsonian Institution) and estimated from Figure 2 in Ogden et al. (2009). Presence data for the ticks were spatially filtered to avoid having multiple presence points within 1km of each other and to include only records located below latitude 47.5˚N so that only records of presumably established tick populations are located (based on recent results from analyses on tick genetics in Chapter 4). The latter was done so as to avoid artificially expanding estimates of current tick distribution due to the transient presence of adventitious ticks beyond the species' established range. A total of 1172 records were obtained for the black-legged tick.

I used a georeferenced set of climatic variables to calibrate my climate niche models, focusing specifically on five winter variables that may limit the distributions of both white-footed mouse and black-legged tick. All values were averaged from records between 1961- 2005. I obtained temperature and precipitation values from the ANUSPLIN dataset version 4.3 that are based on the Natural Resources Canada's historical monthly weather data. Meanwhile, snow depth data was obtained from Environment Canada meteorological stations (http://climate.weatheroffice.gc.ca) and United States NOAA stations (http://www.ncdc.noaa.gov/cdo-web/). I defined the beginning and conclusion of a winter season the time when temperature within a grid cell first falls below 0˚C until it rises above 0˚C in the next calendar year.

I modeled current and future distribution of the white-footed mouse and black-legged tick using seven niche-based modeling approaches: the generalized linear model (GLM), generalized boosting model (GBM), classification tree analysis (CTA), artificial neural network (ANN), random forest (RF), multivariate adaptive regression spline (MARS), and flexible discriminant analysis (FDA). All models were calibrated and evaluated for 10 replicate each, using 70% of species presence data for calibration and the remaining 30% for evaluation. The importance of each climatic variable in the models was evaluated over 5 permutations and the models themselves evaluated using the True Skill Statistic (TSS) and Relative Operating Characteristic (ROC) methods. All model predictions were scaled using a binomial GLM. Since I lack any absence data for white-footed mouse and black-legged ticks in this study, I generated a series of pseudo-absence data for the climate niche-modeling. Five sets of 404 and 1172 pseudo-absence data points were generated for the white-footed mouse and black-legged tick respectively using a surface range envelope (SRE) model. Pseudo-absence and presence points used in all models were weighted equally.

I subsequently combined results from all seven niche-based models to obtain a total consensus model of species occurrence probabilities. Models were first ranked according to TSS scores and any models with TSS scores lower than 0.7 were excluded. A value of 1.6 was used to assign relative weights to each model. The models were assigned relative weights according to their ranks with a decay rate of 1.6.

Structural equation model

My model pathways included interactions among white-footed mouse and black-legged tick occurrence probabilities (extracted from climate-based species distribution models restricted to the study area), predicted mouse density (estimated from habitat and landscape variables), and

observed *B. burgdorferi* prevalence. The SEM analysis was based on data from sites in which both ticks and mice were simultaneously present (N = 26) (Simon et al. 2014). All observed variables were standardized prior to the analysis. The SEM analysis was performed using the *sem* function from the R package *lavaan* (Rosseel et al. 2015). I employed the maximum likelihood estimator under a normal likelihood approach. Standard errors in the analysis were computed using standard bootstrapping. Additionally, the Bollen-Stine bootstrap was used to compute the bootstrap probability value of the test statistic over 1000 draws. I evaluated the performance of my model based on the Bollen-Stine Bootstrap p-values as well as several fit measures such as the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) (see Table D5).

Coefficients obtained from the structural equation models were used to produce a linear equation that predicts *B. burgdorferi* prevalence based on estimated contributions from climate and landscape/habitat conditions on species distribution and abundance of the primary reservoir host and disease vector.

REFERENCES

Dray S, Dufour A-B, Thioulouse J. 2015. ade4: Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences. R package 1.7-2

Jombart T, Collins C, Solymos P, Ahmed I, Calboli F, Cori A. 2014. adegenet: an R package for the exploratory analysis of genetic and genomic data. R package vers 1.4-2

Leo SST, Gonzalez A, Millien V. 2016. Multi-taxa integrated landscape genetics for zoonotic infectious diseases: Deciphering variables influencing disease emergence. *Genome*. doi: 10.1139/gen-2016-0039

Marrotte RR, Gonzalez A, Millien V. 2014. Landscape resistance and habitat combine to provide an optimal model of genetic structure and connectivity at the range margin of a small mammal. *Molecular Ecology* **23**: 3983 – 3998

Ogden NH, Lindsay R, Morshed M, Sockett PN, Artsob H. 2009. The emergence of Lyme disease in Canada. *Canadian Medical Association Journal* **180**: 1221 – 1224

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2015. vegan: Community ecology package. R package vers. 2.2-1

Rosseel Y, Oberski D, Byrnes J, Vanbrabant L, Savalei V, Merkle E, Hallquist M, Rhemtulla M, Katsikatsou M, Barendse M. 2015. lavaan: Latent variable analysis. R package 0.5-18.

