5-18-2017 12:00 AM

Evolutionary genetic aspects of host association in generalist ectoparasites

Benoit Talbot
The University of Western Ontario

Supervisor
Nusha Keyghobadi
The University of Western Ontario Joint Supervisor
Brock Fenton
The University of Western Ontario

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

Despite the use of the host for dispersal by most parasite species, the extremely loose relationship typical between highly mobile hosts and generalist ectoparasites may lead to very different gene flow patterns between the two, leading in turn to different spatial genetic structure, and potentially different demographic history. I examined how similar gene flow patterns are between *Cimex adjunctus*, a generalist ectoparasite of bats present throughout North America, and two of its key bat hosts. I first analyzed the continent-scale genetic structure and demographic history of *C. adjunctus* and compared it to that of two of its hosts, the little brown myotis and the big brown bat, using microsatellite and mitochondrial data. Second, I compared spatial genetic structure of *C. adjunctus* with *Cimex lectularius*, or common bed bug, which associates with a broader range of host species. Third, I compared the effect of land cover on spatial genetic structure of *C. adjunctus* and of the big brown bat in the Great Lakes region. My results support the emerging hypothesis that generalist ectoparasites and their highly mobile hosts display weak, but positive, correlation in spatial genetic structure and demographic history.

Generalist parasites associate with different hosts, which are, in some cases, evolutionarily divergent from each other. In such cases, it is not clear how hosts may affect adaptive genetic variation in the parasites. In the *Cimex* genus, parasite species associate with a range of hosts, including bats, humans, and swallows. I examined how hosts affect adaptive genetic variation in these generalist ectoparasites. I analyzed variation at two salivary protein genes, one

coding for an apyrase and the other for a nitrophorin, in 10 species of *Cimex*. These proteins affect the way parasites feed on their hosts, by preventing clotting and vasoconstriction, and may experience selection depending on host ecology or physiology. I also analyzed allelic divergence at the same two genes in a single species, *C. adjunctus,* associated with several bat species in North America. My results suggest selection and adaptation to the host at genes coding for salivary proteins of blood-feeding ectoparasites across the *Cimex* genus, and also within *C. adjunctus*.

Keywords

Chiroptera, *Cimex adjunctus*, *Cimex brevis*, *Cimex hemipterus*, *Cimex hirundinis*, *Cimex japonicus*, *Cimex latipennis*, *Cimex lectularius*, *Cimex pipistrelli*, *Cimex sp.*, *Cimex vicarius*, Cimicidae, Dispersal, Ectoparasite, *Eptesicus fuscus*, Evolutionary genetics, Gene flow, Generalism, Genetic clustering, Genetic differentiation, Genetic diversity, Genetic drift, Genetic structure, Geographic information systems, Hematophagy, Hemostasis, Host association, Isolation-by-distance, Landscape genetics, Molecular ecology, *Myotis lucifugus*, *Myotis septentrionalis*, Natural selection, Phylogenetics, Phylogeography.

Co-Authorship Statement

All chapters are co-authored with Maarten J. Vonhof, Hugh G. Broders, Brock Fenton and Nusha Keyghobadi. For all chapters, Dr. Keyghobadi coordinated the study, supervised the collection and interpretation of genetic data and revised the writing. Dr. Fenton contributed to the collection of data, coordinated the study, supervised the interpretation of data and revised the writing. Dr. Vonhof and Dr. Broders contributed with most of the sample collection, and helped in the interpretation of data and revision of the writing.

Chapter 5 is co-authored with Ondřej Balvín, in addition to Maarten J. Vonhof, Hugh G. Broders, Brock Fenton and Nusha Keyghobadi. Ondřej Balvín contributed with most of the sample collection for that chapter, and helped in the interpretation of data and revision of the writing.

A version of chapter 2 was published in BMC Evolutionary Biology, a version of chapter 3 was published in Canadian Journal of Zoology, and a version of chapter 4 was published in Ecology and Evolution, all with Maarten J. Vonhof, Hugh G. Broders, Brock Fenton and Nusha Keyghobadi as co-authors. A version of chapter 5 was published in Royal Society Open Science, with Ondřej Balvín, Maarten J. Vonhof, Hugh G. Broders, Brock Fenton and Nusha Keyghobadi as co-authors.

Acknowledgements

First and foremost I thank Nusha Keyghobadi and Brock Fenton for agreeing to undertake this project with me, which was nowhere near their primary research goals. You both spent some of your primary funding to fund this project, and learned about these creatures with me. Your open-mindedness helped me think clearly about the project, your laid-back attitude let me work at my own pace, and your brilliant ideas made me think twice many times. If I do become a researcher one day, you will be examples for me.

I thank Maarten J. Vonhof, Hugh G. Broders and Ondřej Balvín for their continuous support and input into my project. They all offered a large part of the sampling, without which no chapter would have been possible, and then went out of their way to better my ideas and analyses. I thank my two advisors: André Lachance and Graham Thompson. You made me further think through my analyses and results, and most importantly, see the general picture of my whole project and how it fits in the field of biology.

I thank all lab members past and present for their support: Helen Chen, Jenna Siu, Lindsay Crawford, John O'Leary, Maryam Jangjoo, Meghan Murphy, Melissa Lucas and Toby Thorne.

I thank all field assistants who offered their help in collecting samples: Alejandra Ceballos, Benedikt Ségur-Cabanac, Cal Butchkoscki, Charles Brown, Christy Humphrey, David Elliot, Derek Morningstar, Dylan Baloun, Elizabeth Warburton, Jaroslav Cepák, Jenna Siu, Jitka Vilímová, John Whitaker, Jussi Tuomi, Ken Luzynski, Kristin Jonasson, Laura Kaupas, Lee Johnson, Lesley Hale,

Lisa Winhold, Lucas Greville, Lynne Burns, Mark Brigham, Martin Pokorný, Milan Řezáč, Mitsuru Mukohyama, Nobuhiko Kataoka, Paul Faure, Petr Šrámek, Quinn Webber, Rob Valdizon, Robert Barclay, Roger Pearce, Sara McCarthy, Shelby Bohn, Stefano Boscolo, Stephanie Erickson, Toby Thorne, Tomoya Kobayashi, Tony Parr, Vylém Helešic, and all technicians from Abell Pest Control Inc. who contributed in collecting samples.

Funding for my doctoral degree came from the Natural Science and Engineering Council of Canada (NSERC), the Ontario Student Assistance Program (OSAP), Western University, and the Schmeelk Canada Foundation. Funding for my project came from NSERC Discovery Grants separately awarded to Nusha Keyghobadi, Brock Fenton and Hugh G. Broders, from Western Michigan University, which funded Maarten J. Vonhof, and from Ministry of Education, Youth and Sports of the Czech Republic and Charles University, which funded Ondřej Balvín.

Finally, I thank my parents, France Ouellet and Pierre Talbot, my aunt, Josée Talbot, my brother, Dominic Talbot, my sister, Natalie Talbot, and my closest friend and non-biological brother, Mathieu Bourret, for their support and company through this project. I thank my grandfather, Arthur Talbot, who is still here with me and helps me in every step of the way. Finally, I thank Rosa Del Giudice and Tisha. Rosa helped me through ordeals, brought my best qualities to maturity, and overall made me become a better person. Tisha has been an excellent companion and roommate, and I’ve strived to give her the happiest life in return.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Symbols and Abbreviations

±	More or less
#	Number
%	Percent
[Primer]	Primer concentration
1X	Specified reagent concentration in final solution
α	Significance threshold
ΔAICc	Disparity in corrected Akaike Information Criterion
τ	Generation time
µL	Microliter
µM	Micromolar
A	Adenine
ABC	Approximate Bayesian Computations
ABI	Applied Biosystems
ADP	Adenosine diphosphate
AFLP	Amplified Fragment Length Polymorphism
AIC	Akaike Information Criterion
AICc	Corrected Akaike Information Criterion
AMOVA	Analysis of Molecular Variance
ATP	Adenosine triphosphate
bp	Base pair
BIC	Bayesian Information Criterion
C	Cytosine
CO1	Cytochrome c Oxidase 1
CS1	Common Sequence 1
CS2	Common Sequence 2
DDT	Dichlorodiphenyltrichloroethane
df	Degrees of freedom
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate

EBF	Empirical Bayes Factor
EBSP	Extended Bayesian Skyline Plot
EF1α	Elongation Factor 1α
EPFU	*Eptesicus fuscus*
ESRI	Environmental Systems Research Institutes
FEL	Fixed-Effects Likelihood
F	Measure of significance in a model
F_{ST}	Differentiation index
G'_{ST}	Differentiation index
G_{IS}	Inbreeding coefficient
GPS	Geographic Positioning System
HKY	Hasegawa, Kishino and Yano
H_O	Observed heterozygosity
H_S	Expected heterozygosity
IBD	Isolation-by-Distance
K	Number of genetic clusters
km	Kilometer
Ma	Million years
MCMC	Markov chain Monte Carlo
MD	Mismatch Distribution
MEME	Mixed-Effects Model of Evolution
$MgCl_2$	Magnesium chloride
mL	Milliliter
mM	Millimolar
MRM	Multiple Regression on Distance Matrices
MSN	Minimum-Spanning Network
Mya	Million years ago
MYLU	*Myotis lucifugus*
MYSE	*Myotis septentrionalis*
N	Nucleobase
N/A	Non Applicable

N_1	Initial effective population size
N_A	Effective population size after a decline
N_{AL}	Number of alleles
N_B	Effective population size after an expansion
N_E	Effective population size
P	Probability to obtain a result equal or more extreme than observed
PARRIS	Partitioning Approach for Robust Inference of Selection
PCR	Polymerase chain reaction
R^2	Coefficient of determination of a statistic model
REL	Random-Effects Likelihood
RCSB	Research Collaboratory for Structural Bioinformatics
r_W	Relatedness coefficient
SBP	Single Breakpoint Recombination
SD	Standard deviation
SDSC	San Diego Supercomputer Center
SE	Standard error
SNP	Single nucleotide polymorphism
SSCS	Single-Strand Consensus Sequence
t	Time (in years)
T°	Temperature
Taq	*Thermophilus aquaticus*
U	Unit
UCSD	University of California San Diego
USA	United States of America
UV	Ultraviolet
WGS84	World Geodetic System 1984
ZIP	Zone Improvement Plan

List of Software Packages

*BEAST 2.4.2	http://www.beast2.org
ArcGIS 10.3	https://www.arcgis.com/features/index.html
Arlequin 3.5	http://cmpg.unibe.ch/software/arlequin35/
BEAST 1.8.4	http://beast.bio.ed.ac.uk
DIYABC 2.1.0	http://www1.montpellier.inra.fr/CBGP/diyabc/index.php
DNASP 5.1	http://www.ub.edu/dnasp/
Galaxy Tools	https://usegalaxy.org
Geneland 4.0.5	http://www2.imm.dtu.dk/~gigu/Geneland/
GeneMapper 4.0	Applied Biosystems (Forest City, CA)
Genepop 4.2	http://genepop.curtin.edu.au
GenoDive 2.0	http://www.patrickmeirmans.com/software/GenoDive.html
HyPhy 2.2.1	http://www.hyphy.org/w/index.php/Main_Page
MEGA 6.06	http://www.megasoftware.net
ML-Relate	http://www.montana.edu/kalinowski/Software/MLRelate.htm
R 3.x	https://www.r-project.org
SpaGeDi 1.5	http://ebe.ulb.ac.be/ebe/SPAGeDi.html
TCS 1.21	http://darwin.uvigo.es/software/tcs.html

1 Introduction

1.1 Evolutionary processes and molecular markers

Genetic variation is the most fundamental source of variation among living organisms. Genetic variation is ultimately the product of four key evolutionary processes: mutation, gene flow, genetic drift, and selection (Fox and Wolf 2006).

1.1.1 Neutral Genetic Variation

Neutral evolutionary processes are those that are not affected by the fitness of individuals. The three neutral evolutionary processes are mutation, gene flow, and genetic drift.

Mutations occur randomly and are the ultimate source of genetic variation in all species (Fox and Wolf 2006; Loewe and Hill 2010). Once mutations occur, they are subject to the other evolutionary processes, both neutral and adaptive. Thus, a new mutation may become fixed in populations if, through the action of either genetic drift or selection, it replaces the ancestral variant. The replacement of the nucleotide is named a substitution in those cases. Over periods of time on the scale of millions of years, substitutions occur regularly, and can be predicted using various models of DNA evolution (Felsenstein 1981; Hasegawa et al. 1985; Tamura and Nei 1993).

Molecular clocks can be calibrated using the number of substitutions that have occurred between lineages of a species, or between species, along with an independent estimate (e.g., from the fossil record) of the amount of time that has passed since their divergence from a common ancestor. Molecular

clocks can then be applied to lineages or species for which independent information on divergence times are not available, and are thus useful tools for quantifying the timing of past evolutionary events (Brower 1994; Ho and Shapiro 2011; Burns et al. 2014).

Gene flow and genetic drift are two processes that act to structure genetic variation within and among populations (Hutchison and Templeton 1999). Gene flow occurs when movement by individuals is followed by reproduction, and acts to homogenize allele frequencies between populations (Bohonak 1999; Fox and Wolf 2006). High gene flow usually means populations are genetically similar (Blouin et al. 1995; Hanski 1998; Bohonak 1999; Castillo et al. 2014; Ripperger et al. 2014). Genetic drift is the random change in allele frequency that results from the random sampling of alleles at each generation, through unequal reproduction among members of a population. Genetic drift has an opposite effect to gene flow, leading to differentiation of allele frequencies among populations (Levy and Neal 1999; Gandon and Nuismer 2009). The effective population size determines the rate of drift in a population, where populations of smaller effective size experience stronger drift (Fox and Wolf 2006). Several factors, including demographic events, such as bottlenecks or founding events, and sex ratio can affect the effective population size, and in turn affect the extent of genetic drift in a population (Fox and Wolf 2006).

The relative strength of gene flow and genetic drift are often studied through genetic structure, which is the organization of genetic variation among populations (Bohonak 1999; Hutchison and Templeton 1999). Gene flow and genetic drift reduce and increase, respectively, the levels or genetic structure in a

species. High levels of genetic structure in a species means populations tend to be divergent, whereas low levels of genetic structure means populations are genetically similar (Barnes et al. 2007; Barrett et al. 2008; Talbot et al. 2012; Rioux Paquette et al. 2014; Anderson et al. 2015).

Various forms of heritable DNA variation, referred to as different types of molecular markers, are used to characterize genetic structure and study evolutionary processes. If alleles at these genetic loci do not affect individual fitness, then they will be primarily influenced by drift and gene flow, and are considered 'neutral' loci. Although they are not truly neutral, genes from the mitochondrial genome in animals are often used to make phylogenetic and phylogeographic inferences, including inferences about neutral processes (Harpending 1994; Moore 1995; Schneider and Excoffier 1999; Balvín et al. 2012; Bar Yaacov et al. 2012). The mitochondrial genome is usually relatively variable and is not affected by recombination, which makes it useful for drawing evolutionary inferences. Genetic structure is also often studied using microsatellite loci (Pritchard et al. 2000; Guillot et al. 2005; Andreakis et al. 2009; Spice et al. 2012; Talbot et al. 2014), the majority of which are expected to be neutral. These markers are easy to use and can also be very variable, which gives them high resolution to uncover genetic structure.

