ECOLOGICAL CORRELATES OF GENETIC DIVESITY IN BORRELIA BURGDORFERI, THE LYME DISEASE BACTERIUM

by

HOLLY B VUONG

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Ecology and Evolution

written under the direction of

Dr. Peter Morin & Dr. Richard Ostfeld

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2012

ABSTRACT OF THE DISSERTATION

Ecological correlates of genetic diversity in *Borrelia burgdorferi*,

the Lyme disease bacterium

By HOLLY B VUONG

Dissertation Director:
Dr. Peter Morin & Dr. Richard Ostfeld

Lyme disease is the number one reported vector-borne disease in the United States, and this disease continues to spread in the Northeast and Midwest. Lyme disease is caused by *Borrelia burgdorferi*, a bacterium that circulates among vertebrate host species and transmitted among hosts by the black-legged tick (*Ixodes scapularis*). The bacterium has high genetic variation at the outer surface protein C (*ospC*) locus, and past studies suggests that hosts act as ecological niches to the *ospC* genotypes. In particular, five types are known to be human invasive (HIS), making it essential to examine disease risk at the genotypic level. My studies focus on understanding the ecological drivers of *ospC* diversity and frequency profile at the individual, community, and landscape scales. In chapter one, I found that endemic areas of New York State have higher *ospC* richness and diversity than newly invaded areas, and that HIS types are relatively common across the landscape. There is high turnover of genotypes from one population to another population along the invasion scale. In chapter two, host community composition matters when examining *ospC* diversity, and that host composition and host diversity are important in predicting HIS infection prevalence. Contrasting important predictors

between years 2006 and 2009 could be a result of annual variation and/or site variation, since the majority of sites were not sampled in both years. Obtaining better inclusive host community composition and diversity estimates could help with the predictive powers of these metrics on the *ospC* frequency profile, especially with HIS types. In chapter three, variation among nine host species and their associated *ospC* genotype frequency profiles supports the host-niche concept. Short-tailed shrews (*Blarina brevicauda*) and American robins (*Turdus migratorius*) have high proportions of HIS to non-HIS, suggesting they could help contribute to higher disease risk. Lastly, there is support for a trade-off between occurrence frequency and transmission efficiencies of *ospC* types from hosts to ticks feeding on these individuals. This research sheds new light on how host composition and diversity influences disease risk, that HIS types infect all nine host species, and that HIS types occur commonly across NY State.

Acknowledgements

Grad school can be a very trying, surprising, and fulfilling time. In the end, the journey is worth it. However, the adventures of grad school cannot be achieved alone, as speed bumps in life require good friends, intellectual spurts, and even straight, physical help. The adventures also require funding! So, to all those who helped, here is my way of saying thanks.

I would like to first thank my committee for their continued support and expertise on helping me improve my research. Peter Morin, you have been a solid rock that I can lean on, from research trials to ordinary daily life. I will miss your dry sense of humor, your delicious meals, and our one-on-one chatting sessions on the front porch. Rick Ostfeld, your immense knowledge on the Lyme disease system, and on disease ecology in general, has shaped me into a stronger disease ecologist over the years. I am always in awe with what you have accomplished in your career, and I hope I can be as talented as you are over time. Peter Smouse, you are my "Ben Smith." You have always made me feel like I can, and will, go far and beyond in my future career. Thank you for your constant encouragements and for all your stories about life and graduate school. Dina Fonseca, I look up to you as my role model because I see a woman who is outstanding, balances work and family, gives so much of her time to help me and other students out, is funny and beautiful, and maintains a positive outlook about life and work. Thanks for giving me a room in your house and treating like family all those months I stayed at your place! And Dustin Brisson, whose invaluable experiences in the lab helped me accomplish my research projects. Thanks for giving me space in your lab, both at the beginning and end of my graduate stages, and helping me to think critically about the genetic diversity of the Lyme bacterium.

Many thanks to the field crews (year 2006, 2008-2010) that participated in data collection of these massive field projects, especially Shannon Duerr and Kelly Oggenfuss, for their ability to juggle field work, animal handling training, data sorting, people managing, amazing memory, and the two people who can answer just about any question about any of the projects I needed help on. Other people include (postdocs & visiting scientists): Jesse Brunner, Felicia Keesing, May Killilea, Kathleen LoGiudice, Andrea Previtali, and Lisa Schwanz; and (summer assistants): Katie Ansorge, Laura Beard, Laura Bendernagle, Sarah Booker, Alicia Bruce, James Burtis, Elizabeth Carlisle, Amanda Chappell, Laura Cheney, Angie Coates, Rob Davis, Kevin Dougherty, Kira Gilman, Andrea Goth, Jennifer Hoffman, Melanie Klein, Mitch LeSage, Anika Mahoney, Scott Morlando, Laura Muller, Lizz Mulligan, Mike Newhouse, Lindsey Nietmann, Kathleen O'Coonor, Craig Okraska, Suzanne Ostrow, Julia Pulaski, Jill Rasmas, Emily Sechny, Jill Sprance, Buck West, Jay Winiarski, Chad Witko, and Julie Ziemba. Thanks also to the 2012 field crew for help in collating camera trap data, including Andrea Goth, Elizabeth D'Auria, Hannah Clark, Mike Fraatz, Laura Gigliotti, and Nicholas Skaff.

Audrey Beltrani volunteered several weekends in summer 2008 to help collect ticks. Tick collections throughout New York were also assisted by several Rutgers friends, including: Nacho Bartomeus, Kenneth Elgersma, Julia Perzley, Lucas (from Czech Republic), Belen Sanchez, Alicia Shenko, Elena Tartaglia, Ai Wen, and Tom Hughes (Green Lakes State Park).

I thank Felicia Keesing and her Bard lab personnel for positive DNA samples. At Rutgers University, Esther and Rebecca Heiry assisted with some DNA extractions. Thanks to Godefroy Deverey at UPenn for new primer sequences for *ospC*. Statistical

advice was greatly appreciated from Charlie Canham, Ed Green, Michelle Hersh, and David Moriarty.

Thanks to all the Cary Institute staff that have helped me feel welcomed over the past six years and those who helped ease my life from grant submissions to purchasing. Good friends include Kelly Oggenfuss, Shannon Duerr, and Amy Schuler. At UPenn, the Brisson lab members helped me navigate through the laboratory procedures.

At Rutgers, the Morin lab (and visiting lab) members, including Kevin Aagard, Ben Baiser, Aabir Banerji, Josh Caplan, Tim Casey, Jean Deo, Cara Faillace, Nick Lorusso, Jennifer Krumins, Jack Siegrist, Maria Stanko, Paul Walberg, provided their time in reading grant proposals and offering research advice. The Fonseca lab (and visiting lab) members, including Julian Avery, Emilie Cameron, Laran Caplan, George Condon, Kavi Damal, Chenoa de Freece, Ashley deNegre, Andrea Egizi, Becky Heiry, Sébastien Marcombe, Lizz Mulligan, Dana Price, and Jiawu Xu, provided lab assistance and discussions on population genetics. Thanks to Kay Bidle for providing lab space, and David Ehrenfeld for giving me an opportunity to TA Principles of Ecology class, where I became the head TA for two years. Your confidence in me, and your compassion for me as a student, researcher, and friend, made life a bit sweeter at Rutgers.

I also want to recognize Marsha Morin, Allison Cariveau, Christine Tizzano, Georgie DiGiglio, and Priscilla Walsh at Rutgers for helping out from ordering materials to University paperwork, to everything in between. I want to especially thank Marsha for welcoming me to the University, treating me like family, and always offering her and Peter's home to me. Additionally, Karl Kjer and Sheri Scheldorf have been my unofficial uncle and aunt in New Jersey, and I have been lucky to have them as good friends.

There are the friends you rely on for support when research is going badly, when life is kicking you in the behind, and when you have great news to celebrate. My main gals include Maria Stanko, Emilie Cameron, and Belen Sanchez. Other wonderful friends include: Elena Tartaglia, Faye Benjamin, Tina Harrison, Julia Perzley, Dave Smith, Kenneth Elgersma, Ai Wen, Blake Mathys, Jean Deo, Aabir Banerji, Monica Palta, Ileana Perez, Brian Clough, Amy Manning, Nacho Bartomeus, Cara Faillace, and the lovely ladies from the E&E reunion.

From California, I want to thank David Moriarty for his unconditional support, encouragements, and friendship over the years. Lastly, my family, especially my parents, allowed me to take flight and try out new adventures in life. Their continued support in my education and research has made it possible to achieve my goals.

Lastly, without funding, a large part of the research would not have been possible. At Rutgers, I received the Diversity fellowship during my first academic year, and the Bevier fellowship during my last academic year. Teaching assistantships supported me during other academic years. Small grants are important for initial data collection and laboratory reagents, so I want to thank the Graduate School New Brunswick Pre-doctoral Research Award, Ecology and Evolution Excellence Award and Small Mammal Research, and the Hutcheson Memorial Forest research grant. Large research awards to Rick Ostfeld included NSF, EPA, and NIH. These larger awards provided funding for large-scale field projects, rearing facility work, and laboratory analyses.

Dedication

I dedicate my dissertation to my mom, Amy, and my younger sister, Diana. For my mom, because her encouragement and willingness to let me spread my wings away from the nest, allowed me to thrive in my graduate studies. For Diana, because her strength and determination to take on whatever life throws at her, is an inspiration for me to tackle graduate school and life.

Table of Contents

Lists of tables

List of illustrations

Introduction

Approximately 61% of over 1400 pathogens known to be infectious to humans are zoonotic, or arising from animal hosts (Taylor et al. 2001). Of those zoonoses, about 75% are responsible for emerging and re-emerging infectious diseases (Daszak et al. 2000). Our environmental actions and their effects on biodiversity can profoundly influence the emergence and re-emergence of infectious diseases. As we continue to change the landscape, either through physical (e.g., dams and levees that block movement for aquatic organisms, clear cutting, farming in clear cut areas) or non-physical ways (e.g., continual burning of fossil fuel leading to greater concentrations of greenhouse gases and speeding up climate change, sea ice melt, and rising ocean temperatures), we are disturbing wildlife habitats and distribution, and losing wildlife species. The loss of wildlife species may contribute to higher disease risk, because some of the species can buffer the transmission of pathogens from competent host reservoirs to humans (Ostfeld and Keesing 2000).

One important aspect of emerging and re-emerging infectious diseases is the rate of their spread, once they become transmissible in humans. Our ability to travel rapidly with planes, trains, and automobiles, can result in the movement of latent, infected people, who arrive at their destinations before disease onset, leading to outbreaks occurring away from the place of origin. One well-known example is the 2003 outbreak of SARS (severe acute respiratory syndrome), a corona virus suspected of being carried by bats, which had mutated and become infectious to humans (Li et al. 2005). In approximately four months since its outbreak recognition, SARS was contained, but SARS had spread to 27 countries in all continents (Heymann and Rodier 2004). Although this was an extreme case of spread, this shows that our mobility can help spread diseases far and wide.

In the same manner, movement by animals infected with pathogens can help expand the zone encompassed by zoonotic diseases. Birds are the primary hosts for West Nile virus, and mosquitoes (primarily *Culex* spp.), are important bridge vectors that feed on birds and mammals (Hayes 2001). The migratory movement of infected birds, as well as the overwintering populations of infected mosquitoes, may have led to the amplification of West Nile virus across the U.S. in just a few years, following its emergence in New York City (Nasci et al. 2001, Fonseca et al. 2004). Female mosquitoes feed on each host for a relatively short time period, but feed multiple times, whenever proteins are needed for egg-laying (Spielman 2001). Contrast that to the slower spread of Lyme disease in the northeastern U.S., where black-legged ticks (*Ixodes scapularis*) feed once at each life stage, but where they stay attached to any particular host for several days during which they feed continuously, until they have obtained a full blood meal. Thus, host movements during that feeding period can influence the distribution of both the ticks and the Lyme disease bacterium. In general for vector-borne zoonotic diseases, different ecologies of the hosts, pathogens, and vectors may result in interactions that are distinctive to each system. We need continued research to disentangle these interactions and to understand the disease risks.

The community composition of wildlife host species has the potential to influence human disease risk, by affecting transmission rates and/or infection prevalence. Host composition and their relative abundances are particularly important for generalist pathogens. For example, having a community composed of non-competent reservoir species in greater abundances, relative to competent reservoir species, leads to lower Lyme disease risk (Ostfeld and Keesing 2000). The beneficial role of host composition and biodiversity on reducing disease risk has also been demonstrated in other studies that have tested the

hypothesis, including West Nile virus (Swaddle and Calos 2008, Loss et al. 2009, Vuong et al. 2012), hantavirus (Clay et al. 2009, Dizney and Ruedas 2009, Suzán et al. 2009, Voutilainen et al. 2012), schistosomiasis in snails (Johnson et al. 2009), fungal pathogens in planktonic communities (Hall et al. 2009), and even plant pathogens (Mitchell et al. 2002, Haas et al. 2011).

At the core of ecological interactions we can observe are the selective forces that act on both hosts and pathogens, with the host immune system adapting to better detect pathogens, and the pathogen adapting to better evade the immune system (Schulenburg et al. 2009). The coevolutionary relationship between hosts and their pathogens drive genetic polymorphisms in both protagonists at loci involved with host-pathogen interactions (Tellier and Brown 2007). Pathogen adaptations can be even more complex when dealing with multiple host species and their immune systems, perhaps explaining high genetic variation in pathogens such as *Anaplasma phagocytophilum* (Foley et al.) and *Borrelia burgdorferi* (Qiu et al. 2002), two bacterial pathogens transmitted by the same tick vector, which can infect a wide array of host species. For both pathogens, some genotypes appear to evade some host species' immune system better than others. Alternatively, a simpler explanation for genotypic diversity in *B. burgdorferi* could be that multiple niche polymorphism supports helps to maintain variation (Brisson and Dykhuizen 2004). Additionally genotypic differences are important to human disease risk, as some types tend to be more associated with human infections than other types (Seinost et al. 1999, Dykhuizen et al. 2008).

For this dissertation, I examined the ecology of Lyme disease in New York State to understand its range expansion, the role of host community composition and diversity on disease risk, and niche differentiation of genotypes detected within various host species. This

research is significant for several reasons. First, it addresses the ecological underpinnings associated with genotypic diversity of the pathogen, in light of host interactions. Second, the research examines disease risk at the genotypic level, which offers a finer resolution than prevalence and density of infected ticks alone can provide. Third, understanding how genotypes are distributed in space and time is critical for disease risk, as this disease will continue to spread in the northeast. Fourth, this study sheds new light on the contributions to disease risk of host species previously not well known for either their role in cycling this pathogen (Giardina et al. 2000) or for the specific *B. burgdorferi* genotypes associated with human infections.

Lyme disease is the most frequently reported vector-borne disease in the U.S. (Orloski et al. 2000, CDC 2011). The majority of cases occur in the northeast, followed by the Midwest, and a small number of cases occur in northern California. New York State alone had over 2,400 cases in 2010, out of the 22,000 cases reported in 2010 (CDC 2011). Lyme disease was first recognized in the mid 1970s in Lyme, CT when a group of young children were stricken with juvenile arthritis (Steere et al. 1977). It was not until several years later that the bacterium, *Borrelia burgdorferi*, was identified as the etiological agent causing the disease (Burgdorfer et al. 1982).

The black-legged tick (formerly known as deer tick, or *Ixodes dammini*) is the primary vector transmitting *B. burgdorferi* in the Northeast and Midwest (Burgdorfer et al. 1982). Black-legged ticks are found year round in northeastern, deciduous forests (Falco et al. 1999). Their life cycle consist of four stages over two years. Eggs are laid in winter and larvae emerge free of *B. burgdorferi* in late summer. These larvae often feed on small, ground dwelling animals that are often infected with *B. burgdorferi* (e.g., white-footed mice,

Peromyscus leucopus, and eastern chipmunks, *Tamias striatus*), and emerge as infected nymphs the following late spring/early summer (Ostfeld et al. 2006). These infected nymphs can then transmit the pathogen to susceptible hosts, making these newly infected hosts carriers of the pathogen. The nymphal period is most problematic for humans because questing nymphal ticks are small and difficult to detect, have high infection prevalence, and are common in spring and summer, when people enjoy the outdoors (Barbour and Durland 1993). After this initial feed, nymphs undergo another short resting period and molt to become adult ticks in the fall. Although infection prevalence is greater in adult ticks than in nymphs, the larger size of the adult ticks makes them easily detectable and removable (Lane et al. 1991). Additionally, adult tick populations have less overlap with human activity, because their peak activity occurs in late October/early November. Because of the ticks' two- year life cycle, with nymphs emerging and feeding before the larvae, *B. burgdorferi* can be maintained continuously within hosts and vectors.

Several host species are important in *B. burgdorferi* transmission cycles in the wild. Larval and nymphal ticks commonly feed on *P. leucopus* and *T. striatus*, with *T. striatus* providing a larger proportion of blood meals to nymphal ticks compared to other host species (Levine et al. 1985, Brunner and Ostfeld 2008). However, short-tailed (*Blarina brevicauda*) and *Sorex* shrews are also implicated as competent reservoirs (Brisson et al. 2008). The white-tailed deer, *Odocoileus virginianus*, is not important for the cycling of the pathogen, because they are refractory to infections (Magnarelli et al. 1984), but appear to be important as the final blood meal for adult ticks, where the adults find mates (Schmidtmann et al. 1998). Ground-nesting birds (e.g., veery, *Catharus fuscescens*) are important because of their

ability to disperse ticks longer distances than mammals (Ogden et al. 2008), possibly serving as amplifying host when rodent populations are low (Giardina et al. 2000).

In the northeastern U.S., there are 17 major groups of *B. burgdorferi* based on the outer surface protein C (*ospC*) locus (Seinost et al. 1999, Wang et al. 1999, Brisson and Dykhuizen 2004). The *ospC* gene is up-regulated when a tick is feeding, allowing the pathogen to leave the midgut of the tick and enter the salivary gland, facilitating transmission to the hosts (Fingerle et al. 1995, Schwan et al. 1995). The high diversity of these genotypes are indicative of balancing selection, most likely a result of the multiple niche polymorphism encountered with the bacterium infecting a wide array of host species (Brisson and Dykhuizen 2004). Genotypic diversity is also important to human health risk, because not all genotypes infect humans equally. Five of seventeen genotypes are commonly detected in Lyme disease patients (Seinost et al. 1999, Dykhuizen et al. 2008), and are also frequently detected in some small mammalian (Brisson and Dykhuizen 2004) and avian hosts (Ogden et al. 2008). By broadening our understanding of the role of host species on *ospC* composition and frequency profile, we can achieve an improved assessment of the risk of Lyme disease.

The objective here is to understand Lyme disease risk at the genotypic level, over different spatial and ecological scales. For chapter one, I asked how time since invasion of Lyme disease into different regions of New York State has affected the *ospC* composition and frequency profile. The three regions are southeastern New York (lower Hudson Valley), the Capital region (around Albany), and Outskirts (western New York), which correspond as oldest to newest invasion, respectively. I examined *ospC* α, β, γ-diversity, and determined whether the proportions we detected for each genotype in the tick populations occurred at proportions expected by random chance. I also tested whether there were differences in the

proportions of human invasive strains among the five populations, and offer ideas on how *ospC* types are distributed from endemic areas to newly invaded areas.

In chapter two, I focused on how changes in the host community composition influence *ospC* frequency profiles in Dutchess County, New York. I queried the factors associated with the host community that were the best predictors of *ospC* diversity, and also human infection genotype (HIS) prevalence in the ticks. Because higher host diversity has been shown to result in lower Lyme disease risk, based on nymphal infection prevalence (LoGiudice et al. 2003, LoGiudice et al. 2008), I wanted to examine whether the same pattern holds for *ospC* diversity, given that we know competent reservoir hosts also support relatively high proportions of HIS.

In the last chapter, I was interested in how each host species contributes to the *ospC* composition and frequency profile, whether there are distinctions that can be made among host species and among higher taxa (i.e., birds, shrews, rodents), and the genotypic occurrence in the host individuals and the transmission efficiencies of each *ospC* type from infected host to tick vectors. This study elucidates whether hosts can act as ecological niches to *ospC* types (Brisson and Dykhuizen 2004), and offers a novel way of understanding the occurrence of *ospC* types and how they are maintained in the cycle from hosts to ticks. The study highlights the role that birds have in maintaining high *ospC* diversity in the tick population, and their importance is Lyme disease risk, as they also support high proportions of HIS types.

Altogether, the research in this dissertation sheds some new light on how host diversity might affect Lyme disease risk, both in endemic and newly invaded areas. By examining disease risk at the genotypic level, we see that original ideas about small

mammals being the primary important hosts may not be completely accurate. However, the research does not end here, as we are missing some key players in our study that could also affect disease risk. For example, Levi et al. (2012) recently highlighted the importance of the loss of foxes as a bigger driver to increased Lyme disease risk than changes in the deer population, because the loss of foxes releases the small mammals from predation pressures. Nonetheless, even initial disentangling of the web can offer novel insights, hypotheses, and questions, on Lyme disease ecology.

Citation

Barbour, A. G. and F. Durland. 1993. The biological and social phenomenon of Lyme disease. Science **260**:1610-1616.

Brisson, D. and D. E. Dykhuizen. 2004. *ospC* diversity in *Borrelia burgdorferi*: Different hosts are different niches. Genetics **168**:713-722.

Brisson, D., D. E. Dykhuizen, and R. Ostfeld. 2008. Conspicuous impacts of inconspicuous hosts on the Lyme disease epidemic. Proceedings of the Royal Society B: Biological Sciences **275**:227-235.

Brunner, J. L. and R. S. Ostfeld. 2008. Multiple causes of variable tick burdens on small-mammal hosts. Ecology **89**:2259-2272.

Burgdorfer, W., A. G. Barbour, S. F. Hayes, J. L. Benach, E. Grunwaldt, and J. P. Davis. 1982. Lyme disease--A tick-borne spirochetosis? Science **216**:1317-1319.

CDC. 2011. Lyme disease: division of vector-borne infectious diseases. http://www.cdc.gov/lyme/.

Clay, C. A., E. M. Lehmer, S. St Jeor, and M. D. Dearing. 2009. Testing mechanisms of the dilution effect: Deer mice encounter rates, Sin Nombre virus prevalence and species diversity. Ecohealth **6**:250-259.

Daszak, P., A. A. Cunningham, and A. D. Hyatt. 2000. Emerging infectious diseases of wildlife-threats to biodiversity and human health. Science **287**:443-449.

Dizney, L. J. and L. A. Ruedas. 2009. Increased host species diversity and decreased prevalence of Sin Nombre virus. Emerging Infectious Diseases **15**:1012-1018.

Dykhuizen, D. E., D. Brisson, S. Sandigursky, G. P. Wormser, J. Nowakowski, R. B. Nadelman, and I. Schwartz. 2008. The propensity of different *Borrelia burgdorferi sensu stricto* genotypes to cause disseminated infections in humans. Am J Trop Med Hyg **78**:806-810.

Falco, R. C., D. F. McKenna, T. J. Daniels, R. B. Nadelman, J. Nowakowski, D. Fish, and G. P. Wormser. 1999. Temporal relation between *Ixodes scapularis* abundance and risk for Lyme disease associated with erythema migrans. Am. J. Epidemiol. **149**:771-776.

Fingerle, V., U. Hauser, G. Liegl, B. Petko, V. Preac-Mursic, and B. Wilske. 1995. Expression of outer surface proteins A and C of *Borrelia burgdorferi* in *Ixodes ricinus*. J. Clin. Microbiol. **33**:1867-1869.

Foley, J., N. C. Nieto, J. Madigan, and J. Sykes. 2008. Possible differential host tropism in *Anaplasma phagocytophilum* strains in the western United States. Pages 94-97 *in* O. A. E. Sparagano, J. C. Maillard, and J. V. Figueroa, editors. Animal Biodiversity and Emerging Diseases: Prediction and Prevention. Blackwell Publishing, Oxford.

Fonseca, D. M., N. Keyghobadi, C. A. Malcolm, C. Mehmet, F. Schaffner, M. Mogi, R. C. Fleischer, and R. C. Wilkerson. 2004. Emerging vectors in the Culex pipiens complex. Science **303**:1535-1538.

Giardina, A. R., K. A. Schmidt, E. M. Schauber, and R. S. Ostfeld. 2000. Modeling the role of songbirds and rodents in the ecology of Lyme disease. Canadian Journal of Zoology **78**:2184-2197.

Haas, S. E., M. B. Hooten, D. M. Rizzo, and R. K. Meentemeyer. 2011. Forest species diversity reduces disease risk in a generalist plant pathogen invasion. Ecology Letters **14**:1108-1116.

Hall, S. R., C. R. Becker, J. L. Simonis, M. A. Duffy, A. J. Tessier, and C. E. Caaceres. 2009. Friendly competition: evidence for a dilution effect among competitors in a planktonic host-parasite system. Ecology **90**:791-801.

Hayes, C. G. 2001. West Nile virus: Uganda, 1937, to New York city, 1999. Pages 25-37 West Nile Virus: Detection, Surveillance, and Control.

Heymann, D. L. and G. Rodier. 2004. Global surveillance, national surveillance, and SARS. Emerging Infectious Diseases **10**:173-175.

Johnson, P. T. J., P. J. Lund, R. B. Hartson, and T. P. Yoshino. 2009. Community diversity reduces *Schistosoma mansoni* transmission, host pathology and human infection risk. Proceedings of the Royal Society B-Biological Sciences **276**:1657-1663.

Lane, R. S., J. Piesman, and W. Burgdorfer. 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.

Levi, T., A. M. Kilpatrick, M. Mangel, and C. C. Wilmers. 2012. Deer, predators, and the emergence of Lyme disease. Proceedings of the National Academy of Sciences **109**:10942-10947.

Levine, J. F., M. L. Wilson, and A. Spielman. 1985. Mice as reservoirs of the Lyme disease spirochete. The American journal of tropical medicine and hygiene **34**:355-360.

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

LoGiudice, K., S. T. K. Duerr, M. J. Newhouse, K. A. Schmidt, M. E. Killilea, and R. S. Ostfeld. 2008. Impact of host community composition on Lyme disease risk. Ecology **89**:2841-2849.

LoGiudice, K., R. S. Ostfeld, K. A. Schmidt, and F. Keesing. 2003. The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences of the United States of America **100**:567-571.

Loss, S., G. Hamer, E. Walker, M. Ruiz, T. Goldberg, U. Kitron, and J. Brawn. 2009. Avian host community structure and prevalence of West Nile virus in Chicago, Illinois. Oecologia **159**:415-424.

Magnarelli, L., J. Anderson, and W. Chappell. 1984. Antibodies to spirochetes in white-tailed deer and prevalence of infected ticks from foci of Lyme disease in Connecticut. Journal of Wildlife Diseases **20**:21-26.

Mitchell, C. E., D. Tilman, and J. V. Groth. 2002. Effects of grassland and plant species diversity, abundance, and composition of foliar fungal disease. Ecology **83**:1713-1726.

Nasci, R. S., H. M. Savage, D. J. White, J. R. Miller, B. C. Cropp, M. S. Godsey, A. J. Kerst, P. Bennett, K. Gottfried, and R. S. Lanciotti. 2001. West Nile virus in overwintering Culex mosquitoes, New York City, 2000. Emerging Infectious Diseases **7**:742-744.

Ogden, N. H., L. R. Lindsay, K. Hanincova, I. K. Barker, M. Bigras-Poulin, D. F. Charron, A. Heagy, C. M. Francis, C. J. O'Callaghan, I. Schwartz, and R. A.

Thompson. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. Appl. Environ. Microbiol. **74**:1780-1790.

Orloski, K. A., E. B. Hayes, G. L. Campbell, and D. T. Dennis. 2000. Surveillance for Lyme dsease: United States, 1992-1998. Morbidity and Mortality Weekly Report **49**:1-11.

Ostfeld, R., C. D. Canham, K. Oggenfuss, R. J. Winchcombe, and F. Keesing. 2006. Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. Plos Biology **4**:1058-1068.

Ostfeld, R. and F. Keesing. 2000. The function of biodiversity in the ecology of vector-borne zoonotic diseases. Canadian Journal of Zoology **78**:2061-2078.

Qiu, W. G., D. E. Dykhuizen, M. S. Acosta, and B. J. Luft. 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**:833-849.

Schmidtmann, E. T., J. F. Carroll, and D. W. Watson. 1998. Attachment-site patterns of adult blacklegged ticks (Acari: Ixodidae) on White-tailed deer and horses. Journal of Medical Entomology **35**:59-63.

Schulenburg, H., J. Kurtz, Y. Moret, and M. T. Siva-Jothy. 2009. Ecological immunology. Philosophical Transactions of the Royal Society B-Biological Sciences **364**:3-14.