Roy-Dufresne E, Logan T, Simon JA, Chmura GL, Millien V. 2013. Poleward expansion of the white-footed mouse (*Peromyscus leucopus*) under climate change: Implications for the spread of Lyme disease. *PLoS ONE* **8**: e80724

Simon JA, Marrotte RR, Desrosier N, Fiset J, Gaitan J, Gonzalez A, Koffi JK, Lapointe F-J, Leighton PA, Lindsay LR, Logan T, Milord F, Ogden NH, Rogic A, Roy-Dufresne E, Suter D, Tessier N, Millien V. 2014. Climate change and habitat fragmentation drive the occurrence of *Borrelia burgdorferi*, the agent of Lyme disease, at the northeastern limit of its distribution. *Evolutionary Applications* **7**: 750 – 764.

TABLES

Table D1. List of microsatellite primers used to genotype *Peromyscus leucopus* and *Ixodes scapularis*. Information on repeat motifs, primer sequences, observed size ranges and literature references are provided. Polymerase chain reactions (PCR) were performed in 15 µL reactions containing 8.6µL 10X PCR buffer, 0.2mM of each dNTPs, 0.5ρmol/µL of each primer, 0.05U/µL of Taq polymerase and 2.5µL DNA template. All PCR thermocycling conditions included an initial denaturation step at 95°C for 30s, 30 cycles of 95°C for 15s, loci-specific primer annealing temperature for 30s, and 72°C for 45s. PCR amplified products were visualized on an ABI3730 capillary sequencer (Applied Biosystems, Foster City, CA) at Genome Quebec (Montreal, CA) and genotyped using the program GeneMarker (SoftGenetics LLC., State College, PA) with GeneScan 500-LIZ (Applied Biosystems, Foster City, CA) as the size standard.

Loci Name	Repeat Motif	Forward / Reverse Primer Sequences (3' – 5')	Size Ranges	References
White-footed mouse (*Peromyscus leucopus*)				
Pml01	CA	*F:* CATTCAAGACCTGGCTTTTT *R:* TGGGTTTCATCAGTGCTTCT	149 – 205	Chirhart et al. 2005
Pml03	CA	*F:* GCCATTAGTCTATGTGACAG *R:* GCGATGTACCCAGAAAT	231 – 259	Chirhart et al. 2005
Pml04	CA	*F:* CATAAGGTGGCTCGGAATCA *R:* CAGGAAGGGGAAATGACCAT	186 – 244	Chirhart et al. 2005
Pml05	CA	*F:* CTGAGCCAAAAGTGGTCCTT *R:* TGAAGACAGCCCCTCTCTG	182 – 248	Chirhart et al. 2005
Pml06	CA	*F:* CAGGGCTGTAGAGGGAGAAC *R:* ACTGGAGCAGAGGCATTTG	142 – 178	Chirhart et al. 2005
Pml09	CA	*F:* GAATCCATACACCCATGC *R:* TTGCTTTTCGTCAAGTTTT	186 – 226	Chirhart et al. 2005
Pml11	CA	*F:* ACCCCCGAGTGCTGAGATT *R:* TTTGCTGCTTTCCCCAGAGA	218 – 264	Chirhart et al. 2005
Pml12	CA	*F:* GCAGCCTGTATTCTCTCACA *R:* GCCAACCATTTCTTCAAGTG	141 – 167	Chirhart et al. 2005
PLGT58	GT	*F:* GATCTTGTGAACACGCTTCT *R:* TTGATGGCTCTGGAGAGGCT	132 – 184	Schmidt 1999
PLGT66	GT	*F:* CTCTGTCTGCCACACATGCT *R:* GTGCCATCACAGATGTGACA	129 – 165	Schmidt 1999
Black-legged tick (*Ixodes scapularis*)				
IsAC8	AC	F: GAGCTACCCCTTTCATCGTCTTCG R: TCTTCCCGCTGCTGTCTCGTATTC	104 – 142	Fagerberg et al. 2001
IsCTGY17	CTGY	F: TTCTTGTTTTATTGGTTGGGTG R: AATGCAGGGTAAGTTGAGATTG	122 – 154	Fagerberg et al. 2001
IsAC22	AC	F: TATTGTAAGGCCAGTCGCCGCTGC R: CAGCTGGGCCCCTCCTTTTAATCC	149 – 187	Fagerberg et al. 2001
IsAG4	AG	F: ATAAGCAATTCATACGAGATAGT R: AAGAAAATAAAGCGAACAAG	150 – 242	Fagerberg et al. 2001
IsGATA4	GATA	F: CAGACAATGTCATTCAATCGCA R: CGCACAATGCAAAACAAATCTA	187 - 267	Fagerberg et al. 2001
IsGATA3	GATA	F: AGTCCCCTCAGAGCGATTTTCA R: GGCCGCCAGTTTGATGGATA	113 – 165	Fagerberg et al. 2001
IsAG25	AG	F: AAATGTCCGAACAGCCTTAT R: GCCCTTGAGTCTACCCACTA	138 – 202	Fagerberg et al. 2001
IsAC4	AC	F: AAGCGTATCCGATTTGCCCTTCAT R: GGGTCCCAACGATTGCTAAACCAG	130 – 196	Fagerberg et al. 2001
IsCAG12	CAG	F: GAAGAACACCGAGCGAACCAGAAC R: GTGAGCTGAGGGTTGCTGTTGATG	126 – 156	Fagerberg et al. 2001
IsAC20	AC	F: AGAAACACGGAAGGAGAAAGGAGA R: AACTGTGCCAGATGGGAAAGAAGA	117 – 147	Fagerberg et al. 2001