1.1.2 Adaptive Genetic Variation

Selection is an adaptive evolutionary process and is driven by differences in heritable fitness among individuals. Adaptation occurs as a response to environmental, ecological or physiological selection pressures (Fox

and Wolf 2006; Forester et al. 2016). Adaptation is a major factor shaping the degree of genetic, physiological and morphological variation in animals (Hancock et al. 2011; Amato et al. 2011). Positive selection results in greater variation or divergence among species or lineages of a species than expected under only neutral evolutionary processes (mutation, gene flow and genetic drift). It is a significant cause of adaptive radiation at both short timescales (Givnish et al. 2009; Takahashi and Koblmüller 2011) and long, geological timescales (Goldberg et al. 2008; Moen and Morlon 2014; Pincheira-Donoso et al. 2015). Negative selection leads to reduced variation among species or lineages of a species than expected under only neutral evolutionary processes (mutation, gene flow and genetic drift). It is a significant cause for a large variety of conserved genome regions and phenotypic traits (Hermisson et al. 2003; Neff 2004; Amato et al. 2011).

We have known for some time that the degree of adaptation can be affected by genetic drift and gene flow, in addition to selection (Garant et al. 2007; Gandon and Nuismer 2009). Selection is also often not stable and fluctuates through time and space (Pélabon et al. 2010; Ketola et al. 2013). It is thus important to account for these factors when drawing inferences about adaptation.

Adaptation at the molecular level can be studied using two different approaches. The first approach focuses on genes that are known or suspected to code for a trait of interest. In this case, studying factors influencing genetic variation at those candidate genes reveals potential selective pressures acting on those specific genes (Shaw et al. 2007; Mahamdallie and Ready 2012; Dunning

et al. 2013; Du et al. 2015). The second approach is based on assessment of genome-wide variation. This approach can be applied when little information is available about the genes that code for the trait of interest, or if one is interested in uncovering additional genes that affect the trait (Beaumont and Balding 2004; Baute et al. 2015; Talbot et al. 2016). Over the years, various molecular markers have been used to study genome-wide variation, from amplified fragment length polymorphism (AFLP) and derived methods (Meudt and Clarke 2007; Bensch et al. 2008; Fischer et al. 2011), to single nucleotide polymorphisms (SNPs; Davey and Blaxter 2010; Narum and Hess 2011; Schmitt et al. 2012), to whole-genome sequencing (Zimin et al. 2014; International Glossina Genome Initiative et al. 2014) or sequencing of most of the genome (Adelman et al. 2011). These methods allow one to compare genomic variation with environmental or phenotypic variation to assess which factors have potentially driven adaptation over time (Coop et al. 2010; Eckert et al. 2010; Hancock et al. 2011; Frichot et al. 2013; Talbot et al. 2016).

1.2 Types of parasitism

A symbiosis is defined as a physiologically intimate relationship between two species (Fox et al. 2001). Symbioses are widespread across the biological world, and may range from extreme parasitism, characterized by antagonistic interactions, to complete mutualism, characterized by mutually beneficial relationships.

In parasitic symbioses, the parasite benefits from the association while the host is harmed. Parasites are important factors shaping neutral and adaptive

genetic variation of their hosts, and are thus important drivers of organismal, population and community diversity. Using counter strategies, hosts in turn have an effect on genetic variation of their parasites. At least three characteristics can be used to describe the diversity of parasites: their location in or on the body of the host, the proportion of time associated directly with the host, and the number of host species with which they associate.

1.2.1 Ecto- and endoparasitism

Ectoparasites cling to the skin of their host to eat, copulate, or reproduce (Allen 1994; Morand et al. 2006). Most ectoparasites are arthropods, such as mites, ticks and insects (Marshall 1982; McCoy et al. 2002; Christe et al. 2003; Dick and Patterson 2007), but some fish (Hess et al. 2013), plants (Westwood et al. 2012), fungi (Klironomos 2003) and even mammals (Fenton 2001) can also be classified as ectoparasites. Many notorious vectors of pathogens in mammals are ectoparasites, such as dipteran and hemipteran insects (Ewald 1983; Bargues et al. 2006; Lehmann et al. 2009; Olival et al. 2013; International Glossina Genome Initiative et al. 2014). Endoparasites in contrast spend nearly all, if not all, of their life cycle inside the body of a host (Morand et al. 2006). Viruses are endoparasites by definition, needing the internal cell machinery to operate (Davis et al. 2005; Abrams et al. 2013; Kuzmina et al. 2013). Many bacteria will form a mutualistic symbiosis with their host (Schluter and Foster 2012). However, some of these bacteria can opportunistically become parasitic in certain conditions, leading them to become pathogens (Pinheiro et al. 2013). Some endoparasites are Eukaryote taxa, for

example parasitic nematodes (Blouin et al. 1995; Nieberding et al. 2004, 2008) and parasitic alveolates (Allison 2002; Duval et al. 2007; Garamszegi 2009; Schaer et al. 2013). Many notorious diseases in mammals are caused by endoparasites, such as viruses, bacteria and protozoans (Fenner et al. 2013; Morens and Fauci 2013; Pinheiro et al. 2013; Jones 2014).

In mammals, collecting endoparasites is typically easier than collecting ectoparasites, because while endoparasites can usually be collected from internal tissues of their host, collection of ectoparasites from the body of their hosts is opportunistic at best. Partly for that reason, endoparasites have received more attention in the scientific literature. Additional knowledge of ectoparasites is however important because of the ecological, evolutionary and socio-economic roles they play.

1.2.2 Permanent and temporary parasitism

Permanent parasites spend all of their life cycle on or in the body of their host, and are unable to survive when not associated with the host (Balashov 2006, 2011). Temporary parasites spend only part of their life cycle associated with their host (Balashov 2006, 2011). Most ectoparasites of mammals are temporary, being associated with their host only to feed on their blood (Morand et al. 2006). Endoparasites that have different intermediate and definitive hosts, for example some marine acanthocephalans (Goulding and Cohen 2014) and some trematodes (Lively et al. 2004), also are temporary parasites, because they spend only part of their whole life cycle on or in each host (Morand et al. 2006).

1.2.3 Generalist and specialized parasitism

Parasite species that are capable of using two or more host species that are closely related to each other are named weak generalists (Mazé-Guilmo et al. 2016). Parasite species that are capable of using two or more host species that are phylogenetically very different from each other are named strong generalists (Mazé-Guilmo et al. 2016). An example of a strong generalist parasite is the common bed bug, which uses bats, humans and domestic animals, such as chicken, pigeons, rabbits, etc. (Balvín et al. 2012; Booth et al. 2015). On the other hand, specialized parasite species typically are associated with a single host species at any one stage of their life cycle, such as some primate pinworms, some bat flies and some bat rabies strains (Brooks and Glen 1982; Dick and Patterson 2007; Streicker et al. 2010). In those cases, interestingly, we often see a resemblance between the phylogeny of the parasite taxa and the phylogeny of the host taxa, due to co-speciation (Brooks and Glen 1982; Streicker et al. 2010). Temporary parasites, which represent most ectoparasites of mammals, are more likely to be generalists, whereas permanent parasites, which represent most endoparasites of mammals, are more likely to be highly specialized (Sponchiado et al. 2015).

Much research is regularly performed on specialized permanent endoparasites of mammals, while generalist temporary ectoparasites receive considerably less attention. This is surprising, as the latter represent a large and diverse group of organisms. This thesis will focus on a genus of generalist temporary ectoparasites of birds and mammals, the cimicid insects of *Cimex*, because this genus is highly representative of the group as a whole, is relatively

easy to study, and most species may be identified through morphology and genetics.

1.3 Main concepts of dissertation

1.3.1 Use of hosts for movement by generalist ectoparasites

We generally expect parasites and their hosts to show similarity in gene flow and dispersal patterns, due to their potentially strong association. For example, highly specialized permanent endoparasites rely heavily on their hosts for long-distance dispersal, because such parasites typically do not leave the body of their host, sometimes for many generations (Balashov 2011). Less specialized or tightly host-associated parasites also typically rely on their hosts for long-distance dispersal, although to a lesser extent (Mazé-Guilmo et al. 2016). For that reason, we may use patterns of neutral genetic variation of the parasite to gain knowledge about the biology of the host, particularly in cases where the host's genetic diversity is comparatively very low. For example, parasites often offer higher genetic resolution to study continent-scale biogeographical patterns (Nieberding et al. 2004) and contemporary genetic structure (Catalano et al. 2014) of the host. However, it has recently been highlighted that several factors can dissociate gene flow patterns of parasites and their hosts, for example the proportion of free-living life cycle stages, the degree of host-generalism of the parasite, and the mobility of the host (Mazé-Guilmo et al. 2016). For this reason, generalist temporary ectoparasites and highly mobile hosts may display weak or no correlation in their gene flow patterns.

Generalist temporary ectoparasites sometimes disperse on their own, as is seen in parasitic flies (Levin and Parker 2013). In those cases, correlation in gene flow patterns between the parasite and the host would be close to absent (Levin and Parker 2013). However, many other generalist temporary ectoparasites, such as ticks, mites and insects of the Cimicidae family, have poor mobility and can disperse only short distances on their own. Therefore, the greater part of their dispersal potential lies in the use of their hosts. Usinger (1966) hypothesized that cimicid species associated with swallows typically use their hosts to disperse between nests. In this dissertation, I investigated the extent to which a loose parasite-host relationship and high host mobility may drive incongruence of neutral genetic patterns between a parasite and a host. To this end, I used a genus entirely composed of generalist ectoparasites, in which most species parasitize hosts that use one of the most powerful forms of mobility, flight. First, I investigated how similar neutral genetic patterns are between a cimicid ectoparasite species and two of its host species, considering the ectoparasite most likely uses its hosts for the larger part of its long-distance dispersal. Following this, I determined if a weak generalist species, which associates with highly mobile bat hosts, shows different neutral genetic patterns than a strong generalist species, which associates with a broader range of host species. Finally, I determined if landscape composition influences neutral genetic variation similarly in a generalist temporary ectoparasite species and in one of its key hosts.

1.3.2 Adaptation of generalist ectoparasites to their hosts

We generally expect specialized parasites to be highly adapted to their hosts, as a result of evolutionary arms races and co-adaptation (Hanifin et al. 2008; Hall et al. 2011; Jacobs et al. 2014). One of the most extreme examples is between bacteriophages and their bacterial hosts, where phenotypic shifts resulting from association to a local host species can appear within the phage in only a few generations (Hall et al. 2011; Leggett et al. 2013). However, it is not clear how host-driven selective pressures operate in generalist ectoparasites, which frequently go off the host and often associate with a range of other host species. For example, gene flow from populations using alternate hosts may impede adaptation to any given host species (Garant et al. 2007).

Despite the loose relationship generalist ectoparasites may have with their hosts, we sometimes observe narrow host ranges of such parasite species, such as in bat flies (Dick and Patterson 2007). Also, despite the possibility of gene flow among them, populations of generalist ectoparasite species associated with different hosts sometimes show neutral genetic divergence. For example, both the seabird tick and the common bed bug show divergence between populations associated with different host species (Kempf et al. 2009; Booth et al. 2015). Furthermore, individuals of generalist ectoparasite species sometimes even display higher fitness in some potential host species than others. For example, individuals of the seabird tick associated with different host species were shown to feed for longer on their preferred host species than on other potential host species (McCoy et al. 2002). In this dissertation, I investigated if individuals across a whole genus of blood-feeding ectoparasites, associating with

one of a few phylogenetically very different hosts, show molecular adaptation to the host with which they associate. I analyzed candidate loci coding for proteins that are useful in blood feeding, because these proteins are likely an important driver of fitness for these species. Following this, using the same candidate loci in one species of the genus that associates with phylogenetically similar but ecologically different hosts. I examined whether parasite populations show adaptation to the host species from which they were sampled.

1.4 Dissertation structure

My thesis consists of five data chapters that were designed as separate studies for independent publication. Chapter 2-5 are published, and Chapter 6 will soon be submitted.

Chapters 2-4 primarily pertain to the use of hosts for movement and gene flow by generalist temporary ectoparasites. In those chapters, I use *Cimex adjunctus* as a model species representing weak generalist ectoparasites, and compare its neutral genetic variation to that of two of its hosts and to that of a closely related species, the more well-known and strong generalist *Cimex lectularius*, or common bed bug.

In Chapter 2, I study *C. adjunctus* throughout most of its range in North America, and I use tools and approaches to analyze spatial (*i.e.* genetic structure) as well as temporal (*i.e.* demographic history) neutral genetic variation. I compare observed patterns to those previously observed in other studies on two of *C. adjunctus*' main hosts, the little brown bat and the big brown bat. My purpose was to investigate the similarity of large-scale spatial genetic structure

and past demography between a generalist ectoparasite and two of its highly mobile hosts. In Chapters 3-4, I study the same ectoparasite species, but in a smaller geographic scale, the Great Lakes region, that is more likely to represent contemporary gene flow patterns. In Chapter 3, I compare patterns of neutral genetic variation of *C. adjunctus* to the better-known and strong generalist *C. lectularius* in the same geographic area. My purpose was to investigate the similarity of neutral genetic patterns between *C. adjunctus*, a species that is usually associated with highly mobile hosts (*i.e.* bats) and *C. lectularius*, a species that may associate with a larger variety of types of hosts (*i.e.* bats, humans, chicken and other domestic animals). In Chapter 4, I compare patterns of neutral genetic variation in *C. adjunctus* to those of one of its key hosts, the big brown bat, and measure the influence of landscape (land cover) on their respective spatial genetic structure. My purpose was to investigate how the response of gene flow patterns to intricate patterns of landscape structure differs between a highly mobile host and a generalist ectoparasite.

Chapters 5-6 primarily pertain to molecular adaptation of generalist temporary ectoparasites to their hosts. In those chapters, I analyze two genes coding for salivary proteins used by members of the Cimicidae family during blood feeding to maximize blood intake.

In Chapter 5, I tested for signals of selection at these two genes, using various species in the genus *Cimex* that associate with one of three types of hosts: birds, humans or bats (except *Cimex lectularius* that may associate with at least humans and bats). My goal was to determine if there is evidence of adaptation to phylogenetically different hosts in this group of of ectoparasites. In

Chapter 6, I tested for signals of selection and adaptation at the same two genes, using *C. adjunctus* individuals collected throughout the species range, and associated with one of three different bat host species. Here, my goal was to determine if there is evidence for local adaptation to ecologically different, but phylogenetically similar, hosts in a single species of ectoparasite.

1.5 Literature cited

Abrams AJ, Cannatella DC, Hillis DM, Sawyer SL (2013) Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. *Journal of General Virology*, **94**, 2082–2093.

Adelman ZN, Kilcullen KA, Koganemaru R et al. (2011) Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a single population. *Plos One*, **6**, e26228.

Allen JR (1994) Host resistance to ectoparasites. *Scientific and Technical Review of the Office International des Epizooties*, **13**, 1287–1303.

Allison AC (2002) The discovery of resistance to malaria of sickle-cell heterozygotes. *Biochemistry and Molecular Biology Education*, **30**, 279–287.

Amato R, Miele G, Monticelli A, Cocozza S (2011) Signs of selective pressure on genetic variants affecting human height. *Plos One*, **6**, e27588.