Schwan, T. G., J. Piesman, W. T. Golde, M. C. Dolan, and P. A. Rosa. 1995. Induction of an outer surface protein on *Borrelia burgdorferi* during tick feeding. Proceedings of the National Academy of Sciences of the United States of America **92**:2909-2913.

Seinost, G., D. E. Dykhuizen, R. J. Dattwyler, W. T. Golde, J. J. Dunn, I. N. Wang, G. P. Wormser, M. E. Schriefer, and B. J. Luft. 1999. Four clones of *Borrelia burgdorferi sensu stricto* cause invasive infection in humans. Infection and Immunity **67**:3518-3524.

Spielman, A. 2001. Structure and seasonality of nearctic *Culex pipiens* populations. Pages 220-234. West Nile Virus: Detection, Surveillance, and Control.

Steere, A. C., S. E. Malawista, D. R. Snydman, R. E. Shope, W. A. Andiman, M. R. Ross, and F. M. Steele. 1977. Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. Arthritis and Rheumatism **20**:7-17.

Suzán, G., E. Marcé, J. T. Giermakowski, J. N. Mills, G. Ceballos, R. S. Ostfeld, B. Armién, J. M. Pascale, and T. L. Yates. 2009. Experimental evidence for reduced rodent diversity causing increased Hantavirus prevalence. PLoS ONE **4**:e5461.

Swaddle, J. P. and S. E. Calos. 2008. Increased avian diversity is associated with lower incidence of human West Nile infection: Observation of the dilution effect. PLoS ONE **3**.

Taylor, L. H., S. M. Latham, and M. E. J. Woolhouse. 2001. Risk factors for human disease emergence. Philosophical Transactions: Biological Sciences **356**:983-989.

Tellier, A. and J. K. M. Brown. 2007. Stability of genetic polymorphism in host–parasite interactions. Proceedings of the Royal Society B: Biological Sciences **274**:809-817.

Voutilainen, L., S. Savola, E. R. Kallio, J. Laakkonen, A. Vaheri, O. Vapalahti, and H. Henttonen. 2012. Environmental change and disease dynamics: effects of intensive forest management on Puumala hantavirus infection in boreal Bank Vole populations. PLoS ONE **7**:e39452.

Vuong, H. B., M. Remmenga, R. Creamer, and D. F. Caccamise. 2012. Ecological associations of West Nile virus and avian hosts in an arid environment. Pages 3-21 *in* E. Paul, editor. Emerging Avian Disease. University of California Press.

Wang, I., D. E. Dykhuizen, W. Qui, J. J. Dunn, E. M. Bosler, and B. J. Luft. 1999. Genetic diversity of *ospC* in a local population of *Borrelia burgdorferi sensu stricto*. Genetics **151**:15-30.

Chapter 1: *Borrelia burgdorferi* genotypic distribution along a relative invasion time scale in New York State

Abstract

Our research examines how time since invasion influences the genotypic diversity of the Lyme disease bacterium, *Borrelia burgdorferi*. Seventeen *B. burgdorferi* genotypes, based on variation at the outer surface protein C (*ospC*) locus, are known to circulate among the wide array of host species. We asked whether time since invasion would positively correlate with higher *ospC* richness, greater *ospC* genotypic diversity, and increased infection prevalence in ticks. Our research encompassed three regions in New York State – Southeastern (SENY), Capital, and Outskirts, which corresponds with oldest to most recent invasion times. We had five populations (Capital, SENY08, SENY11, Outskirts08, Outskirts11), due to sampling in two years for SENY and Outskirts, with multiple sites within a population. We tried to obtain 30 ticks per site, but this was not possible at some sites. We tested the ticks for *B. burgdorferi* using PCR amplification, and tested *B. burgdorferi* positive samples for *ospC* genotypes, using a reverse line blot method. We calculated Shannon-Weiner diversity estimates of *ospC*, and jackknifed the data to obtain beta diversity estimates. We also use jackknife sampling to determine whether the *ospC* proportions in each population were expected by random chance. We compared the three most commonly occurring *ospC* proportions in our populations against previous studies to evaluate *ospC* genotypic dispersal to the Outskirts. We found general support for the conclusion that more recently invaded areas have lower *ospC* richness, diversity, and infection prevalence. Turnover of *ospC* was large and significant between Capital and SENY08, and between Capital and SENY11, partly due to higher *ospC* richness in the

Capital population. Four of fifteen genotypes were detected more often than expected by chance in the Capital population, but all of those genotypes occurred exclusively in this population. SENY08 and SENY11 had five genotypes, which occurred more frequently than expected, with little overlap in the most frequent genotypes. Similarly, Outskirts08 and Outskirts11 had two non-overlapping genotypes that were frequently occurring in these populations, potentially suggesting some inherent bias in the movement of those genotypes to the "advancing front." Human invasive genotypes were generally part of the top three genotypes occurring in four of the five populations, comparable to other studies that examine local cycling and spread of *ospC* genotypes. Although the Capital region had low incidence rates only 20 years ago, it currently has similar incidence rates and *ospC* diversity and higher *ospC* richness than SENY, suggesting that both the Capital and SENY regions should be considered as one large endemic Lyme disease zone. The relatively high frequency of human invasive genotypes occurring in four of five populations signals that these genotypes may be present in vagile host species such as birds or larger mammals, which can readily spread these genotypes, and not restricted to competent, small mammal reservoirs. Lyme disease will likely continue to spread in the state, hence understanding the distribution of *ospC* genotypes, particularly those associated with increased human infection, offers better insights into disease risk.

Introduction

Lyme disease is the most frequently reported vector-borne disease in the U.S. (Orloski et al. 2000, CDC 2011). The majority of cases occur in the northeast, followed by the Midwest, and a small number of cases occur in northern California. In 2010, over 22,000 cases were reported in the US, with over 2,400 cases reported from New York State (NY State) alone (CDC 2011). Lyme disease continues to spread in New York State (White et al.

1991, Glavanakov et al. 2001), in addition to other areas of the northeast (Mathers et al. 2011, MacQueen et al. 2012), Midwest (Gatewood et al. 2009), and Canada (Ogden et al. 2008, 2011). Although increases in infection prevalence in ticks and incidence rates in humans are important metrics of disease risk, we should also be considering how pathogenic strains of the Lyme bacterium, *Borrelia burgdorferi s.s.*, are spreading in relation to disease risk. By examining disease risk at the genotypic level, we can obtain finer resolution on how the disease is spreading, rather than just the occurrence of the pathogen. For this study, we examined how relative time since invasion influences the genotypic distribution of *B. burgdorferi* in NY State.

Previous research on Lyme disease ecology enables us to make broader connections on how risk may change from endemic areas of Lyme disease to newly invaded areas. First, *B. burgdorferi* is a well-studied pathogen, and we have a wealth of information on host reservoir competence (LoGiudice et al. 2003, Keesing et al. 2009), tick survival (Brunner et al. *in prep*), disease expansion (White et al. 1991, Dennis et al. 1998, Hoen et al. 2009), and general ecology (Ostfeld 1997, Jones et al. 1998). Second, *B. burgdorferi* can be classified into multiple genotypes, based on allelic diversity at the outer surface protein C (*ospC*) locus, with only a few genotypes highly associated with human cases (Seinost et al. 1999, Dykhuizen et al. 2008, Strle et al. 2011). Third, there is some evidence that *ospC* genotypes are associated with particular host species (Brisson and Dykhuizen 2004). The composition of host communities may then influence the distribution and circulation of these genotypes, as well as the derivative patterns of human infection. Fourth, while there are no well-established invasion dates for Lyme disease in different areas of NY State, the incidence rates of the disease recorded by the state health department over the last 16 years permit

relative ordering of invasion dates among areas (NYSDOH 2012). The combination of long-term research on Lyme disease, relative invasion times, and host-*ospC* genotype associations allow us to test several hypotheses concerning the spread of this disease.

Borrelia burgdorferi (Burgdorfer et al. 1982, Johnson et al. 1984), a spirochete bacterium, can infect a variety of wildlife host species, and is transmitted among wildlife and humans via the blacklegged tick (*Ixodes scapularis*) vector (Steere et al. 1978, Burgdorfer et al. 1982). The bacterium can also be transmitted at low infection levels via other tick vectors in the US (Salkeld and Lane 2010, Brinkerhoff et al. 2011). The life history of *I. scapularis* consists of four life stages: egg, larva, nymph, and adult. In the northeast, larval ticks hatch from early to mid August, are born free of the pathogen, but can pick up the bacterium by feeding on an infected host. In the following spring, these larvae have reached the nymphal stage, and if infected, can transmit the pathogen to a non-infected host. If not infected, these nymphs can acquire the pathogen from an infected host, while feeding. Several months later, these nymphs emerge as adults, take their final blood meal, reproduce and die. The nymphal stage is considered the most dangerous infective stage, as it is small, present when people are active outdoors, with some proportion of the tick population infected (Ostfeld et al. 1996, Ostfeld 1997).

Lyme disease was first diagnosed in Lyme, Connecticut, in a handful of children and adults with inflammatory arthritis (Steere et al. 1977). Later, the etiological agent was detected in ticks (Burgdorfer et al. 1982) and shown to occur in wildlife. The disease subsequently spread throughout the northeast, and in NY State, Westchester County was considered a hot-spot for Lyme disease (Glavanakov et al. 2001). Lyme disease incidence continues to spread in a northerly and westerly direction, perhaps tracking the potential range

expansion of *I. scapularis* (White et al. 1991). Although deer populations are more widespread across the U.S., we do not see complete overlap in the tick and LD incidence with the deer populations (Dennis et al. 1998, Brownstein et al. 2003). It may be that habitat is present, but abiotic features of the habitat are not conducive to the population maintenance of the vector, and hence of the bacterium. Conversely, there may be good habitat available for the tick population to be sustained, but detection of these ticks is limited.

The transmission efficiencies of various genotypes of *B. burgdorferi* can also influence disease risk (Dykhuizen et al. 2008). The *ospC* locus is under diversifying selection, since it is the target of antibody response from the wide array of host species that can become infected with the bacterium (Wang et al. 1999, Qiu et al. 2002, Brisson and Dykhuizen 2004, Travinsky et al. 2010). There are 22 known major groups of *ospC* genotypes (henceforth *ospC* types) in the United States, with 15 groups occurring commonly in the Northeast (Wang et al. 1999, Qiu et al. 2002, Brisson and Dykhuizen 2004). There is approximately 2-8% sequence variation within any major *ospC* group, but greater than 8% divergence among major groups (Qiu et al. 2002).

Brisson and Dykhuizen (2004) suggested that host species act as separate ecological niches for *ospC* types. They show that certain *ospC* types were more frequently detected in the four host species they examined (white-footed mouse – *Peromyscus leucopus*, eastern chipmunk – *Tamias striatus*, gray squirrel – *Sciurus carolinensis*, and short-tailed shrew – *Blarina brevicauda*) than would be expected if the genotypes infected these four host species equally. This may be due to differential evasion of *ospC* types by the host immune systems. White-footed mice, and secondarily, chipmunks, are good at supporting human invasive genotypes (hereafter HIS), compared with short-tailed shrews and gray squirrels. HIS

genotypes are types detected in relatively higher frequencies from infected people, and they include genotypes A, B, I, K, and sometimes N (Seinost et al. 1999, Dykhuizen et al. 2008, Strle et al. 2011). For example, larval ticks feeding on white-footed mice had detectable genotypes A, B, I, and K, but larval ticks feeding on squirrels were primarily detected with genotypes A and K. Qualitatively, most *ospC* types infect most animal host species (Hanincova et al. 2006), but the difference in relative frequencies of each type from host species to host species, suggests that there may be genotypic biases in which types are associated with which host species (Brisson and Dykhuizen 2006).

There is increasing interest in avian host-*ospC* genotype associations because of the potential for migratory birds to carry infected ticks to newly invaded areas. Research in Canada initially showed that birds carried a smaller subset of *ospC* genotypes compared to areas of endemic cycling in northeastern U.S. (Ogden et al. 2008), but after several years of studies in the same area, all 17 known *ospC* genotypes in the northeast had been encountered. Recent invasions into Maine also show similarly low numbers of *ospC* genotypes in the tick population (Mathers et al. 2011, MacQueen et al. 2012). In three of the studies, *ospC* genotypes A and K, and sometimes N, tend to dominate the infections detected in ticks from these hosts within a few years of detecting LD in the area (Ogden et al. 2008, Ogden et al. 2011, MacQueen et al. 2012). Since the dispersal of *ospC* is limited by the movement of their tick vector and host species, and tick dispersal depends on attachment to a host, the host-*ospC* associations and vagility of host species could influence the distribution and diversity of *ospC* types at both local and broader, regional scales. We would expect that *ospC* genotypes more associated with mobile hosts species would be found in more geographic areas, due to the dispersal capabilities of these host species.

Our primary research question here is: How does time since invasion of *B. burgdorferi* influence the genotypic diversity of *ospC* detected in the tick population? Secondarily, we ask whether there are particular genotypes that may be dispersed more readily than others. We hypothesize that areas with greater time since invasion will exhibit greater prevalence, number of *ospC* types, and diversity of *ospC* types than more recently invaded areas. We also examine the top three genotypes in each region and offer potential conclusions on the influence of host movement on *ospC* genotype dispersal.

Methods

Site selection & Tick collection – We obtained LD incidence rates in all NY State counties from the New York State Department of Health's (NYSDOH 2012) website for records from 1994 – 2010. Based on the endemic levels of LD at the beginning of recorded incidence in the state, we separated the counties into three regions – southeastern (SENY), Capital, and Outskirts. Counties in SENY had high incidence rates in the early 1980s, whereas counties in the Capital region had low incidence rates until about 2003, after which rates have increased to the present time. Counties in the Outskirts had incidence near zero until about 2007, but they remain considerably lower than the other two regions, suggesting an "advancing front" for LD, as it spreads throughout the state. We sampled ticks from 13 counties within the three regions SENY: (Columbia, Dutchess, Orange, Rockland, and Westchester); Capital (Albany, Rensselaer, and Saratoga); Outskirts (Oneida, Onondaga, Otsego, Seneca, and Tompkins). Incidence values within each region were binned into five or six year intervals, and the rates were log transformed after adding 1.0 to each value to deal with zeros in the data. Box plots were generated to examine the mean and spread of the incidence rates for each region and time interval.

We sampled questing nymphal ticks from May 2008 to early July 2008, and in June 2011, from selected state parks and town/private properties, distributed across the three regions. We visited 23 sites, spread across 13 counties from Westchester County in the southeast to Saratoga County in upper Hudson valley, and to Seneca County in the Finger Lakes region of the state (Figure 1. 1). In similar deciduous forest, we collected at least 30 ticks per site, except in areas of low abundance. Our smallest tick collection was of 8 ticks at one site. We collected questing nymphal ticks by dragging a 1 m^2 white, corduroy cloth haphazardly on the forest floor and stopping to collect ticks from the cloth approximately every 30 steps. All nymphal and adult ticks were collected in 100% ethanol and labeled for future lab work and stored at -20°C.

DNA extraction – Individual nymphal ticks were crushed in 120 µl DNAzol and 0.1 mg/µl proteinase K and allowed to sit overnight in 1.5 ml microcentrifuge tubes. Samples were then incubated at 95°C for 10 minutes in a dry bath to inactivate any pathogen in the samples. The samples were cooled to room temperature and centrifuged at 10,000 g for 10 minutes. We transferred the supernatant into newly labeled vials with 50 µl of 100% ethanol. The samples were mixed vigorously followed by a 10 minute waiting period. Samples were centrifuged at 16,000 g for 10 minutes to generate a DNA pellet, and the pellet was washed with 50µl of 75% ethanol, with gentle inversions, and centrifuged at 16,000 g, and repeated once. We decanted the supernatant onto kimwipes and inverted the tubes for approximately 10 minutes to allow the pellet to dry partially. We hydrated and re-suspended the DNA with 35 µl of T.E (pH 7.5) and stored the samples in -20°C freezer until testing time. Ticks collected in summer 2011 were extracted using DNeasy 96 blood and tissue kit (Qiagen cat # 69581), following the manufacturer's protocol for spin column on plates when extracting

tissue. Samples were hydrated with 100µl buffer AE and stored in -20°C freezer until testing time.

qPCR – The *Borrelia* 23s rRNA gene was amplified with a real-time PCR (Courtney et al. 2004). A 75bp fragment was amplified, using Tamra probe Bb23Sp and primers Bb23Sf and Bb23Sr (Table 1. 1A), using the Applied Biosystems 7500 system. Each sample was tested in triplicates in a 20 µl reaction with a final concentration of 1x TaqMan Environmental Master Mix 2.0 (Life Technologies, Foster City, CA), 0.7 µM of each primer, 0.175 µM of the probe, and 2µl of DNA. The reaction procedure is as follows: 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Each test plate consisted of five *B. burgdorferi* standards, starting with 10^{-5} mg/µl of a 2mg/µl to 10^{-10} mg/µl of cloned 23s rRNA obtained from the Keesing lab at Bard College. Positive samples were then subjected to nested PCR for amplification of the *ospC* gene. Ticks collected in summer 2011 were not subjected to qPCR reactions.

Borrelia Nested *PCR* – We amplified a 596 bp fragment of *ospC* using a semi-nested PCR with outer primers OC6+ and OC623-, followed by internal primers OC6+F and OC602- (Table 1. 1A). The outer PCR reaction final concentrations in a 25 µl reaction are as follows: 1x Buffer, 3 mM MgCl2, 200 µM of each dNTP, 0.4 µM of each primer, 2.5 units of Ampli Taq (Applied Biosystems, Branchburg, NJ), and 2.5 µl of DNA. The Veriti thermocycler conditions were as follows: 95°C at 10min, 25 cycles of 95°C for 45sec, 56°C for 35sec, and 72°C for 1min, and 72°C for 10min for elongation, and a 10°C hold. 1 µl of PCR product was added to the second round of PCR amplification along with internal primers. For the nested PCR, 30 cycles of the same thermocycler conditions were used with the same concentrations, but with a final volume of 40 µl. To determine the presence of *ospC*

fragments, we ran the samples on 1% agarose gels. All samples were placed in -20°C until ready for testing different genotypes using the reverse line blot. Due to loss of fluorescein detection from early amplification, we re-amplified the positive samples using product from the outer PCR run to make new nested products, using 36 cycles rather than 30, just before testing for *ospC* types in the reverse line blot.

Reverse Line Blot (RLB) - Borrelia burgdorferi major *ospC* groups were detected using a reverse line blot procedure (hereafter RLB) (Qiu et al. 2002, Brisson and Dykhuizen 2004). This method is useful and more straightforward in separating multiple genotypes infecting a single tick than is traditional Sanger sequencing and SSCP (Qiu et al. 2002). The method utilizes short oligonucleotide probes that are specific to *ospC* major groups for hybridization of single stranded PCR products. Major groups are determined based on sequence similarity, with 2-8% sequence variation within a group, but more than 8% sequence variation among groups.

The RLB utilizes a positively-charged nylon membrane with tightly bound oligos that are specific for different major *ospC* groups, crosslinked at 125 kJ of energy in a stratalinker. These oligo probe sequences (Table 1. 1B) had been used previously by Qiu et al. (2002) and Brisson & Dykhuizen (2004), except for probe I+6-1. One nucleotide was removed from the probe sequence developed by Brisson & Dykhuizen (2004) to better match the *ospC* I major group sequence. Oligos were purchased from IDT and tailed with extra Thymidine bases, using a terminal transferase kit (Roche cat # 03333566001). The following procedure follows Brisson & Dykhuizen (2004): In a 50 µl reaction volume, we included 10 µl of 5X TdT reaction buffer, 10 µl of 25 mM CoCl2, 5 µl of 10 mM dTTP, 350U of terminal transferase (TdT), and 300 pmoles probes that were initially diluted in HPLC water to avoid inhibition of

EDTA in the terminal transferase reaction, and HPLC water to volume. The mixes were placed in a 37°C incubator for three hours, followed by a 10 min heat shock of 65°C water bath to kill the TdT. The tailed probes were stored in -20°C until ready for use.

Positive controls for the *ospC* major groups were obtained from D. Brisson at the University of Pennsylvania (Philadelphia, PA). Frozen clones in the -80°C freezer were quickly scraped and smeared onto imMedia Amp Blue (Life Technologies, Grand Island, NY) agar plates. The samples grew overnight at 37°C, and two colonies from each major group were picked for growth in luria broth containing 50 µg/ml ampicillin. Each colony was placed in 4 ml of luria broth in sterile culture tubes and allowed to shake overnight at 200 rpm at 37°C. One ml of luria broth with *E. coli* growth was set aside with approximately 250 µl of glycerol for storage in -80 C for positive control stocks. The remaining 3 ml were subject to a plasmid extraction protocol from the Qiagen Mini Plasmid Extraction kit (Qiagen cat # 12123). The DNA was eluted with a final volume of 50 µl. Due to high concentrations of plasmid DNA, we only used the nested PCR protocol to amplify the DNA for these positive controls. Samples were run on a 1% agarose gel to inspect for the presence of the correct band size. The samples were cleaned using ExoSAP and sent for sequencing at the UPenn sequencing facility. The sequences were compared to sequences from Wang et al. (1999), downloaded from GenBank, as well as those from D. Brisson, to ensure that correct sequence matches with the major groups.

To prepare the membrane for probe linking, we placed a new membrane on the slot side of the Immunogen Minislot 30 manifold and marked the back corner of the membrane to determine orientation of the slots. We wet the membrane with dH_2O, placed two sheets of thick, square cut filter paper on the back of the membrane and wet both sheets, then replaced

the back side of the Minislot 30. The manifold was secured together with plastic screws to ensure that liquids in each lane (or slot) do not spread between lanes. Each lane corresponds to a probe specific for the *ospC* major groups. We diluted 10 µl of tailed probe into 2 mL of TE and added this mix to a lane in the manifold with a gentle vacuum suction. When the liquid in the lanes was removed, we washed the lanes twice with 2 mL of TE. Upon removal of the liquids, we removed the membrane from the Minislot 30 and immediately crosslinked the probes to the membrane. The membrane was rinsed with dH_2O and stored at 4 C with 2x SSC buffer until ready for use.

The membrane was rotated 90 degrees and placed in the Immunogen Miniblotter 45, slot side down. Two square foamy pads were placed behind the membrane followed by the back side of the miniblotter and secured with plastic screws. Residual liquid in the lanes was removed with vacuum suction. Each lane was filled with 100 µl of DIG Easy hybridization solution with 0.1 mg/µl polyA (polyA from Sigma cat # P9403) and rocked slowly in a 42°C incubator for 2 hours. During this rocking period, the samples and water negative control were prepared to be tested by mixing 10-15 µl of PCR product (or water) into 85-90 µl of DNEasy hybridization solution containing 0.1 mg/ µl polyA. We mixed five *ospC* positive controls of 1 µl each into 95 µl of DNEasy hyb solution with polyA. All samples were denatured by boiling at 99°C in an eppendorf thermocycler for 10 min and quickly shocked on ice water to maintain single stranded PCR products. The hybridization solution previously placed in the manifold was removed with vacuum suction and the new hybridization solution with DNA products was placed into the individual lanes and rocked gently for another two hours at 42°C. Upon removal of the solution, 100 µl of 2x SSC/0.5% SDS was added to wash each lane by rocking the miniblotter at room temperature for 5 minutes and repeated once

more. Next, 100 µl of 0.1x SSC/0.5% SDS was added for another wash for 15 minutes at 50°C and repeated. After vacuum suction of the liquids from the manifold, the membrane was placed in a 2.6 L rectangular Pyrex dish and rinsed once for 5 min with 50 ml washing buffer, rocked at room temperature. All steps following were conducted at room temperature. The washing buffer included 0.1M Maleic Acid/0.15M Sodium Chloride and 0.3v/v of Tween20. After removing the washing buffer, we added 75 ml of Blocking buffer solution (1:10 Blocking solution to Maleic Acid buffer) to rock for 45 min at room temperature. This was followed by 25 mL of Blocking buffer with 1:5000 anti-fluorescein AP fragment goat anti-body and rocked at 45 min. Next, the membrane was washed twice with 70 mL of washing buffer, rocking for 15 min each. After removing the washing buffer, we added 30 mL of blocking buffer (1:10 Blocking solution: maleic acid) and rocked for 5 minutes. The membrane was then placed in a Food Saver vacuum bag and 1.5 ml of CSPD was added immediately to the membrane and spread evenly across the membrane. The reaction sat for five minutes before the CSPD liquid was squeezed out of the bag and sealed. To increase the reaction rate, we placed the sealed membrane in a 37°C incubator for 10 min. Finally, in a dark room, we placed a chemiluminescent x-ray film on top of the membrane and allowed it to be exposed for 10-20 minutes, before processing. The CSPD cleaves the AP fragment from DNA that has bound to its correct probe and allows the fluorescein label to light up, resulting in dark squares on the x-ray film when developed. The Roche DIG starter kit II (cat # 11585614910) provided the Easy hybridization solution, Blocking solution, and CSPD.

Statistical Analyses - Due to the inefficiencies of re-amplification of positive tick samples when we conducted the RLB, we ultimately tested 135 out of 169 positive tick samples. We obtained reliable *ospC* detections from 104 of those 135 samples, so all

analyses and discussions pertaining to *ospC* community diversity are based on these 104 samples. Although we tested the samples against 17 specific genotypic probes in addition to a control probe, we excluded *ospC* C from our analyses, as *ospC* C is a hybrid of *ospC* E and I, rendering it difficult to differentiate between single, double, and triple infections (Qiu et al. 2002).

We estimated the strain richness of *ospC* in each region using the Chao2 estimator in EstimateS version 8.2 (Colwell 2011). The Chao2 estimator uses incidence data to estimate richness, based on singletons and doubletons encountered in the data. We ran 1000 permutations and allowed for sampling with replacement. We used the bias correction estimator for all three regions. The Chao2 estimates and 95% confidence intervals were used to generate rarefaction curves.

GenAlEx (or Genetics Analysis in Excel, version 6.4) is a tool used in population genetics to examine population structure (Peakall and Smouse 2006). We used GenAlEx to estimate Shannon-Weiner (SW) diversity, and to examine the differences in *ospC* frequency profile among our years (β_{year} diversity), and within each population (αdiversity for Capital, SENY08, Outskirts08, SENY11, and Outskirts11). A jackknife permutation function within GenAlEx calculated the effective *ospC* richness, and pairwise turnover of *ospC* types (Δ_β diversity). Pairwise comparisons (Δ_β) between populations were calculated with Fisher's Exact tests. Before calculating α, β, and γ (across whole study) diversities in GenAlEx, we classified multiply infected ticks into ticks with single infections (i.e., if a tick is infected with *ospC* genotype A and B, it is treated as a tick with the A genotype and another tick with the B genotype) in order to obtain 14 different haplotypes, each of which correspond to a single *ospC* genotype, rather than having a greater number of multiply-infected combinations

for single ticks. This allows us to examine the αand ß diversity of *ospC* genotypes within and among the regional levels. This classification increased our sample size from 104 single (and multiply) infected ticks, to 243 occurrences of a single *ospC* type infection.

We tested for differences in prevalence of *B. burgdorferi* genotypes among the regions using contingency table analyses and used a Bonferroni correction to make pairwise comparisons of prevalence. To determine whether the proportions of each genotype we detected in a region were within a 95% confidence interval of the observed proportion, we conducted a jackknife analysis on each genotype individually. We ran 1000 random permutations, without replacement, based on the total number of ticks with reliable *ospC* data (104 ticks). The permutations generated presence (1) and absence (0) data, based on the total numbers of ticks present with a particular *ospC* type and the total number of ticks without that same particular *ospC* type in each region and year. These permuted numbers were then reassigned to each region based on the numbers of *ospC* infections tested for those sampling strata. For example, in the 104 randomly drawn numbers of '0's and '1's in the first permutation, the first 48 generated numbers would be assigned to the Capital region, as there were 48 ticks with reliable *ospC* data from that region. The next 5 generated numbers would be assigned to the Outskirts08, 26 to SENY08, 5 to Outskirts11, and the last 20 numbers are assigned to SENY11 for a total of 104 ticks. This was repeated another 999 times to obtain reliable expectations and 95% confidence limits. All statistical analyses were carried out in program R (R 2008) (version 2.15) unless otherwise noted.

Finally, we were interested in determining the potential mechanism for differences in the geographic spread of *ospC* genotypes. The vagility of different host species might have the potential to influence the distribution and frequency profile of *ospC* given the

transmission of specific *ospC* types associated with those host species to the tick vectors that get dispersed to new areas (Brisson and Dykhuizen 2004, Hanincova et al. 2006). Host species that are more vagile should spread *ospC* types associated with them farther than types associated with less vagile species. To make predictions on which genotypes would spread farther, based on host vagility, we searched the literature for known host species-*ospC* associations, or for broad host group-genotype associations (i.e., avian hosts). We used the authors' data to determine which three *ospC* genotypes were most frequently detected in each host species (or group), to establish how each species might contribute to the dispersal of *ospC* genotypes (Brisson and Dykhuizen 2004, Hanincova et al. 2006, Ogden et al. 2008, Mathers et al. 2011, Ogden et al. 2011, MacQueen et al. 2012). We considered avian hosts as the most vagile, followed by larger mammals (e.g., raccoons and opossums), with the least vagile being small mammals (e.g. mice and shrews). Although deer are vagile hosts, they are fed upon more by adult ticks than by larvae or nymphs, and they are refractory to *Borrelia* infections (Magnarelli et al. 1984). Hence, they would not contribute much to the distribution of *B. burgdorferi*. We compared the three most detected genotypes associated with these host species against the top three genotypes we detected for our entire study at the regional level.