Table D2. Variable importance scores for each model employed in the climate niche models predicting white-footed mouse and black legged tick species distributions. Importance scores were averaged over five replicates of each model approach and are positively correlated with the importance of the variable in determining species distribution. The models used were the generalized linear model (GLM), generalized boosting model (GBM), classification tree analysis (CTA), artificial neural network (ANN), random forest (RF), multivariate adaptive regression spline (MARS), and flexible discriminant analysis (FDA). The highest variable contribution values from each model are underlined.

White-footed mouse (*Peromyscus leucopus*)							
	GLM	**GBM**	**CTA**	**ANN**	**RF**	**MARS**	**FDA**
Snow	0.1198	0.0179	0.0893	0.0076	0.0883	0.0469	0.0566
Winter Length	<u>0.9048</u>	0.1580	0.3046	0.5369	0.2044	<u>0.8162</u>	<u>0.9858</u>
Minimum temperatures	0.2384	0.0211	0.0282	0.2607	0.0281	0.3201	0.2230
Maximum temperatures	0.3849	<u>0.4375</u>	<u>0.6084</u>	<u>0.8763</u>	<u>0.2460</u>	0.4467	0.1714
Precipitation	0.0140	0.0772	0.1295	0.2290	0.0867	0.0884	0.0270
Black-legged tick (*Ixodes scapularis*)							
	GLM	**GBM**	**CTA**	**ANN**	**RF**	**MARS**	**FDA**
Snow	0.2346	0.0672	0.0932	0.0017	0.0534	0.0299	0.0098
Winter Length	<u>0.7055</u>	0.0159	0.1224	0.5524	0.0913	0.4144	0.1782
Minimum temperatures	0.5971	<u>0.3088</u>	0.3873	0.5568	<u>0.1940</u>	0.2772	0.6315
Maximum temperatures	0.3125	0.2191	<u>0.5281</u>	<u>0.8824</u>	0.1570	<u>0.4925</u>	<u>0.6415</u>
Precipitation	0.1182	0.0870	0.1007	0.3543	0.0538	0.1438	0.0760

Table D3. Results from Analyses of Molecular Variance (AMOVA) on white-footed mouse and black-legged tick populations in southern Quebec, Canada.

White- footed mouse (*Peromyscus leucopus*)				
	df	Sum of squares	Variance components	Percentage variance
Among populations	8	11898.101	23.102	15.31
Within populations	527	67373.270	127.843	84.69
Total	535	79271.371	150.945	
Black-legged tick (*Ixodes scapularis*)				
	df	Sum of squares	Variance components	Percentage variance
Among populations	8	6570.058	7.757	4.35
Within populations	775	132287.863	170.694	95.65
Total	783	138857.921	178.451	

Table D4. Landscape genetics results from nine white-footed mouse and black-legged tick populations in southern Quebec, Canada. Distributions of mouse and tick allele frequencies were examined relative to site connectivity estimated from a heterogeneous landscape made up of roads, water bodies, agriculture areas, urban development, and forest patches. Statistically significant values are bolded.

Variables examined	**White-footed mouse**		**Black-legged ticks**	
	F	**p-values**	**F**	**p-values**
Patch connectivity (roads)	**1.6893**	**0.037**	0.5863	0.793
Patch connectivity (water)	0.8938	0.577	0.6907	0.716
Patch connectivity (agriculture)	**1.6361**	**0.047**	1.6769	0.146
Patch connectivity (developed)	**1.7006**	**0.037**	0.5976	0.795
Patch connectivity (forest)	0.9989	0.475	0.8515	0.553

Table D5. Summary of statistical significance and fit indices of the structural equation model.