Anderson SJ, Kierepka EM, Swihart RK, Latch EK, Rhodes OE (2015) Assessing the permeability of landscape features to animal movement: using genetic structure to infer functional connectivity. *Plos One*, **10**, e0117500.

Andreakis N, Kooistra WHCF, Procaccini G (2009) High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Molecular Ecology*, **18**, 212–226.

Balashov YS (2006) Types of parasitism of acarines and insects on terrestrial vertebrates. *Entomological Review*, **86**, 957–971.

Balashov YS (2011) Parasitism and ecological parasitology. *Entomological Review*, **91**, 1216–1223.

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Bar Yaacov D, Arbel-Thau K, Zilka Y et al. (2012) Mitochondrial DNA variation, but not nuclear DNA, sharply divides morphologically identical chameleons along an ancient geographic barrier. *Plos One*, **7**, e31372.

Bargues MD, Klisiowicz DR, Panzera F et al. (2006) Origin and phylogeography of the Chagas disease main vector *Triatoma infestans* based on nuclear

rDNA sequences and genome size. *Infection, Genetics and Evolution*, **6**, 46–62.
Barnes I, Shapiro B, Lister A et al. (2007) Genetic structure and extinction of the woolly mammoth, *Mammuthus primigenius*. *Current Biology*, **17**, 1072–1075.
Barrett LG, Thrall PH, Burdon JJ, Linde CC (2008) Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in Ecology and Evolution*, **23**, 678–685.
Baute GJ, Kane NC, Grassa CJ, Lai Z, Rieseberg LH (2015) Genome scans reveal candidate domestication and improvement genes in cultivated sunflower, as well as post-domestication introgression with wild relatives. *New Phytologist*, **206**, 830–838.
Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.
Bensch S, Åkesson S, Irwin DE (2008) The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. *Molecular Ecology*, **11**, 2359–2366.
Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, **141**, 1007–1014.
Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.
Booth W, Balvín O, Vargo EL, Vilímová J, Schal C (2015) Host association drives genetic divergence in the bed bug, *Cimex lectularius*. *Molecular Ecology*, **24**, 980–992.
Brooks DR, Glen DR (1982) Pinworms and primates: a case study in coevolution. *Proceedings of the Helminthological Society of Washington*, **49**, 76–85.
Brower AV (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–6495.
Burns LE, Frasier TR, Broders HG (2014) Genetic connectivity among swarming sites in the wide ranging and recently declining little brown bat (*Myotis lucifugus*). *Ecology and Evolution*, **4**, 4130–4149.
Castillo JA, Epps CW, Davis AR, Cushman SA (2014) Landscape effects on gene flow for a climate-sensitive montane species, the American pika. *Molecular Ecology*, **23**, 843–856.
Catalano SR, Whittington ID, Donnellan SC, Gillanders BM (2014) Parasites as biological tags to assess host population structure: guidelines, recent genetic advances and comments on a holistic approach. *International Journal for Parasitology: Parasites and Wildlife*, **3**, 220–226.
Christe P, Giorgi MS, Vogel P, Arlettaz R (2003) Differential species-specific ectoparasitic mite intensities in two intimately coexisting sibling bat species: resource-mediated host attractiveness or parasite specialization? *Journal of Animal Ecology*, **72**, 866–872.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.
Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416–423.
Davis PL, Holmes EC, Larrous F et al. (2005) Phylogeography, population dynamics, and molecular evolution of European bat lyssaviruses. *Journal of Virology*, **79**, 10487–10497.
Dick CW, Patterson BD (2007) Against all odds: explaining high host specificity in dispersal-prone parasites. *International Journal for Parasitology*, **37**, 871–876.
Du Q, Gong C, Wang Q et al. (2015) Genetic architecture of growth traits in *Populus* revealed by integrated quantitative trait locus (QTL) analysis and association studies. *New Phytologist*, **209**, 1067–1082.
Dunning LT, Dennis AB, Thomson G et al. (2013) Positive selection in glycolysis among Australasian stick insects. *BMC Evolutionary Biology*, **13**, 215.
Duval L, Robert V, Csorba G et al. (2007) Multiple host-switching of Haemosporidia parasites in bats. *Malaria Journal*, **6**, 157.
Eckert AJ, Bower AD, GonzáLez-MartíNez SC et al. (2010) Back to nature: ecological genomics of Loblolly pine (*Pinus taeda*, Pinaceae): ecological genomics of Loblolly pine. *Molecular Ecology*, **19**, 3789–3805.
Ewald PW (1983) Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics*, **14**, 465–485.
Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.
Fenner L, Egger M, Bodmer T et al. (2013) HIV infection disrupts the sympatric host–pathogen relationship in human tuberculosis. *Plos Genetics*, **9**, e1003318.
Fenton MB (2001) *Bats*. Checkmark Books, New York, NY, USA.
Fischer MC, Foll M, Excoffier L, Heckel G (2011) Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Molecular Ecology*, **20**, 1450–1462.
Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2016) Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, **25**, 104–120.
Fox CW, Roff DA, Fairbairn DJ (2001) *Evolutionary ecology: concepts and case studies*. Oxford University Press, Oxford, NY, USA.
Fox CW, Wolf JB (2006) *Evolutionary genetics: concepts and case studies*. Oxford University Press, Oxford, NY, USA.
Frichot E, Schoville SD, Bouchard G, Francois O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.
Gandon S, Nuismer SL (2009) Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *The American Naturalist*, **173**, 212–224.
Garamszegi L (2009) Patterns of co-speciation and host switching in primate malaria parasites. *Malaria Journal*, **8**, 110.

Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, **21**, 434–443.
Givnish TJ, Millam KC, Mast AR et al. (2009) Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B: Biological Sciences*, **276**, 407–416.
Goldberg J, Trewick SA, Paterson AM (2008) Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3319–3334.
Goulding TC, Cohen CS (2014) Phylogeography of a marine acanthocephalan: lack of cryptic diversity in a cosmopolitan parasite of mole crabs. *Journal of Biogeography*, **41**, 965–976.
Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.
Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host-parasite coevolutionary arms races give way to fluctuating selection: bacteria-phage coevolutionary dynamics. *Ecology Letters*, **14**, 635–642.
Hancock AM, Witonsky DB, Alkorta-Aranburu G et al. (2011) Adaptations to climate-mediated selective pressures in humans. *Plos Genetics*, **7**, e1001375.
Hanifin CT, Brodie ED, Brodie ED (2008) Phenotypic mismatches reveal escape from arms-race coevolution. *Plos Biology*, **6**, e60.
Hanski I (1998) Metapopulation dynamics. *Nature*, **396**, 41–49.
Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.
Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.
Hermisson J, Hansen TF, Wagner GP (2003) Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *The American Naturalist*, **161**, 708–734.
Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology*, **22**, 2898–2916.
Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.
Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898.
International Glossina Genome Initiative, Attardo GM, Abila PP et al. (2014) Genome sequence of the tsetse fly (*Glossina morsitans*): vector of African trypanosomiasis. *Science*, **344**, 380–386.
Jacobs FMJ, Greenberg D, Nguyen N et al. (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, **516**, 242–245.
Jones B (2014) Pathogen genetics: exploring gene function and parasite–host protein interactions. *Nature Reviews Genetics*, **15**, 646–646.

Kempf F, Boulinier T, De MeeûS T, Arnathau C, McCOY KD (2009) Recent evolution of host-associated divergence in the seabird tick *Ixodes uriae*. *Molecular Ecology*, **18**, 4450–4462.
Ketola T, Mikonranta L, Zhang J et al. (2013) Fluctuating temperature leads to evolution of thermal generalism and preadaptation to novel environments. *Evolution*, **67**, 2936–2944.
Klironomos JN (2003) Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology*, **84**, 2292–2301.
Kuzmina NA, Lemey P, Kuzmin IV et al. (2013) The phylogeography and spatiotemporal spread of South-Central skunk rabies virus. *Plos One*, **8**, e82348.
Leggett HC, Buckling A, Long GH, Boots M (2013) Generalism and the evolution of parasite virulence. *Trends in Ecology and Evolution*, **28**, 592–596.
Lehmann T, Hume JCC, Licht M et al. (2009) Molecular evolution of immune genes in the malaria mosquito *Anopheles gambiae*. *Plos One*, **4**, e4549.
Levin II, Parker PG (2013) Comparative host–parasite population genetic structures: obligate fly ectoparasites on Galapagos seabirds. *Parasitology*, **140**, 1061–1069.
Levy F, Neal CL (1999) Spatial and temporal genetic structure in chloroplast and allozyme markers in *Phacelia dubia* implicate genetic drift. *Heredity*, **82**, 422–431.
Lively CM, Dybdahl MF, Jokela J, Osnas EE, Delph LF (2004) Host sex and local adaptation by parasites in a snail- trematode interaction. *The American Naturalist*, **164**, S6–S18.
Loewe L, Hill WG (2010) The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1153–1167.
Mahamdallie SS, Ready PD (2012) No recent adaptive selection on the apyrase of Mediterranean *Phlebotomus*: implications for using salivary peptides to vaccinate against canine leishmaniasis: adaptive evolution of sandfly apyrase. *Evolutionary Applications*, **5**, 293–305.
Marshall AG (1982) *The ecology of ectoparasitic insects*. Academic Press, Cambridge, MA, USA.
Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G (2016) Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, **19**, 336–347.
McCoy KD, Boulinier T, Schjørring S, Michalakis Y (2002) Local adaptation of the ectoparasite *Ixodes uriae* to its seabird host. *Evolutionary Ecology Research*, **4**, 441–456.
Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, **12**, 106–117.
Moen D, Morlon H (2014) From dinosaurs to modern bird diversity: extending the time scale of adaptive radiation. *Plos Biology*, **12**, e1001854.
Moore WS (1995) Inferring Phylogenies from mtDNA Variation: Mitochondrial-Gene Trees Versus Nuclear-Gene Trees. *Evolution*, **49**, 718.

Morand S, Krasnov BR, Poulin R (2006) *Micromammals and macroparasites: from evolutionary ecology to management*. Springer Publishing, New York, NY, USA.
Morens DM, Fauci AS (2013) Emerging infectious diseases: threats to human health and global stability. *Plos Pathogens*, **9**, e1003467.
Narum SR, Hess JE (2011) Comparison of F_{ST} outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.
Neff BD (2004) Stabilizing selection on genomic divergence in a wild fish population. *Proceedings of the National Academy of Sciences*, **101**, 2381–2385.
Nieberding CM, Durette-Desset M-C, Vanderpoorten A et al. (2008) Geography and host biogeography matter for understanding the phylogeography of a parasite. *Molecular Phylogenetics and Evolution*, **47**, 538–554.
Nieberding C, Morand S, Libois R, Michaux JR (2004) A parasite reveals cryptic phylogeographic history of its host. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2559–2568.
Olival KJ, Dick CW, Simmons NB et al. (2013) Lack of population genetic structure and host specificity in the bat fly, *Cyclopodia horsfieldi*, across species of *Pteropus* bats in Southeast Asia. *Parasites and Vectors*, **6**, 231–248.
Pélabon C, Hansen TF, Carter AJR, Houle D (2010) Evolution of variation and variability under fluctuating, stabilizing, and disruptive selection. *Evolution*, **64**, 1912–1925.
Pincheira-Donoso D, Harvey LP, Ruta M (2015) What defines an adaptive radiation? Macroevolutionary diversification dynamics of an exceptionally species-rich continental lizard radiation. *BMC Evolutionary Biology*, **15**.
Pinheiro A, Woof JM, Abi-Rached L, Parham P, Esteves PJ (2013) Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens. *Plos One*, **8**, e73934.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Rioux Paquette S, Talbot B, Garant D, Mainguy J, Pelletier F (2014) Modelling the dispersal of the two main hosts of the raccoon rabies variant in heterogeneous environments with landscape genetics. *Evolutionary Applications*, **7**, 734–749.
Ripperger SP, Tschapka M, Kalko EKV, Rodríguez-Herrera B, Mayer F (2014) Resisting habitat fragmentation: high genetic connectivity among populations of the frugivorous bat Carollia castanea in an agricultural landscape. *Agriculture, Ecosystems and Environment*, **185**, 9–15.
Schaer J, Perkins SL, Decher J et al. (2013) High diversity of West African bat malaria parasites and a tight link with rodent *Plasmodium* taxa. *Proceedings of the National Academy of Sciences*, **110**, 17415–17419.
Schluter J, Foster KR (2012) The evolution of mutualism in gut microbiota via host epithelial selection. *Plos Biology*, **10**, e1001424.

Schmitt MW, Kennedy SR, Salk JJ et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proceedings of the National Academy of Sciences*, **109**, 14508–14513.
Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.
Shaw KL, Parsons YM, Lesnick SC (2007) QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket *Laupala*. *Molecular Ecology*, **16**, 2879–2892.
Spice EK, Goodman DH, Reid SB, Docker MF (2012) Neither philopatric nor panmictic: microsatellite and mtDNA evidence suggests lack of natal homing but limits to dispersal in Pacific lamprey. *Molecular Ecology*, **21**, 2916–2930.
Sponchiado J, Melo GL, Landulfo GA et al. (2015) Interaction of ectoparasites (Mesostigmata, Phthiraptera and Siphonaptera) with small mammals in Cerrado fragments, western Brazil. *Experimental and Applied Acarology*, **66**, 369–381.
Streicker DG, Turmelle AS, Vonhof MJ et al. (2010) Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science*, **329**, 676–679.
Takahashi T, Koblmüller S (2011) The adaptive radiation of cichlid fish in Lake Tanganyika: a morphological perspective. *International Journal of Evolutionary Biology*, **2011**, 1–14.
Talbot B, Chen T-W, Zimmerman S et al. (2016) Combining genotype, phenotype, and environment to infer potential candidate genes. *Journal of Heredity*, **108**, 207–216.
Talbot B, Garant D, Raquette SR, Mainguy J, Pelletier F (2014) Genetic structure and diversity among rabid and nonrabid raccoons. *Ecoscience*, **20**, 345–351.
Talbot B, Garant D, Rioux Paquette S, Mainguy J, Pelletier F (2012) Lack of genetic structure and female-specific effect of dispersal barriers in a rabies vector, the striped skunk (*Mephitis mephitis*). *Plos One*, **7**, e49736.
Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Westwood JH, dePamphilis CW, Das M et al. (2012) The parasitic plant genome project: new tools for understanding the biology of *Orobanche* and *Striga*. *Weed Science*, **60**, 295–306.
Zimin A, Stevens KA, Crepeau MW et al. (2014) Sequencing and assembly of the 22-Gb Loblolly pine genome. *Genetics*, **196**, 875–890.