Results

General patterns – The trend in LD incidence rates from 1994 to 2010 are consistently high in SENY, but the Outskirts and Capital regions show increasing incidence over time, with a slower increase in the Outskirts (Figure 1. 2). By the (2006-2010) time period, incidence rates in the Capital and SENY regions appear quite similar, whereas those in the Outskirts remain considerably lower.

We tested 796 nymphal ticks across 23 sites in 13 counties in three regions, and detected 168 ticks infected with *B. burgdorferi* (21.3%) within 20 sites in 11 counties. The

three sites with no positive ticks were all in the Outskirts region. We found a significant difference in *B. burgdorferi* prevalence over time, with a prevalence of 18.7% (124/664) in 2008 and 34.4% (45/131) in 2011 ($\chi^2 = 15.14$, df = 1, $p < 0.001$, Yates correction). Due to the year effect, we examined the data as five populations rather than three regions to test for population differences, and found that prevalence of *B. burgdorferi* differed among the populations ($\chi^2 = 21.60$, df = 4, $p < 0.001$). SENY11 had the greatest prevalence (36.0%, 32/89) followed by Outskirts11 (31.0%, 13/42), Capital (22.6%, 63/279), SENY08 (21.3%, 51/239), and Outskirts08 (6.8%, 10/146). None of the sites sampled in 2008 were sampled in 2011, so inter-annual comparisons of prevalence for particular sites across years are not possible.

ospC genotype detection – There were 48 reliable tick samples in the Capital region for *ospC* data and 46 tick samples in SENY. The rarefaction curves for these regions both asymptoted, indicating that *ospC* types in the region had been sufficiently sampled (Figure 1. 3). Although the curve did not yield an asymptote in the Outskirts with only 10 tick samples, the curve does appear to be decelerating (Figure 1. 3). These curves also show that Outskirts has lower *ospC* genotype richness than do SENY or Capital regions, although the 95% CI are large with the small sample size.

ospC α, β, and γ-diversities: The *ospC* γ-diversity in our study was H' = 2.573, which translates into 13.108 effective *ospC* types detected. We detected slightly higher *ospC* diversity in 2008 (β_{2008}: H' = 2.511) than in 2011 (β_{2011}: H' = 2.317) based on Shannon-Weiner (SW) diversity calculations (Table 1. 2A). The [0, 1] scaled divergence between the two years was $\Delta_{\beta year} = 0.143$ ($p = 0.011$, Fisher's Exact test), indicating large turnover of genotypes between 2008 and 2011. Thus, we conducted all subsequent diversity analyses

based on five populations rather than three (Capital, Outskirts08, SENY08, Outskirts11, and SENY11). The Capital had the highest α diversity (2.54), followed by the two SENY populations (SENY11 = 2.41, SENY08 = 2.31), and the lowest occurred with the two Outskirts populations (Outskirts11 = 1.75, Outskirts08 = 1.49). This also resulted in a lower effective *ospC* richness, compared to the actual observed richness (Table 1. 2A).

The pairwise comparisons using jackknife permutations between our populations showed that most population pairs had some effective non-overlapping genotypes (Table 1. 2B). The largest effective β_{pop} was between Capital and SENY11 (effective β_{pop} = 13.485 genotypes) while the smallest effective β_{pop} was between Outskirts08 and Outskirts11 (effective β_{pop} = 8.094 genotypes). The maximum β_{pop} diversity is less than a value of two, due to unequal sample sizes in our data (Table 1. 2C). The largest $\Delta_{\beta pop}$ diversity value was between the Outskirts08 and Outskirts11 ($\Delta_{\beta pop}$ = 0.737) due to high turnover of genotypes between the years, and the smallest $\Delta_{\beta pop}$ was between SENY08 and SENY11 ($\Delta\beta_{pop}$ = 0.070) due to having the same types occurring in both years, but at different frequencies. Nonetheless, the significant differences in the $\Delta_{\beta pop}$ diversity between any pair of populations were those between Capital and SENY08 (p = 0.007), and between Capital and SENY11 (p = 0.012) (Table 1. 2D).

Frequency profile of *ospC*: We tested whether the relative frequencies of each *ospC* type differed from random associations (Table 1. 3). In the Capital region, almost half of the genotypes were less frequent than expected by chance, and the four genotypes that were more frequent were the only genotypes exclusive to that population. For the SENY populations, most genotypes occurred at frequencies expected by chance, while five occurred at greater frequency. The five genotypes that occurred at higher frequencies in SENY08 were

not consistently high in SENY11. In the Outskirts, which had fewer *ospC* types detected, most of the types occurred at frequencies expected by chance. Similar to SENY, the most frequent types in Outskirts08 were not the most frequent types in Outskirts11.

Small mammals (squirrels, mice, shrews, and chipmunks) tended to have greater proportions of *ospC* K, A, D, and E (Brisson and Dykhuizen 2004, Hanincova et al. 2006, MacQueen et al. 2012) (Table 1. 4A, B, C). Medium sized mammals (raccoons and opossums) had greater *ospC* variability and yielded types A, B, D, I, and N (Hanincova et al. 2006). Avian communities have more often yielded *ospC* A, B, K, and to a lesser extent, G and N (Ogden et al. 2008, Mathers et al. 2011) (Table 1. 4D). There appears to be substantial genotypic overlap among the diverse host communities, suggesting that these genotypes may be more of a 'host generalist' than we had realized at the outset. Based on our sampling, type K was one of the top three genotypes in most sites (Table 1. 4E), while types B, G, and N were also relatively common in three of the five regions. Outskirts11 was most different in its top three genotypes detected, with high proportions of types D, E, and F, rather than high proportions of K and N, as was the case for the other sites.

We did not detect differences in the frequencies of HIS (human invasive genotypes) to non-HIS across the regions by year ($\chi^2 = 7.527$, df = 4, $p = 0.111$). In four of five cases, HIS types made up at least 40% of the *ospC* community. For Outskirts08, 75% of the genotypes were HIS compared with 20% in 2011. This contrasts with SENY, where *B. burgdorferi* has presumably been circulating in the tick and animal populations for a much longer period of time. For both 2008 and 2011, HIS types continued to make up 50% of the *ospC* communities. The Capital region was only sampled in 2008 and HIS made up 40% of the genotype types detected.

Discussion

Our results show that infection prevalence, *ospC* strain richness, and *ospC* diversity were lower in recently invaded areas (Outskirts) than in relatively more endemic areas of Lyme disease (Capital and SENY) (Figure 1. 3, Table 1. 2). There was relatively large turnover ($\Delta_{\beta pop}$) of genotypes among our five populations, due to differences in effective genotypic richness and frequencies. We also found that some human invasive genotypes (HIS) were the most frequently occurring genotypes in four of five populations, suggesting that vagile host species may be important for the spread of *B. burgdorferi* into newly invaded areas. Although there is overall support that there is a negative relationship between time since invasion and *ospC* frequency profiles at the broad scale, there are two facets of the data that do not adhere to this pattern.

First, since the ticks and/or bacterium arrived earlier in SENY (Barbour and Fish 1993) before the Capital region, we expected higher infection prevalence in the SENY08 and SENY11, but found that the Capital population had a prevalence value between the values detected in the two SENY populations. The Capital region may have reached an endemic level of bacterial cycling, like SENY, by the time we sampled, resulting in similarly high infection prevalence and *ospC* richness and diversity between these two regions. The rapid rate of change in incidence in the Capital region is surprising. Within a time span of about 15-20 years, the incidence rate, which was once as low as the current rate for the Outskirts, has become similar to that of the SENY region. This range expansion had been previously predicted for NY State, based on past incidence rate reports and spatial autocorrelation analyses (Glavanakov et al. 2001). Recently, researchers examining the expansion of LD in Canada suggest that the expansion rate is about 46 km yr^{-1} and that it can be faster in a warmer climate (Leighton et al. 2012). Given this rate, we should have seen that LD spread

from SENY to the Capital region in about three years. Most likely, the interactions between hosts and tick populations would have a larger, and more realistic, effect on *B. burgdorferi* expansion than climate warming (White et al. 1991, Rand et al. 1998).

Second, when we compared each of the regions over time, we found large increases in infection prevalence in both SENY and Outskirts from 2008 to 2011. The higher prevalence in SENY11 would provide more support for our hypothesis that older invaded areas have higher prevalence, but higher prevalence in Outskirts11, a newly invaded area, would not. The higher prevalence in Outskirts11 was unforeseen, given that human incidence rates are still lower in the Outskirts than in the Capital and SENY regions (Figure 1. 2). The increase in infection prevalence in both SENY11 and Outskirts11 could be potentially due to several factors. First, year 2011 could have had anomalously high infection prevalence. With a relatively steady incidence rate in the SENY region, we would have not expected large changes in infection prevalence in the ticks. Conversely, we would expect at least some increase in the Outskirts region, but the infection prevalence increased substantially from 6.8% to 31%. We are unclear what factors might have influenced this higher prevalence value. Two, the sites sampled in both regions in 2008 and 2011 are different, and the 2011 collections may have had generally higher infection prevalence. Because we sampled different sites in 2008 and 2011, we cannot determine whether site or year as the factor causing increases in infection prevalence. Another conflating factor is that the incidence rates for both SENY and Outskirts have large confidence intervals over the (2004 – 2010) time period. Three, there may be more transmission of the pathogen among vertebrate hosts and tick vectors by 2011 compared to 2008, leading to higher prevalence.

The number of different *ospC* genotypes detected in the Capital region was higher (16) compared to SENY11 (12), SENY08 (12), Outskirts11 (7), and Outskirts08 (5), and a similar pattern was observed from the effective richness based on Shannon-Wiener diversity estimates with Capital having higher richness (12.67) compared to SENY11 (11.15), SENY08 (10.07), Outskirts11 (5.74), and Outskirts08 (4.46) (Table 1. 2B). The Capital region had several rare genotypes detected, including *ospC* types J and O, which were detected only once and twice, respectively. Interestingly, *ospC* type O was not detected in questing nymphal ticks sampled from Long Island (Qiu et al. 2002) or in Millbrook, NY (Brisson and Dykhuizen 2004), nor from fed larval ticks collected from four common host species in Millbook, NY (Brisson and Dykhuizen 2004). Conversely, *ospC* type J was found in questing nymphs in both these studies, but again, was not as commonly detected as most other genotypes (Qiu et al. 2002, Brisson and Dykhuizen 2004). Our recent investigation on host-genotype associations (Vuong – chapter 3) detected *ospC* type J from red squirrels (*Tamiasciurus hudsonicus*), while *ospC* type O was detected in three bird species we sampled. Yet, both genotypes were detected from ticks collected from companion animals and humans in Canada, most likely distributed there by migratory birds (Ogden et al. 2011). The low circulation of these genotypes suggests that ticks may not feed as commonly on red squirrels or birds as they would on more abundant species in the community.

In questing nymphal ticks, *ospC* types L and M are also not very common (Qiu et al. 2002, Brisson and Dykhuizen 2004, Ogden et al. 2008, Ogden et al. 2011), but they are most commonly detected from ticks feeding on avian hosts (Vuong – chapter 3). The *ospC* type M detected in our study was primarily associated with one site in the Capital region, a site that also supported the highest infection prevalence (70%) of all sites in this study. Although not

a true island, the long and narrow park is bordered by approximately seven miles of the Hudson River and Schodack Creek, around the majority of the park. This could make it a good stopover or breeding area for long-distance migrants like Veeries and Wood Thrushes, but we do not have knowledge on the stopover sites of these migratory birds. Although we detect neither *ospC* L or M types in SENY, they have both been previously detected in the area (Qiu et al. 2002, Brisson and Dykhuizen 2004), and their absence may reflect *ospC* frequency changes over time. For example, Qiu et al. (1997) examined gene flow and migration of *ospA*, another locus of *B. burgdorferi*, and found that although the array of *ospA* types from year to year were similar, the relative frequencies of the types were annually quite variable.

α, β and γ-Diversity: We found that α-diversities were quite different among our five populations. High diversities in the Capital (2.539), SENY08 (2.310), and SENY11 (2.412) are suggestive that these areas have had *B. burgdorferi* circulating in the tick and wildlife population for a relatively long time. The presence of many different *ospC* types, coupled with an absence of highly dominant genotypes in the population, may be attributable to a diverse host community. The infection of the many vertebrate host species by the pathogen can help to maintain genotypic diversity of *B. burgdorferi* through balancing selection (Brisson and Dykhuizen 2004). In the Outskirts, α-diversities were much lower (Outskirts08: 1.494, Outskirts11: 1.748), hinting that this region is at the "advancing front" of LD. There are lower numbers of genotypes, and the preeminence of one or two of the initial *ospC* type invaders, suggest that only certain host species (that support these genotypes) were able to spread their *B. burgdorferi* genotypes to the Outskirts. Alternatively, all genotypes could be present, but the small population of ticks/and or *B. burgdorferi* could undergo stochastic

extinction, changing the *ospC* type and frequency profiles we detected in the tick population. The low richness and dominance of a few genotypes is not unusual for a newly invaded area. On a costal Maine island, where annual sampling for *I. scapularis* and small mammals were occurring, researchers were able to identify when *Borrelia* first arrived (MacQueen et al. 2012). In the first year of detection, *B. burgdorferi ospC* type H was the predominant founder genotype. Over time, greater numbers of *ospC* types emerged, which the authors attribute to gene flow from migratory birds dropping off infected ticks which carry other *ospC* types, leading to a progressive decrease in the dominance of *ospC* H. Four years after the first detection of *B. burgdorferi*, *ospC* type H was nearly absent, while types A, K, and U had become more common. These changes will likely continue until long-term circulation of these genotypes in the tick and host community stabilizes and maintains a strong balancing selection for *ospC* variation.

The β-diversities and [0,1] standardized $\Delta_{\beta pop}$ measures were quite large for most pairwise population comparisons, except for the SENY08/SENY11 ($\Delta_{\beta pop} = 0.07$) (Table 1. 2B, 2C). The low turnover of genotypes between SENY08/SENY11is probably an indicator that the host community composition in the area is similar enough to support similar *ospC* types and small changes in the frequency profiles of the *ospC* types. The SENY region is also quite old, compared with the Capital and Outskirts regions, in when *B. burgdorferi* invaded into this area. This long time cycling of the pathogen in the tick and host populations may signal some equilibrium in *ospC* dynamics. The larger turnover ($\Delta_{\beta pop} \geq 0.20$) between all other pairwise comparisons illustrates the potential of differing host community composition among these populations, thereby resulting in large turnovers of *ospC* types detected in each population.

Conversely, the large divergence ($\beta = 1.99$, $\Delta_{\beta pop} = 0.74$) between the two Outskirts populations strongly suggests that the Outskirts region represents an "advancing front" zone for LD. Although the Fisher's Exact test did not find the β-diversity between these two populations to be significant, this may be due to a small sample size in these two populations, which made it difficult to detect a significant difference. Being a newly invaded area, and having small tick and/or bacterial populations, stochastic processes could lead to the spotty occurrence of *B. burgdorferi* in the Outskirts. These spotty occurrences could affect the local *ospC* genotypic profiles, possibly explaining different frequencies of commonly detected *ospC* types between Outskirts08 and Outskirts11. Large changes in *ospC* genotypes in newly invaded areas have also recently been demonstrated in Maine (MacQueen et al. 2012). Periodic sampling in the same area would be beneficial in assessing invasion success of the ticks and/or bacterium.

Although the divergence between Capital and SENY08 ($\beta = 1.95$, $\Delta_{\beta pop} = 0.20$) and Capital and SENY11 ($\beta = 1.88$, $\Delta_{\beta pop} = 0.21$) are toward the lower end of the scale, relative to other pairwise comparisons, these values are not trivial with respect to their *ospC* array and frequency profile differences they represent. The higher effective α-diversity in the Capital ($\alpha = 2.54$) compared with SENY08 ($\alpha = 2.31$) and SENY11 ($\alpha = 2.41$) suggest that differences in host community composition, tick population dynamics, site conditions, and/or tick dispersal via vagile hosts could all be contributing to the large difference in α-diversity. More in-depth and long-term studies to examine change in the host community and in the *ospC* arrays of the bacterium would help to elucidate some of these differences.

The lower β-diversity value in 2011 ($\beta_{2011} = 2.32$), compared to 2008 ($\beta_{2008} = 2.51$), is most likely a result of not sampling in the Capital region in year 2011. We were able to

detect three genotypes that were exclusive to the Capital region. Still, the large difference in β-diversity between years ($\Delta_{\beta year} = 0.14$) may have been affected by the big turnover of genotypes between Outskirts08 and Outskirts11 ($\Delta_{\beta pop} = 0.74$). Even though the same sites were not sampled across years, changes in relative frequencies and *ospC* composition suggest that there is continual change in which *ospC* types most frequently occur. This may be due to changes in the host composition, stochastic losses of rare *ospC* genotypes, or even small tick populations. These continual fluctuations can make it difficult to assess disease risk locally, but our data show that broad patterns, especially in terms of human invasive genotype dispersal discussed below, may still be valid.

Frequency profile of *ospC*: The Jackknife analysis showed great variation in which genotype was more common, less common, or about the same as expected by chance. Within the Capital population, many of the *ospC* types were found less frequently than expected, averaged over the whole study, except for the genotypes that were exclusively detected in the Capital, which were more common than expected. However, due to the study-wide rarity of two of the four genotypes that were exclusive to the Capital region, it is not unexpected that the Capital proportions for those two types would be above the greater 95% CI. In both the SENY and Outskirts region, the genotypes that occurred more commonly in one year were not necessarily the common genotypes in another year (Table 1. 3). This could be supportive of temporal fluctuations of *ospC* types (Qiu et al. 1997), site variation due to sampling in non-overlapping sites between 2008 and 2011, or differences in the host community composition giving rise to different frequency profiles of *ospC* (Brisson and Dykhuizen 2004).

Based on the literature on genotypic associations for mammalian (Brisson and Dykhuizen 2004, Hanincova et al. 2006) and avian hosts (Ogden et al. 2008, Mathers et al. 2011, Ogden et al. 2011, MacQueen et al. 2012), the most prevalent genotypes associated with these hosts include A, B, D, G, K, and N (Table 1. 4A - D). This suggests that these genotypes may be supported by a wide variety of vertebrate hosts, which would permit these genotypes to more easily distribute to new areas, as well as circulate among vertebrate hosts and ticks more often in endemic areas. Indeed, we do see that these genotypes are often detected in high proportions across the landscape. Interestingly, four (A, B, K, and N) of the six commonly detected genotypes from the mammal and avian communities are also human invasive types (HIS) (Seinost et al. 1999, Dykhuizen et al. 2008). The proportion of HIS to non-HIS types were not different from one another across our five populations, which further suggests that these genotypes have an inherent bias in being distributed because of their ability to infect many host species. Although the similar proportions of HIS to non-HIS is detected across the regions and years and would suggest similar disease risk across all these regions, it should be noted that the Outskirts have lower tick populations and lower infection prevalence, especially Outskirts08, making these sites of lower risk compared to SENY, a longer invaded region of LD.

Caveats: We sampled in year 2011 in the Outskirts and SENY regions to increase our sample sizes and geographic spread of sites within these two regions. We combined the data for each region across the years in order to have a sufficient sample size to obtain reliable estimates for the rarefaction curve. Even so, we have a small sample size in the Outskirts region, but this is a reflection of low prevalence of ticks in these areas (Hoen et al. 2009), rather than poor sampling effort. As a region, the Outskirts had significantly lower infection

prevalence (12.2%, 23/188) than did the SENY (36.4%, 83/228) or Capital (22.6%, 63/279) regions, resulting in fewer samples with reverse line blot data and a poorer estimate of *ospC* strain richness detected with the Chao2 estimator. The rarefaction curve did not asymptote, but is clearly decelerating with increasing sample size, suggesting that the data need to be interpreted with some caution (Figure 1. 2). At the very least however, the Outskirts data to provide a glimpse of the recent infection history, for contrast with regions having longer histories of infection (Capital and SENY).

Our study was able to show that the spread of Lyme disease from endemic to newly invaded areas differs in terms of the genotypic variation detected along the continuum. We found greater prevalence of *B. burgdorferi*, higher numbers of *ospC* types, and higher diversity of *ospC* types in older invaded areas than in newly invaded areas. Our data also show that some HIS types are commonly occurring in both endemic areas and in the "advancing front," suggesting that these genotypes seem to evade the immune systems well and can be supported by a variety of host species, particularly vagile species like birds, which can potentially disperse infected ticks to areas that are more separates spatially. In this study, we focused only on the *ospC* genotypic distribution within questing nymphal ticks, but are left without knowledge on the role of host communities in influencing *ospC* dynamics. In the next chapter, I explore the role of host community composition and diversity to understand how they affect *ospC* diversity and HIS infection prevalence.

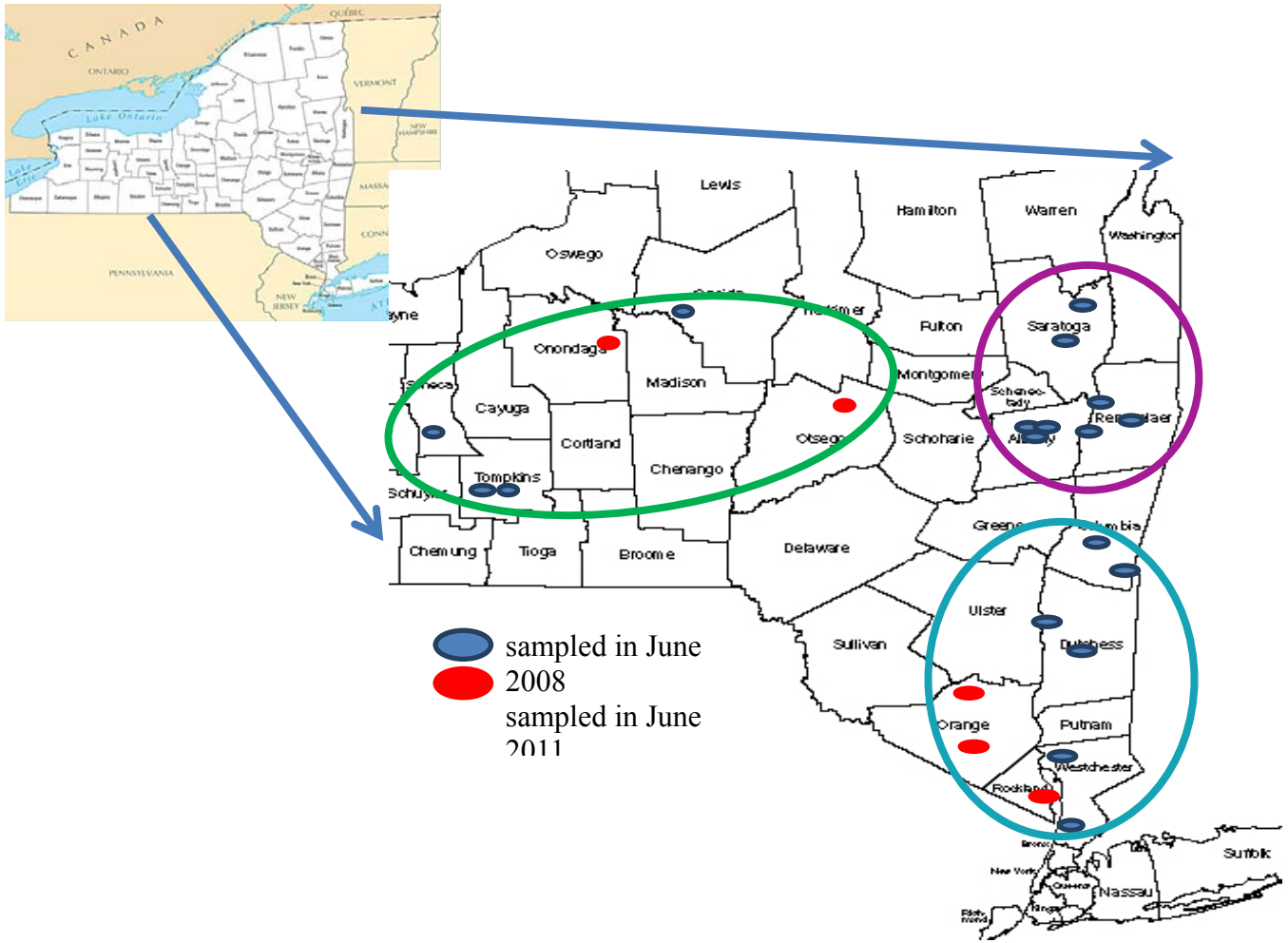


Figure 1. 1. Map of sampled sites within New York State. The blue circle represents the SENY region, purple is the Capital region, and the two green ovals are the Outskirts. We did not include the site with only four ticks in our data.

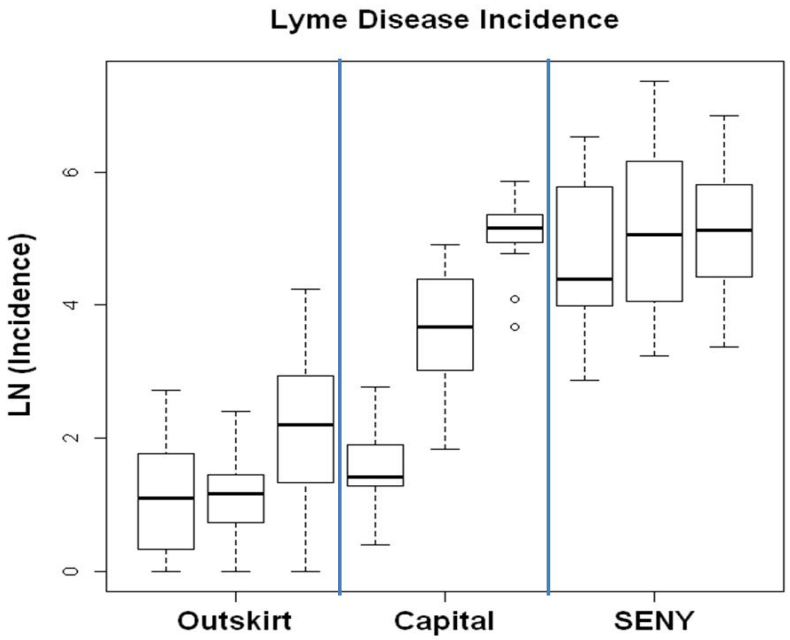


Figure 1. 2. Box and whisker plots of LD incidence rates from New York State Department of Health from 1994-2010. The rates are reported for each NY State county, with county data subsumed within regions in our analysis. Within each region, the plots are based on the following time intervals: 1994–1998, 1999–2003, and 2004–2010.

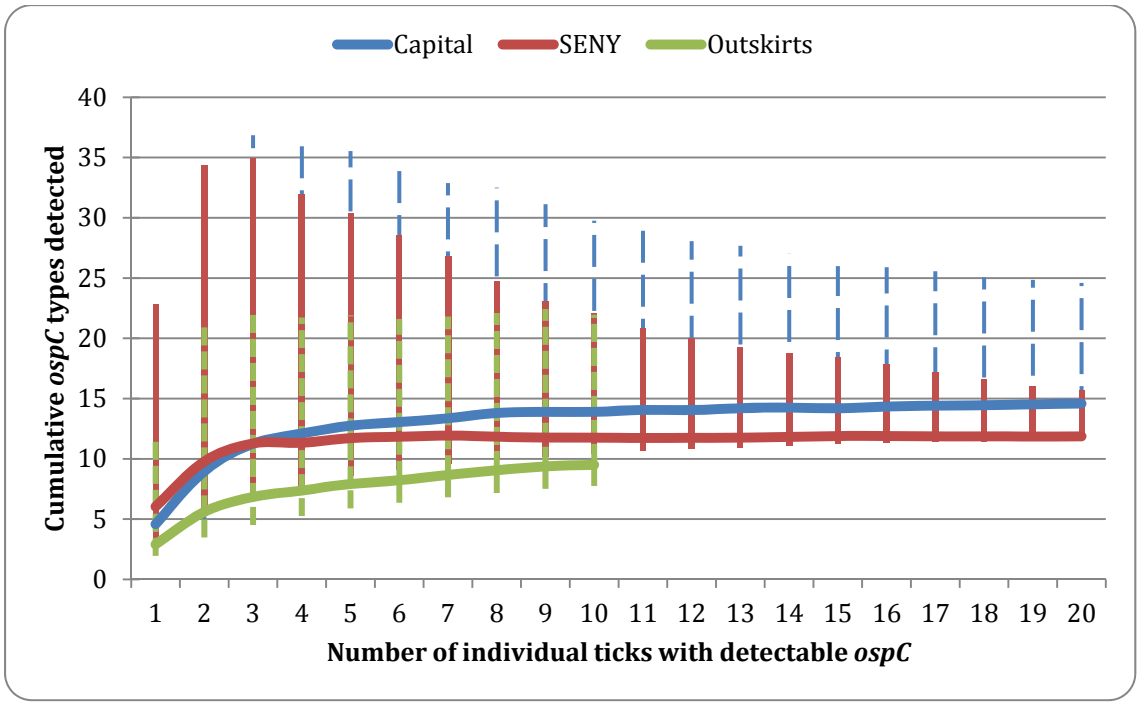


Figure 1. 3. Rarefaction curves with 95% CI for each region based from EstimateS. Although the intervals are large, the pattern shows that the Capital region has the most *ospC* genotypes, followed by SENY, and Outskirts. There were 46 ticks with detectable *ospC* in SENY and 48 in the Capital, but we only plotted

Table 1. 1. Table of primer and probe sequences used for *ospC* characterization. Brisson and Dykhuizen (2004) redesigned four probes for better binding to the DNA product compared to the same probes as Qiu et al. (2002). We also redesigned probe OC-I+1-1 for better binding to the DNA product compared to OC-I+1 from Brisson and Dykhuizen (2004).