Measures of model performance	
P-value (Bollen-Stine Bootstrap)	0.126
Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000
Log likelihood user model (H0)	-136.188
Log likelihood unrestricted model (H1)	-136.188
Akaike Information Criterion (AIC)	280.376
Bayesian Information Criterion (BIC)	285.408
Sample size adjusted BIC	272.992

FIGURES

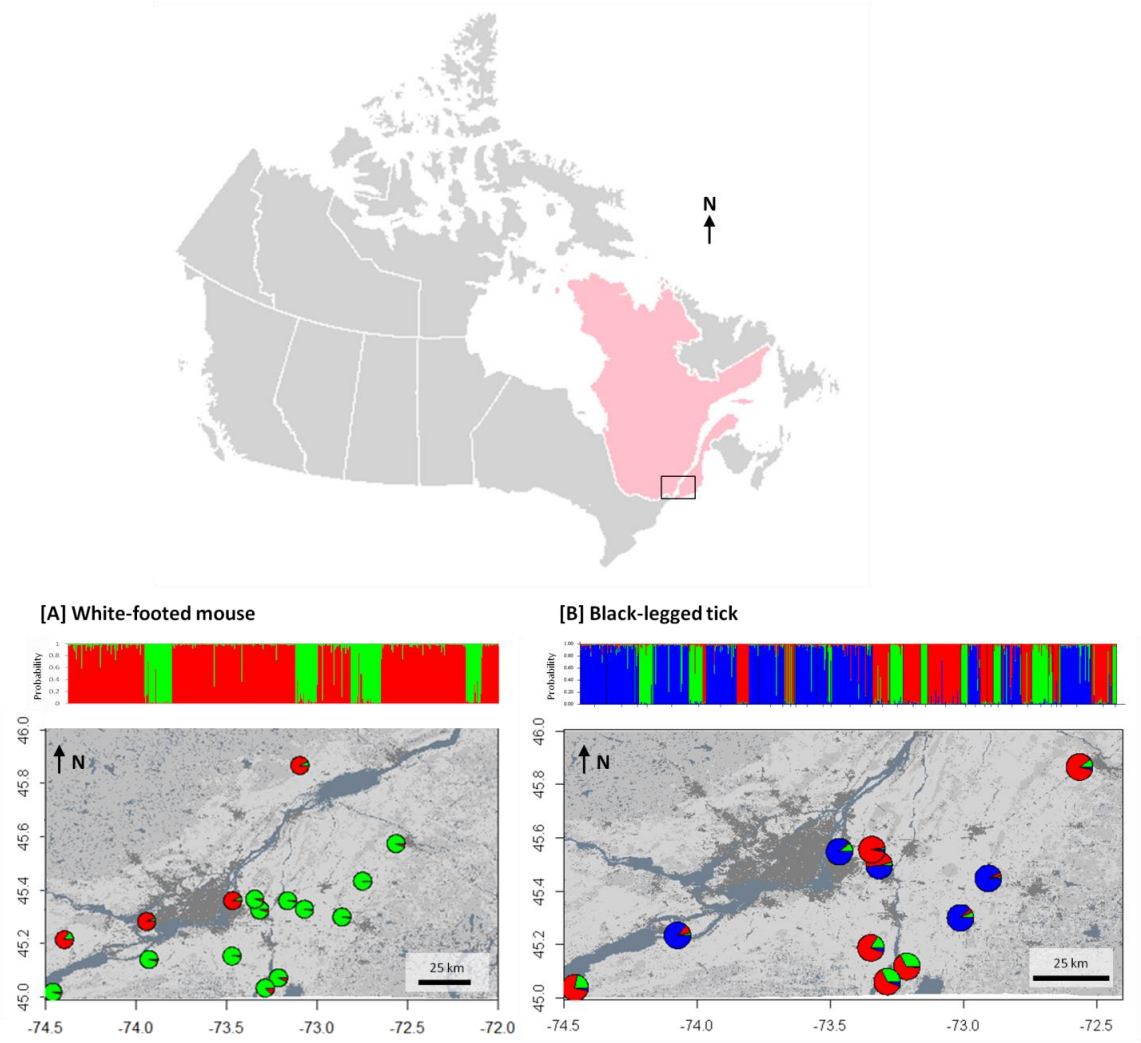


Figure D1. Results from individual-based clustering analysis on [A] 16 white-footed mouse populations and [B] 11 black-legged tick populations. Pie charts in the maps show proportion of membership of each white-footed mouse population in each cluster. Bar-plots indicating probability of individual assignment for each species are provided above each map. Analyses were performed using no admixture and independent allele frequency models. Ten iterations for each number of populations (k) equaling 1 through 10 were initially analyzed for 100,000 Markov Chain Monte Carlo (MCMC) generations with an initial burnin of 10,000 generations. The most likely number of genetic clusters was next calculated from likelihood outputs produced by structure using the approach by Evanno et al. (2005). Based on this result, a more thorough analysis with 'k' defined was completed with MCMC running for 500,000 generations with an initial burnin of 50,000.