2 Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*

2.1 Introduction

Parasites, through effects on host survival and reproduction, can modify the morphology, life history and behavior of their hosts. Parasites may also influence the dynamics of host populations thereby shaping communities (Poulin 1999). Hosts in turn may also have important effects on their parasites. Many parasite species, whether endoparasites or ectoparasites, remain closely associated with their hosts through much of their life cycle (Marshall 1981), and often rely on their hosts for dispersal. Dispersal, in turn, influences gene flow and therefore genetic structure and diversity of a species; across a broad range of taxa, less dispersal is associated with increased spatial structure and differentiation (Bohonak 1999). Not surprisingly, spatial genetic structure of a parasite frequently reflects dispersal of its host. For example, population genetic structure of parasitic nematodes of cattle, sheep and white-tailed deer is explained by host movements (Blouin et al. 1995). However, relative to their hosts, parasites often show higher levels of genetic differentiation. As such, analysis of the trematode parasite (*Pagioporus shawi*) permitted more detailed information on population assignments in its host, the steelhead trout (*Oncorhynchus mykiss*) than could be obtained by examining genetic variation in the host itself (Criscione et al. 2006). In addition to dispersal, parasites and hosts may have experienced correlated demographic and range dynamics (Anderson and Gordon 1982; Thrall and Burdon 1997) which will also be reflected in their population genetic structure; for instance, patterns of genetic variation among

populations of the parasitic nematode *Heligmosomoides polygyrus* have revealed demographic and historic events affecting its host, the field mouse *Apodemus sylvaticus* (Nieberding et al. 2004). Furthermore, differences in regional abundance of two *Apodemus* species likely caused differentiation of both the *Apodemus* host and their *Heligmosomoides* parasite species (Nieberding et al. 2008).

However, it has recently been shown that a strong link between host dispersal and parasite genetic structure is not ubiquitous, and depends on factors that include the degree of association with the host and host mobility (Mazé-Guilmo et al. 2016). Here, we investigated spatial genetic structure and past demography of an ectoparasite that is associated with highly mobile flying hosts, and would be considered a weak generalist based on its association with a number of different host species that are closely related to each other; Mazé-Guilmo et al. 2016). Our study complements a body of work on spatial genetic structure and phylogeography of various ectoparasites associated with hosts having higher mobility (McCoy et al. 2003; van der Mescht et al. 2015; Engelbrecht et al. 2016).

Insects in the genus *Cimex* (Order: Hemiptera) are temporary ectoparasites of homeothermic animals. They do not remain on their host at all times but rather remain in nests or roosts between blood meals (Usinger 1966). Most *Cimex* species are associated exclusively with bats, while a few associate with a more diverse range of hosts (Usinger 1966; Goddard 2009; Criado et al. 2011). *Cimex adjunctus* is a widespread ectoparasite of bats in North America, occurring from the eastern seaboard to the Rocky Mountains, and from Labrador

and the Northwest Territories south to Texas (Usinger 1966). It parasitizes a number of bat species, including the big brown bat (*Eptesicus fuscus*) and the little brown myotis (*Myotis lucifugus*), two species that often roost in buildings (Furlonger et al. 1987; Ellison et al. 2007; Pearce and O'Shea 2007). The generation time of *C. adjunctus* is unknown, but is likely similar to that of the common bed bug *C. lectularius*, which can range from two to 12 generations a year depending on monthly temperatures (Usinger 1966), and is certainly much shorter than that of its hosts.

Usinger (1966) proposed that *Cimex* species have a very low inherent capacity for dispersal over longer distances, on the scale of kilometers. He thought it unlikely that adult *Cimex* species disperse on their own. He therefore hypothesized that *Cimex* species can disperse occasionally attached to a host's body (Usinger 1966). Previous studies of genetic diversity of the big brown bat and little brown myotis in North America have reported high within-site genetic variation and generally low among-site differentiation, although there are differences between patterns at nuclear and mitochondrial markers (*E. fuscus*, Vonhof et al. 2008; Turmelle et al. 2011; *M. lucifugus*, Burns et al. 2014; Johnson et al. 2015; McLeod et al. 2015; Vonhof et al. 2015). Overall, these studies indicate that high levels of gene flow are maintained over long distances in both bat species, while genetic structuring of mitochondrial variation suggests a higher degree of female than male philopatry. For *C. adjunctus*, likely only a fraction of host dispersal events result in successful parasite dispersal so gene flow may be lower in *C. adjunctus* relative to these two host species. Furthermore, *C. adjunctus* may experience frequent extirpation and recolonization events.

Bartonička and Růžičková (2013) identified bat bug load as a possible cause of roost-switching in bats, with numbers of bats dropping as the population of *C. pipistrelli* reaches a high. They also found the appearance of *C. pipistrelli* 21 to 56 days after the first bat visit in any given roost. Since *C. adjunctus*, like *C. pipistrelli,* does not stay on the host between blood meals, sudden host population decreases within roosts might drive local extirpation events.

Although different ectoparasite races are often associated with different host species (Tomisawa and Akimoto 2004; McCoy et al. 2005; Dick and Patterson 2007), high gene flow among populations associated with different host species has also been documented. In Europe, *Cimex pipistrelli* is morphologically, but not genetically, differentiated among bat host species (Balvín et al. 2013). This suggests possible morphological plasticity, but high gene flow, among individuals associated with different host species. In North America, we might also expect gene flow among *C. adjunctus* populations on different host species. Many different North American bat species temporarily roost together for short intervals during the night, such as many *Myotis* species including *M. lucifugus* and *E. fuscus* (Adam and Hayes 2000), potentially facilitating host switching by *C. adjunctus*.

Much of North America was unsuitable for many bat species during the last Pleistocene glacial maximum, and both *M. lucifugus* and *E. fuscus* are hypothesized to have expanded their ranges from glacial refugia. Dixon (2011a) suggested that little brown myotis populations currently in Minnesota have dispersed from a single large southeastern US glacial refugium, and Turmelle et al. (2011) suggested that big brown bat populations have dispersed from several

eastern and western US glacial refugia into what is now Colorado. Range and demographic expansion in little brown myotis has also been proposed on the east coast of Canada (Burns et al. 2014). We expect that the potential dependence of *C. adjunctus* on its host species for long-distance dispersal and colonization may have contributed to broadly congruent patterns of historical range expansion over large spatial scales.

We investigated the spatial genetic structure and phylogeography of *C. adjunctus* across its range in North America. Because of its comparatively shorter generation time, the likelihood that only a fraction of bat dispersal events may result in ectoparasite gene flow, and the potential for local extirpations, we predicted stronger spatial genetic structure in *C. adjunctus* relative to its hosts. Because of the potential for movement among host species, we also examined differentiation among populations found on different host species. Finally, based on the hypothesis that post-Pleistocene climate warming had similar effects on the demographic history of *C. adjunctus* as that of its hosts, we predicted genetic signatures of demographic and range expansion in *C. adjunctus*.

2.2 Materials and Methods

2.2.1 Sample collection

We collected *C. adjunctus* across much of its North American range. Most samples are from mist-netted host individuals of *E. fuscus*, *M. lucifugus* or *M. septentrionalis* (Table A.1). Mist net capture locations were adjacent to a known summer roost (house, barn, cabin, church, school or abandoned mine) of either of the three bat species, or within forested national, provincial, state or

territorial lands (Table A.2). Most mist-netted bats and the *C. adjunctus* individuals they harboured likely came from the adjacent known roost, although it is possible that a small proportion came from different roosts in the area. Overall, between 3 and 15% of mist-netted bats harboured a parasite, depending on the location. We also sampled *C. adjunctus* individuals from the interior of two summer roosts. One roost was in a church attic inhabited by *M. lucifugus*, and one was in a house attic inhabited by *E. fuscus* (Table A.2). Because we could be certain of the roost site in these cases, we considered these two sampling locations as distinct from their adjacent mist-netting capture locations. Upon collection, we stored samples immediately in a 95% ethanol solution until further analyses. We then generated *CO1* mitochondrial DNA sequence data and nine nuclear microsatellite genotype data for all individuals. All samples included in this study were confirmed as being *C. adjunctus* using a DNA barcoding approach (Hebert et al. 2003). We compared the *CO1* sequence for each sample to known *CO1* sequences for *Cimex* species from published sources (Balvín et al. 2015).

2.2.2 Genetic analyses

We extracted DNA from the whole insect for all samples using the DNeasy Blood and Tissue Kit (QIAGEN, Germantown, Maryland, United States). We then amplified a 576-bp fragment of the *CO1* gene from each individual using the primers: F 5'- TATGAGCAGGCATGTTAGGG and R 5'- ATAGATGTTGATAAAGAATTGGG (Designed by our group based on published sequences of Balvin et al. (2015). We used a DNAEngine PTC-200 Thermal

Cycler (Bio-Rad, Hercules, California, United States) to execute the polymerase chain reaction (PCR) amplification. We performed PCR in 25 µL final volume using the following recipe: 1X Taq polymerase buffer excluding $MgCl_2$ (Applied Biosystems, Foster City, California, United States), 1.5 mM of $MgCl_2$, 0.2 mM of each type of dNTP, 0.3 µM of each primer, 1 U of Taq polymerase (ABI), and 1 µL of DNA extraction product. We used the following PCR program: an initial denaturation step of 1 min at 94°C, followed by 36 cycles of 30 sec of denaturation at 94°C, 45 sec of annealing at 49°C and 45 sec of extension at 72°C, finished by a final extension step of 5 min at 72°C. We visualized PCR products by 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad) on a UV transluminator to check the quality and size of amplified fragments. Then, we sequenced the amplified gene fragment for every sample using Sanger sequencing with BigDye terminator chemistry (ABI) and analyzed the fragments on a 3730xl DNA Analyzer (ABI). We aligned all sequences using MEGA 6.06.

We also genotyped all individuals at nine microsatellite loci originally designed for *Cimex lectularius* (Cle002, Cle003, Cle013, Cle015, from Fountain et al. 2014, and Clec21, Clec48, Clec15, Clec104 and BB28B, from Booth et al. 2012; Table A.3). We used a DNAEngine PTC-200 Thermal Cycler (Bio-Rad) to execute PCR amplification. For all microsatellite markers, we performed PCR using the following recipe: 1X Taq polymerase buffer excluding $MgCl_2$ (ABI), 2.175 mM of $MgCl_2$, 0.216 mM of each type of dNTP, 0.25 to 1.2 µM (Table A.3) of each primer, 1 U of Taq polymerase (ABI), 2 µL of DNA extraction product, in total volume of 12 µL. For markers from Fountain et al. (2014), we used the following thermal cycling: an initial denaturation step of 15 min at 95°C, followed

by 11 cycles of 30 sec of denaturation at 94°C, 1 min and 30 sec of annealing (initially at 65°C and reduced 1°C at every cycle) and 1 min of extension at 72°C, followed by 26 cycles of 30 sec of denaturation at 94°C, 1 min and 30 sec of annealing at 55°C and 1 min of extension at 72°C, finished by a final extension step of 10 min at 72°C. For markers from Booth et al. (2012), we used the following thermal cycling: an initial denaturation step of 3 min at 95°C, followed by 35 cycles of 30 sec of denaturation at 95°C, 30 sec of annealing at 59 to 61°C (Table A.3) and 30 sec of extension at 72°C, and a final extension step of 5 min at 72°C. We amplified each locus individually. PCR products were visualized by 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad) on a UV transluminator to check the quality and size of amplified fragments. We then sized products on a 3730xl DNA Analyzer (ABI). We called all microsatellite genotypes for each species using GeneMapper Software 4.0 (ABI), and we checked all calls manually.

2.2.3 Statistical analyses

Microsatellite diversity, and Hardy-Weinberg and linkage disequilibrium

For sites with data for at least five sampled *C. adjunctus* individuals, and for genetic clusters (see next section), we calculated average number of alleles, expected and observed heterozygosity, and inbreeding coefficient. For microsatellite loci, we tested for Hardy-Weinberg and linkage disequilibrium within each site with data for at least two sampled *C. adjunctus* individuals, using Genepop 4.2. For each type of test, we corrected for multiple tests using Bonferroni correction, with a threshold α of 0.05.

Range-wide genetic structure

We tested our prediction of range-wide genetic structure and an effect of geographic distance in *C. adjunctus* using genetic clustering, tests of isolation-by-distance (IBD), and an analysis of molecular variance (AMOVA). We conducted a Bayesian clustering analysis using Geneland 4.0.5, which takes into account geographic coordinates of individual samples. We used 100,000 iterations, thinned every 100th iteration, and a post-process burn-in of 200 (of the 1,000 left after thinning), for *K* values between 1 and 20. We executed 10 runs, and kept the one with the highest posterior mean density, after burn-in. We attempted to identify the population to which each individual was assigned the most often, defined here as the population where the majority of Markov Chain Monte Carlo (MCMC) chains converged for any given individual. We also conducted a *K*-Means clustering analysis using GenoDive 2.0 on allele frequencies, for *K* values between 1 and 20, and using 50,000 simulation steps, to validate results obtained with the Geneland method. We used Bayesian Information Criterion (BIC) values to determine the most likely *K* value.

We conducted an individual-level analysis of IBD, using the estimate of genetic relatedness, r_W (Wang 2002), calculated with SpaGeDi 1.5. We calculated $1 - r_W$ for each pairwise relationship, in order to obtain genetic distances. We calculated geographic distance (in km) between sample sites, corrected for sphericity of the earth, using the 'rdist.earth' function from the 'fields' package (Fields Development Team 2006) in R 3.1.3. We then fit pairwise genetic distance to geographic distance using Multiple regression on distance matrices (MRM), in the 'MRM' function from the 'ecodist' package in R 3.1.3

(Goslee and Urban 2007), which uses a Mantel test derived linear regression model. We assessed significance through a permutation procedure (9,999 replicates). An assumption of the r_W relatedness index, and most other relatedness indices, is that individuals are in a large random mating population without population structure (Wang 2011). In an attempt to correct for the population structure present in our dataset, we subsequently conditioned IBD models for genetic clustering. For each pair of individuals assigned to the same population in clustering analyses, we assigned a value of 0, and for each pair of individuals assigned to different populations, we assigned a value of 1. We then tested the effect of geographic distance, together with genetic clustering, on genetic distance in an MRM model.

For all sites with at least two sampled individuals, we used AMOVA to examine the proportion of genetic variation among sites, and among individuals associated with different host species. AMOVA was executed in GenoDive 2.0 for microsatellite data, and Arlequin 3.5 for mitochondrial data.

Demographic history

We tested the prediction that *C. adjunctus* would show signals of demographic and range expansion, similar to some of its bat hosts, with a suite of methods for investigating demographic history using either mitochondrial data alone, or both mitochondrial and microsatellite data. First, we produced a minimum-spanning network of mitochondrial haplotypes (MSN) using TCS 1.21, with a 95% connection limit. MSNs can indicate past range expansions if they

show starburst like patterns (Yuan et al. 2010; Pulgarín-R and Burg 2012). We expected to find such evidence pointing towards range expansion in *C. adjunctus*.

We executed a Mismatch Distribution (MD) analysis with DNASP 5.1. The purpose of this analysis is to compare the distribution of the frequency of each number of pairwise mitochondrial sequence mismatches in the dataset to the expected distributions under demographic expansion or constant population size through time. A unimodal peak at a non-zero number of pairwise mismatches is associated with demographic expansion, which we expected to observe, whereas more than one non-zero number of pairwise mismatches is usually associated with a constant population size through time (Pereira et al. 2001).