A.

NAMES	LENGTH	PRIMER SEQUENCE (5' TO 3')	CITATION
		quantitative PCR	
Bb23Sf	24	CGAGTCTTAAAAGGGCGATTTAGT	Courtney et al. 2004
Bb23Sr	22	GCTTCAGCCTGGCCATAAATAG	Courtney et al. 2004
Bb23Sp	26	AGATGTGGTAGACCCGAAGCCGAGTG-TAMRA	Courtney et al. 2004
		nested PCR	
OC6+	24	AAAGAATACATTAAGTGCGATATT	Qiu et al. 2002
OC623-	24	TTAAGGTTTTTTTTGGACTTTCTGC	Qiu et al. 2002
OC602-	22	GGGCTTGTAAGCTCTTTAACTG	Qiu et al. 2002
OC6+F	24	Fluorescein-AAAGAATACATTAAGTGCGATATT	Qiu et al. 2002

Table 1. Continued.

B.

NAMES	LENGTH	PRIMER SEQUENCE (5' TO 3')	CITATION
		Reverse Line Blot probes	
OC-ALL	26	AGATTAGGCCCTTTAACAGACTCATC	Qiu et al. 2002
OC-A	22	ATTGTGATTATTTTCGGTATCC	Qiu et al. 2002
OC-B	16	CTCGTTGCGATTTGCT	Qiu et al. 2002
OC-D	36	ATGATTATTTAGAGTGCCTAAAGCATTGTTTTGATC	Qiu et al. 2002
OC-E	39	TGTGTTTTTACTCTGATTGGCCTCTAAACCATTATTGCC	Qiu et al. 2002
OC-F	27	CGCCTGAACGCCTAAACCATTTGCATC	Qiu et al. 2002
OC-H	30	GCCCCCATCGTCACCCAAAGTGCCATTTTG	Qiu et al. 2002
OC-K	33	CCCCGCTTCGCTACCTAAACCAGCATTTTGTTG	Qiu et al. 2002
OC-L	27	ATCGCTACCTAAAGTACCACCTGCTTC	Qiu et al. 2002
OC-M	31	ACCGGCATTTAAACCATTTTGGGCTATCAAA	Qiu et al. 2002
OC-N	30	GTTTTGCACATCATCTAAACCATTATTATT	Qiu et al. 2002
OC-O	22	TTGGTTAACTAAGCCATTTGCC	Qiu et al. 2002
OC-T	18	ATGGCCTGCATCGACACT	Qiu et al. 2002
OC-U	19	CTGCCCTTGCAAGTCCTGT	Qiu et al. 2002
OC-C182+2	20	TGCAAGTAAGGTCTCAACTT	Brisson & Dykhuizen 2004
OC-C - 7	25	TCCGTTGTTATCTGCCTCATTATCT	Brisson & Dykhuizen 2004
OC-G + 3	19	GGTGTTGTGATTCGCATCA	Brisson & Dykhuizen 2004
OC-J + 7 - 6	17	TTGACCCACTTCAGCAC	Brisson & Dykhuizen 2004
OC-I+1-1	23	GTTGAAATTAAATATGCTCCTGA	This study

Table 1. 2. Shannon-Weiner diversity and permutations of ospC genotypes with corresponding Fisher's Exact test. A) Sample sizes, Shannon-Wiener diversity values, and effective *ospC* richness given jackknife resampling of the array of *ospC* detected in each population; B) Jackknife tests of β_{pop} diversity values between each pairwise comparison are below the diagonal, and the effective *ospC* richness between each pairwise comparisons are above the diagonal; C) Jackknife tests of maximum β diversity values are given below the diagonal, and the [0, 1]-scaled $\Delta_{\beta pop}$ values, are given above the diagonal. These values are corrected for the different sample sizes in our population x year levels; D) Pairwise diversity values and outcomes of the permuational testing using natural Logarithms. The weighted Shannon-Wiener values between any two populations are below the diagonal, and the *p*-values that the randomly generated values are equal to or greater than the observed values (Fisher's Exact tests) are above the diagonal.

A.

Year	**2008**			**2011**	
N	182			61	
β_{year} Diversity	2.511			2.317	
Population	**Capital**	**Outskirts**	**SENY**	**Outskirts**	**SENY**
n	107	8	67	10	51
α Diversity	2.539	1.494	2.310	1.748	2.412
Effective *ospC* richness	12.670	4.456	10.074	5.743	11.153
Observed *ospC* richness	16	5	12	7	12

B.

	Capital	Outskirts08	SENY08	Outskirts11	SENY11
Capital		12.569	12.854	12.906	13.485
Outskirts08	1.067		9.848	8.094	11.192
SENY08	1.108	1.066		10.519	10.906
Outskirts11	1.090	1.578	1.123		11.108
SENY11	1.109	1.136	1.036	1.110	

C.

	Capital	Outskirts08	SENY08	Outskirts11	SENY11
Capital		0.280	0.201	0.326	0.211
Outskirts08	1.287		0.216	0.737	0.367
SENY08	1.947	1.404		0.342	0.070
Outskirts11	1.339	1.988	1.471		0.276
SENY11	1.876	1.487	1.982	1.562	

D.

	Capital	Outskirts08	SENY08	Outskirts11	SENY11
Capital		0.395	0.007	0.108	0.012
Outskirts08	0.065		0.738	0.096	0.226
SENY08	0.103	0.064		0.08	0.721
Outskirts11	0.086	0.456	0.116		0.488
SENY11	0.103	0.128	0.035	0.105	

Table 1. 3. Jacknife tests on proportion of *ospC* types detected in each population. For example, the Capital region had 48 ticks that were tested positive for at least one *ospC* type. There were 6 ticks out of 48 with detectable *ospC* A in this region. Therefore, the proportion of positive to total ticks tested was 0.125 for *ospC* A. Each *ospC* type was jackknifed separately with 1000 permutations. Yellow highlights indicate the proportions were significantly greater than expected (above the upper 95% confidence interval). Green highlights indicate the proportions were significantly lower than expected (below the lower 95% confidence interval). Blue highlight indicates the proportion was at the lower 95% confidence interval.

Year	2008	2008	2008	2011	2011
Region	Capital	Outskirt	SENY	Outskirt	SENY
A	0.125	-	0.192	-	0.250
B	0.146	-	0.231	-	0.400
D	0.104	-	0.231	0.200	0.250
E	0.146	-	0.115	0.800	0.200
F	0.146	-	0.192	0.200	0.250
G	0.146	0.200	0.346	-	0.100
H	0.125	0.200	0.231	-	0.200
I	0.042	0.200	0.077	-	0.100
J	0.021	-	-	-	-
K	0.292	0.600	0.577	0.200	0.300
L	0.146	-	-	-	-
M	0.375	-	-	-	-
N	0.271	0.400	0.231	0.200	0.200
O	0.042	-	-	-	-
T	0.021	-	0.077	0.200	0.150
U	0.083	-	0.077	0.200	0.150

Table 1. 4. Top three *ospC* types detected from ticks in other studies.

A. Brisson and Dykhuizen (2004) (Table 4)[a]			
White-footed Mouse	Eastern Chipmunk	Gray Squirrel	Short-tailed Shrew
K 68.0	K 65.3	K 52.0	K 84.0
D 57.3	D 64.0	A 38.0	D 48.0
A 42.7	T 44.0	E 20.0	E 40.0

B. Hanincova et al. (2006) (Table 4)[ab]					
White-footed Mouse	Eastern Chipmunk	Gray Squirrel	Pine Vole	Raccoon	Opossum
K 55.0	D 29.0	N 62.0	A 52.0	D 31.0	N 11.0
A 52.0	U 20.0	U 22.0	K 14.0	I 16.0	B 10.0
B 40.0	K 7.0		D, I 14.0	A 10.0	A, D 2.0

C. MacQueen et al. (2012) (Figure 2)[c]	
Year 2003	*H* > A
Year 2005	*H* > A > B
Year 2007	A > K > U

D. Ogden et al. (2008)[d] (Table 4)	Mathers et al. (2011)[e] (Figure 4)	Ogden et al. (2011)[f] (Figure 4)
A 23.7	A 26.7	A > K > N
K 18.4	B 26.7	
B 13.2	G 13.3	

E. This study's *ospC* frequencies[g]									
Capital08		Outskirt08		SENY08		Outskirt11		SENY11	
M	17.3	K	37.5	K	22.4	E	40.0	B	15.7
K	13.5	N	25.0	G	13.4	D	10.0	K	11.8
N	12.5	G	12.5	B	9.0	F	10.0	A	9.8

[a]Percentages are number of ticks tested positive for each *ospC* type, out of the total numbers tested per species. Values add up to greater than 100% because there can be more than one *ospC* type per tick.
[b]The *ospC* types correspond to genotypes based on the 16s rDNA amplification.
[c]Tick samples collected primarily from questing tick collections and white-footed mouse.
[d]Percent tick positive. Tick samples collected from migrating birds.
[e]Percent tick positive. Tick samples collected from birds, primarily from Common Yellowthroats.
[f]Based on frequency graph. Tick samples were collected from human and companion pets, although the study focused on birds as transporter of these infected ticks.
[g]The rank-order for each region is based on the greatest proportion of each *ospC* type in that region. *ospC* K is a common genotype being one of the top three genotypes present in four of five regions. Most genotypes in the top three are HIS types (A, B, K, and N).

Citation

Barbour, A. G. and D. Fish. 1993. The biological and social phenomenon of Lyme disease. Science **260**:1610-1616.

Brinkerhoff, R. J., C. M. Folsom-O'Keefe, H. M. Streby, S. J. Bent, K. Tsao, and M. A. Diuk-Wasser. 2011. Regional variation in immature *Ixodes scapularis* parasitism on North American songbirds: Implications for transmission of the Lyme pathogen, *Borrelia burgdorferi*. Journal of Medical Entomology **48**:422-428.

Brisson, D. and D. E. Dykhuizen. 2004. *ospC* diversity in *Borrelia burgdorferi*: Different hosts are different niches. Genetics **168**:713-722.

Brisson, D. and D. E. Dykhuizen. 2006. A modest model explains the distribution and abundance of *Borrelia burgdorferi* strains. American Journal of Tropical Medicine and Hygiene **74**:615-622.

Brownstein, J. S., T. R. Holford, and D. Fish. 2003. A climate-based model predicts the spatial distribution of the Lyme disease vector *Ixodes scapularis* in the United States. Environmental Health Perspectives **111**:1152-1157.

Burgdorfer, W., A. G. Barbour, S. F. Hayes, J. L. Benach, E. Grunwaldt, and J. P. Davis. 1982. Lyme disease--A tick-borne spirochetosis? Science **216**:1317-1319.

CDC. 2011. Lyme disease: division of vector-borne infectious diseases. http://www.cdc.gov/lyme/.

Colwell, R. K. 2011. EstimateS, ver 8.2, Statistical estimation of species richness and shared species from samples.

Courtney, J. W., L. M. Kostelnik, N. S. Zeidner, and R. F. Massung. 2004. Multiplex real-time PCR for detection of *Anaplasma phagocytophilum* and *Borrelia burgdorferi*. Journal of Clinical Microbiology **42**:3164-3168.

Dennis, D. T., T. S. Nekomoto, J. C. Victor, W. S. Paul, and J. Piesman. 1998. Reported distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the United States. Journal of Medical Entomology **35**:629-638.

Dykhuizen, D. E., D. Brisson, S. Sandigursky, G. P. Wormser, J. Nowakowski, R. B. Nadelman, and I. Schwartz. 2008. The propensity of different *Borrelia burgdorferi sensu stricto* genotypes to cause disseminated infections in humans. Am J Trop Med Hyg **78**:806-810.

Gatewood, A. G., K. A. Liebman, G. Vourc'h, J. Bunikis, S. A. Hamer, R. Cortinas, F. Melton, P. Cislo, U. Kitron, J. Tsao, A. G. Barbour, D. Fish, and M. A. Diuk-Wasser. 2009. Climate and tick seasonality are predictors of *Borrelia burgdorferi* genotype distribution. Applied and Environmental Microbiology **75**:2476-2483.

Glavanakov, S., D. J. White, T. Caraco, A. Lapenis, G. R. Robinson, B. K. Szymanski, and W. A. Maniatty. 2001. Lyme disease in New York State: spatial pattern at a regional scale. Am J Trop Med Hyg **65**:538-545.

Hanincova, K., K. Kurtenbach, M. Diuk-Wasser, B. Brei, and D. Fish. 2006. Epidemic spread of Lyme borreliosis, northeastern United States. Emerging Infectious Diseases **12**:604-611.

Hoen, A. G., G. Margos, S. J. Bent, M. A. Diuk-Wasser, A. Barbour, K. Kurtenbach, and D. Fish. 2009. Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. Proceedings of the National Academy of Sciences:-.

Johnson, R. C., G. P. Schmid, F. W. Hyde, A. G. Steigerwalt, and D. J. Brenner. 1984. *Borrelia burgdorferi* sp. nov.: Etiologic agent of Lyme disease. International Journal of Systematic Bacteriology **34**:496-497.

Jones, C. G., R. S. Ostfeld, M. P. Richard, E. M. Schauber, and J. O. Wolff. 1998. Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk. Science **279**:1023-1026.

Keesing, F., J. Brunner, S. Duerr, M. Killilea, K. LoGiudice, K. Schmidt, H. Vuong, and R. S. Ostfeld. 2009. Hosts as ecological traps for the vector of Lyme disease. Proceedings of the Royal Society B-Biological Sciences **276**:3911-3919.

Leighton, P. A., J. K. Koffi, Y. Pelcat, L. R. Lindsay, and N. H. Ogden. 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.

LoGiudice, K., R. S. Ostfeld, K. A. Schmidt, and F. Keesing. 2003. The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences of the United States of America **100**:567-571.

MacQueen, D. D., C. Lubelczyk, S. P. Elias, B. K. Cahill, A. J. Mathers, E. H. Lacombe, P. W. Rand, and R. P. Smith Jr. 2012. Genotypic diversity of an emergent population of *Borrelia burgdorferi* at a costal Maine island recently colonized by *Ixodes scapularis*. Vector-Borne and Zoonotic Diseases **12**:456-461.

Magnarelli, L., J. Anderson, and W. Chappell. 1984. Antibodies to spirochetes in white-tailed deer and prevalence of infected ticks from foci of Lyme disease in Connecticut. Journal of Wildlife Diseases **20**:21-26.

Mathers, A., R. P. Smith, B. Cahill, C. Lubelczyk, S. P. Elias, E. Lacombe, S. R. Morris, C. P. Vary, C. E. Parent, and P. W. Rand. 2011. Strain diversity of *Borrelia burgdorferi* in ticks dispersed in North America by migratory birds. Journal of Vector Ecology **36**:24-29.

NYSDOH. 2012. New York State Department of Health: Communicable Diseases http://www.health.ny.gov/statistics/diseases/communicable/.

Ogden, N. H., L. R. Lindsay, K. Hanincova, I. K. Barker, M. Bigras-Poulin, D. F. Charron, A. Heagy, C. M. Francis, C. J. O'Callaghan, I. Schwartz, and R. A. Thompson. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. Appl. Environ. Microbiol. **74**:1780-1790.

Ogden, N. H., G. Margos, D. M. Aanensen, M. A. Drebot, E. J. Feil, K. Hanincová, I. Schwartz, S. Tyler, and L. R. Lindsay. 2011. Investigation of genotypes of *Borrelia burgdorferi* in *Ixodes scapularis* ticks collected during surveillance in Canada. Applied and Environmental Microbiology **77**:3244-3254.

Orloski, K. A., E. B. Hayes, G. L. Campbell, and D. T. Dennis. 2000. Surveillance for Lyme dsease: United States, 1992-1998. Morbidity and Mortality Weekly Report **49**:1-11.

Ostfeld, R. S. 1997. The ecology of Lyme-disease risk. American Scientist **85**:338-346.

Ostfeld, R. S., K. R. Hazler, and O. M. Cepeda. 1996. Temporal and spatial dynamics of *Ixodes scapularis* (Acari: Ixodidae) in a rural landscape. Journal of Medical Entomology **33**:90-95.

Peakall, R. and P. E. Smouse. 2006. GenAlEx 6: genetic analysis is Excel. Populaiton genetic software for teaching and research. Molecular Ecology Notes **6**:288-295.

Qiu, W. G., E. M. Bosler, J. R. Campbell, G. D. Ugine, I. N. Wang, B. J. Luft, and D. E. Dykhuizen. 1997. A population genetic study of *Borrelia burgdorferi sensu stricto* from eastern Long Island, New York, suggested frequency-dependent selection, gene flow and host adaptation. Hereditas **127**:203-216.

Qiu, W. G., D. E. Dykhuizen, M. S. Acosta, and B. J. Luft. 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**:833-849.

R. 2008. R: A language and environment for statistical computing. Version 2.15. R Foundation for Statistical Computing, Vienna, Austria.

Rand, P. W., E. H. Lacombe, R. P. Smith, and J. Ficker. 1998. Participation of birds (Aves) in the emergence of Lyme disease in southern Maine. Journal of Medical Entomology **35**:270-276.

Salkeld, D. J. and R. S. Lane. 2010. Community ecology and disease risk: lizards, squirrels, and the Lyme disease spirochete in California, USA. Ecology **91**:293-298.

Seinost, G., D. E. Dykhuizen, R. J. Dattwyler, W. T. Golde, J. J. Dunn, I. N. Wang, G. P. Wormser, M. E. Schriefer, and B. J. Luft. 1999. Four clones of *Borrelia burgdorferi sensu stricto* cause invasive infection in humans. Infection and Immunity **67**:3518-3524.

Steere, A., T. Broderick, and S. Malawista. 1978. Erythema chronicum migrans and Lyme arthritis: epidemiologic evidence for a tick vector. American Journal of Epidemiology **108**:312-321.

Steere, A. C., S. E. Malawista, D. R. Snydman, R. E. Shope, W. A. Andiman, M. R. Ross, and F. M. Steele. 1977. Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. Arthritis and Rheumatism **20**:7-17.

Strle, K., K. L. Jones, E. E. Drouin, X. Li, and A. C. Steere. 2011. *Borrelia burgdorferi* RST1 (*OspC* Type A) genotype is associated with greater inflammation and more severe Lyme disease. American Journal of Pathology **178**:2726-2739.

Travinsky, B., J. Bunikis, and A. G. Barbour. 2010. Geographic differences in genetic locus linkages for *Borrelia burgdorferi*. Emerging Infectious Diseases **16**:1147-1150.

Wang, I., D. E. Dykhuizen, W. Qui, J. J. Dunn, E. M. Bosler, and B. J. Luft. 1999. Genetic diversity of *ospC* in a local population of *Borrelia burgdorferi sensu stricto*. Genetics **151**:15-30.

White, D. J., H. Chang, and J. L. Benach. 1991. The geographic spread and temporal increase of the Lyme disease epidemic. JAMA: The Journal of the American Medical Association **266**:1230-1236.

Chapter 2: Influences of host community composition on *Borrelia burgdorferi* genotypic diversity

Abstract

Our study examined the relationship between vertebrate host and *Borrelia burgdorferi* diversities and composition, using the outer surface protein C (*ospC*) for our bacterial marker. Previous research suggests that hosts act as ecological niches to the array of *ospC* genotypes, hence higher frequencies of particular genotypes may be detected in different host species. Given the *a priori* concept that competent vertebrate host reservoirs are good hosts for human invasive strains (HIS), we examined the relationship between host community metrics (host diversity and the relative abundances of white-footed mouse, eastern chipmunk, and short-tailed shrew) and *ospC* diversity, as well as prevalence of HIS, within infected ticks. We sampled the host community using live traps, camera traps, and avian surveys in Dutchess County, NY, in summer 2006 and 2009, with few overlapping sites between the two years. We collected questing nymphal ticks in 2007 and 2010, because these ticks represent the larval population that fed on the host community in the previous year. We tested the ticks for *B. burgdorferi* using PCR, and tested *B. burgorferi* positive ticks for their *ospC* genotypes, using the reverse line blot method. *ospC* diversity varied across the sites within each year, with some significantly large turnover among sites, within each year. We found that host diversity is not a significant predictor of *ospC* diversity in either year, but in 2009, the combination of the relative abundances of mice, chipmunks, and shrews provided a useful predictive model for *ospC* diversity. On the other hand, host diversity and mouse relative abundance were the best predictors of HIS prevalence in 2006, but chipmunk relative abundance was the best predictor of HIS prevalence in 2009. The results were

opposite to our expectations, with host diversity having a positive association with HIS prevalence, and mouse and chipmunk relative abundances having a negative association with HIS prevalence. The contrasting best models between years, and contradictory results from our predictions, underscore the need to further investigate how disease risk varies across time and space, especially at the genotypic level because of human invasive strains. In order to make good predictions, we require a combination of understanding the role of each host species in contributing *ospC* genotypes, their reservoir competency, and the interplay between infection probabilities by infected ticks, and transmission efficiencies of *ospC* types from hosts to ticks.

Introduction

Without genetic variation, organisms cannot adapt to changing environments, potentially leading to their demise. This is just as true for miniscule viruses as it is for the largest animals and plants on the planet. In a rapidly changing environment, genetic variation will allow some species to adapt and persevere even while others do not.

In the world of zoonotic disease ecology, host individuals represent immunological environments, whose immune systems the pathogens must evade in order to reproduce. Although we typically define a pathogen as a single species, we often find that named species can be subdivided into distinct genotypes or strains, whose genetically-coded immune evasion capabilities are both variable and adaptively consequential. These microorganisms have very high replication rates, resulting in potentially high rates of mutational substitution. Additionally, strong selective pressure from the host immune systems can drive the pathogen to adapt more quickly to evade detection (Liang et al. 2004), resulting in adaptive genetic changes.

The interplay between multi-host and multi-genotype systems can lead to some hosts faring better than others in reducing disease onset. The evolutionary pressure to evade the immune system may result in strains with different pathogenicities, with pathogenicities unlikely to be equal across host species. For example, there are two primary lineages of West Nile virus. Lineage I is identified as the more virulent strain in humans, compared with lineage II, although there seems to be increasing pathogenicity associated with strains in lineage II (Murray et al. 2010). Similar patterns have also been shown for Lyme disease (Kurtenbach et al. 1998, Girard et al. 2011), anaplasmosis (Barbet et al. 2006), canine distemper (Kapil et al. 2008), foot and mouth disease (Sangula et al. 2012), chytridiomycosis (Velo-Anton et al. 2012), and toxoplasmosis (Wendte et al. 2011), to name just a few.

In this study, we focus on a widespread zoonotic bacterium that infects many wildlife host species throughout the temperate zone, worldwide (Kurtenbach et al. 2002). In the northeastern United States, *Borrelia burgdorferi s.s.*, the bacterium that causes Lyme disease, is transmitted among wildlife hosts and to humans by the blacklegged tick vector (*Ixodes scapularis*). *Borrelia burgdorferi s.s.* is part of the larger *B. burgdorferi sensu lato* complex, within which 18 genospecies are currently described (Rudenko et al. 2011, Stanek and Reiter 2011). Most of these genospecies are distributed in Eurasia, while a smaller subset occurs in the US. Of the three common genospecies that infect humans (*B. burgdorferi* s.s., *B. afzelii*, and *B. garinii*), only *B. burgdorferi s.s.* causes Lyme disease in the US and Europe (Rudenko et al. 2011). However, within *B. burgdorferi s.s.* (henceforth *B. burgdorferi*), the bacterium can also be differentiated into multiple genotypes, based on genetic variation at the outer surface protein C (*ospC*) locus, a highly variable part of the genome (Wang et al. 1999, Qiu et al. 2002). There are 17 alleles described in the Northeast (Wang et al. 1999, Qiu et al.

2002), which can be used to assert genetic variation within and among host populations (Brisson and Dykhuizen 2004). Variation at the *ospC* locus is hypothesized to be traceable to strong selective pressures, mounted by multiple host immune systems attacking the pathogen (Wang et al. 1999, Qiu et al. 2002). The constant pressures from the hosts' immune systems attacking and removing the pathogen, and the pathogen adapting to evade these attacks, has probably resulted in multiple *B. burgdorferi* genotypes that are able to evade some host species better than others. However, a multiple-niche polymorphism hypothesis has been shown to be a sufficient explanation for the genotypic diversity of *B. burgorferi* (Brisson and Dykhuizen 2004)

Heterogeneity among individuals, and among species, may help drive the multiple niche polymorphism for *B. burgdorferi*, leading to the suggestion that different hosts act as ecological niches for different *ospC* genotypes (Brisson and Dykhuizen 2004). This is indicated by greater relative frequencies of particular *ospC* types in some vertebrate hosts than in other hosts, although genotypes are evidently not confined to any particular host species (Brisson and Dykhuizen 2004, Hanincova et al. 2006). Of importance, with respect to human health risk, are the *ospC* genotypes known to be relatively invasive in people diagnosed with Lyme disease, which include types A, B, I, K, and N. Invasiveness was measured as higher infection proportions in people, compared with ticks (Dykhuizen et al. 2008). The high prevalence of these human invasive strains/genotypes (HIS) in humans suggests that these genotypes are more pathogenic in humans than are non-HIS genotypes.

To understand Lyme disease risk, we must understand how differences in the vertebrate community composition can affect the genotypic diversity of *B. burgdorferi*. We first address the role of biodiversity and host community metrics on the genotypic diversity

of *B. burgdorferi*. If the wildlife host-genotype associations are strong, we would expect that greater host community diversity would lead to greater *ospC* diversity in the pathogen. Second, we test which host community metrics correlate well with the probability of HIS infection, given an infected tick. Based on past research on host-genotype associations (Brisson and Dykhuizen 2004), we predict that the HIS infection probability should increase with the proportions of mice, chipmunks, and shrews in the community, but that increasing host community diversity should be negatively correlated with HIS probability. By taking a community ecology approach to understanding the evolutionary influences on genotypic diversity of *B. burgdorferi*, we hope to improve our understanding of how local biodiversity influences disease risk.

Methods

Field Collections – We sampled the mammalian and avian communities throughout Dutchess County NY in 2006 and 2009, at 30 and 19 sites, respectively, with 7 of the sites sampled in both years. We used live traps and camera traps to survey the mammal community, and visual and auditory surveys for bird hosts. Trapping was conducted from 30 May – 19 September 2006, and 2 June – 2 October 2009 (see Brunner et al. *submitted* for details). In 2006, we sampled each site rotationally for two-consecutive nights (= 1 trapping session) every other week, whereas in 2009, this was done every week. We used Sherman live traps to capture mice and chipmunks, and used Tomahawk live traps to capture gray squirrels. We used an 8 x 8 grid system for live trapping, placing one Sherman trap at 15 m intervals, and Tomahawks at 30 m intervals, for a maximum of 16 Tomahawks and 64 Sherman traps on a full grid. Grid dimensions were adjusted to maximize trapping coverage in smaller forest fragments, but maintained the same distance separation. Larger fragments in year 2009 also had a buffer strip of up to 30 meters with 3 to 30 traps, depending on trap type

(Brunner et al. *submitted*). Animals captured in the buffer traps were not included in the analyses. Traps were set between 15:30 – 17:30 in the afternoon, and were checked the following morning from 08:30 – 12:00.