Then, we constructed an extended Bayesian skyline plot (EBSP) using mitochondrial data in BEAST 1.8.4. We used a linear EBSP model, and random local clock, which reportedly performs better than strict and relaxed clocks for most situations using intraspecific data (Drummond and Suchard 2010; Brown and Yang 2011). In trial runs, we found the HKY substitution model (Hasegawa et al. 1985) to be the best-fitting model, as has also been shown for *Triatoma infestans* (Bargues et al. 2006), a species in a genus closely related to *Cimex*. We used the gamma sites model to account for heterogeneity of substitution rate among individual loci. We used the default value of 10,000,000 MCMC chains, logging every 1,000 chains. We set the substitution rate to 0.575%/Ma, or half of 1.15%/Ma, which is the standard Arthropod mitochondrial pairwise substitution rate as reported by (Brower 1994). All other parameters were kept at default value. EBSPs allow one to visualize effective population size (N_E) multiplied by

generation time (τ) since some time in the past. In the case of highly structured populations, Heller et al. (2013) suggested that a pooled sampling scheme, where several individuals are taken from about ten populations, was ideal to avoid a confounding effect of population structure, as opposed to all samples taken from the same population or one sample taken for each of a large number of populations. The sampling scheme used in our analysis fits well with the described pooled scheme. We expected to see an increase in effective population size over time, corresponding with a post-Pleistocene climate warming timeline.

Finally, we executed approximate Bayesian computations (ABC) on both mitochondrial and microsatellite data, using DIYABC 2.1.0. ABCs allow one to compare posterior probabilities of different demographic scenarios (Bertorelle et al. 2010). As per the method of Chakraborty et al. (2014), we input three scenarios in the analyses (for population sizes $N_A > N_1 > N_B$), mimicking an increase in effective population size from N_1 to N_A at time *t*, a decrease in population size from N_1 to N_B at time *t*, and finally constancy in population size at N_1. Boundaries for N_A, N_B, N_1 and *t* priors are available in Table 2.1. We set the potential time for the population size change event between 10,000 and 10 million years ago, to encompass a broad period of major climatic changes in the northern hemisphere (Zachos 2001). We set the upper boundary of effective population after an increase (N_A) to 10 times the initial upper boundary of effective population size (N_1), to limit our analysis to population size increases of at least an order of magnitude. Similarly, we set the lower boundary of effective population size after a decrease (N_B) to 1/10 of the initial lower boundary of

Table 2.1 Parameter values used in the approximate Bayesian computation analysis of demographic history of *Cimex adjunctus*. The set lower and upper boundaries of the three effective population size parameters are shown : N_1 is the effective population size before population size change, N_A is the effective population size after demographic expansion, and N_B is the effective population size after demographic decline. The time period over which a population size change potentially occurred is *t* (in years).

Parameter	Lower boundary	Upper boundary
N_A	500,000	50,000,000
N_B	50,000	5,000,000
N_1	500,000	5,000,000
t	10,000	10,000,000

effective population size (N_1) to limit our analysis to population size decreases of at least an order of magnitude. If no change in population size of at least an order of magnitude occurred in the analysis timeframe, or if both a population size decrease and increase of similar magnitude occurred, then the scenario of constant population size would be most likely. We conducted a series of initial trial runs to determine the effective population size parameters for ABC analysis where we could achieve convergence between priors and observed values (Table 2.1). As an example, our final effective population size parameter values are large in comparison with those in a study on the invasive ladybird *Harmonia axyridis* (Lombaert et al. 2011). Trial runs also indicated the best fit was achieved when we used a mutation rate per site between 10^{-7} and 10^{-9}, with a mean at 10^{-8}, for both mitochondrial and microsatellite markers, and the Stepwise Mutation Model (setting coefficient P at 0) for microsatellite markers. The estimated mutation rate for microsatellites is on the low end for such markers, but consistent with the relatively low variability observed for markers originally developed in a different species (*i.e.*, ascertainment bias; Li and Kimmel 2013). We used “Mean number of alleles” and “Mean genic diversity” as summary statistics for microsatellite loci, and “Number of haplotypes”, “Mean of pairwise differences” and “Private segregating sites” for the mitochondrial locus. We computed 3,000,000 simulated datasets to compare with the observed dataset. First, we pre-evaluated the fit of observed values to prior distributions of scenarios, using a Principal Component Analysis implemented with the software. In a graph of the first two principal components, a good prior assessment is reflected in the observed values being approximately in the centre of the prior

values for all three scenarios. Second, we calculated posterior probabilities for all three scenarios using a logarithmic regression, to determine which scenario is the most likely given the data. We expected strong support for a scenario mimicking a demographic expansion.

2.3 Results

We collected 160 *Cimex adjunctus* samples from throughout its range in North America (108 from *E. fuscus*, 36 from *M. lucifugus* and 16 from *M. septentrionalis*; Fig. 2.1; Table A.1), from 45 sites (Table A.2). We successfully amplified a fragment of the cytochrome c oxidase 1 (*CO1*) gene for 154 *C. adjunctus*, and identified 41 haplotypes with 46 polymorphic sites (data available in Table A.1). We also genotyped 150 of the *C. adjunctus* at nine microsatellite loci (data available in Table A.4). We successfully obtained both mitochondrial and microsatellite data for 144 *C. adjunctus* (94% of the *CO1* dataset and 96% of the microsatellite dataset; Table A.1). For approximate Bayesian computation (ABC), which uses both types of markers, we used the overlapping dataset (144 individuals). For analyses using only microsatellite loci (genetic clustering and isolation-by-distance, IBD), we used the whole microsatellite dataset (150 individuals) and for analyses using only mitochondrial DNA (minimum spanning network, MSN; mismatch distribution, MD; and extended Bayesian skyline plot, EBSP), we used the complete *CO1* dataset (154 individuals). Finally, we used only sites with data for at least two individuals, and for which we obtained both mitochondrial and microsatellite data, for the analysis of molecular variance (AMOVA) analysis, which resulted in a dataset of 127 individuals from 26 sites

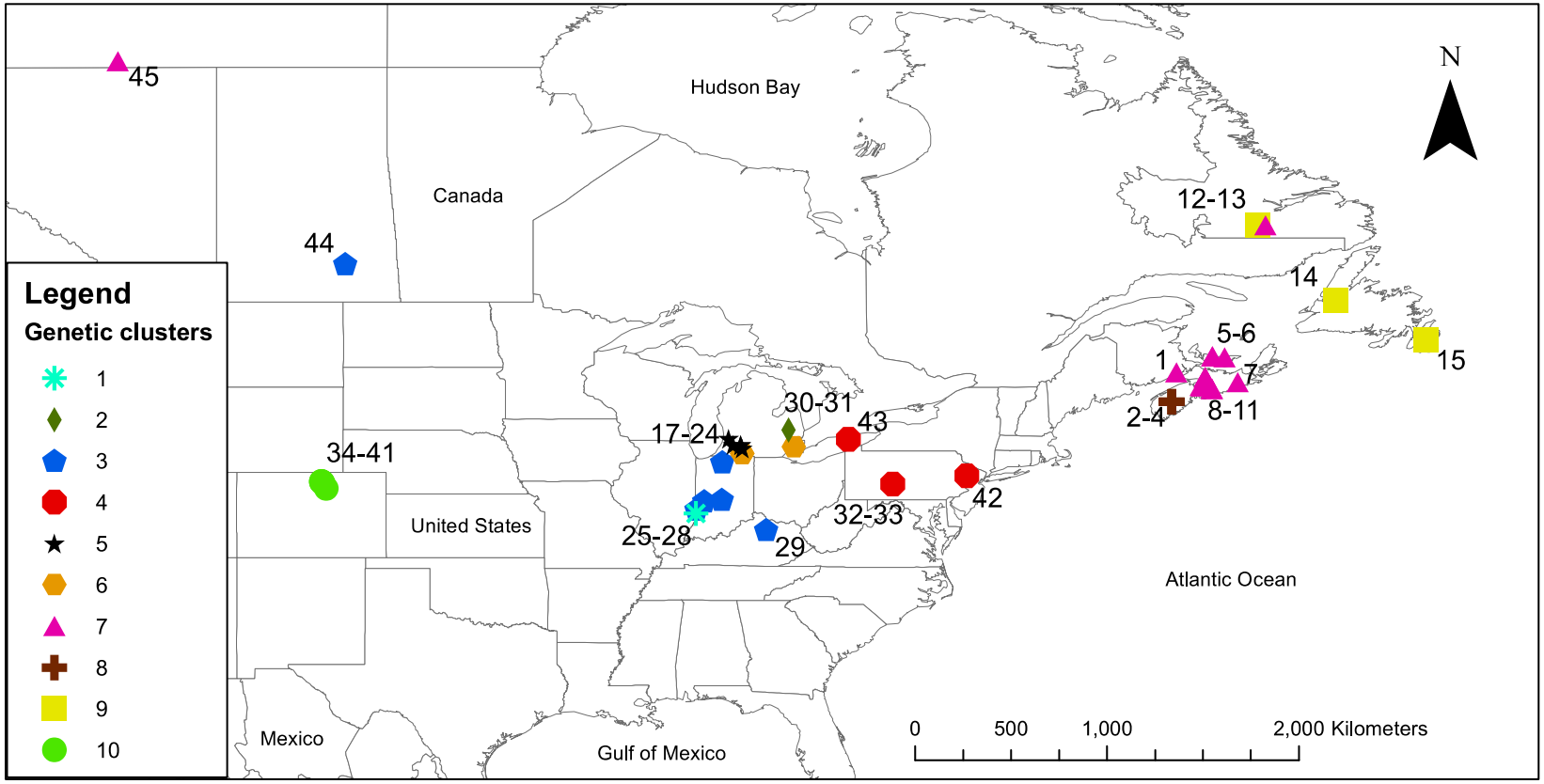

Figure 2.1. Sampling locations of *Cimex adjunctus* in North America. Created with ArcGIS 10.3 (ESRI, Redlands, United States). Numbers on the map correspond to site numbers in Table 2.2, Table A.1 and Table A.2. Membership to each of ten genetic clusters, defined using microsatellite data in Geneland 4.0.5, is shown with a unique shade and shape. Cluster numbers are given in the Legend and correspond to those in Table 2.2 and Table A.1.

(82 % of the CO1 dataset and 85% of the microsatellite dataset).

2.3.1 Microsatellite diversity, and Hardy-Weinberg and linkage disequilibrium

Among the nine microsatellite loci, we observed between two and 31 alleles. Across different sites and genetic clusters (identified by Geneland 4.0.5), average number of alleles ranged from 1.5 to 4, expected heterozygosity ranged from 0.18 to 0.62, observed heterozygosity ranged from 0.09 to 0.25, and the inbreeding coefficient varied between 0.00 and 0.77 (Table 2.2). Variation in genetic diversity and inbreeding coefficients did not show any obvious spatial pattern. We found three significant cases of deviation from Hardy-Weinberg equilibrium (one site at the loci Clec104 and Cle015, and another population at Clec104). Since these incidences of deviation from Hardy-Weinberg equilibrium were not systematic across loci or populations, we retained these two markers and two populations for our analyses. We did not find any evidence of significant linkage disequilibrium in any marker.

2.3.2 Range-wide genetic structure

Genetic clustering analyses using the Geneland method revealed 10 genetic clusters (Table 2.3), which were generally concordant with geographic location (Fig. 2.1). One interesting exception was that individuals from the Northwest Territories and Saskatchewan clustered with individuals from distant regions (Clusters 3 and 7; Fig. 2.1). There was no association between genetic clusters identified by Geneland 4.0.5 and any major geographic barriers that might knowingly impact dispersal. The sampling year and host species did not

Table 2.2 Genetic diversity estimates for *C. adjunctus*, averaged across nine microsatellite markers, for sites with five or more sampled individuals and for genetic clusters identified by Geneland 4.0.5 (with the exception of Cluster 1, in which there was only one individual; Table A.1). Site and cluster numbers correspond to those in Fig. 2.1, Fig. 2.2, Table A.1 and Table A.2.

Site/Cluster	Average number of alleles	Expected heterozygosity	Observed heterozygosity	Inbreeding coefficient G_{IS}
Site 17	2.000	0.275	0.278	-0.009
Site 19	1.889	0.363	0.093	0.745
Site 30	2.778	0.346	0.201	0.420
Site 31	2.111	0.327	0.254	0.223
Site 32	3.000	0.193	0.193	0.369
Site 36	2.444	0.325	0.224	0.310
Site 39	2.556	0.293	0.241	0.178
Site 40	2.222	0.239	0.145	0.393
Site 41	1.778	0.184	0.160	0.129
Cluster 2	3.000	0.349	0.225	0.431
Cluster 3	2.556	0.394	0.246	0.375
Cluster 4	3.222	0.332	0.235	0.291
Cluster 5	2.778	0.309	0.186	0.399
Cluster 6	2.222	0.327	0.247	0.245
Cluster 7	3.889	0.349	0.225	0.354
Cluster 8	3.000	0.615	0.143	0.768
Cluster 9	1.667	0.250	0.194	0.222
Cluster 10	3.556	0.301	0.196	0.349

Table 2.3 Results of clustering and isolation-by-distance analyses of *Cimex adjunctus*, estimated using microsatellite markers. Most likely number of genetic clusters (K) estimated using the Geneland method, isolation-by-distance (IBD) and IBD while correcting for population genetic structure (IBD + K) are shown. *Statistically significant at α = 0.05.

Statistic		Value
Most likely K		10
IBD (r_W)	*P*	0.001*
	R^2	0.19
IBD (r_W) + K	*P* (IBD)	0.001*
	P (K)	0.001*
	R^2	0.21

seem to strongly affect clustering, as individuals associated with different host species or sampled at different years were frequently assigned to the same cluster (Table A.5). Using *K*-means clustering, we obtained the lowest BIC value at $K = 11$, and the second lowest BIC value was at $K = 10$. Moreover, we observed significant IBD calculated on individual genetic relatedness values ($P = 0.001$, $R^2 = 0.19$; Table 2.3). Conditioning for genetic structure slightly improved the fit of the IBD model ($R^2 = 0.21$; Table 2.3).

AMOVA results were very different between the two types of markers. For microsatellites, considerably less of the total variation was explained by among than within sampling site differences (22.8%; Table 2.4) than within sites (37.0%), but for mitochondrial data the variation among (48.8%) and within sites (43.7%) were similar. The proportion of genetic variation among host species was high for microsatellite data (40.2%; Table 2.4), but quite low for mitochondrial data (7.4%).

2.3.3 Demographic history

Considering those haplotypes represented by four or more individuals, there was some degree of spatial structuring in their distribution. Specifically, distinct haplotypes were associated with the western and eastern ends of *C. adjunctus'* range (Fig. 2.2). One interesting observation was that individuals from Northwest Territories and Saskatchewan had very similar haplotypes to individuals from the Midwest of the United States. The MSN did not show a well-defined starburst pattern (Fig. 2.3). Also, the mismatch distribution (MD) showed multiple peaks rather than a single peak that would have indicated potential

Table 2.4 Results of analysis of molecular variance (AMOVA) on *Cimex adjunctus*, using mitochondrial and microsatellite data. Percentage of total variation among host species, among sample sites (population), and within sample sites are shown.

Source of Variation	mitochondrial	microsatelite
Among host species	7.4	40.2
Among populations	48.8	22.8
Within populations	43.7	37.0

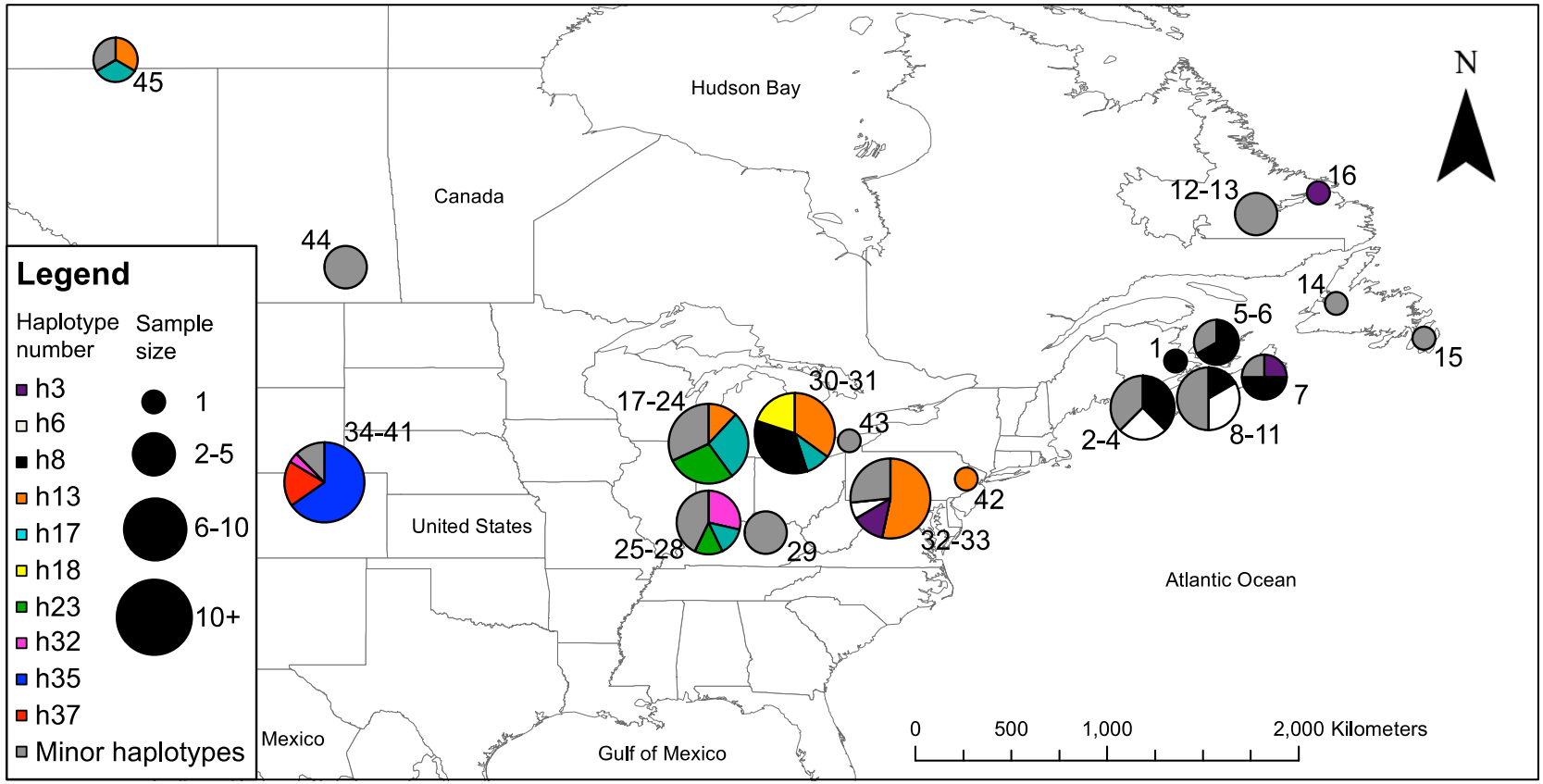

Figure 2.2 Frequencies of mitochondrial DNA haplotypes of *Cimex adjunctus* across its range. Data for nearby sampling sites are combined in a single pie chart. Rare haplotypes represented by fewer than four individuals in the entire data set are shown in grey. Haplotypes represented by four or more individuals in the entire dataset are identified with unique shades as indicated in the Legend, and corresponding to haplotype colors in Fig. 2.3. Site numbers correspond to those in Table 2.2, Table A.1 and Table A.2. Sizes of circles indicate sample sizes.

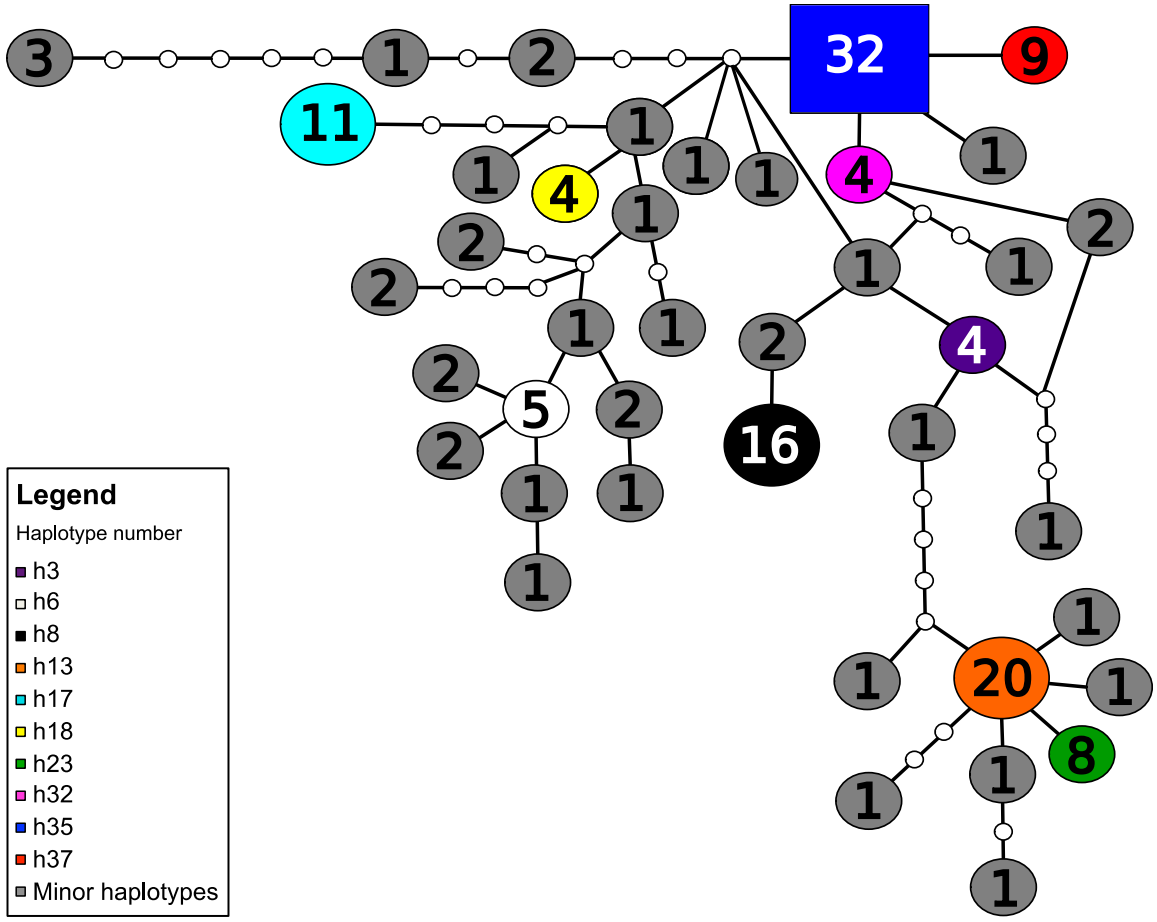

Figure 2.3 Minimum-spanning network of mitochondrial cytochrome c oxidase 1 (*CO1*) sequences of *Cimex adjunctus*. Haplotypes represented by fewer than four individuals are shown in grey. Haplotypes represented by four or more individuals are identified with unique shades as indicated in the Legend, and corresponding to those in Fig. 2.2. Each circle represents a unique sequence, each line segment is a mutational step, numbers are sample sizes for each unique sequence, small circles without a sample size are intermediate, unsampled haplotypes, and the square represents the putative ancestral sequence.

demographic expansion in the past (Fig. 2.4). Evolution of N_{ET} through time, estimated using EBSP, showed mostly constant population size with a possible gradual decrease from about 200,000 to 30,000 years ago to about half of the initial population size, followed by a small increase to the present (Fig. 2.5). Finally, ABC analysis gave strongest support to a scenario mimicking a decrease in effective population size of at least an order of magnitude between 10 million years ago and 10,000 years ago (Table 2.5; See Fig. A.6 for pre-evaluation of prior distributions of scenarios with the observed values).

2.4 Discussion

2.4.1 Range-wide genetic structure

Analyses of mitochondrial and microsatellite genetic markers supported our prediction of high range-wide genetic structure, mediated by geographic distance, in *C. adjunctus*, an ectoparasite of bats. Across the range of *C. adjunctus*, we found significant genetic structure, a large proportion of which was explained by geographic distance. Whereas IBD has not been previously investigated in most bat parasites (but see Olival et al. 2013), it has been investigated in two of the key hosts of *C. adjunctus*, the big brown bat and the little brown myotis. A relationship between genetic and geographic distance has been observed in both the big brown bat (Vonhof et al. 2008) and little brown myotis across a considerably smaller spatial scale (Johnson et al. 2015) than examined here. Range-wide IBD has also been described for little brown myotis (Vonhof et al. 2015), based on population-level analyses using F_{ST}. Thus, geographic distance explains a lot of the variation in genetic structure of

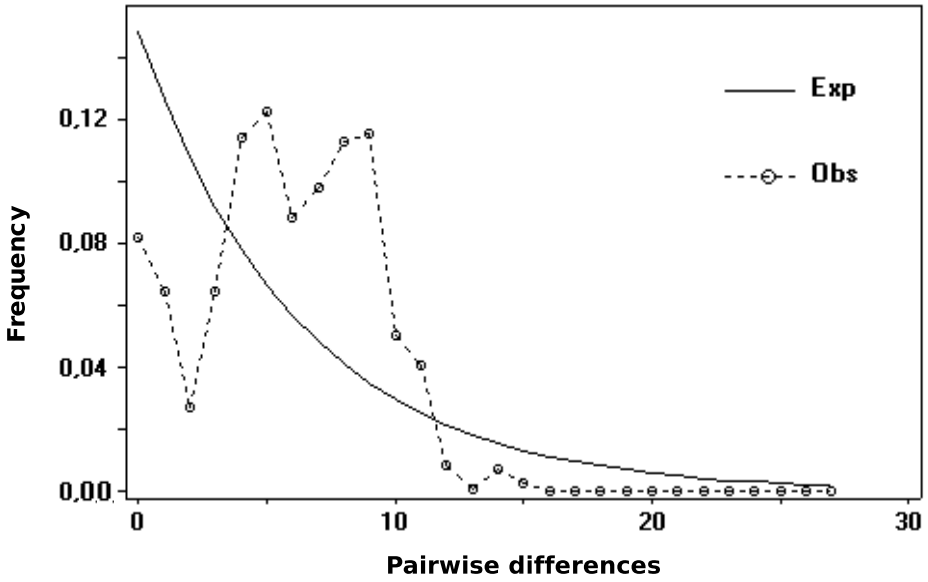

Figure 2.4 Frequency of pairwise mismatches among cytochrome c oxidase 1 (*CO1*) sequences of *Cimex adjunctus* in North America.

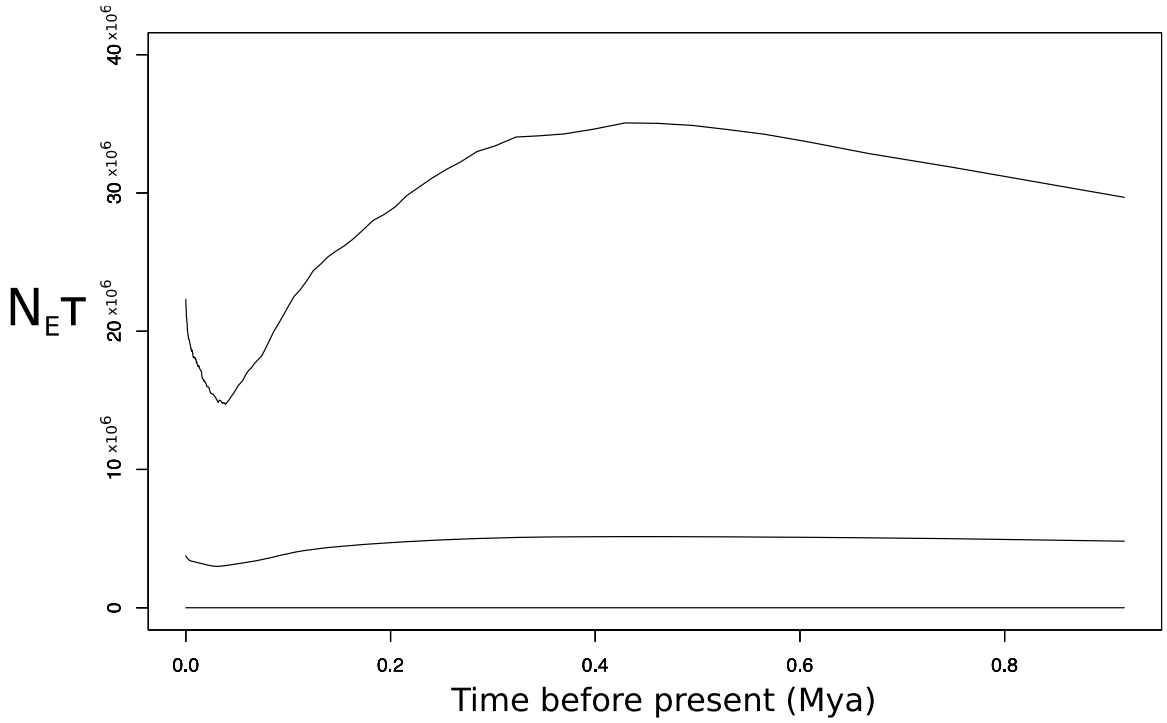

Figure 2.5 Extended Bayesian skyline plot estimated using cytochrome c oxidase 1 (*CO1*) data of *Cimex adjunctus*. Shown are the mean and 95% highest posterior density interval of the product of effective population size (N_E) and generation time (τ) through time (in million years ago; Mya).

Table 2.5 Results of approximate Bayesian computation analysis of effective population size (N_E) history of *Cimex adjunctus*. Posterior probabilities of each scenario (with confidence interval in parentheses) are shown.

Scenario	Posterior probability
N_E Increase	0.297 (0.285 – 0.308)
N_E Decrease	0.522 (0.508 – 0.536)
N_E Constant	0.181 (0.167 – 0.196)

C. adjunctus as it does in two of its hosts, which could potentially reflect the reliance of *C. adjunctus* on their hosts for dispersal.