Each mammal was identified to species, ear tagged with a unique code, sexed, weighed, and evaluated for reproductive status. Although the trapping occurred throughout the summer, our small mammal diversity measures are based on data from August through early October, which coincides with peak larval tick abundances. The larvae that feed on these animals in the current year become next year's questing nymphs. We calculated the minimum number alive (MNA) for the three most common host species captured at each site: white-footed mouse (*Peromyscus leucopus*), eastern chipmunk (*Tamius striatus*), and short-tailed shrew (*Blarina brevicauda*). MNA is based on mark-recapture data, where individuals are marked upon initial capture and recorded as present or absent in subsequent trapping sessions. We had four trapping sessions in 2006 and eight trapping sessions in 2009. We averaged the MNA values across these trapping sessions within a year and used the averaged value to estimate population densities, based on grid size.

Our estimates of avian diversity for year 2006 are based on point-count surveys conducted for 24 sites in year 2005 and six sites in year 2004. We assume that diversity remained relatively consistent over these years. In year 2009, avian surveys were conducted in 17 of 19 sites. Each site was visited thrice each in 2004 and 2005, and twice in 2009. All visits were in June, during the peak period of nymphal tick activity (Ostfeld 1997). Birds within a 100 m radius of the observer were identified by sight and sound. Most counts were conducted at the center of the grids, but some counts were conducted off center, due to the irregular shapes of the grid. Counts were conducted between 05:00 to 10:00 AM, to

maximize avian detection during early morning activity. We only included the American Robin (*Turdus migratorius*), Veery (*Catharus fuscescens*), Wood Thrush (*Hylocichla mustelina*), and Ovenbird (*Seiurus aurocapilla*) in calculating Shannon-Wiener diversity index measures, as these are the four primary ground dwelling birds that can be important hosts in our community (Giardina et al. 2000). In these point-counts, male birds are more likely to be detected than female birds, both visually and by sound, because during the breeding season, male birds are more vocal and active. To correct for observing approximately half of the population, we doubled the counts of these birds, using the maximum number of birds heard across the surveys in each year, and calculated densities based on the 100 m radius plots.

To obtain quasi-quantitative estimates of densities for mammals detected only by camera traps, we employed a procedure used previously (LoGiudice et al. 2003, LoGiudice et al. 2008). We placed motion-detecting wildlife cameras (DeerCam and GotchaCritter) at the sites, starting in early October 2006 and in mid October 2009. Sites were baited with carnivore scent lures during week one, and raw chicken and corn cob in a non-rewarding fashion, during week two in 2006. The order of lures was switched in 2009. The variety of lures maximized the number of species detected in our sites. The number of identifiable individuals in each picture and the number of pictures provide an index of 'activity level' for those animals at the site. The number of trap nights was corrected for malfunctions to standardize the data. Based on the 'activity level' and live trapping data, we assigned the animals into one of three 'activity density' values. These values are based on published density estimates of these animal species in similar habitats (see LoGiudice et al. 2008 for more details). Briefly, the site with the highest quartile of 'activity levels' for each species

was assigned the 'most common' density values, while lower quartile values were placed under the 'common' density values. If the animal was absent or rare, the density was either zero or some low value, depending on the species. The quasi-Shannon-Wiener diversity values were based on the most commonly detected species of the host community. The Shannon-Wiener calculations incorporated values based on 'activity density' estimates, averaged weekly minimum number of live densities of mice, chipmunks, and short-tailed shrews, and density estimates of avian hosts (LoGiudice et al. 2003, 2008). Hereafter, all densities will be called 'activity density' for simplicity.

Questing nymphs were collected by dragging a 1 m^2 white, corduroy cloth on the forest floor during the nymphal peak period of June/July 2007 and 2010 (Falco and Fish 1992). We randomly dragged four 30 m transects in our trapping grid to obtain a density estimate of the tick population. These questing nymphs represent the previous summer's larvae that fed on the host community in 2006 and 2009, respectively. To estimate *B. burgdorferi* infection prevalence of nymphs with the largest possible sample of ticks, we conducted additional tick drags on many of the sites, following the second density drags. These supplemental drags were not used for calculations of tick density.

Lab Analyses – Laboratory analyses were conducted as for Chapter 1. Briefly, we extracted DNA using Qiagen DNEasy kits. The samples were tested for the presence of *B. burgdorferi* by amplifying the *ospC* gene with a semi-nested PCR procedure. We used primers OC6F/OC623R as the outer primer and then used OC6+F/OC602R for the semi-nested PCR for year 2006 ticks (Brisson and Dykhuizen 2004, Vuong - Chapter 1). For year 2009, new primers were developed for better amplification. The outer primers were OC-368F/OC693R and the nested primers were OC4+F/OC643 (Deverey et al. *in prep*). We used

the following protocol with the new primers: In a 20 µl reaction for the outer PCR, our final concentrations included 1x buffer, 2.5 mM MgCl2, 200 µM dNTP each, 0.5 µM of each primer, 0.5 units of Amplitaq, and 1 µl of DNA. The thermocycler conditions were as follows: 95°C for 1 min, 39 cycles of 95°C for 40 sec, 54°C for 35 sec, and 72°C for 1 min, and a 10 min elongation at 72°C. Then, 1 µl of PCR product was added to the second round of PCR amplification along with the internal primers. For the nested PCR, all concentrations and cycle numbers were the same, but we used a 53°C annealing temperature and ran 25 µl reactions. We tested for *ospC* genotypes with a reverse line blot (RLB) analysis that utilizes a DNA-DNA hybridization technique with specific probes for each *ospC* genotype (Qiu et al. 2002, Brisson and Dykhuizen 2004). Single-stranded PCR products that bind to these probes undergo a chemiluminiscent reaction and become dark spots on the developed x-ray film. Genotype *ospC* C is a hybrid of genotype E and genotype I, making double and triple co-infections with these genotypes difficult to separate. Therefore, *ospC* C was removed from the dataset. Genotype *ospC* J was found once in one year and absent the other year, so it was also removed from the analyses, resulting in a total of 15 *ospC* genotypes to compare.

Statistical Analysis – Due to differences in animal trapping and variation with infection prevalence and *ospC* frequencies (Wang et al. 1999), we analyzed the data separately for years 2006 and 2009. We converted all multiply infected ticks into single-infected samples (e.g., a tick infected with *ospC* A and B is now two samples, one with *ospC* A infection and another with *ospC* B infection). This allowed examination of relative frequency differences of the 15 genotypic *ospC* types, independent of their co-infections. This increased our sample size from 171 infected ticks to 349 single infected samples for year 2006, and from 103 infected ticks to 248 single infected samples for year 2009. We

tested Shannon diversity for *ospC* at each site with 1000 permutations, using GenAlEx (version 6.4) (Peakall and Smouse 2006). This analysis provided Fisher's Exact tests between all pairwise site comparisons, as well as estimates of alpha and beta diversity within populations and beta diversity measures among the populations. We tested whether host community metric covariates (Shannon diversity, and the relative abundances of mice, chipmunks, and short-tailed shrews) were good predictors of *ospC* diversity, using a generalized linear model (GLM) regression analysis. Models were tested for significance against an intercept-only model, using analysis of deviance.

Because human infections are attributed to a subset of *ospC* genotypes, labeled human invasive strains (HIS) (Dykhuizen et al. 2008), our objective was to determine how the host community influences the frequency split between HIS and non-HIS genotypes. Mice, chipmunks, and short-tailed shrews are competent reservoirs of *B. burgdorferi* (Brisson et al. 2008) and these species are known to support several or all of the HIS types (Brisson and Dykhuizen 2004). We expected that communities with greater proportions of mice, chipmunks, and short-tailed shrews to have a greater tick infection prevalence with an HIS type. We coded HIS/non-HIS as a binary value for logistic regression and tested whether any of the covariates were good predictors of HIS infection probability. Significance for the logistic regression models was assessed by testing against an intercept-only model, using analysis of deviance. All analyses were conducted in R 2.15 (R 2008), unless otherwise noted.

Results

In 2006, 171 of 250 (68.4%) tick samples hybridized efficiently with specific probes in the reverse line blots, whereas all 103 samples that amplified in year 2009 hybridized with the probes efficiently. Thus, *ospC* data for year 2006 are based on samples in which

hybridization was successful. The *ospC* Shannon α diversity ranged from 0.760 – 1.896 in 2006 and 0.673 – 2.476 in 2009. However, the overall α diversity was greater for 2006 (α_{2006} = 2.558) compared to 2009 (α_{2009} = 2.387), which may have to do with greater turnover over *ospC* types in 2006 relative to 2009 ($\Delta_{\beta 2006}$ = 0.443; $\Delta_{\beta 2009}$ = 0.327) (Table 2. 1). Because we used different primer sets between the 2006 and 2009 data, we tested whether primers had an effect on the proportions of *ospC* types detected each year. We found that the proportions of each *ospC* genotypes in each year were not significantly correlated with one another (r = 0.49, df = 13, p = 0.06), suggesting that there may have been a small potential bias in primer binding to *B. burgdorferi*, or that the PCR products bound differentially in the reverse line blot. However, the mean number of genotypes per tick (2006: 2.05 ± 1.50 sd; 2009: 2.41 ± 1.78 sd) was not significantly different from each other across the years (t = 1.70, df = 187, p = 0.09). On balance, we concluded that cautious comparisons between the years may be worthwhile.

When we examined the turnover of *ospC* types between any two sites, we found several significant differences in the pairwise comparisons using Fisher's Exact tests for 2006 (Table 2. 2). There were multiple sites in which there was sufficient turnover of *ospC* types ($\Delta_{\beta site:site}$) to lead to significant pairwise comparisons. For 2009, only one site had large turnover of *ospC* types ($\Delta_{\beta site:site}$) from a few of the other sites, leading to significant pairwise comparisons (Table 2. 3).

For both years, linear regressions indicated that host community diversity was not a good predictor of *ospC* diversity; for 2006 (F = 0.29, df = 1, 26, p =0.60, Table 2. 4a); for 2009 (F = 1.54, df = 1, 15, p = 0.23, Table 2. 4b). Shannon α diversity estimates of the host community ranged from 0.88 to 2.09 in 2006 and 0.58 to 1.52 in 2009. No other host

community metrics were significant predictors of *ospC* for year 2006. In 2009, the relative abundances of chipmunks, and shrews, were significant single covariates. However, the best model included the relative abundances of all three common small mammal host species. Mouse and chipmunk relative abundances had a positive association with *ospC* diversity, while shrew relative abundance had a negative association.

We found significant negative correlations between host diversity and mouse relative abundance with HIS infection prevalence for 2006, using logistic regression (Table 2. 5a, Figure 2. 1). Although more complex models supporting a combination of these covariates were significant predictors, the best model was the single covariate model of host diversity. In 2009, we saw contrasting results; host diversity was not a significant predictor of HIS infection prevalence, but the combination of relative abundances of mice, chipmunks, and short-tailed shrews were much better predictors (Table 2. 5b). In this case, the relative abundances of mice and chipmunks had a negative relationship with HIS infection prevalence, while the relative abundances of short-tailed shrews had a positive relationship. However, model comparisons between chipmunk alone, compared with the combination of the three host species, were not significantly different from one another, so the only predictor necessary for HIS prevalence in 2009 is the relative abundance of chipmunks.

Discussion

The focus of our study was to determine whether *B. burgdorferi* genotypic diversity, and HIS (human invasive strain) prevalence, could be predicted by vertebrate host diversity and composition. This was based on the *a priori* assumption that hosts act as ecological niches to *ospC* types (Brisson and Dykhuizen 2004), so that we would expect a positive relationship between vertebrate hosts and *ospC* diversities. Also, given that HIS types were detected in relatively high frequencies in mice, chipmunks, and shrews (Brisson and

Dykhuizen 2004), we expected a positive relationship between the hosts' relative abundances and HIS prevalence.

We found that neither host diversity nor the relative abundance of mice were good predictors of *ospC* diversity using linear regressions. This was surprising, given that there are detectable differences in the relative frequencies of genotypes detected within particular host species (Brisson and Dykhuizen 2004, Hanincova et al. 2006, Vuong – Chapter 3). The effects of relative abundances of chipmunks and shrews on *ospC* diversity are more equivocal, because these covariates were significant predictors in 2009, but not in 2006. Chipmunks had a positive relationship with *ospC* diversity, whereas there was a negative relationship between shrews and *ospC* diversity. Within the host-genotype associations tested by Brisson and Dykhuizen (2004), there was a lower detectability of *ospC* genotypes in shrews than in chipmunks. If this relationship also holds true at these forest sites, then higher relative abundances of shrews might be associated with lower *ospC* diversity, as compared to chipmunk relative abundance. However, it should be noted that these significant relationships in the single predictor models were not significant when a single data point with the largest residual (potential outlier) for these two species was removed from the data. This may suggest that we happened to sample a few sites with distinctively high relative abundances for these two species, relative to the other sites in our study. Then again, *ospC* frequencies change rapidly from year to year, even in the same area (Qiu et al. 2002, MacQueen et al. 2012), so given little overlap in the sites between 2006 and 2009, it would not be surprising if the forces acting on each site were idiosyncratic, leading to inconsistent outcomes.

There are several reasons why we may not have seen a strong relationship between *ospC* diversity and host community metrics. First, these genotypes are not mutually exclusive

for particular hosts. Although there are biases in the relative frequencies of genotypes detected within each host species; a 'rare' genotype in one vertebrate host species may not be so 'rare' in another vertebrate host species. For example, *ospC* T and U detections are relatively uncommon in mice, but more common in chipmunks (Brisson and Dykhuizen 2004). Consequently, across the whole host community, no genotype becomes truly rare in the tick population. In order to know how each species contributes to the *ospC* profiles detected in the tick populations, this would require an excessive amount of trapping of the majority of individuals in the communities, and testing for the *ospC* profiles from each individual to obtain a baseline comparison between vertebrate host communities and tick populations.

Second, influxes of pathogens into the host community via vagile host species can potentially have large effects on the genotype dynamics. Birds are able to support a wide array of *ospC* genotypes, especially high proportions of HIS to non-HIS types (Vuong – chapter 3). Avian migration into the northeast in the spring could help re-introduce genotypes that were lost via stochastic processes the year before, or lost from population crashes of host or tick populations (Ogden et al. 2008, Ogden et al. 2011).

Third, our host community diversity measure does not capture all animal species at our sites, because obtaining reliable density estimates of rare species, and non-distinct host individuals of species, is inefficient with current camera trapping technologies. For example, we included red foxes in our activity densities for year 2006 because we detected red foxes in six of our sites, and their activity levels differed among the sites. But, we could not include them in year 2009, because we only detected one fox in all of the photos taken. A recent paper by Levi et al. (2012) proposed that fox densities can be a good predictor of spatial

Lyme disease. Higher incidence of Lyme disease occurred in areas where fox densities were reduced, most likely from interference competition between foxes and coyotes. Both foxes and coyotes attack small mammals, but foxes are less generalized in their kills, they are surplus killers, and their densities are much higher than those of coyotes. Hence, small mammals experience predator release, in the absence of foxes. Since these small mammals tend to be competent reservoirs of *B. burgdorferi*, the increased density and relative abundances of these small mammals in the community can increase infection rates. Coyotes were not included in our activity density estimates, because they were detected at only three sites (by camera) in 2006 and one site in 2009.

We found different significant predictors of HIS prevalence for each year. In 2006, the relative abundance of mice and host Shannon diversity were significant predictors, but in 2009, the significant predictor was the relative abundance of chipmunks. The relative abundance of shrews was not a significant predictor of HIS prevalence in either year, but they tend to show a positive relationship between abundance and prevalence (Figure 2. 1). We had expected a negative relationship between host diversity and HIS prevalence, and a positive relationship between the relative abundances of mice, chipmunks, and shrews with HIS prevalence. Our results were the opposite of what we expected. This probably implies that HIS are being supported by other host species in the community and that the relationship between these other vertebrate host species and *ospC* genotypes is at least as strong as those of the three primary reservoir species. Indeed, we found that some or all five of the HIS types can be detected in a suite of host species, from American robins and veeries to masked shrews and red squirrels (Vuong – Chapter 3). Hanincova et al. (2006) detected relatively high *ospC* B in opossums, adding to the species repertoire within which HIS can be detected.

Our host-genotype study (Vuong – Chapter 3) also shows that robins and short-tailed shrews had the greatest proportion of HIS types detected, more so than the mice and chipmunks. This may be why there is a positive trend in the relative abundances of short-tailed shrews with HIS prevalence for both years.

There are several main themes that emerge from this study. First, host composition and diversity matter. We found that higher host diversity could potentially lead to higher disease risk because many host species can support larger proportions of HIS types (Vuong – chapter 3), which leads to increasing disease risk at the genotypic level when we examine the full community. We also found that while competent reservoirs, such as the white-footed mouse and eastern chipmunk, may feed quite a few ticks (LoGiudice et al. 2008, Keesing et al. 2009), their increasing relative abundances does not necessarily imply greater genotypic disease risk. The directions of these relationships were not as expected, suggesting that we need to rethink how we assess disease risk given differences in host community diversity and composition. Second, annual and/or site variation potentially influences the *ospC* genotypes we can detect. This may be due to changes in the vertebrate host or tick population dynamics that could affect the *ospC* frequency profiles. Although most of our sites were not sampled in both years, Qiu et al. (1997) showed that there can be large variation in the frequency profiles of *ospC* types even given multi-year sampling in the same area. Last, there is a need to incorporate the greater host community. We were able to estimate a quasi-Shannon vertebrate host diversity, but to obtain a more refined community, this would require greater sophistication in identifying individuals of species that are difficult to differentiate in camera traps.

There are several factors that might influence *ospC* diversity. Our study primarily focused on the role of host composition and diversity as the drivers of *ospC* diversity. In an effort to highlight some deterministic and stochastic factors that would affect *ospC* genotype composition, frequencies, and diversity, we created a flow chart of how factors might be connected to the end product (Figure 2. 2). We suggest that host identity, transmission efficiencies of genotypes from host to tick, immune responses of host, and detection strength of *ospC* types are deterministic factors that can affect *ospC* diversity at a site. Stochastic factors might include host migration influencing the dispersal of *ospC* types, loss of *ospC* types due to small populations that can blink in and out, as well as stochastic loss of small tick populations. By knowing which variables are important and can be estimated better through empirical work, compared to variables that are difficult to estimate, may help future investigators improve our understanding of the Lyme disease system.

Table 2. 1. *ospC* Shannon-Weiner diversity values and α and β diversities among sites within year. There is slightly greater turnover of *ospC* types in year 2006 than in year 2009. For year 2006, only 28 of 30 sites had *ospC* data, and in year 2009, there were 17 of 19 sites with *ospC* data.

Year	Source of Info	DF	sH	Percent	Standardized Divergence
	Among Sites	27	0.547	21.39	0.443
2006	**Within Sites**	321	2.010	78.60	0.921
	Total	348	2.558	100.00	0.925
	Among Sites	16	0.359	15.03	0.327
2009	**Within Sites**	231	2.028	84.98	0.915
	Total	247	2.387	100.00	0.912

Table 2. 2. Pairwise comparisons between sites for year 2006. Below the diagonal are delta Δ_β between any two pairs, and above the diagonal are the *p* values. Sites with significant Δ_β have their *p* values bolded; sites that were near significant ($0.05 < p < 0.06$) have their *p* values underlined.

Site	#19	#32	#36	#37	#39	#40	ANDE
#19		0.841	0.669	1.000	0.292	0.382	**0.001**
#32	0.166		0.927	0.920	0.938	0.891	0.504
#36	0.161	0.150		0.946	0.910	0.753	0.932
#37	0.286	0.227	0.146		0.715	0.932	0.531
#39	0.284	0.097	0.160	0.230		0.691	0.300
#40	0.693	0.166	0.145	0.286	0.143		0.649
ANDE	0.611	0.220	0.148	0.471	0.298	0.195	
BEEK	0.472	0.366	0.189	0.471	0.456	0.420	0.284
BONT	0.219	0.296	0.133	0.210	0.204	0.133	0.205
DECH	0.693	0.131	0.108	0.334	0.105	0.144	0.143
FRIED	0.328	0.358	0.223	0.447	0.292	0.637	0.584
GREEN	0.311	0.201	0.152	0.297	0.165	0.106	0.100
HACK	0.683	0.476	0.204	0.521	0.360	0.683	0.529
HEIER	0.181	0.244	0.105	0.131	0.181	0.112	0.190
HILDE	0.287	0.284	0.174	0.148	0.284	0.287	0.371
HOME	0.373	0.459	0.227	0.353	0.365	0.408	0.512
HUNT	0.238	0.191	0.244	0.273	0.220	0.331	0.339
IBM	0.220	0.184	0.149	0.299	0.261	0.238	0.278
JOHN	0.131	0.135	0.080	0.159	0.182	0.094	0.105
LEON	0.272	0.200	0.184	0.296	0.209	0.195	0.235
N PARK	0.673	0.175	0.123	0.362	0.282	0.291	0.105
NDRGC	0.260	0.124	0.215	0.303	0.241	0.185	0.225
RAMA	0.662	0.310	0.222	0.413	0.291	0.142	0.333
REDH	0.412	0.508	0.212	0.520	0.681	0.586	0.436
ROCK	0.235	0.130	0.148	0.271	0.131	0.117	0.128
SBROL	0.693	0.223	0.126	0.334	0.204	0.144	0.281
SH MA	0.237	0.293	0.277	0.395	0.367	0.303	0.348
VASS	0.462	0.223	0.173	0.286	0.247	0.375	0.420

Table 2. Continue

Site	BEEK	BONT	DECH	FRIED	GREEN	HACK	HEIER
#19	0.273	**0.027**	1.000	1.000	**0.013**	0.418	**0.048**
#32	0.247	**0.022**	0.920	0.399	0.517	**0.048**	0.053
#36	0.767	**0.707**	0.948	0.682	0.758	0.528	0.796
#37	0.753	0.183	0.952	1.000	0.200	0.523	0.542
#39	0.069	0.145	0.842	0.544	0.692	0.174	0.174
#40	0.303	0.308	1.000	0.511	0.729	0.224	0.436
ANDE	1.000	0.164	0.712	0.442	0.923	0.064	0.145
BEEK		**0.027**	0.278	0.821	0.390	**0.001**	0.254
BONT	0.280		0.858	0.606	0.075	0.769	0.071
DECH	0.472	0.093		1.000	0.844	0.653	0.293
FRIED	0.477	0.139	0.482		0.290	1.000	**0.031**
GREEN	0.219	0.201	0.103	0.241		0.475	0.111
HACK	0.655	0.116	0.485	0.343	0.193		0.132
HEIER	0.173	0.177	0.133	0.241	0.188	0.175	
HILDE	0.321	0.190	0.327	0.369	0.252	0.499	0.119
HOME	0.311	0.235	0.408	0.328	0.244	0.211	0.160
HUNT	0.290	0.263	0.288	0.321	0.309	0.511	0.274
IBM	0.310	0.198	0.161	0.202	0.202	0.341	0.233
JOHN	0.139	0.186	0.067	0.152	0.146	0.214	0.191
LEON	0.306	0.223	0.225	0.351	0.192	0.530	0.132
N PARK	0.222	0.160	0.396	0.562	0.078	0.637	0.116
NDRGC	0.254	0.316	0.247	0.435	0.166	0.580	0.205
RAMA	0.405	0.194	0.250	0.563	0.186	0.687	0.127
REDH	0.321	0.357	0.586	0.683	0.387	0.521	0.195
ROCK	0.278	0.215	0.086	0.247	0.056	0.196	0.184
SBROL	0.334	0.166	0.231	0.637	0.159	0.683	0.072
SH MA	0.137	0.311	0.272	0.252	0.310	0.453	0.320
VASS	0.195	0.232	0.462	0.482	0.211	0.683	0.120

Table 2. Continue

Site	HILDE	HOME	HUNT	IBM	JOHN	LEON	N PARK
#19	0.961	0.448	0.677	0.748	0.332	0.296	1.000
#32	0.719	0.060	0.749	0.804	0.571	0.528	0.712
#36	0.914	0.462	0.422	0.902	0.895	0.817	0.711
#37	0.978	0.517	0.815	0.749	0.377	0.353	0.896
#39	0.531	0.110	0.385	0.524	0.267	0.586	0.397
#40	0.843	0.091	0.508	0.610	0.682	0.486	1.000
ANDE	0.627	**0.022**	0.210	0.901	0.839	0.404	1.000
BEEK	1.000	0.301	0.473	0.611	0.519	0.290	1.000
BONT	0.311	0.081	**0.032**	0.315	0.064	0.100	**0.042**
DECH	0.874	0.144	0.385	0.895	1.000	0.523	1.000
FRIED	0.783	0.539	0.600	0.983	0.457	0.185	1.000
GREEN	0.281	0.172	0.055	0.551	0.402	0.399	1.000
HACK	0.399	0.792	0.061	0.599	**0.024**	**0.024**	0.466
HEIER	0.628	0.308	**0.008**	0.070	**0.020**	0.450	0.174
HILDE		0.542	0.863	0.847	0.298	0.805	0.285
HOME	0.311		0.053	0.267	**0.009**	**0.016**	0.175
HUNT	0.217	0.435		0.331	**0.042**	0.085	0.108
IBM	0.264	0.360	0.340		0.998	0.673	0.820
JOHN	0.171	0.284	0.229	0.047		0.474	0.874
LEON	0.155	0.439	0.311	0.223	0.145		0.406
N PARK	0.335	0.410	0.301	0.199	0.067	0.203	
NDRGC	0.193	0.403	0.241	0.222	0.144	0.144	0.176
RAMA	0.324	0.443	0.454	0.190	0.092	0.233	0.325
REDH	0.379	0.396	0.579	0.382	0.192	0.466	0.362
ROCK	0.212	0.277	0.264	0.149	0.136	0.159	0.114
SBROL	0.327	0.316	0.403	0.190	0.072	0.225	0.396
SH MA	0.267	0.343	0.164	0.236	0.182	0.368	0.257
VASS	0.287	0.281	0.195	0.220	0.109	0.303	0.396

Table 2. Continue

Site	NDRGC	RAMA	REDH	ROCK	SBROL	SH MA	VASS
#19	0.492	0.093	0.754	0.131	1.000	0.636	1.000
#32	0.967	0.586	0.083	0.872	0.850	0.394	0.847
#36	0.637	0.551	0.621	0.762	0.878	0.249	0.615
#37	0.615	0.811	0.439	0.390	0.946	0.262	1.000
#39	0.788	0.450	**0.001**	0.871	0.550	0.125	0.328
#40	0.787	1.000	0.139	0.755	1.000	0.353	1.000
ANDE	0.531	0.544	0.409	0.842	0.495	0.272	0.303
BEEK	0.604	0.360	0.854	0.210	0.529	0.952	0.748
BONT	**0.013**	0.133	**0.012**	0.080	0.129	**0.011**	**0.014**
DECH	0.400	0.666	0.264	0.969	1.000	0.437	1.000
FRIED	0.458	0.699	0.251	0.290	1.000	0.625	1.000
GREEN	0.674	0.567	**0.036**	0.991	0.609	0.059	0.247
HACK	0.052	0.240	0.328	0.450	0.417	0.087	0.463
HEIER	0.103	0.421	0.131	0.134	0.833	**0.005**	0.336
HILDE	1.000	0.807	0.667	0.558	0.877	0.733	0.964
HOME	0.098	0.104	0.259	0.148	0.242	0.176	0.451
HUNT	0.774	0.127	**0.030**	0.152	0.174	0.867	0.804
IBM	0.778	0.917	0.440	0.873	0.839	0.550	0.755
JOHN	0.499	0.867	0.137	0.510	0.988	0.203	0.565
LEON	0.677	0.432	**0.041**	0.629	0.520	0.054	0.248
N PARK	0.894	1.000	0.731	0.833	1.000	0.376	1.000
NDRGC		0.590	0.168	0.912	0.413	0.787	0.667
RAMA	0.242		0.293	0.628	1.000	0.368	0.666
REDH	0.407	0.560		0.071	0.586	0.169	0.607
ROCK	0.096	0.177	0.343		0.584	0.149	0.408
SBROL	0.247	0.076	0.460	0.162		0.488	1.000
SH MA	0.197	0.284	0.380	0.271	0.272		1.000
VASS	0.223	0.250	0.460	0.201	0.231	0.125	

Table 2. 3. Pairwise comparisons between sites for year 2009. Below the diagonal are Δ_β between any two pairs, and above the diagonal are the *p* values. Sites with significant Δ_β have their p values bolded; sites that were near significant ($0.05 < p < 0.06$) have their p values underlined.