However, the overall degree of genetic structuring appears to be higher in *C. adjunctus* than in its hosts. Analysis of microsatellite genotypes has revealed only two genetic clusters in both big brown bat (Nadin-Davis et al. 2010) and little brown myotis (Vonhof et al. 2015), both at continental spatial scales, whereas our results point to ten genetic clusters in *C. adjunctus*. Likewise, very little genetic variation (< 10% with microsatellite data, and < 20% with mitochondrial data) occurs among spatially separate sites in big brown bat (Vonhof et al. 2008) and in little brown myotis (Burns et al. 2014; Johnson et al. 2015; McLeod et al. 2015; Vonhof et al. 2015). In *C. adjunctus*, about one third of the microsatellite variation and about one half of mitochondrial variation occur among sites (after taking out variation among host species). These observations suggest that *C. adjunctus* is more subdivided within its range than at least two of its hosts, and that its genetic structure does not entirely reflect the dispersal patterns of its hosts. Interestingly however, both genetic clustering and MSN results also offer some evidence of possible continent-scale long-distance movement in *C. adjunctus*, as reflected in the relationships among individuals from the Northwest Territories, Saskatchewan, Maritime Canada and the US Midwest. Relationships among *C. adjunctus* samples from these locations echo a pattern that was observed in *M. lucifugus*, where a set of sites in the central United States and central to northwestern Canada are connected by high gene flow (Vonhof et al. 2015).

Spatial structuring of genetic diversity can arise when gene flow is not sufficiently high to homogenize allele frequencies throughout the study area, and across a broad range of animal species dispersal ability is correlated with both gene flow and population genetic structure (Bohonak 1999). This has led to the prediction that genetic structure of many parasites will reflect host dispersal and genetic structure (Mazé-Guilmo et al. 2016). However, the association between host dispersal and parasite genetic structure has recently been shown to be generally weak (Mazé-Guilmo et al. 2016). Furthermore, genetic structure in parasites is often found to be stronger than that of their host, as we have observed here for *C. adjunctus*. For example, a finer genetic structure was found in an endoparasitic nematode H. *polygyrus* than in its host, the field mouse *A. sylvaticus* (Nieberding et al. 2004). One reason for stronger genetic structuring in parasites than their hosts could be that, for parasites using their host as a means of dispersal, not every host dispersal event will result in dispersal by the parasite. This is likely to be the case for *C. adjunctus*, which spends a considerable proportion of time living off of its hosts within cracks and crevices in roosting sites. First, only a small subset of dispersing bats is likely to be accompanied by *C. adjunctus.* Second, dispersal mortality in the parasite may be very high due to grooming behaviour of bats that can cause the parasites to fall off (ter Hofstede and Fenton 2005). Additionally, parasites that have a generation time that is much shorter than that of their hosts, that are associated with more than one host species, or that are associated with highly mobile hosts typically show a much stronger genetic structure than their host, as highlighted by Mazé-Guilmo et al.

(2016). All of these factors are true for *C. adjunctus*, and could explain the much stronger genetic structure we observed for relative to two of its key hosts.

In addition to gene flow and dispersal, genetic structure may also be influenced by genetic drift in small populations, which acts to increase differentiation (Levy and Neal 1999). Bat-associated *Cimex* populations might be much smaller than populations of their hosts, although information on *C. adjunctus* population sizes is limited. In addition, it is possible that *C. adjunctus* experiences localized extirpations and recolonizations when roosts are abandoned by bats and subsequently re-occupied. The resulting founder events would further reduce effective population sizes and lead to higher genetic differentiation in *C. adjunctus* via genetic drift.

We also examined the proportion of genetic variance among samples of *C. adjunctus* associated with different host species. Interestingly, we found a sharp difference between mitochondrial DNA and microsatellite markers in this regard. Mitochondrial data suggested considerably less variation among populations associated with different host species compared to microsatellite data. At the same time, microsatellite data showed less variation among populations than did the mitochondrial data, indicating that the difference we observed with respect to host species does not reflect a generally poorer ability of the mitochondrial data to detect differentiation in *C. adjunctus*.

Our mitochondrial data are consistent with an earlier study on *C. pipistrelli* that found no genetic differentiation among individuals associated with different host species, using mitochondrial *CO1* and four nuclear loci (Balvín et al. 2013). Our microsatellite results contradict these results from *C. pipistrelli*,

although it is important to point that all nuclear loci in the study of Balvin et al. (2013) showed almost no variation. Mitochondrial DNA is maternally inherited and will therefore variation in it will reflect dispersal and history of the maternal lineage only. It is possible therefore that sex-biased behaviour in *C. adjunctus* could be the reason for our results. Male-biased dispersal among roosts could lead to the higher proportion of genetic variation among sites in mitochondrial data than in microsatellite data. On the other hand, female-biased switching of hosts within roosts could be responsible for the lower proportion of genetic variation among host species observed in the mitochondrial versus microsatellite data. Autonomous (*i.e.*, not host-assisted) female-biased movements over short distances, such as between neighbouring apartment units, have been described in the common bed bug, *C. lectularius* (Cooper et al. 2015). If female *C. adjunctus* also move more readily at short distances within roosts, that could explain both a higher rate of host-switching among females and a lower rate of transport among roosts by their hosts (since females might spend more time off of the hosts while they engage in exploratory behaviour). However, there is currently no information available on sex-biased dispersal or host switching in *C. adjunctus*. Our results not only suggest sex-biased dispersal or host switching in *C. adjunctus*, but also highlight the need to use more than one type of marker when investigating genetic diversity in an understudied species.

The most well studied member of the genus *Cimex* is the human associated common bed bug, *C. lectularius*. Several studies have examined genetic structure in *C. lectularius* across a range of spatial scales (Balvín et al. 2012; Booth et al. 2012; Saenz et al. 2012; Fountain et al. 2014; Booth et al.

2015). However, most such studies focus on a considerably smaller scale than we do here, making direct comparisons of genetic structure difficult. For example, Saenz et al. (2012) describe a weaker IBD pattern in *C. lectularius* than we observed for *C. adjunctus*, which could be due in part to the smaller spatial scale of their sampling (eastern United States only). On the other hand, our genetic diversity estimates for *C. adjunctus* were strikingly similar to those found in one study on *C. lectularius* (Booth et al. 2012), although we report slightly higher average numbers of alleles. In an interesting parallel, a study of *C. lectularius* populations associated with bats and humans found higher average numbers of alleles in the bat-associated populations than human-associated populations (Booth et al. 2015). Another study of *C. lectularius* in Europe (Balvín et al. 2012) found higher mitochondrial DNA variation among bat and human associated populations than we observed among populations of *C. adjunctus* associated with different bat species. One likely reason for this dissimilarity between *C. adjunctus* and *C. lectularius* is that the former is a weak generalist, associated with closely related species (Mazé-Guilmo et al. 2016), while the former is a strong generalist, associated with phylogenetically very different species. Overall, sample sizes and the number of microsatellite markers used were lower in our study than in several studies of *C. lectularius* genetic structure (Booth et al. 2012; Fountain et al. 2014; Booth et al. 2015), but were nonetheless appropriate given the much broader spatial and temporal scale of resolution of our analyses (Nieberding et al. 2004; Criscione et al. 2006; James et al. 2011; van der Mescht et al. 2015).

2.4.2 Demographic history

We predicted signals of range and demographic expansion in *C. adjunctus*, based on the fact that there are widespread signatures of historic population expansion in many vertebrates, invertebrates and plant populations, including in the bat hosts of *C. adjunctus*. Such patterns are most probably attributable to postglacial climate warming (Grant 2015), However, we found that the history of this ectoparasite is marked most strongly by demographic decline, with only a weak signal of recent demographic expansion, and no clear pattern of range expansion. For example, typical starburst patterns were previously observed in the haplotypic networks of *E. fuscus* and *M. lucifugus* (Turmelle et al. 2011; Burns et al. 2014; McLeod et al. 2015), indicative of range expansion. However, we found no clear starburst pattern for *C. adjunctus*. This is unlikely to be a result of inadequate spatial sampling since our samples cover most of the known range of this species (Usinger 1966).

We also found evidence of population decline in the demographic history of *C. adjunctus* using a variety of approaches. According to EBSP results, a gradual decline might have started at around 200,000 years ago, corresponding roughly to the Illinoian glaciation, a time of likely very harsh climate for most species in North America (Swenson and Howard 2005). A small demographic recovery may have started at around 30,000 years ago. Our ABC results confirmed a population decline as the most likely historical scenario. Two previous studies found signals of demographic expansion in *M. lucifugus* in eastern Canada (Burns et al. 2014) and Minnesota, United States (Dixon 2011b). A small potential increase in *C. adjunctus* effective population size indicated in the EBSP starting

30,000 years ago is in a similar timeframe as, but is of much smaller amplitude than, the demographic expansion found in both *M. lucifugus* studies. Relative to those studies, our analysis was able to span a larger amount of time, probably due to the larger spatial scale of our sampling.

2.4.3 Conclusions

Parasites that are mostly free-living and have hosts that are highly mobile, such as ectoparasites of bats, may be expected to show a genetic structure that contrasts with the dispersal patterns and genetic structure of their hosts (Mazé-Guilmo et al. 2016). These same factors may also lead to a difference in historic patterns of change in host and parasite ranges and population sizes. We have found exactly this pattern in *C. adjunctus*, an insect ectoparasite associated with a number of bat species in North America. This free-living parasite moves off the host between blood meals and could be actively removed by the host through anti-parasitism behaviour. Our results highlight that the genetic structure and demographic history of a weak generalist ectoparasite, particularly one that has a loose relationship with its hosts, can be very different from that of its hosts.

2.5 Literature Cited

Adam MD, Hayes JP (2000) Use of bridges as night roosts by bats in the Oregon Coast Range. *Journal of Mammalogy*, **81**, 402–407.

Anderson RM, Gordon DM (1982) Processes influencing the distribution of parasite numbers within host populations with special emphasis on parasite-induced host mortalities. *Parasitology*, **85**, 373–398.

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Balvín O, Roth S, Vilímová J (2015) Molecular evidence places the swallow bug genus *Oeciacus* Stål within the bat and bed bug genus *Cimex* Linnaeus (Heteroptera: Cimicidae). *Systematic Entomology*, **40**, 652–665.
Balvín O, Vilímová J, Kratochvíl L (2013) Batbugs (*Cimex pipistrelli* group, Heteroptera: Cimicidae) are morphologically, but not genetically differentiated among bat hosts. *Journal of Zoological Systematics and Evolutionary Research*, **51**, 287–295.
Bargues MD, Klisiowicz DR, Panzera F et al. (2006) Origin and phylogeography of the Chagas disease main vector *Triatoma infestans* based on nuclear rDNA sequences and genome size. *Infection, Genetics and Evolution*, **6**, 46–62.
Bartonička T, Růžičková L (2013) Recolonization of bat roost by bat bugs (*Cimex pipistrelli*): could parasite load be a cause of bat roost switching? *Parasitology Research*, **112**, 1615–1622.
Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.
Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, **141**, 1007–1014.
Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.
Booth W, Balvín O, Vargo EL, Vilímová J, Schal C (2015) Host association drives genetic divergence in the bed bug, *Cimex lectularius*. *Molecular Ecology*, **24**, 980–992.
Booth W, Saenz VL, Santangelo RG et al. (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.
Brower AV (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6491–6495.
Brown RP, Yang Z (2011) Rate variation and estimation of divergence times using strict and relaxed clocks. *BMC Evolutionary Biology*, **11**, 271.
Burns LE, Frasier TR, Broders HG (2014) Genetic connectivity among swarming sites in the wide ranging and recently declining little brown bat (*Myotis lucifugus*). *Ecology and Evolution*, **4**, 4130–4149.
Chakraborty D, Sinha A, Ramakrishnan U (2014) Mixed fortunes: ancient expansion and recent decline in population size of a subtropical montane primate, the Arunachal macaque *Macaca munzala*. *Plos One*, **9**, e97061.
Cooper R, Wang C, Singh N (2015) Mark-release-recapture reveals extensive movement of bed bugs (*Cimex lectularius* L.) within and between apartments. *Plos One*, **10**, e0136462.
Criado PR, Belda Junior W, Criado RFJ, Vasconcelos e Silva R, Vasconcellos C (2011) Bedbugs (Cimicidae infestation): the worldwide renaissance of an old partner of human kind. *The Brazilian journal of infectious diseases*, **15**, 74–80.

Criscione CD, Cooper B, Blouin MS (2006) Parasite genotypes identify source populations of migratory fish more accurately than fish genotypes. *Ecology*, **87**, 823–828.
Dick CW, Patterson BD (2007) Against all odds: explaining high host specificity in dispersal-prone parasites. *International Journal for Parasitology*, **37**, 871– 876.
Dixon MD (2011a) Post-Pleistocene range expansion of the recently imperiled eastern little brown bat (*Myotis lucifugus lucifugus*) from a single southern refugium. *Ecology and Evolution*, **1**, 191–200.
Dixon MD (2011b) Population genetic structure and natal philopatry in the widespread North American bat *Myotis lucifugus*. *Journal of Mammalogy*, **92**, 1343–1351.
Drummond AJ, Suchard MA (2010) Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, **8**, 114.
Ellison LE, O'Shea TJ, Neubaum DJ, Bowen RA (2007) Factors influencing movement probabilities of big brown bats (*Eptesicus fuscus*) in buildings. *Ecological applications: a publication of the Ecological Society of America*, **17**, 620–627.
Engelbrecht A, Matthee S, du Toit N, Matthee CA (2016) Limited dispersal in an ectoparasitic mite, *Laelaps giganteus* , contributes to significant phylogeographic congruence with the rodent host, *Rhabdomys*. *Molecular Ecology*, **25**, 1006–1021.
Fields Development Team (2006) *Fields: tools for spatial data*. National Center for Atmospheric Research, Boulder, CO, USA.
Fountain T, Duvaux L, Horsburgh G, Reinhardt K, Butlin RK (2014) Human-facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*. *Molecular Ecology*, **23**, 1071–1084.
Furlonger CL, Dewar HJ, Fenton MB (1987) Habitat use by foraging insectivorous bats. *Canadian Journal of Zoology*, **65**, 284–288.
Goddard J (2009) Bed bugs (*Cimex lectularius*) and clinical consequences of their bites. *The Journal of the American Medical Association*, **301**, 1358–1366.
Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**.
Grant WS (2015) Problems and cautions with sequence Mismatch Analysis and Bayesian Skyline Plots to infer historical demography. *Journal of Heredity*, **106**, 333–346.
Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.
Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.
Heller R, Chikhi L, Siegismund HR (2013) The confounding effect of population structure on Bayesian Skyline Plot inferences of demographic history. *Plos One*, **8**, e62992.

ter Hofstede HM, Fenton MB (2005) Relationships between roost preferences, ectoparasite density, and grooming behaviour of neotropical bats. *Journal of Zoology*, **266**, 333–340.
James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH (2011) Spatial genetic structure of a symbiotic beetle-fungal system: toward multi-taxa integrated landscape genetics. *Plos One*, **6**, e25359.
Johnson LNL, McLeod BA, Burns LE et al. (2015) Population genetic structure within and among seasonal site types in the little brown bat (*Myotis lucifugus*) and the northern long-eared bat (*M. septentrionalis*). *Plos One*, **10**, e0126309.
Levy F, Neal CL (1999) Spatial and temporal genetic structure in chloroplast and allozyme markers in *Phacelia dubia* implicate genetic drift. *Heredity*, **82**, 422–431.
Li B, Kimmel M (2013) Factors influencing ascertainment bias of microsatellite allele sizes: impact on estimates of mutation rates. *Genetics*, **195**, 563–572.
Lombaert E, Guillemaud T, Thomas CE et al. (2011) Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: case study of the invasive ladybird *Harmonia axyridis*. *Molecular Ecology*, **20**, 4654–4670.
Marshall AG (1981) *The ecology of ectoparasitic insects*. Academic Press, London, UK.
Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G (2016) Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, **19**, 336–347.
McCoy KD, Boulinier T, Tirard C, Michalakis Y (2003) Host-dependant genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution*, **57**, 288–296.
McCoy KD, Chapuis E, Tirard C et al. (2005) Recurrent evolution of host-specialized races in a globally distributed parasite. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2389–2395.
McLeod BA, Burns LE, Frasier TR, Broders HG (2015) Effect of oceanic straits on gene flow in the recently endangered little brown bat (*Myotis lucifugus*) in maritime Canada: implications for the spread of white-nose syndrome. *Canadian Journal of Zoology*, **93**, 427–437.
van der Mescht L, Matthee S, Matthee CA (2015) Comparative phylogeography between two generalist flea species reveal a complex interaction between parasite life history and host vicariance: parasite-host association matters. *BMC Evolutionary Biology*, **15**.
Nadin-Davis SA, Feng Y, Mousse D, Wandeler AI, Aris-Brosou S (2010) Spatial and temporal dynamics of rabies virus variants in big brown bat populations across Canada: footprints of an emerging zoonosis. *Molecular Ecology*, **19**, 2120–2136.
Nieberding CM, Durette-Desset M-C, Vanderpoorten A et al. (2008) Geography and host biogeography matter for understanding the phylogeography of a parasite. *Molecular Phylogenetics and Evolution*, **47**, 538–554.