Site	BOST	COOK	#37	PEACE	#19	COON	#32	VASS	#40
BOST		0.979	1.000	0.935	1.000	0.755	0.438	0.902	0.525
COOK	0.097		0.983	0.774	0.915	0.634	0.808	0.901	0.187
#37	0.127	0.109		0.503	0.878	0.204	0.679	0.776	0.819
PEACE	0.138	0.168	0.270		0.665	0.848	0.415	0.473	**0.048**
#19	0.243	0.142	0.235	0.228		0.649	0.837	0.747	0.118
COON	0.227	0.173	0.333	0.114	0.270		0.276	0.447	**0.043**
#32	0.402	0.164	0.359	0.309	0.144	0.261		0.474	**0.022**
VASS	0.163	0.124	0.290	0.274	0.288	0.232	0.370		**0.017**
#40	0.269	0.281	0.224	0.433	0.431	0.493	0.599	0.584	
SPRAG	0.299	0.195	0.252	0.432	0.342	0.515	0.393	0.418	0.280
STURG	0.146	0.115	0.125	0.176	0.064	0.203	0.152	0.169	0.355
#39	0.114	0.133	0.238	0.222	0.199	0.310	0.293	0.145	0.435
TOMM	0.105	0.121	0.068	0.266	0.195	0.224	0.269	0.225	0.200
FELL	0.172	0.073	0.212	0.250	0.209	0.252	0.212	0.204	0.320
#36	0.079	0.053	0.159	0.132	0.221	0.164	0.261	0.162	0.275
POUGH	0.282	0.173	0.290	0.311	0.227	0.437	0.264	0.303	0.522
WILD	0.046	0.074	0.049	0.090	0.068	0.088	0.067	0.075	0.103

Table 3. Continue

Site	SPRAG	STURG	#39	TOMM	FELL	#36	POUGH	WILD
BOST	0.719	0.972	0.872	0.932	0.836	0.998	0.640	0.951
COOK	0.684	0.917	0.877	0.888	0.980	0.995	0.669	0.636
#37	0.923	0.976	0.670	1.000	0.645	0.950	0.800	0.936
PEACE	0.174	0.815	0.432	0.258	0.503	0.972	0.315	0.379
#19	0.860	0.978	0.771	0.803	0.783	0.847	1.000	0.782
COON	**0.040**	0.539	0.304	0.453	0.404	0.911	0.124	0.285
#32	0.553	0.793	0.440	0.455	0.683	0.525	1.000	0.712
VASS	0.464	0.731	0.787	0.423	0.715	0.845	0.510	0.559
#40	0.604	0.135	0.079	0.696	0.225	0.350	0.053	0.216
SPRAG		0.595	0.764	0.567	0.550	0.738	1.000	0.567
STURG	0.261		0.854	0.857	0.559	0.882	0.863	0.929
#39	0.237	0.121		0.676	0.746	0.827	0.888	0.907
TOMM	0.236	0.119	0.171		0.522	0.920	0.360	0.893
FELL	0.222	0.205	0.188	0.216		0.831	0.405	0.347
#36	0.197	0.171	0.141	0.143	0.143		0.787	0.572
POUGH	0.141	0.140	0.116	0.238	0.253	0.175		0.809
WILD	0.082	0.044	0.051	0.046	0.093	0.085	0.060	

Table 2. 4a. Linear regression estimates from host community metrics on *ospC* diversity in 2006. The covariate estimates, deviance, F, Pr(>F) are from model comparisons with an intercept only model. There were no significant predictors of *ospC* diversity in 2006.

Predictors	**Covariates**	**Deviance**	**F**	**Pr(>F)**	**AIC**
H' community	0.205	0.077	0.289	0.596	46.312
Mouse Rel. Abund.	0.398	0.157	0.596	0.447	45.987
Chipmunk Rel. Abund.	0.340	0.027	0.100	0.755	46.514
Short-tailed Shrew Rel. Abund.	-1.075	0.076	0.285	0.598	46.316
Mouse+Chipmunk+Shrew	0.539, 0.804, -1.123	0.355	0.428	0.735	49.163
Mouse+Chipmunk	0.539, 0.766	0.272	0.507	0.609	47.509
Mouse+Shrew	0.392, -1.041	0.228	0.421	0.661	47.694
Chipmunk+Shrew	0.378, -1.120	0.109	0.198	0.822	48.182

Table 2. 4b. Linear regression estimates from host community metrics on *ospC* diversity in 2009. The top table provides covariate estimates, deviance, F, and Pr(>F) are from model comparisons with an intercept model only. The best model includes mouse, chipmunk, and shrew relative abundances as predictors of *ospC* diversity. The best model included the relative abundances of all three common host species. Only significant complex model comparisons are shown.

Predictors	**Covariates**	**Deviance**	**F**	**Pr(>F)**	**AIC**
H' community	-0.632	0.389	1.736	0.208	26.697
Mouse Rel. Abund.	0.893	0.599	2.848	0.112	25.603
Chipmunk Rel. Abund.	2.563	1.166	6.763	**0.020**	22.232
Short-tailed Shrew Rel. Abund.	-4.546	1.277	7.742	**0.014**	21.484
Mouse+Chipmunk+Shrew	0.825, 2.849, -4.040	2.984	16.837	**<0.001**	**5.592**
Mouse+Chipmunk	1.078, 2.858	2.023	8.189	**0.004**	17.389
Mouse+Shrew	0.639, -4.060	1.569	5.032	**0.023**	21.35
Chipmunk+Shrew	2.632, -4.658	2.506	14.084	**<0.001**	11.814

Higher Order Model Comparisons	**Deviance**	**F**	**Pr(>F)**
Mouse+Chipmunk+Shrew vs Mouse+Chipmunk	-0.961	16.270	0.001
Mouse+Chipmunk+Shrew vs Mouse+Shrew	-1.415	23.950	<0.001
Mouse+Chipmunk+Shrew vs Chipmunk+Shrew	-0.478	8.086	0.014
Mouse+Chipmunk vs Mouse	-1.424	11.530	0.004
Mouse +Chipmunk vs Chipmunk	-0.857	6.938	0.020
Mouse+Shrew vs Mouse	-0.970	6.224	0.026
Chipmunk+Shrew vs Chipmunk	-1.340	15.064	0.002
Chipmunk+Shrew vs Shrew	-1.229	13.813	0.002

Figure 2. 1. Host community metrics on HIS infection prevalence. The bottom histogram indicates how many samples were not infected with an HIS type (HIS = 0), and the top histogram indicates how many samples were infected with HIS (HIS =1). The red line is the relationship between the covariates and infection probability. The first y-axis is the HIS prevalence and the second y-axis is the number of samples for the histograms.

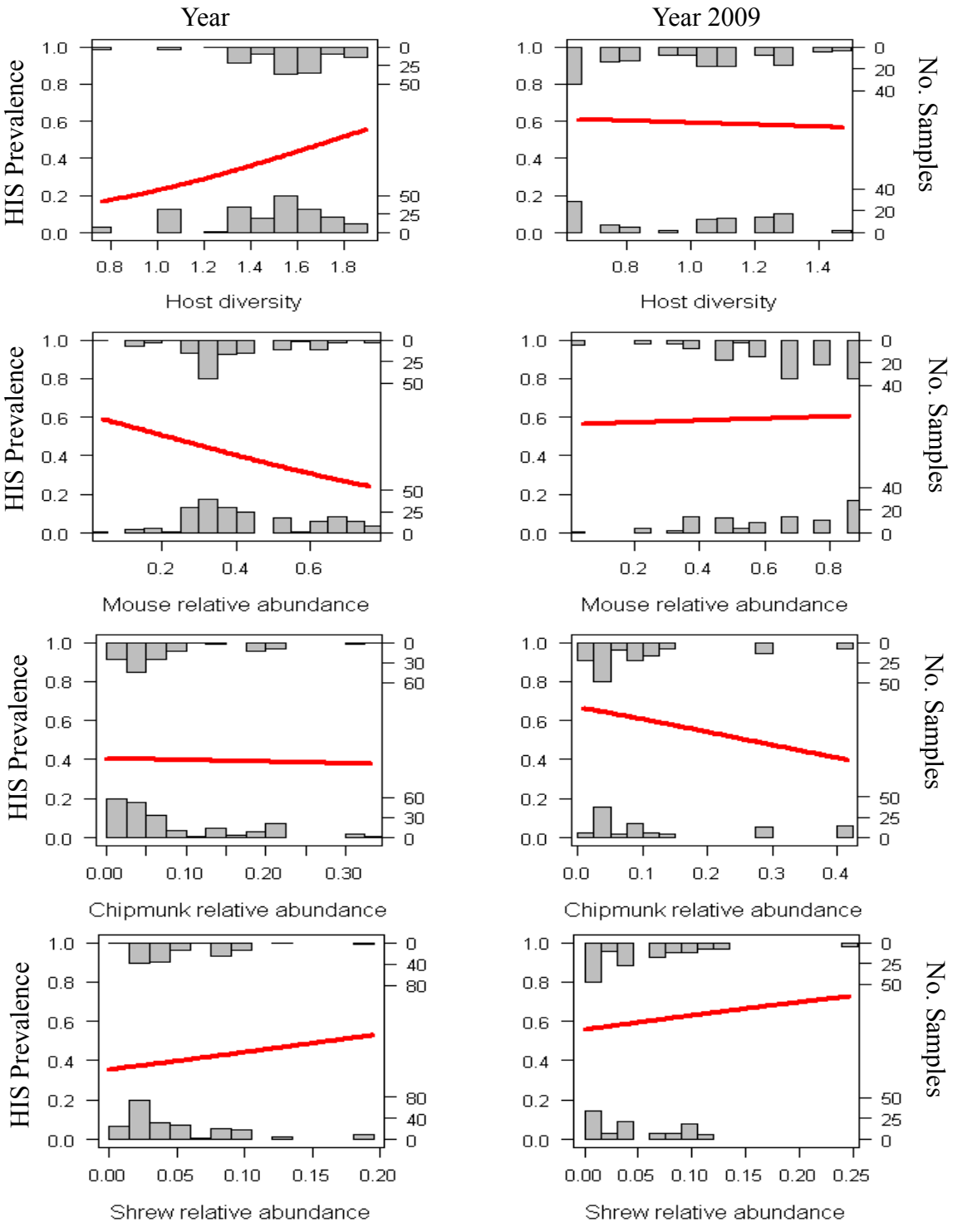

Year
Year 2009
HIS Prevalence
No. Samples
Host diversity
Mouse relative abundance
Chipmunk relative abundance
Shrew relative abundance

Table 2. 5a. Logistic regression on HIS infection probability for 2006. The top table provides z scores, df, deviance, and Pr (>Chi) from model comparisons with an intercept only model. Host community diversity and mouse relative abundance were the two significant predictors. Significant models with two or more covariates tended to included H' and mouse relative abundance. The bottom table provides only significant higher model comparisons.

Predictors	**z scores**	**df**	**Deviance**	**Pr (>Chi)**	**AIC**
H' community	3.354	1	12.364	**<0.001**	**460.91**
Mouse Rel. Abund.	-2.919	1	8.937	**0.003**	464.33
Chipmunk Rel. Abund.	-0.255	1	0.065	0.798	473.21
Short-tailed Shrew Rel. Abund.	1.319	1	1.740	0.187	471.53
Mouse+Chipmunk+Shrew	-3.203, -1.494, 1.324	3	12.692	**0.005**	464.58
Mouse+Chipmunk	-3.218, -1.399	2	10.935	**0.004**	464.34
Mouse+Shrew	-2.873, 1.208	2	10.402	**0.006**	464.87
Chipmunk+Shrew	-0.379, 1.349	2	1.885	0.390	473.39

Higher Order Model Comparison	**df**	**Deviance**	**Pr (>Chi)**
Mouse+Chipmunk+Shrew vs Chipmunk+Shrew	1	10.808	0.001
Mouse+Chipmunk vs Chipmunk	1	10.869	<0.001
Mouse+Shrew vs Shrew	1	8.662	0.003

Table 2. 5b. Logistic regression on HIS infection probability for 2009. The top table provides z scores, df, deviance, Pr (>Chi) from model comparisons with an intercept model only. All complex higher order models included chipmunk relative abundances, but model comparisons suggest that the single covariate of chipmunk relative abundances is sufficient in predicting HIS infection prevalence. The bottom table provides significant higher order model comparisons.

Predictors	**z scores**	**df**	**Deviance**	**Pr (>Chi)**	**AIC**
H' community	-0.435	1	0.189	0.664	339.03
Mouse Rel. Abund.	0.327	1	0.107	0.744	339.11
Chipmunk Rel. Abund.	-2.492	1	6.336	**0.012**	**332.88**
Short-tailed Shrew Rel. Abund.	1.121	1	1.290	0.256	337.93
Mouse+Chipmunk+Shrew	-0.221, -2.702, 1.516	3	9.829	**0.020**	333.39
Mouse+Chipmunk	-1.019, -2.645	2	7.413	**0.025**	333.81
Mouse+Shrew	1.011, 1.454	2	2.319	0.314	338.90
Chipmunk+Shrew	-2.879, 1.786	2	9.780	**0.008**	331.44

Higher Order Model Comparison	**df**	**Deviance**	**Pr (>Chi)**
Mouse+Chipmunk+Shrew vs Mouse+Shrew	1	7.510	0.006
Mouse+Chipmunk vs Mouse	1	7.306	0.007
Chipmunk+Shrew vs Shrew	1	8.490	0.004

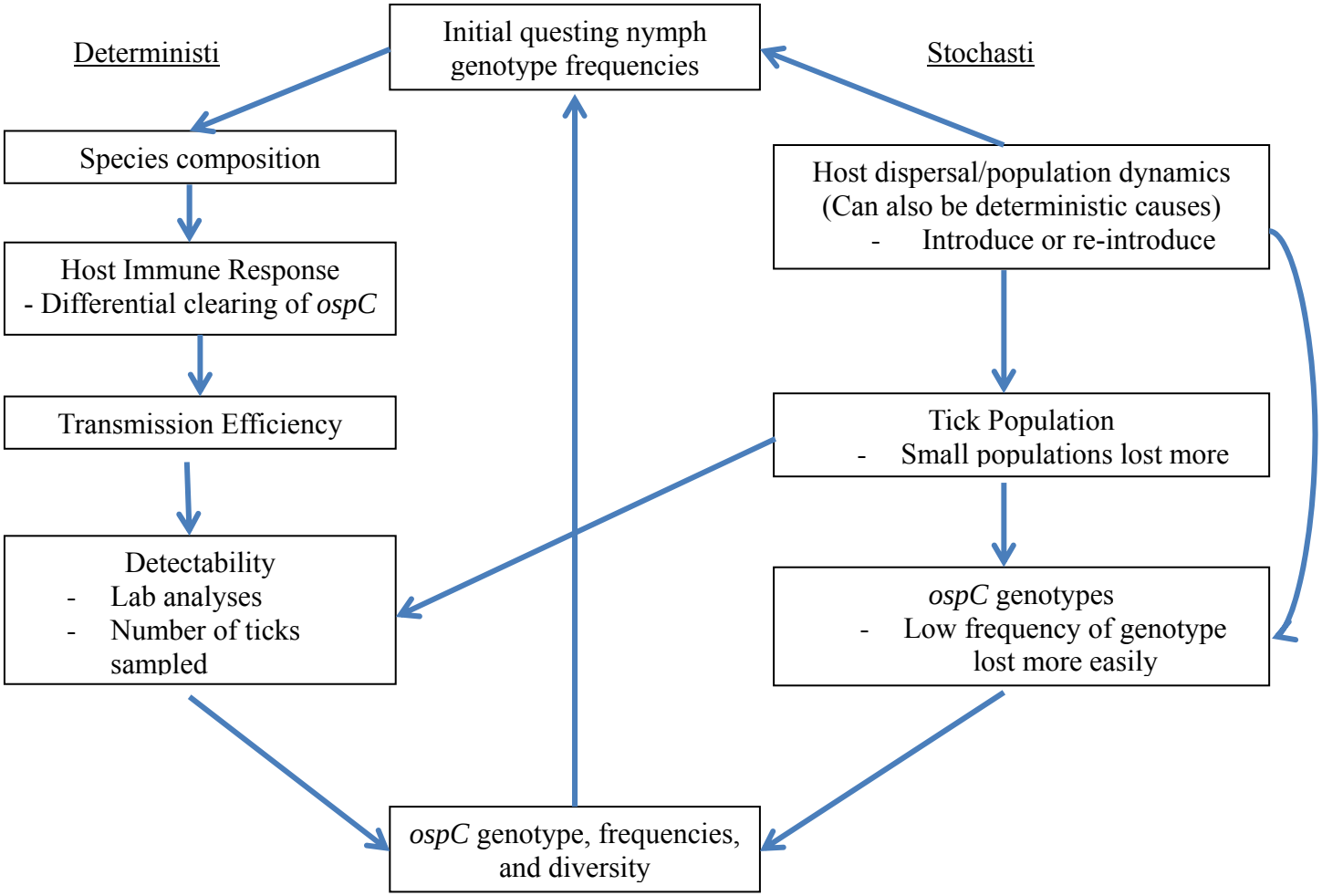

Initial questing nymph genotype frequencies
Deterministi
Stochasti
Species composition
Host Immune Response
- Differential clearing of ospC
Transmission Efficiency
Detectability
- Lab analyses
- Number of ticks sampled
Host dispersal/population dynamics
(Can also be deterministic causes)
- Introduce or re-introduce
Tick Population
- Small populations lost more
ospC genotypes
- Low frequency of genotype lost more easily
ospC genotype, frequencies, and diversity

Figure 2. 2. Flow diagram on deterministic and stochastic factors that might influence *ospC* genotypic profiles. The diagram can be used as a guide in determining where variation can be captured the most to have more predictive power examining how host biodiversity influences *ospC* genotypes.

Citation

Barbet, A. F., A. M. Lundgren, A. R. Alleman, S. Stuen, A. Bjoersdorff, R. N. Brown, N. L. Drazenovich, and J. E. Foley. 2006. Structure of the expression site reveals global diversity in MSP2 (P44) variants in *Anaplasma phagocytophilum*. Infection and Immunity **74**:6429-6437.

Brisson, D. and D. E. Dykhuizen. 2004. *ospC* diversity in *Borrelia burgdorferi*: Different hosts are different niches. Genetics **168**:713-722.

Brisson, D., D. E. Dykhuizen, and R. Ostfeld. 2008. Conspicuous impacts of inconspicuous hosts on the Lyme disease epidemic. Proceedings of the Royal Society B: Biological Sciences **275**:227-235.

Dykhuizen, D. E., D. Brisson, S. Sandigursky, G. P. Wormser, J. Nowakowski, R. B. Nadelman, and I. Schwartz. 2008. The propensity of different *Borrelia burgdorferi sensu stricto* genotypes to cause disseminated infections in humans. Am J Trop Med Hyg **78**:806-810.

Falco, R. C. and D. Fish. 1992. A comparison of methods for sampling the deer tick, *Ixodes dammini*, in a Lyme disease endemic area. Experimental and Applied Acarology **14**:165-173.

Giardina, A. R., K. A. Schmidt, E. M. Schauber, and R. S. Ostfeld. 2000. Modeling the role of songbirds and rodents in the ecology of Lyme disease. Canadian Journal of Zoology **78**:2184-2197.

Girard, Y. A., N. Fedorova, and R. S. Lane. 2011. Genetic diversity of *Borrelia burgdorferi* and detection of *B. bissettii*-like DNA in serum of north-coastal California residents. Journal of Clinical Microbiology **49**:945-954.

Hanincova, K., K. Kurtenbach, M. Diuk-Wasser, B. Brei, and D. Fish. 2006. Epidemic spread of Lyme borreliosis, northeastern United States. Emerging Infectious Diseases **12**:604-611.

Kapil, S., R. W. Allison, L. Johnston, B. L. Murray, S. Holland, J. Meinkoth, and B. Johnson. 2008. Canine distemper virus strains circulating among North American dogs. Clinical and Vaccine Immunology **15**:707-712.

Keesing, F., J. Brunner, S. Duerr, M. Killilea, K. LoGiudice, K. Schmidt, H. Vuong, and R. S. Ostfeld. 2009. Hosts as ecological traps for the vector of Lyme disease. Proceedings of the Royal Society B-Biological Sciences **276**:3911-3919.

Kurtenbach, K., M. Peacey, S. G. T. Rijpkema, A. N. Hoodless, P. A. Nuttall, and S. E. Randolph. 1998. Differential transmission of the genospecies of *Borrelia burgdorferi sensu lato* by game birds and small rodents in England. Applied and Environmental Microbiology **64**:1169-1174.

Kurtenbach, K., S. M. Schafer, S. de Michelis, S. Etti, and H.-S. Sewell. 2002. *Borrelia burgdorferi sensu lato* in the vertebrate host.*in* J. S. Gray, O. Kahl, R. S. Lane, and G. Stanek, editors. Lyme Borreliosis: Biology, Epidemiology, and Control. CABI International, Oxford, UK.

Levi, T., A. M. Kilpatrick, M. Mangel, and C. C. Wilmers. 2012. Deer, predators, and the emergence of Lyme disease. Proceedings of the National Academy of Sciences **109**:10942-10947.

Liang, F. T., J. Yan, M. L. Mbow, S. L. Sviat, R. D. Gilmore, M. Mamula, and E. Fikrig. 2004. *Borrelia burgdorferi* changes its surface antigenic expression in response to host immune responses. Infection and Immunity **72**:5759-5767.

LoGiudice, K., S. T. K. Duerr, M. J. Newhouse, K. A. Schmidt, M. E. Killilea, and R. S. Ostfeld. 2008. Impact of host community composition on Lyme disease risk. Ecology **89**:2841-2849.

LoGiudice, K., R. S. Ostfeld, K. A. Schmidt, and F. Keesing. 2003. The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences of the United States of America **100**:567-571.

MacQueen, D. D., C. Lubelczyk, S. P. Elias, B. K. Cahill, A. J. Mathers, E. H. Lacombe, P. W. Rand, and R. P. Smith Jr. 2012. Genotypic diversity of an emergent population of *Borrelia burgdorferi* at a costal Maine island recently colonized by *Ixodes scapularis*. Vector-Borne and Zoonotic Diseases **12**:456-461.

Murray, K. O., E. Mertens, and P. Despres. 2010. West Nile virus and its emergence in the United States of America. Veterinary Research **41**.

Ogden, N. H., L. R. Lindsay, K. Hanincova, I. K. Barker, M. Bigras-Poulin, D. F. Charron, A. Heagy, C. M. Francis, C. J. O'Callaghan, I. Schwartz, and R. A. Thompson. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. Appl. Environ. Microbiol. **74**:1780-1790.

Ogden, N. H., G. Margos, D. M. Aanensen, M. A. Drebot, E. J. Feil, K. Hanincová, I. Schwartz, S. Tyler, and L. R. Lindsay. 2011. Investigation of genotypes of *Borrelia burgdorferi* in *Ixodes scapularis* ticks collected during surveillance in Canada. Applied and Environmental Microbiology **77**:3244-3254.

Ostfeld, R. S. 1997. The ecology of Lyme-disease risk. American Scientist **85**:338-346.

Peakall, R. and P. E. Smouse. 2006. GenAlEx 6: genetic analysis is Excel. Populaiton genetic software for teaching and research. Molecular Ecology Notes **6**:288-295.

Qiu, W. G., E. M. Bosler, J. R. Campbell, G. D. Ugine, I. N. Wang, B. J. Luft, and D. E. Dykhuizen. 1997. A population genetic study of *Borrelia burgdorferi sensu stricto* from eastern Long Island, New York, suggested frequency-dependent selection, gene flow and host adaptation. Hereditas **127**:203-216.

Qiu, W. G., D. E. Dykhuizen, M. S. Acosta, and B. J. Luft. 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**:833-849.

R. 2008. R: A language and environment for statistical computing. Version 2.15. R Foundation for Statistical Computing, Vienna, Austria.

Rudenko, N., M. Golovchenko, L. Grubhoffer, and J. Oliver Jr. 2011. Updates on *Borrelia burgdorferi sensu lato* complex with respect to public health. Ticks and Tick-borne Diseases **2**:123-128.

Sangula, A. K., G. J. Belsham, V. B. Muwanika, R. Heller, S. N. Balinda, and H. R. Siegismund. 2012. Co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya highlights challenges to foot-and-mouth disease control. Archives of Virology **155**:1625-1630.

Stanek, G. and M. Reiter. 2011. The expanding Lyme *Borrelia* complex--clinical significance of genomic species? Clinical and Microbiological Infections **17**:487-493.

Velo-Anton, G., D. Rodriguez, A. E. Savage, G. Parra-Olea, K. R. Lips, and K. R. Zamudio. 2012. Amphibian-killing fungus loses genetic diversity as it spreads across the New World. Biological Conservation **146**:213-218.

Wang, I., D. E. Dykhuizen, W. Qui, J. J. Dunn, E. M. Bosler, and B. J. Luft. 1999. Genetic diversity of *ospC* in a local population of *Borrelia burgdorferi sensu stricto*. Genetics **151**:15-30.

Wendte, J. M., A. K. Gibson, and M. E. Grigg. 2011. Population genetics of *Toxoplasma gondii*: New perspectives from parasite genotypes in wildlife. Veterinary Parasitology **182**:96-111.

Chapter 3: Contribution of *Borrelia burgdorferi* genotypes between mammalian and avian hosts

Abstract

Borrelia burgdorferi, the bacterium that causes Lyme disease, infects a wide array of host species. Previous research argued that hosts act as ecological niches for *ospC* (outer surface protein C) genotypes, with particular genotypes occurring at higher frequencies in some host species than others. However, this study was based on four commonly infected small mammal species. We include several additional mammalian and avian host species to test whether there is support for the niche concept. Because Lyme disease risk is primarily associated with five *ospC* types, known as human invasive strains (HIS), we also examined how HIS types distribute themselves among host species, and among higher taxa (birds, shrews, rodents). We adapted the patch occupancy model used for species detection, to test for the occurrence probabilities (ψ) and transmission efficiencies (ε) associated with each *ospC* type. We found support for the niche hypothesis, with differing *ospC* frequencies detected within each host species. A principal components analysis showed that birds clustered together, while rodents formed a looser cluster, and the two shrew species were far apart from each other. Examining only HIS types, the species clustered more by the HIS types they support than by the higher taxa level. Types A and K occurred frequently in our sample (high ψ), while types H, L, and U were more infrequent (low ψ), and all other types were intermediate. High occurrence did not signify high transmission, nor does low occurrence signify low transmission. Hence types T and U had the highest (ε), followed by several of the HIS types. Taken together, this shows that genotypes do vary across the host species, indicating support for the concept that hosts act as ecological niches. Differences in

transmission efficiencies, where infrequent genotypes have high efficiencies, might suggest that these genotypes are better at being taken up by the tick vectors. Additionally, our study highlights the importance of avian hosts in contributing HIS types to ticks.

Introduction

Lyme disease is caused by an infection with the spirochete bacterium, *Borrelia burgdorferi s.s.* (Burgdorfer et al. 1982). This bacterium primarily circulates within vertebrate host species, vectored by the blacklegged tick (*Ixodes scapularis*). The bacterium is under balancing selection, which is likely a result of multiple niche polymorphisms due to heterogenous environments (variation in individuals within and among species) (Brisson and Dykhuizen 2004) rather than due to immunological pressures from the wide host species the bacterium infects (Wang et al. 1999, Qiu et al. 2002). Hence, the *B. burgdorferi* population is not genetically homogenous, but is rather composed of distinct genotypes. Lyme disease risk is primarily estimated using diagnostic testing of B. *burgdorferi* infection within host or ticks, and often ignores the genotypic variation of the pathogen. But, estimating risk based on prevalence of the pathogen alone can be misleading, if these genotypes differ in their ability to invade humans (and other vertebrate hosts) and/or cause disease. Therefore, it is fundamental that we understand the genotypic composition of B. *burgdorferi* detected in the tick populations, to better estimate disease risk.

In the northeastern U.S., small- and medium sized mammals, and birds, are the two main host groups of *B. burgdorferi* (Kurtenbach et al. 2002). White-tailed deer (*Odocoileus virginianus*), as well as lizards in western and southeastern US, can host ticks (Eisen et al. 2004, Giery and Ostfeld 2007, Salkeld and Lane 2010), but they do not become infected with *B. burgdorferi* because these hosts have antibodies to clear the infections (Magnarelli et al. 1984, Lane et al. 2006). Small mammal species, especially the white-footed mouse

(*Peromyscus leucopus*), eastern chipmunk (*Tamius striatus*), and short-tailed shrew (*Blarina brevicauda*), are known to be competent reservoirs of the Lyme bacterium (Lane et al. 1991, LoGiudice et al. 2003, Brisson et al. 2008, Keesing et al. 2009). Medium-sized animals like raccoons (*Procyon lotor*) and opossums (*Didelphis virginianus*) can become infected, but are not very competent host species (Keesing et al. 2009). Birds are also considered to play a role in Lyme disease (Giardina et al. 2000), with American robins (*Turdus migratorius*) as a primary avian reservoir (Battaly and Fish 1993, Richter et al. 2000, Ginsberg et al. 2005). However, other studies on birds and tick attachment rates also indicate that other thrushes, blackbirds, and other ground foraging birds can be important hosts in the ecology of Lyme disease (Rand et al. 1998, Taragel'ova et al. 2008, Mitra et al. 2010). It would beneficial to understand how these species contribute to genotypic variation of ***B. burgdorferi***, in conjunction with their role as competent or incompetent host species.

The outer surface protein C (*ospC*) is one of the most studied ***B. burgdorferi*** genetic loci (Ohnishi et al. 2001, Liang et al. 2004). Genotypic variation at this locus can be differentiated into 22 *ospC* genotypes, of which 17 are known to occur in the northeastern United States (Wang et al. 1999, Qiu et al. 2002). The surface protein is a target for antibodies produced by the host immune responses (Fikrig et al. 1991). Variation of the *ospC* locus may provide the pathogen to differentially evade particular host immune systems it might encounter, but this has not been currently shown to occur. Nonetheless, studies have found differential infection rates of vertebrate hosts by particular genotypes (Brisson and Dykhuizen 2004, Hanincova et al. 2006). In humans, there are five of seventeen types considered to be invasive, which were defined by the authors as having greater frequency in

humans compared to their frequencies in the tick populations (Seinost et al. 1999, Dykhuizen et al. 2008). These human invasive genotypes (HIS) include *ospC* types A, B, I, K, and N.

Due to vertebrate host heterogeneity, this sets up the potential for hosts to act as ecological niches for certain genotypes (Brisson and Dykhuizen 2004). Current data on host-genotype associations, based on *ospC* associations, are drawn from a handful of studies primarily associated with small mammals and migratory birds (Brisson and Dykhuizen 2004, Alghaferi et al. 2005, Anderson and Norris 2006, Hanincova et al. 2006, Ogden et al. 2008, Mathers et al. 2011, MacQueen et al. 2012). For both host taxa, the studies show support for the ecological niche concept. However, Swanson and Norris (2008) argue that changes in frequencies can occur within an individual over time as these animals are continually infected with any or all genotypes from the ticks they host.

In these previously cited studies on host-*ospC* genotype associations, the studies focused exclusively on mammals or birds, but not both together. Moreover, the mammals used in the study often came from areas where Lyme disease is endemic (except MacQueen et al. 2012), whereas the birds in the studies were captured in relatively recent invaded areas of Lyme disease, making comparisons of between these two host taxa difficult. Additionally, although these studies were able to detect HIS types from ticks feeding on animals, their emphasis was not on the HIS contribution of these species (see Brisson and Dykhuizen (2004) and Alghaferi et al. (2005) for exceptions).

Our study focused on how particular host species, and higher taxa of hosts (birds, shrews, rodents), present in endemic areas of Lyme disease, differ in the *B. burgdorferi ospC* genotypic composition and frequency profiles they support. We specifically highlight HIS detections in each species, and higher taxa, to determine whether there is differential

contribution of HIS types among these species and taxa. As Brinkerhoff et al. (2011) state in their conclusion on the role of migratory birds dispersing *ospC* genotypes:

"If bird-specialized *B. burgdorferi* genotypes cannot persist in mammalian hosts and do not cause disease in humans, the role of birds in spreading *B. burgdorferi*-infected ticks may be inconsequential to human health. However, if birds commonly carry *B. burgdorferi* strains that are infectious to humans and other mammalian species, their impact on Lyme disease eco-epidemiology could be profound."

Additionally, we estimated the occurrence probability of each genotype, and their support intervals, from our data using a likelihood approach to determine the commonness of the genotypes (MacKenzie et al. 2002). In order for the bacterium to circulate from infected host to a non-infected host, it must pass through the tick vector. Hence, the transmission efficiency of each genotype also matters. The likelihood model also estimates transmission efficiencies, and their support intervals, to establish how well the genotype is at entering a tick, given an infected host.

Methods

Small and medium-sized mammals were live-trapped and avian hosts were mist netted during the summers of 2008 – 2010 at the Cary Institute of Ecosystem Studies in Millbrook, NY (IACUC # 06-01 and 09-01). White-footed mice and chipmunks were captured on long-term Lyme disease research grids that focus on oak masting and population dynamics of small mammals. The animals were captured in folding aluminum Sherman traps, and only dispersing juveniles and adults, but not lactating females, were temporarily removed for sampling. Other vertebrate host species were trapped to examine reservoir competency (Keesing et al. 2009). We used pit-fall traps to sample masked and short-tailed shrews, but some were caught in Sherman traps set for mice and chipmunks. To sample flying squirrels

(*Glaucomys volans*), we placed Sherman traps inside small squirrel-sized Tomahawk traps on platforms located on tree trunks ~2 m above ground. The opening of the Sherman was facing the back of the Tomahawk to prevent gray squirrels (*Sciurus carolinensis*) from entering the Sherman trap. We captured red (*Tamiasciurus hudsonicus*) and gray squirrels, and Virginia opossums (*Didelphis virginianus*) opportunistically, by placing small and medium-sized Tomahawks haphazardly on the Cary property. Gray squirrels were also sampled on the long-term grids, as well as from Poughkeepsie, NY, in order to increase our sample size, due to low trapping success on the Cary property. Large tomahawk traps were placed at the burrow entrances of groundhog (*Marmota monax*) nests to capture groundhogs, and in various locations on Cary Institute property to capture raccoons (*Procyon lotor*) and skunks (*Mephitis mephitis*). Mist nets were set up in several locations to capture American Robins (*Turdus migratorius*), Gray Catbirds (*Dumetella carolinensis*), Veeries (*Catharus fuscescens*), Ovenbirds (*Seiurus aurocapilla*), and Wood Thrushes (*Hylocichla mustelina*). We focused on these birds, because they are common, nest on or near the ground, and feed on the ground, making them potential hosts for tick vectors (Giardina et al. 2000).

The animal hosts were taken to the Rearing Facility at the Cary Institute and held in cages for 4-7 days. All animals were fed and watered twice a day, but checked periodically during the day and replenished with food and water *ad libitum*. We expected ticks feeding on the hosts at time of capture to feed to repletion and fall off by day 4. Because feeding to repletion takes longer than 24 hour for newly attached ticks, this allowed us to inoculate the animals at the end of day 3. We inoculated each individual with 100 unfed larval ticks, a mix of ticks collected from Cary Institute property, an off-site property known for high larval densities, and as well as with laboratory-raised ticks. Inoculations of the hosts were

conducted over a table draped with a white cloth, so that we could find ticks that had fallen from the host during the inoculation period and account for the numbers of ticks that stayed on the host. The animals were placed into appropriate sized and ventilated PVC pipes for 3-4 hours to restrict animal movement, to allow the larvae the opportunity to find a suitable attachment site. We placed fruit pieces in the pipe to serve as a food and water source. After this time period, we returned the animals to their cages. We checked the PVC pipe for larvae that did not attach or that were partially eaten to account for the true number of ticks that attached to the host individual. For three days post inoculation, we recorded the number of ticks that fell into the collecting pan located under each animal's cage to determine feeding success (see Keesing et al. 2009 for more methodological details). The animals were released into their original capture location at the end of day 7. Fully fed larval ticks were collected into plaster of paris vials, moistened with de-ionized water and labeled by host species, individual tag number, and date of collection. The ticks were monitored for several weeks to determine molting success. The newly molted nymphs were collected and flash frozen for future *B. burgdorferi ospC* characterization.

We used Qiagen DNEasy extraction kits to extract DNA from each individual tick, following the animal tissue protocol. We had a final elution volume of 50-100 ul with buffer AE. Most samples in this study were tested for *B. burgdorferi* using qPCR at Bard College (Hersh et al. *In Press*) to amplify a 75 bp fragment of the 23S rDNA with Tamra probe Bb23Sp and primers Bb23Sf and Bb23Sr (Courtney et al. 2004). Positive eluted DNA samples were sent to the University of Pennsylvania lab to amplify the *ospC* gene, using newly developed outer primers OC-368F/OC693R and nested primers OC4+F/OC643 (Deverey et al. *in prep*). These samples were then subjected to the reverse line blot (RLB) to

test for specific *ospC* genotypes that were infecting these host species (Qiu et al. 2002, Brisson & Dykhuizen 2004, Vuong - Chapter 1). A small subset of whole tick samples were extracted and tested for the presence of *ospC*, using the same primer set at Rutgers University. All samples were subjected to 1% gel electrophoresis to determine the presence of the gene, before being tested with the RLB.

Our dataset represents the following 10 host species (and the number of positive individuals): white-footed mouse (12), eastern chipmunk (10), short-tailed shrew (10), masked shrew (3), eastern gray squirrel (4), red squirrel (7), striped skunk (1), American robin (13), Veery (16), and Wood Thrush (4). We tested a minimum of three positive ticks per host individual and up to seven positive random ticks if there were more than seven positive ticks per individual. We removed our one skunk from the analyses, but will present the strain detected in our results. The inclusion of birds in our study allowed us to examine the *ospC* frequencies across a larger host species range compared to past studies, and separated by host type (birds, shrews, and rodents) when appropriate.

Statistical Analyses – We examined the pattern of distribution for all 17 *ospC* types within hosts species with Principal Component Analysis (PCA). To correct for differences in the numbers of individuals tested per species, we obtained the proportion of the 17 *ospC* types within a species by dividing the number of times we detected *ospC* by the total number of detections of all *ospC* types for that species. These proportion values were then used in the PCA. We also calculated species-specific relative human invasive strains (HIS) proportions, based on the five HIS types within each species and used these proportions in a PCA. We tested for a difference in the proportion of HIS to non-HIS genotypes across all the species, and separately by host type, using a contingency analysis.

We adopted the occupancy and detection likelihood analysis of MacKenzie et al. (2002) to test for the frequency of occurrence of *ospC*$_t$ (where t = types A, B… U), in our host individuals, and to estimate the transmission efficiency of *ospC* types from an infected host to a tick. Because every species will not be positive for all *ospC* types, nor will all ticks feeding on that host individual always pick up the same genotypes, the non-detection of a type in the host individual does not imply the absence of that *ospC* type, unless the detection probability is equal to one (MacKenzie et al. 2002). Because we expected *a priori* that there should be some host-genotype associations, the probability of detection of all *ospC* types in the host individuals would be less than one. The use of this model is beneficial, because it allows us to estimate probabilities associated with the occurrence of the *ospC* types and their transmission efficiencies across all the samples, and also to estimate species-specific probabilities. This model is also robust for smaller replicates, as long as the occurrence probability is greater than 0.3 (MacKenzie et al. 2002).

Individual ticks from the hosts are considered replicate observations from that particular host. We ran 16 null and species-specific models, with 2500 iterations each, and compared the models using AIC_c criteria (Burnham and Anderson 2002). *ospC* type J was not analyzed, because we detected this genotype in a single host species in our study, rendering comparisons of different hosts pointless. In the species-specific model, the species are the covariates. Species with at least one positive individual for a particular *ospC* type were included in those analyses pertaining to that particular genotype. For example, we did not include masked shrews in the *ospC* type A analysis, because we did not detect this genotype in any of the three masked shrew individuals, so our sample size for *ospC* type A was reduced from 79 to 76. The null model provided an average probability of occurrence

(ψ) and transmission efficiency (ε), along with the respective support limits. The species-specific models provided (ψ) values and support limits for each host species, and a common (ε) value and its support limits. These probabilities provide information on the commonality of these *ospC* types in our study and how well each type is transmitted from host to tick, which is an important component of continual cycling for these genotypes.

Results

ospC distribution and HIS proportions among hosts – Masked shrews had the lowest *ospC* richness (*ospC* types = 4) and veeries had the highest (*ospC* types = 16), but this was affected by the number of individuals sampled (Table 3. 1) ($r_s = 0.90$, n = 9, $p < 0.001$). Vertebrate host individuals averaged 4.05 (± 2.29 sd) *ospC* types, while each tick individual averaged 2.07 (± 1.24 sd) *ospC* types. There are also divergent *ospC* proportions in different host species (Table 3. 2). Most host species tend to have higher numbers of individual infected with particular *ospC* types. For example, *ospC* types T and U are common in the squirrels and chipmunks, while *ospC* types G is detected more often in mice, robins, and veeries. The PCA, which is based on species-specific *ospC* proportions, show that birds tend to cluster together, while rodents have a loose cluster, and short-tailed shrews are more similar to the rodents than to the masked shrew (Figure 3. 1). The first two axes of the PCA explained approximately 56.3% of the *ospC* variation. Masked shrews are positively associated with axis 1, due to higher proportions of *ospC* types B, E, G, and H, and the rest of the species were negatively associated with axis 1 due to higher proportions of *ospC* types A, F, M, N, T, and U (Table 3. 2). The second axis separated the birds from the rodents, with rodents positively associated and birds negatively associated with axis 2. Rodents tend to have greater proportions of *ospC* types F, H, T, and U, and birds were more associated with *ospC* types A, D, I, K, and O.

Examining only HIS types, the first two axes of the PCA explained 82.0% of the variation. Axis 1 was positively associated with larger proportions of *ospC* type B and negatively associated with larger proportions of *ospC* type A. Axis 2 was positively associated with larger proportions of *ospC* type N and negatively associated with *ospC* types I and K. We see a similar pattern in the spatial arrangement of the species, based only on HIS types, as we did when we examined the species using all the *ospC* types. In the PCA for HIS only, the strain separation of masked shrew, gray squirrel, and wood thrush become more pronounced (Figure 3. 2). Masked shrew is again on its own, along axis 1, because we detected only *ospC* type B in masked shrew. Gray squirrels and wood thrushes are far apart from one another along axis 2 because we detected only *ospC* types A and N in squirrels but we detected *ospC* types A, I, and K in wood thrushes. The other species had similar proportions of the HIS types they supported (Figure 3. 3), which resulted in the PCA clustering of those species.

The ratio of HIS to non-HIS was significantly different among the nine host species ($\chi^2 = 18.557$, df = 8, $p = 0.0167$). At the high end, HIS made up at least 60% of the detection in short-tailed shrews and American robins, compared to the low end of ~20-25% HIS proportion detected in gray and red squirrels. All other host species had an HIS to non-HIS ratio between 30~45% (Figure 3. 4A). When we combined species into the broader host types, we again detected a significant difference in the ratio of HIS to non-HIS types ($\chi^2 = 6.734$, df = 2, $p = 0.0345$). Rodents and birds were more similar in their HIS proportions than were either to the shrew group (Figure 3. 4B).

Occurrence and transmission efficiencies of ospC types– The null models had consistently smaller AIC_c values than did species-specific models, except for *ospC* types C

and T (Table 3. 3), suggesting that additional parameters do not, in general, provide a better fitting model. Three *ospC* types (H, L, and U) had occurrence probabilities that were low ($\psi < 0.2$), two had high probabilities ($\psi > 0.5$; types A and K), and the remaining types had intermediate occurrence probabilities ($\sim 0.3 < \psi < \sim 0.4$) (Figure 3. 5A). Although the support intervals are rather wide, there is sufficient support to separate the common from the infrequent genotypes in our samples.

The estimates for transmission efficiency (ε) of $ospC_t$ from an infected host to a tick individual also vary, but generally have tighter support intervals, than for the (ψ) estimates (Figure 3. 5B). There were two *ospC* types (T and U) with high transmission probabilities ($\varepsilon \approx 0.7$), with a few more types with medium probabilities ($0.55 < \varepsilon < 0.6$), and with the remaining types with lower transmission probabilities ($0.35 < \varepsilon < 0.45$). Again, there is enough support to differentiate *ospC* types with high transmission efficiencies (T and U) from the types that have low transmission probabilities.

Discussion

Our study shows that birds, shrews, and rodents make different contributions to the composition and frequency profile of *B. burgdorferi ospC* genotypes detected from ticks feeding on these host species. Our data also support the idea that hosts act as different ecological niches based on differences in the relative proportions of *ospC* types detected in each species (Brisson and Dykhuizen 2004). This was especially true for *ospC* type J, which was a rare genotype, and detected only in red squirrels. We also found that human invasive strains (HIS) are detected in all host species, and sometimes as commonly in avian hosts as it was in mammalian hosts. Finally, we found that the occurrence (ψ) varies among host species and transmission efficiencies (ε) of *ospC* types varies among *ospC* genotypes.

However, the commonness (ψ) of an *ospC* type does not translate directly into its transmission efficiency (ε).

ospC distribution and HIS proportions among hosts – Most (16 out of 17) *ospC* types were detected in at least three host species, with species showing differences in which ospC type occurred more often in the vertebrate host species (Table 3. 2). For example, *ospC* types T and U, are common in squirrels and chipmunks, but less so in the other host species. *ospC* types I, N, and sometimes M, were relatively more abundant in avian hosts than in either mammalian group. Variation in relative proportions of *ospC* types is seen in all three host groups, and across the nine host species, suggesting support for the niche concept, but further investigation of the strength of host-genotype assocations is needed, due to incongruities between our results and those of previous studies (Brisson and Dykhuizen 2004, Hanincova et al. 2006).

The genotypes that were commonly detected in white-footed mice and short-tailed shrews in our study have commonly been detected in these species in previous studies (Brisson and Dykhuizen 2004, Hanincova et al. 2006). We detected some differences in which genotypes were more common for the chipmunks compared to past studies, but the differences were most pronounced for the gray squirrels. In our study, *ospC* types A, N, and T were detected in relatively high proportions in chipmunks, whereas Brisson and Dykhuizen (2004) found *ospC* types D, K, and T to be most common, and Hanincova et al. (2006) detected more of types D, K, and U in ticks that fed on chipmunks. For the gray squirrel, we found relatively high frequencies of types F, M, and T, but Brisson and Dykhuizen (2004) showed high frequencies of types A, E, and K, and Hanincova et al. (2006) detected types N and U from squirrels. Temporal and spatial differences among these three studies may be

contributing to some of these differences, but they may also indicate that the host-genotype association is dependent on where and when the ticks were collected. Further investigation of hosts and their *ospC* types would help to solidify the niche concept or dispel it convincingly. We will need larger samples, of course, collected over both time and space, to put this question to rest.

We also found strong *ospC* genotype associations with the American robin and wood thrush, and less so with the veery. For the first two birds, we again see that some genotypes are detected more often than others; types A, G, K, and N (robin), and types I and M (wood thrush). It may be that the veery is a host that is permissive to more tick feeding and hence, potentially leading to more *B. burgdorferi* genotypes that infect this species. If so, this would suggest that veeries are important hosts in helping to circulate *B. burgdorferi* genotypes, including HIS types that are detected as commonly has non-HIS types.

Could permissiveness of the host immune systems between birds, shrews, and mammals differ enough to garner differences in the number of ospC they support, and the frequencies of which these genotypes occur? The existing ecological immunology research for Lyme disease may not be able to answer that yet. The combination of genotypes, host species, as well as variation of host immune systems within a species, may be too great to conquer in order to understand the ecological immunology that is taking place with *B. burgorferi.*

Occurrence and transmission efficiencies of ospC$_t$ – We found that the majority of null models had a better fit than the species-specific model. This may have to with the small sample sizes, in both tick replicates, and number of individuals, for the species-specific models. Obtaining species-specific estimates would have provided either more support, or

greater, rejection to the host as ecological niches concept, but obtaining solid species-specific estimates would require an immense amount of animal trapping, rearing, and tick collection. Nonetheless, our null model was able to show that genotypes occur at different probabilities across our samples (Figure 3. 5A). These differences in occurrence potentially imply that particular genotypes are better at evading the immune systems of host species than other genotypes. In our study, types A and K were the most commonly occurring genotypes across the samples, and they were commonly occurring in other studies that examined small mammal (Brisson and Dykhuizen 2004, Hanincova et al. 2006) and avian hosts (Ogden et al. 2008). Their commonness suggests that either these genotypes easily infect these host species, and/or that they have evolved mechanisms to circumvent multiple immune systems (Liang et al. 2002). One evasion mechanism may be to down regulate the *ospC* gene with the help of an *ospC* operator that is upstream in the genome (Xu et al. 2007). Currently, we do not know whether *ospC* specific types are better at down regulating the *ospC* gene, but given the high occurrence probabilities for types A and K, and low occurrence probabilities for types H, L, and U, we might predict that types A and K would have a greater propensity to down regulate the *ospC* gene.

Transmission efficiencies can also play a role in disease risk, since a genotype that does not transmit well from a host to a tick would not be able to circulate and be sustained in the host community. Most genotypes were transmitted easily from host to tick, but there were a few genotypes that were better at being transmitted (Figure 3. 5B). For example, *ospC* types T and U are the two most efficiently transmitting genotypes from host to individual ticks in our samples. These genotypes were most commonly detected in the red squirrel, gray squirrel, and chipmunk, all Sciurids. Thus, in terms of *ospC* type U, which has a relatively

low occurrence, but a high transmission efficiency, it may be that type U is a good competitor at leaving the host and entering the ticks, compared to other genotypes that might be present in Sciurids. Another example is *ospC* type H, which was most commonly detected in masked shrews. Although only a few genotypes were detected from masked shrew, *ospC* type H had relatively high transmission efficiency, which will likely maintain this genotype in the tick populations, even with a low occurrence probability in the mammalian host. Recently, (Haven et al. 2012) showed through modeling that there may be a trade-off in rapidly cleared strains to transmit more efficiently. This trade-off would allow the rapidly cleared strain to circulate more often among the vertebrate hosts and tick vectors, compared to persistent strains that can infect the host individual for longer periods of time.

The relatively high proportion of HIS types detected in our host species, and higher host taxa, suggests that these genotypes are capable of existing, and proliferating, across multiple host species. Indeed, host species clustered more together based on the HIS proportions, than they do as a host taxa in the PCA analysis (Figure 3. 2). High infection probabilities by these HIS types can be especially dangerous for people, given that type A, and secondarily, type I, are shown to cause major inflammation, and severe disease onset in humans (Strle et al. 2011). Our study was able to show that avian hosts can play a big role in Lyme disease risk, especially at the genotypic level, with relatively high HIS proportions (40-60%). This is not to say that the white-footed mouse is not an important host, because although HIS types make up only about 45% of the genotypes detected in mice, compared to about 60% in robins, the mice populations are generally denser in the forests than the robins. Our data does highlight that inclusion of birds, in addition to small mammals, can provide a more complete picture on how each species, and the greater host community, might influence

disease risk. In essence, the combination of population size, reservoir competency, and HIS proportions, need to be factored into the role of each host species when determining human Lyme disease risk.

Table 3. 1. Sample size for each host species and the corresponding number of *ospC* types detected in each host species.

Host Species*	BLBR	SOCI	PELE	SCCA	TAHU	TAST	AMRO	VEER	WOTH
# Host Indiv.	10	3	12	4	7	10	13	16	4
ospC Richness	11	4	14	6	13	13	13	16	9

* Host species acronyms:
BLBR (*Blarina brevicauda* – short-tailed shrew),
SOCI (*Sorex cinereus* – masked shrew),
PELE (*Peromyscus leucopus* – white-footed mouse),
SCCA (*Sciurus carolinensis* – eastern gray squirrel),
TAHU (*Tamiasciurus hudsonicus* – red squirrel),
TAST (*Tamius striatus* – eastern chipmunk),
AMRO (*Turdus migratorius* – American robin),
VEER (*Catharus fuscescens* - veery),
WOTH (*Hylocichla mustelina* - wood thrush)

Table 3. 2. Number of host individuals for which on *ospC* type was detected in each species. The row sums represent the total number of *ospC* detected across all individuals of that species. The frequency of a particular *ospC* type for a species, divided by the row sum total of that species, provides the proportion of that *ospC* type, which was used in the Principal Component Analysis in Figure 1.

Species*	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	T	U	Sum
BLBR n=10	6	4	-	2	1	2	1	1	-	-	4	-	1	3	-	-	1	26
SOCI n=3	-	2	-	-	1	-	1	2	-	-	-	-	-	-	-	-	-	6
PELE n=12	9	5	7	4	3	6	7	2	1	-	8	-	1	2	-	1	1	57
SCCA n=4	1	-	-	-	-	3	-	-	-	-	-	-	2	1	-	2	1	10
TAHU n=7	4	-	3	2	1	3	3	2	2	2	1	-	-	1	-	3	4	31
TAST n=10	5	-	1	2	2	3	3	-	4	-	2	2	4	5	-	8	2	43
AMRO n=13	8	4	1	2	3	-	6	1	4	-	7	1	3	6	2	-	-	48
VEER n=16	7	6	4	7	8	6	10	1	5	-	11	1	8	5	6	2	1	88
WOTH n=4	2	-	-	1	1	1	1	-	2	-	2	-	3	-	1	-	-	14
MEME[$] n=1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
Sum	42	22	16	20	20	24	32	9	18	2	35	4	22	23	9	16	10	324

* Host Species acronym: BLBR (short-tailed shrew), SOCI (masked shrew), PELE (white-footed mouse), SCCA (gray squirrel), TAHU (red squirrel), TAST (chipmunk), AMRO (American robin), VEER (veery), WOTH (wood thrush).
[$] MEME (*Mephitis mephitis* – striped skunk) was not included in any analyses due to only having one positive individual for the species.

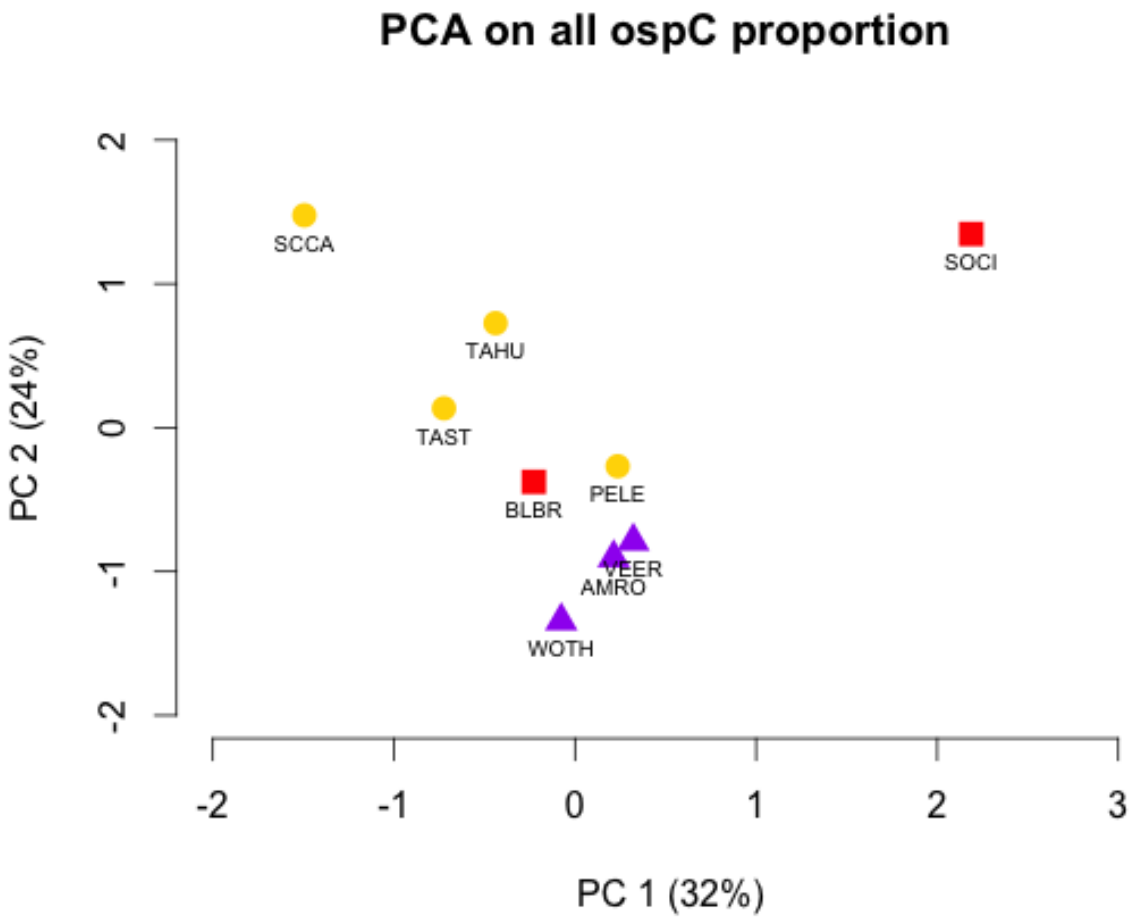


Figure 3. 1. Host species distribution based on standardized factor scores on all *ospC* types from a principal component analysis. Host types are differentiated by symbols: Triangle for birds, squares for shrews, circles for rodents. AMRO = American Robin, VEER = Veery, WOTH = Wood Thrush, BLBR = short-tailed shrew, *SOREX* = masked shrew, PELE = white-footed mouse, SCCA = eastern gray squirrel, TAHU = red squirrel, TAST = eastern chipmunk

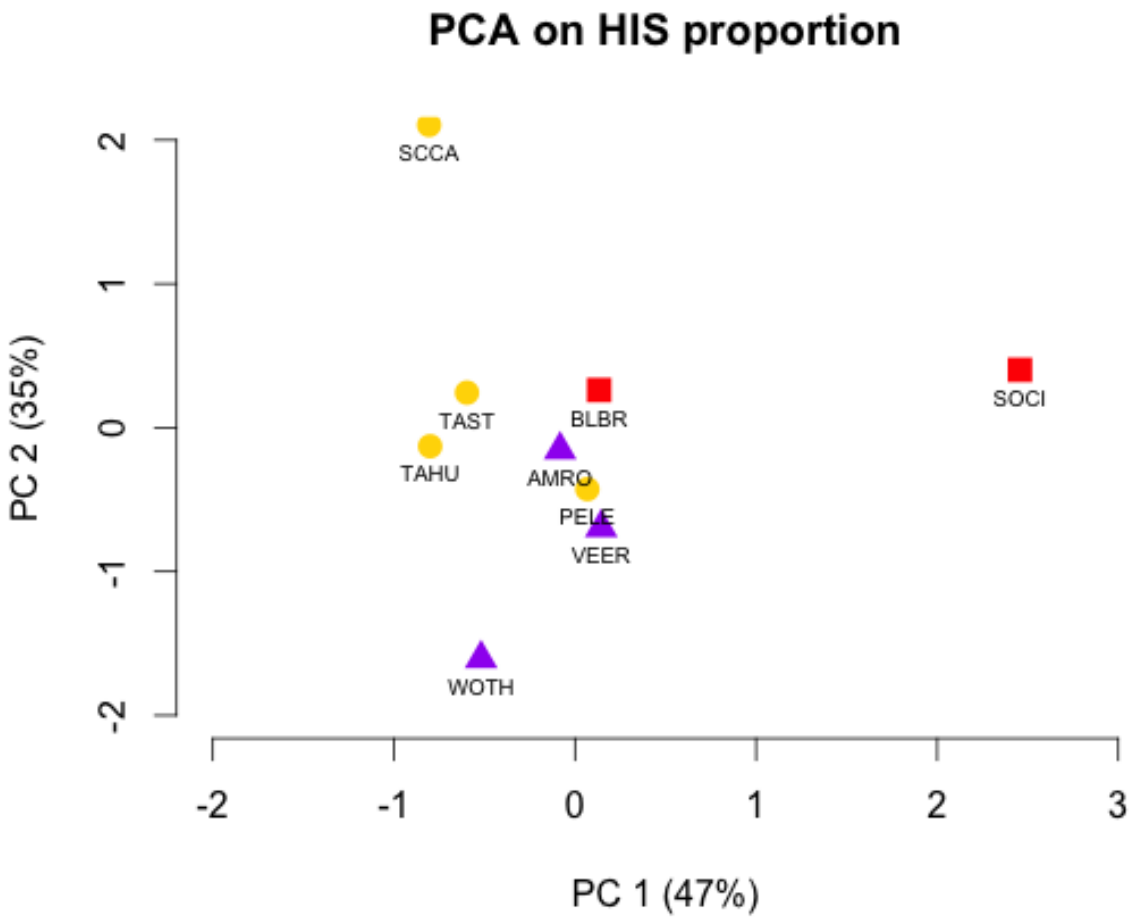


Figure 3. 2. Host species distribution based on the standardized factor scores of only HIS types from a principal component analysis. Higher taxa are differentiated by the following symbols: Triangle for birds, squares for shrews, circles for rodents. AMRO = American Robin, VEER = Veery, WOTH = Wood Thrush, BLBR = short-tailed shrew, SOCI = masked shrew, PELE = white-footed mouse, SCCA = eastern gray squirrel, TAHU = red squirrel, TAST = eastern chipmunk

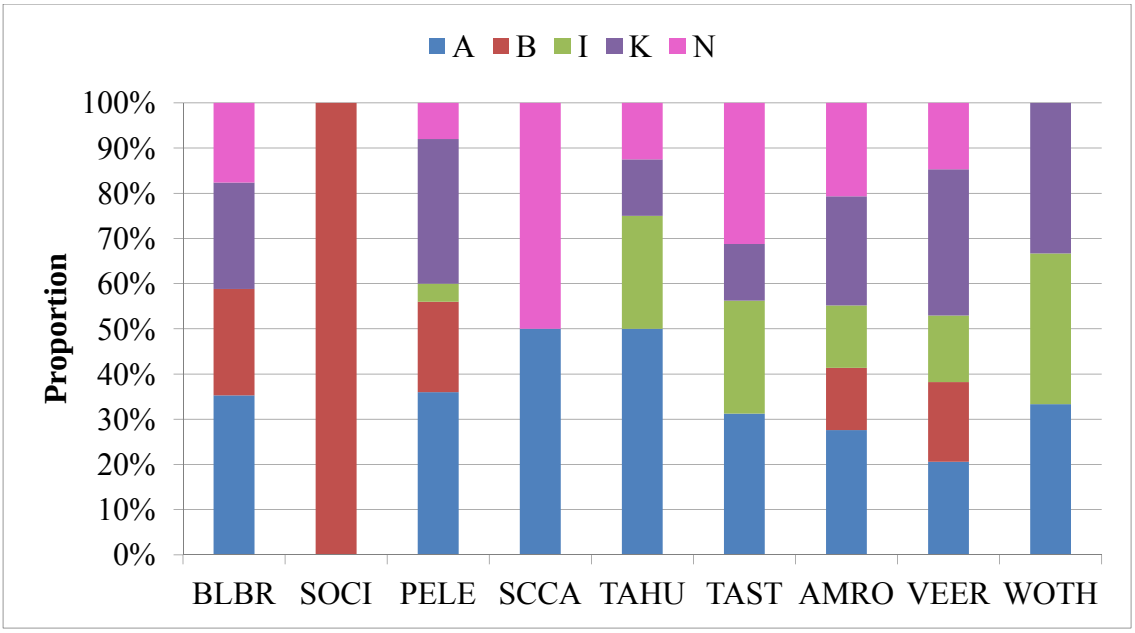
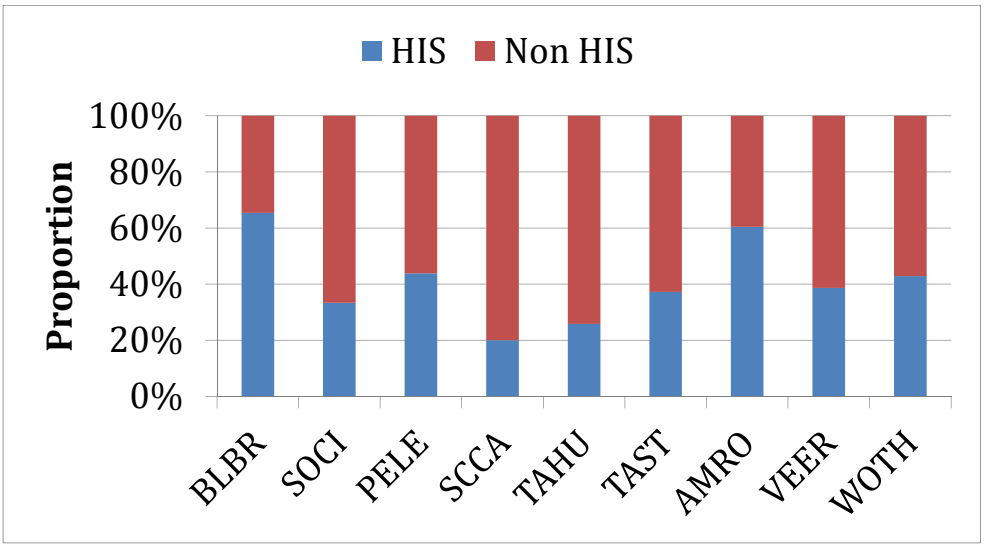


Figure 3. 3. Proportion of HIS types within each host species. Species codes: BLBR = short-tailed shrew, SOCI = masked shrew, PELE = white-footed mouse, SCCA = gray squirrel, TAHU = red squirrel, TAST = eastern chipmunk, AMRO = american robin, VEER = verry, WOTH = wood thrush

A.



B.

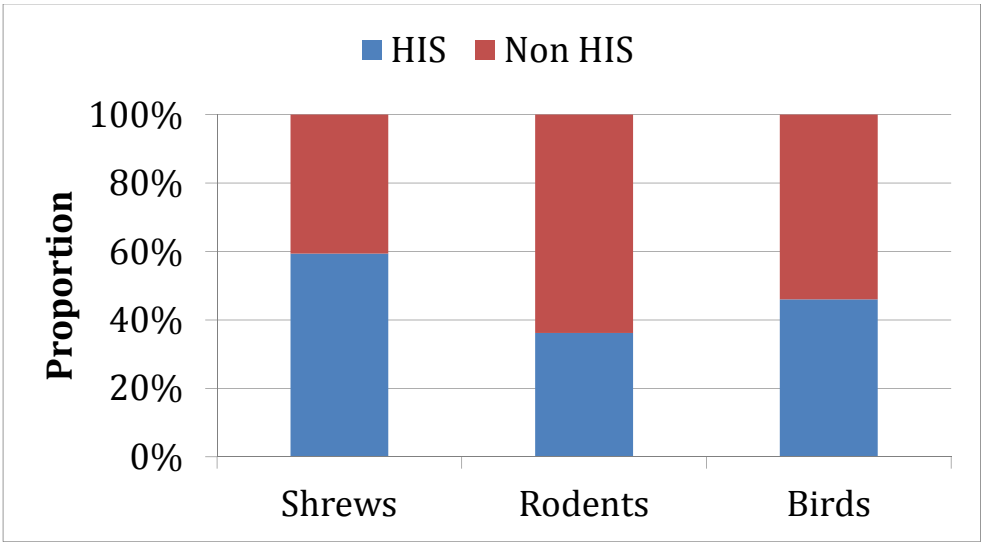
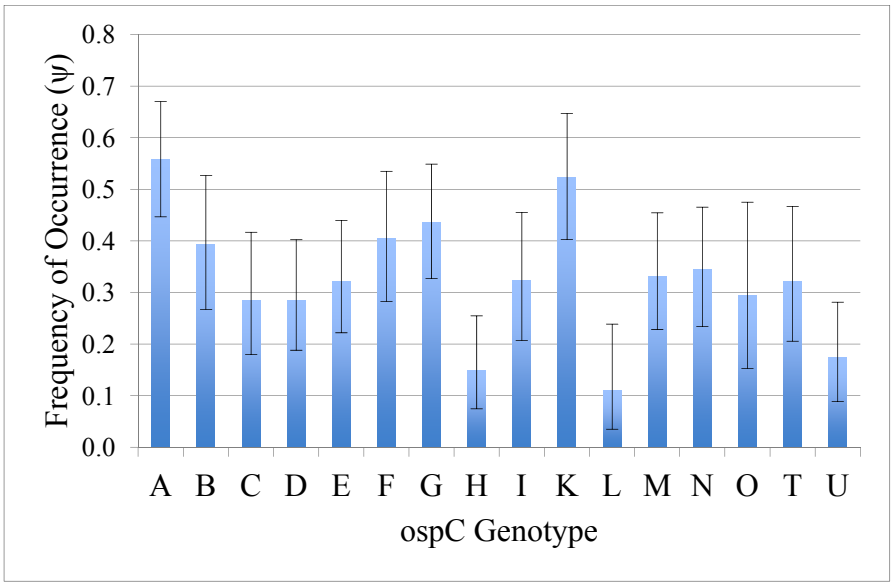


Figure 3. 4. Proportion of HIS to non-HIS types among species (A) and among higher taxa (B). Species codes: BLBR = short-tailed shrew, SOCI = masked shrew, PELE = white-footed mouse, SCCA = gray squirrel, TAHU = red squirrel, TAST = eastern chipmunk, AMRO = american robin, VEER = verry, WOTH = wood thrush

Table 3. 3 Maximum likelihood estimates and AIC_c values from the null and species-specific models. Overall, the null model had lower AICc values, except for *ospC* C and T, compared to the species-specific model.

	Null Model			Species-Specific Model		
ospC type	MLE	No. Parameters	AIC_c	MLE	No. Parameters	AIC_c
A	-229.83	2	463.82	-229.25	9	479.22
B	-122.92	2	250.08	-122.68	6	259.14
C	**-105.97**	**2**	**216.16**	**-100.81**	**6**	**215.27**
D	-126.36	2	256.90	-125.53	8	269.34
E	-149.39	2	302.95	-146.04	9	312.85
F	-142.26	2	288.73	-140.05	8	298.77
G	-191.76	2	387.69	-187.63	9	396.02
H	-61.53	2	127.27	-58.53	7	133.17
I	-114.00	2	232.21	-112.07	7	240.21
K	-196.04	2	396.25	-190.71	8	399.70
L	-28.98	2	62.29	-28.47	4	66.11
M	-138.53	2	281.25	-132.61	8	283.62
N	-137.87	2	279.92	-135.85	8	289.98
O	-56.66	2	117.72	-55.89	4	121.20
T	**-92.47**	**2**	**189.20**	**-83.78**	**6**	**181.56**
U	-65.41	2	135.04	-61.40	7	138.99

A.



B.

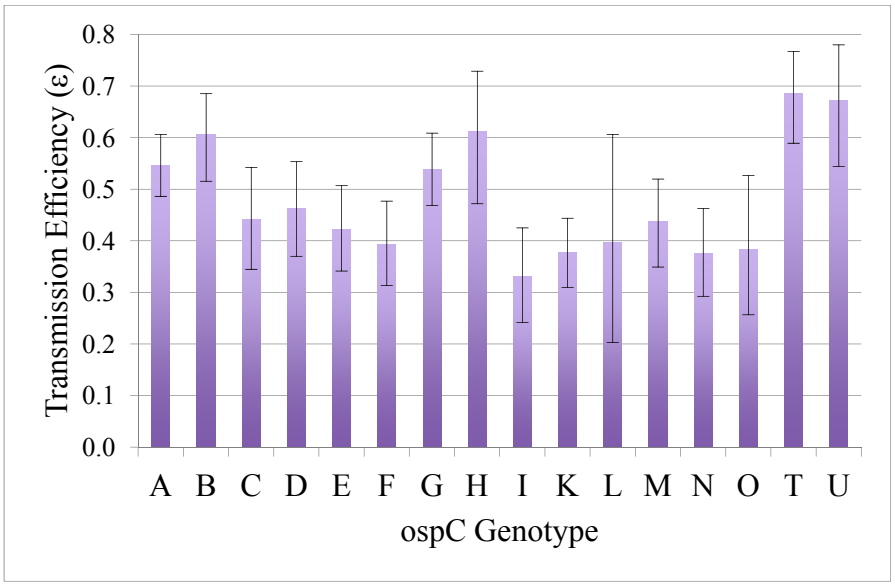


Figure 3. 5. Occurence frequencies (A) and transmission efficiencies (B) of the *ospC* types detected from the null model likelihood analysis. Transmission efficiences estimate the probability that the *ospC* type will be transmitted from an infected host to an individual tick.

Citation

Alghaferi, M. Y., J. M. Anderson, J. Park, P. G. Auwaerter, J. N. Aucott, D. E. Norris, and J. S. Dumler. 2005. *Borrelia burgdorferi ospC* heterogeneity among human and murine isolates from a defined region of northern Maryland and southern Pennsylvania: Lack of correlation with invasive and noninvasive genotypes. Journal of Clinical Microbiology 43:1879-1884.

Anderson, J. M. and D. E. Norris. 2006. Genetic diversity of *Borrelia burgdorferi sensu stricto*, in *Peromyscus leucopus*, the primary reservoir of Lyme disease in a region of endemicity in southern Maryland. Applied and Environmental Microbiology 72:5331-5341.

Battaly, G. R. and D. Fish. 1993. Relative importance of bird species as hosts for immature *Ixodes dammini* (Acari: Ixodidae) in a suburban residential landscape of southern New York State. Journal of Medical Entomology 30:740-747.

Brinkerhoff, R. J., C. M. Folsom-O'Keefe, K. Tsao, and M. A. Diuk-Wasser. 2011. Do birds affect Lyme disease risk? Range expansion of the vector-borne pathogen *Borrelia burgdorferi*. Frontiers in Ecology and the Environment 9:103-110.

Brisson, D. and D. E. Dykhuizen. 2004. *ospC* diversity in *Borrelia burgdorferi*: Different hosts are different niches. Genetics 168:713-722.

Brisson, D., D. E. Dykhuizen, and R. Ostfeld. 2008. Conspicuous impacts of inconspicuous hosts on the Lyme disease epidemic. Proceedings of the Royal Society B: Biological Sciences 275:227-235.

Burgdorfer, W., A. G. Barbour, S. F. Hayes, J. L. Benach, E. Grunwaldt, and J. P. Davis. 1982. Lyme disease--A tick-borne spirochetosis? Science 216:1317-1319.

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York.

Courtney, J. W., L. M. Kostelnik, N. S. Zeidner, and R. F. Massung. 2004. Multiplex real-time PCR for detection of *Anaplasma phagocytophilum* and *Borrelia burgdorferi*. Journal of Clinical Microbiology 42:3164-3168.

Dykhuizen, D. E., D. Brisson, S. Sandigursky, G. P. Wormser, J. Nowakowski, R. B. Nadelman, and I. Schwartz. 2008. The propensity of different *Borrelia burgdorferi sensu stricto* genotypes to cause disseminated infections in humans. Am J Trop Med Hyg 78:806-810.

Eisen, L., R. J. Eisen, and R. S. Lane. 2004. The roles of birds, lizards, and rodents as hosts for the western black-legged tick *Ixodes pacificus*. Journal of Vector Ecology 29:295-308.

Fikrig, E., S. W. Barthold, F. S. Kantor, and R. A. Flavell. 1991. Protection of mice from Lyme borreliosis by oral vaccination with *Escherichia coli* expressing OspA. Journal of Infectious Diseases 164:1224-1227.

Giardina, A. R., K. A. Schmidt, E. M. Schauber, and R. S. Ostfeld. 2000. Modeling the role of songbirds and rodents in the ecology of Lyme disease. Canadian Journal of Zoology 78:2184-2197.

Giery, S. T. and R. S. Ostfeld. 2007. The role of lizards in the ecology of lyme disease in two endemic zones of the northeastern United States. Journal of Parasitology 93:511-517.

Ginsberg, H. S., P. A. Buckley, M. G. Balmforth, E. Zhioua, S. Mitra, and F. G. Buckley. 2005. Reservoir competence of native North American birds for the Lyme disease spirchete, *Borrelia burgdorferi*. Journal of Medical Entomology 42:445-449.

Hanincova, K., K. Kurtenbach, M. Diuk-Wasser, B. Brei, and D. Fish. 2006. Epidemic spread of Lyme borreliosis, northeastern United States. Emerging Infectious Diseases 12:604-611.

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

Hersh, M. H., M. Tibbetts, M. Strauss, R. Ostfeld, and F. Keesing. *In Press*. Quantifying reservoir competence of wildlife host species for *Babesia microti*. Emerging Infectious Diseases.

Keesing, F., J. Brunner, S. Duerr, M. Killilea, K. LoGiudice, K. Schmidt, H. Vuong, and R. S. Ostfeld. 2009. Hosts as ecological traps for the vector of Lyme disease. Proceedings of the Royal Society B-Biological Sciences 276:3911-3919.

Kurtenbach, K., S. M. Schafer, S. de Michelis, S. Etti, and H.-S. Sewell. 2002. *Borrelia burgdorferi sensu lato* in the vertebrate host.*in* J. S. Gray, O. Kahl, R. S. Lane, and G. Stanek, editors. Lyme Borreliosis: Biology, Epidemiology, and Control. CABI International, Oxford, UK.

Lane, R. S., J. Mun, L. Eisen, and R. J. Eisen. 2006. Refractoriness of the western fence lizard (*Sceloperus occidentalis*) to the Lyme disease group spirochete *Borrelia bissettii*. Journal of Parasitology 92:691-696.

Lane, R. S., J. Piesman, and W. Burgdorfer. 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.

Liang, F. T., M. B. Jacobs, L. C. Bowers, and M. T. Philipp. 2002. An immune evasion mechanism for spirochetal persistence in Lyme borreliosis. Journal of Experimental Medicine 195:415-422.

Liang, F. T., J. Yan, M. L. Mbow, S. L. Sviat, R. D. Gilmore, M. Mamula, and E. Fikrig. 2004. *Borrelia burgdorferi* changes its surface antigenic expression in response to host immune responses. Infection and Immunity 72:5759-5767.

LoGiudice, K., R. S. Ostfeld, K. A. Schmidt, and F. Keesing. 2003. The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences of the United States of America 100:567-571.

MacQueen, D. D., C. Lubelczyk, S. P. Elias, B. K. Cahill, A. J. Mathers, E. H. Lacombe, P. W. Rand, and R. P. Smith Jr. 2012. Genotypic diversity of an emergent population of *Borrelia burgdorferi* at a costal Maine island recently colonized by *Ixodes scapularis*. Vector-Borne and Zoonotic Diseases 12:456-461.

Magnarelli, L., J. Anderson, and W. Chappell. 1984. Antibodies to spirochetes in white-tailed deer and prevalence of infected ticks from foci of Lyme disease in Connecticut. Journal of Wildlife Diseases 20:21-26.

Mathers, A., R. P. Smith, B. Cahill, C. Lubelczyk, S. P. Elias, E. Lacombe, S. R. Morris, C. P. Vary, C. E. Parent, and P. W. Rand. 2011. Strain diversity of *Borrelia burgdorferi* in ticks dispersed in North America by migratory birds. Journal of Vector Ecology 36:24-29.

Mitra, S. S., P. A. Buckley, F. G. Buckley, and H. S. Ginsberg. 2010. Highly variable acquisition rates of *Ixodes scapularis* (Acari: Ixodidae) by birds on an atlantic barrier island. Journal of Medical Entomology 47:1019-1027.

Ogden, N. H., L. R. Lindsay, K. Hanincova, I. K. Barker, M. Bigras-Poulin, D. F. Charron, A. Heagy, C. M. Francis, C. J. O'Callaghan, I. Schwartz, and R. A. Thompson. 2008. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. Appl. Environ. Microbiol. 74:1780-1790.

Ohnishi, J., J. Piesman, and A. M. de Silva. 2001. Antigenic and genetic heterogeneity of *Borrelia burgdorferi* populations transmitted by ticks. Proceedings of the National Academy of Sciences 98:670-675.

Qiu, W. G., D. E. Dykhuizen, M. S. Acosta, and B. J. Luft. 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 160:833-849.

Rand, P. W., E. H. Lacombe, R. P. Smith, and J. Ficker. 1998. Participation of birds (Aves) in the emergence of Lyme disease in southern Maine. Journal of Medical Entomology 35:270-276.

Richter, D., A. Spielman, N. Komar, and F. R. Matuschka. 2000. Competence of American robins as reservoir hosts for Lyme disease spirochetes. Emerging Infectious Diseases 6:133-138.

Salkeld, D. J. and R. S. Lane. 2010. Community ecology and disease risk: lizards, squirrels, and the Lyme disease spirochete in California, USA. Ecology 91:293-298.

Seinost, G., D. E. Dykhuizen, R. J. Dattwyler, W. T. Golde, J. J. Dunn, I. N. Wang, G. P. Wormser, M. E. Schriefer, and B. J. Luft. 1999. Four clones of *Borrelia burgdorferi sensu stricto* cause invasive infection in humans. Infection and Immunity 67:3518-3524.

Strle, K., K. L. Jones, E. E. Drouin, X. Li, and A. C. Steere. 2011. *Borrelia burgdorferi* RST1 (OspC Type A) genotype is associated with greater inflammation and more severe Lyme disease. American Journal of Pathology 178:2726-2739.

Swanson, K. I. and D. E. Norris. 2008. Presence of multiple variants of *Borrelia burgdorferi* in the natural reservoir *Peromyscus leucopus* throughout a transmission season. Vector-Borne and Zoonotic Diseases 8:397-405.

Taragel'ova, V., J. Koci, K. Hanincova, K. Kurtenbach, M. Derdakova, N. H. Ogden, I. Literak, E. Kocianova, and M. Labuda. 2008. Blackbirds and song thrushes constitute a key reservoir of *Borrelia garinii*, the causative agent of Borreliosis in Central Europe. Applied and Environmental Microbiology 74:1289-1293.

Wang, I., D. E. Dykhuizen, W. Qui, J. J. Dunn, E. M. Bosler, and B. J. Luft. 1999. Genetic diversity of *ospC* in a local population of *Borrelia burgdorferi sensu stricto*. Genetics 151:15-30.

Xu, Q., K. McShan, and F. T. Liang. 2007. Identification of an opsC operator critical for immune evastion of *Borrelia burgdorferi*. Molecular Microbiology 64:220-231.

Conclusion

The core ecology of Lyme disease encompasses a pathogen, a variety of vertebrate hosts, and a primary tick vector. The interactions among these three entities can be intricate, and ongoing research improves our understanding, allowing us to refute old claims, and to test new hypotheses. In this dissertation, I have shed some light on Lyme disease expansion in New York State, offered new insights on how host biodiversity may increase disease risk, and explored the relative contributions of avian and mammalian hosts to the transmission of *Borrelia burgdorferi*, the Lyme disease bacterium, by examining disease risk at the genotypic level of the pathogen. Using a genotypic approach, I have provided a more refined interpretation of disease risk than infection prevalence or density of infected ticks provide alone.

In the first chapter, I showed that time since invasion of *B. burgdorferi* influences the *ospC* composition and frequency profiles. In newly invaded areas, *ospC* richness, *ospC* diversity, and infection prevalence were lower than in longer occupied, more endemic areas. The spread of Lyme disease is the result of the distribution of ticks and the bacterium being dispersed by vagile host species, such as birds and large mammals, and general population range expansion of vertebrate hosts and ticks. Dispersal has a chance element, which may be why our two Outskirts populations differ in their *ospC* composition and frequencies, more so than the two endemic (and larger) populations of southeastern New York. Interestingly, four of five populations in this study had relatively high frequencies of human invasive strains (HIS), suggesting that such genotypes infect many host species, thus allowing easier transmission of these genotypes to ticks feeding on infected hosts. Although newly invaded areas seem to have lower tick abundance, the

possibility that disease may arise from an HIS type would be elevated. As Lyme disease continues to spread, it would be useful to determine the temporal lag before the "advancing front" will reach similar *ospC* profiles to endemic areas.

The second chapter focused on whether host community metrics (diversity, and the relative abundances of white-footed mouse, chipmunk, and short-tailed shrew) were good predictors of *ospC* diversity and of HIS prevalence. I found that host composition and diversity can matter, but the contrasting results between years show that annual, and/or site, variation can lead to different significant predictors. Also, the direction of associations between the community metrics and *ospC* diversity and HIS prevalence were opposite to those expected. There was a positive relationship between host diversity and *ospC* diversity, but a negative relationship between relative abundances of mice and *ospC* diversity. Increased relative abundances of chipmunks and mice were also negatively related to HIS infection. These results imply that we are missing important host species in our host diversity estimates, if we seek to examine disease risk at the genotypic level. Birds, voles, and other large mammals, may be helping to maintain high *ospC* diversity in tick populations.

The object of chapter three was to understand the host-genotype associations for a large number of species, including birds, determine whether host species or higher taxa (birds, shrews, and rodents) differ in their *ospC* associations, and to examine *ospC* occurrences and transmission efficiencies from host to tick. Multiple niche polymorphism is likely maintaining the balancing selection on *B. burgdorferi*, and leading to high genotypic diversity. Indeed, I found differences in the relative frequencies of each *ospC* type within a species, supporting the concept that hosts act as ecological niches to *ospC*

genotypes. Birds, shrews, and rodents were more different from one another when all *ospC* types were examined in the PCA, but there was greater mixing of host types in a cluster when only HIS types were examined. This suggests that HIS types are good at infecting, and possibly evading, most or all host species, and that birds should also be considered important hosts in the role of Lyme disease risk. Using a likelihood approach and adopting a species detection model, I was able to show that *ospC* occurrences differ across the host individuals, with some genotypes occurring much more often than other genotypes. Additionally, high occurrence does not always lead to high transmission, nor does low occurrence always lead to low transmission. The trade-offs between occurrence and transmission efficiencies may be why rare genotypes can continue to circulate within the host community and tick populations.

I chose to examine Lyme disease risk at the genotypic level, because risk is not equal across all genotypes. Hence, it is important to understand the dynamics of these genotypes to better comprehend disease risk at the finer scale. Indeed my studies shed light on disease risk at the large spatial scale across New York State, at the host community level, and down to the role that individual host species play in contributing these *ospC* genotypes. However, there remains much work to be done. We still need to determine how larger mammals and mesopredators contribute to *ospC* richness and frequency profiles; whether and how HIS genotypes infect, and possibly evade, the immune system better than do non-HIS genotypes; and whether there is competition among genotypes within hosts or within the ticks that would affect transmission efficiencies. These are just a few of the directions that I hope will be addressed in future Lyme disease research.