Nieberding C, Morand S, Libois R, Michaux JR (2004) A parasite reveals cryptic phylogeographic history of its host. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2559–2568.
Olival KJ, Dick CW, Simmons NB et al. (2013) Lack of population genetic structure and host specificity in the bat fly, *Cyclopodia horsfieldi*, across species of Pteropus bats in Southeast Asia. *Parasites and Vectors*, **6**, 231–248.
Pearce RD, O'Shea TJ (2007) Ectoparasites in an urban population of big brown bats (*Eptesicus fuscus*) in Colorado. *The Journal of parasitology*, **93**, 518–530.
Pereira L, Dupanloup I, Rosser ZH, Jobling MA, Barbujani G (2001) Y-chromosome mismatch distributions in Europe. *Molecular Biology and Evolution*, **18**, 1259–1271.
Poulin R (1999) The functional importance of parasites in animal communities: many roles at many levels? *International Journal for Parasitology*, **29**, 903–914.
Pulgarín-R PC, Burg TM (2012) Genetic signals of demographic expansion in downy woodpecker (*Picoides pubescens*) after the last North American glacial maximum. *Plos One*, **7**, e40412.
Saenz VL, Booth W, Schal C, Vargo EL (2012) Genetic analysis of bed bug populations reveals small propagule size within individual infestations but high genetic diversity across infestations from the eastern United States. *Journal of Medical Entomology*, **49**, 865–875.
Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.
Thrall PH, Burdon JJ (1997) Host-pathogen dynamics in a metapopulation context: the ecological and evolutionary consequences of being spatial. *The Journal of Ecology*, **85**, 743–753.
Tomisawa R, Akimoto S (2004) Host range and host preference of a flea weevil, *Orchestes hustachei*, parasitizing aphid galls. *Entomological Science*, **7**, 21–30.
Turmelle AS, Kunz TH, Sorenson MD (2011) A tale of two genomes: contrasting patterns of phylogeographic structure in a widely distributed bat. *Molecular Ecology*, **20**, 357–375.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Vonhof MJ, Russell AL, Miller-Butterworth CM (2015) Range-wide genetic analysis of little brown bat (*Myotis lucifugus*) populations: estimating the risk of spread of White-Nose Syndrome. *Plos One*, **10**, e0128713.
Vonhof MJ, Strobeck C, Fenton MB (2008) Genetic variation and population structure in big brown bats (*Eptesicus fuscus*): is female dispersal important? *Journal of Mammalogy*, **89**, 1411–1419.
Wang J (2002) An estimator for pairwise relatedness using molecular markers. *Genetics*, **160**, 1203–1215.
Wang J (2011) Unbiased relatedness estimation in structured populations. *Genetics*, **187**, 887–901.

Yuan M-L, Wei D-D, Zhang K et al. (2010) Genetic diversity and population structure of *Panonychus citri* (Acari: Tetranychidae), in China based on mitochondrial COI gene sequences. *Journal of Economic Entomology*, **103**, 2204–2213.
Zachos J (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.

3 Population structure in two geographically sympatric and congeneric ectoparasites: *Cimex adjunctus* and *Cimex lectularius*, in the North American Great Lakes region

3.1 Introduction

Natural populations are often subdivided, most commonly as a result of landscape heterogeneity (Storfer et al. 2007). Individuals may move from one patch to the other, but usually will not settle or breed in intervening areas. Different species often exhibit contrasting levels of connectivity among subpopulations, as well as different local dynamics (Mimet et al. 2013). These, in turn, affect population persistence, and genetic diversity and differentiation (Neel 2008, Andreakis et al. 2009). The fundamental organisation and dynamics of spatially subdivided populations are described by models (Harrison 1991) that provide predictions of population genetic structure and differentiation (Mayer et al. 2009; Fig. 3.1). At one end of a continuum of population structure, patchy populations are characterised by high connectivity among subpopulations, effectively constituting a single panmictic population. Genetic differentiation among subpopulations in a patchy population is essentially non-existent, and the subpopulations would be expected to form a single genetic cluster and not display isolation-by-distance (IBD). At the other extreme, in non-equilibrium metapopulations, subpopulations are disconnected from each other. Non-equilibrium metapopulations are characterized by high differentiation among subpopulations, with almost all subpopulations predicted to each form a separate genetic cluster. IBD is also not expected in this case because of the predominance

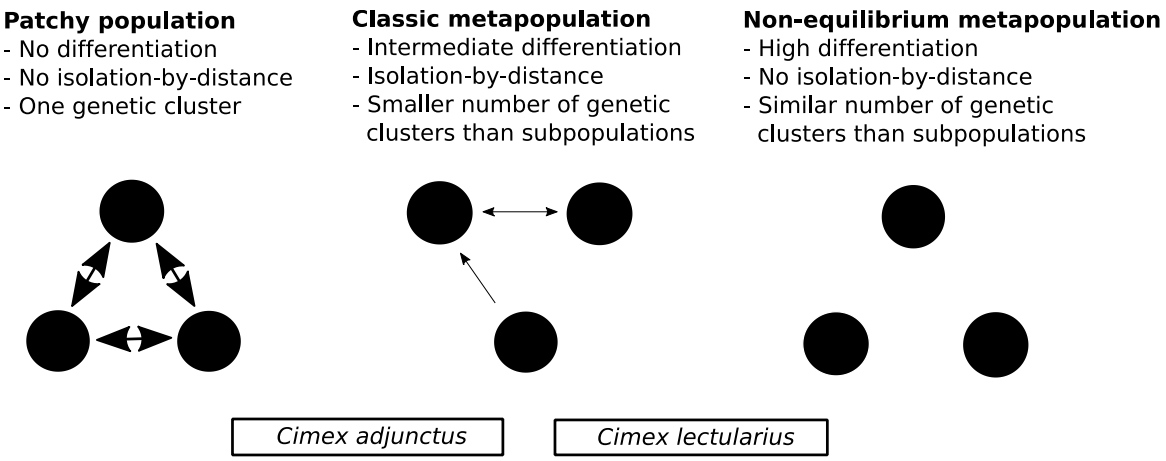

Figure 3.1 Schematic to illustrate a continuum among three population structure models on which we based our study (from left to right): patchy population, classic metapopulation and non-equilibrium metapopulations (based on Mayer et al. 2009); and illustration of where *C. adjunctus* and *C. lectularius* might lie in the continuum, following our study's results.

of genetic drift over gene flow in determining genetic differentiation (Hutchison and Templeton 1999). Finally, metapopulations are intermediate in a continuum between patchy and isolated populations. Finally, classic metapopulations are intermediate in the continuum between patchy populations and non-equilibrium metapopulations (Hanski 1998). Classic metapopulations have some, but limited, connectivity among subpopulations, and connectivity is usually distance dependent. In classic metapopulations genetic differentiation is intermediate between patchy populations and non-equilibrium metapopulations; there should be several distinct genetic clusters, but the number of such clusters is expected to be less than the number of occupied habitat patches. Also, because more distant subpopulations are connected by lower dispersal and gene flow, isolation-by-distance should be present in a classic metapopulation (Bohonak 1999), but not in populations following the other two models (Mayer et al. 2009). Real populations in nature may show characteristics of a classic metapopulation in combination with some attributes of either a patchy population or a non-equilibrium metapopulation (Rasic and Keyghobadi 2012).

Understanding population structure is important because it predicts regional dynamics and persistence (Harrison 1991; Hanski 1998). However, key ecological variables determining population structure may be difficult to quantify directly. For example, movements among subpopulations can be very be difficult to track in species that are small or cryptic, such as ectoparasites. Because each population model makes specific genetic predictions (Mayer et al. 2009), a snapshot of patterns of genetic differentiation can provide an assessment of population structure. In our study, we used genetic data to investigate population

structure of two congeneric parasite species living in sympatry, but associated with different hosts, in the Great Lakes region of North America.

The genus *Cimex* (Order: Hemiptera, Class: Insecta) is characterized by species that are temporary ectoparasites of warm-blooded animals, mostly bats. Species of *Cimex* typically remain in the hosts' roosts, emerging from cracks in the walls only to obtain blood meals (Usinger 1966; Cooper et al. 2015). *Cimex* species are hypothesized to have low inherent capacity for dispersal between contiguous structures, rather depending on their hosts for dispersal (Usinger 1966). In central and eastern North America, *Cimex adjunctus* is a widespread ectoparasite of North American bats, although it is also known to bite people visiting or residing near bat roosting sites (Goddard et al. 2012). This species occurs from the east coast to the Rocky Mountains, and from Labrador and Northwest Territories to Texas (Usinger 1966). Talbot et al. (2016; Chapter 2) found high levels of continent-wide spatial genetic structure in *C. adjunctus*, although with evidence of multiple potential instances of long-distance dispersal. The big brown bat (*Eptesicus fuscus*) and the little brown myotis (*Myotis lucifugus*) are two key hosts of *C. adjunctus* that frequently roost in buildings (Furlonger et al. 1987; Ellison et al. 2007; Pearce and O'Shea 2007). The big brown bat is known to frequently switch roosts during the summer due to temperature and parasite density (Ellison et al. 2007). The common bed bug (*Cimex lectularius*) is a congener that is a public health concern in many countries (Goddard 2009; Criado et al. 2011). This ectoparasite feeds primarily on humans and is most commonly found associated with humans in their dwellings, although it is also known to feed on a range of other animals, including chickens and bats (Usinger 1966). *C.*

lectularius is regularly observed in association with bats in Europe (Balvín et al. 2012; Booth et al. 2015), although has never been recorded with bats in North America. Pesticides have been used for many decades on bed bug infestations around the world. The effect of DDT was particularly strong on *C. lectularius* populations, effectively eliminating them from households (Adelman et al. 2011). Use of DDT on populations of other *Cimex* species that associate with bats was likely not as intense or widespread as it was for *C. lectularius* populations associated with humans. In recent years, *C. lectularius* has experienced resurgence in many parts of the world (Davies et al. 2012).

In *C. lectularius,* limited human-mediated gene flow and colonization, along with local extinctions driven by pest control practices, result in a mixture of classic and non-equilibrium metapopulation attributes (Harrison 1991, Fountain et al. 2014). There is significant genetic differentiation among infestation locations, and typically either no or very weak isolation-by-distance (Saenz et al. 2012). Interestingly, genetic differentiation indices (F_{ST}) are much higher among human-associated *C. lectularius* subpopulations than among bat-associated conspecifics collected from roosts (Booth et al. 2015), indicating an effect of host-association on population structure. While the genetic attributes and population structure of *C. lectularius* have been addressed in several studies (Booth et al. 2012; Saenz et al. 2012; Fountain et al. 2014; Booth et al. 2015), the characteristics of populations of other *Cimex* species that typically associate with bats have received little attention (but see Talbot et al. 2016, Chapter 2). An understanding of the structure and genetics of populations of these insects could provide insight into their propagation and potential impact on bat populations. *C. adjunctus* is of particular importance

because it is the most widespread cimicid parasite of bats in North America (Usinger 1966), and its key hosts, the big brown bat and the little brown myotis, are currently threatened by the fungus causing White-Nose Syndrome (Blehert et al. 2009) .

We first investigated the population structure of *C. adjunctus*, an understudied and fairly abundant species of *Cimex* in North America. *C. adjunctus* is concentrated in bat roosts, usually within man-made structures. We then wanted to compare population structure of *C. adjunctus* with that of its well-known and sympatric congener, *C. lectularius*. We therefore examined household infestations of *C. lectularius* in the same geographic area and at a similar spatial scale. We predicted that population structure would differ between the two *Cimex* species and that *C. lectularius* would show higher levels of genetic structure and differentiation than *C. adjunctus* due to possibly more limited movement and more frequent extinctions.

3.2 Materials and Methods

3.2.1 Data collection

We examined *C. adjunctus* population structure over a spatial scale that is representative of contemporary dispersal of their hosts (Penczykowski et al. 2016). Population structure at a larger spatial scale (e.g., range wide) could reflect historic demographic processes, such as post-glacial recolonizations and secondary contact zones (Swenson and Howard 2005), and has been described in Talbot et al. (2016; Chapter 2). From a larger *C. adjunctus* dataset (Talbot et al. 2016; Chapter 2), we selected individuals from sites around the Great Lakes

region of the United States and Canada (Fig. 3.2) with a median distance corresponding to the estimated mean translocation distances in bat hosts (Norquay et al. 2013). Most samples are from mist-netted host individuals of *E. fuscus* and *M. lucifugus* between 2005 and 2014 (Table B.1). Mist net capture locations were adjacent to a known summer roost (house, barn, church, or school) of either of the two bat species, or within forested provincial or state lands (Table B.1). Most mist-netted bats and the *C. adjunctus* individuals they harboured likely came from the adjacent known roost, although it is possible that a small proportion came from different roosts in the area. Overall, between 3 and 15% of mist-netted bats harboured a parasite, depending on the location. We also sampled *C. adjunctus* individuals from the interior of two summer roosts. One roost was in a church attic inhabited by *M. lucifugus*, and one was in a house attic inhabited by *E. fuscus* (Table B.1). Because we could be certain of the roost site in these cases, we considered these two sampling locations as distinct from their adjacent mist netting capture locations. Talbot et al. (2016; Chapter 2) showed, in a range-wide study of *C. adjunctus* that included the samples used in this study, that there is limited or no effect of sampling year or host species on genetic clustering.

We collected *C. lectularius* samples from infested homes in the same geographic area (Fig. 3.2) and at a comparable spatial scale, with the help of Abell Pest Control Inc. (Toronto, Ontario, Canada) in 2014 (Table B.2). Due to privacy reasons, we only obtained postal codes for each *C. lectularius* individual, from which we obtained centroid geographical coordinates in WGS84 datum using the public CivicSpace USA ZIP Code Database (Schuyler Erle, CivicSpace Labs Inc., San Francisco, California, United States) and the crowd-sourced

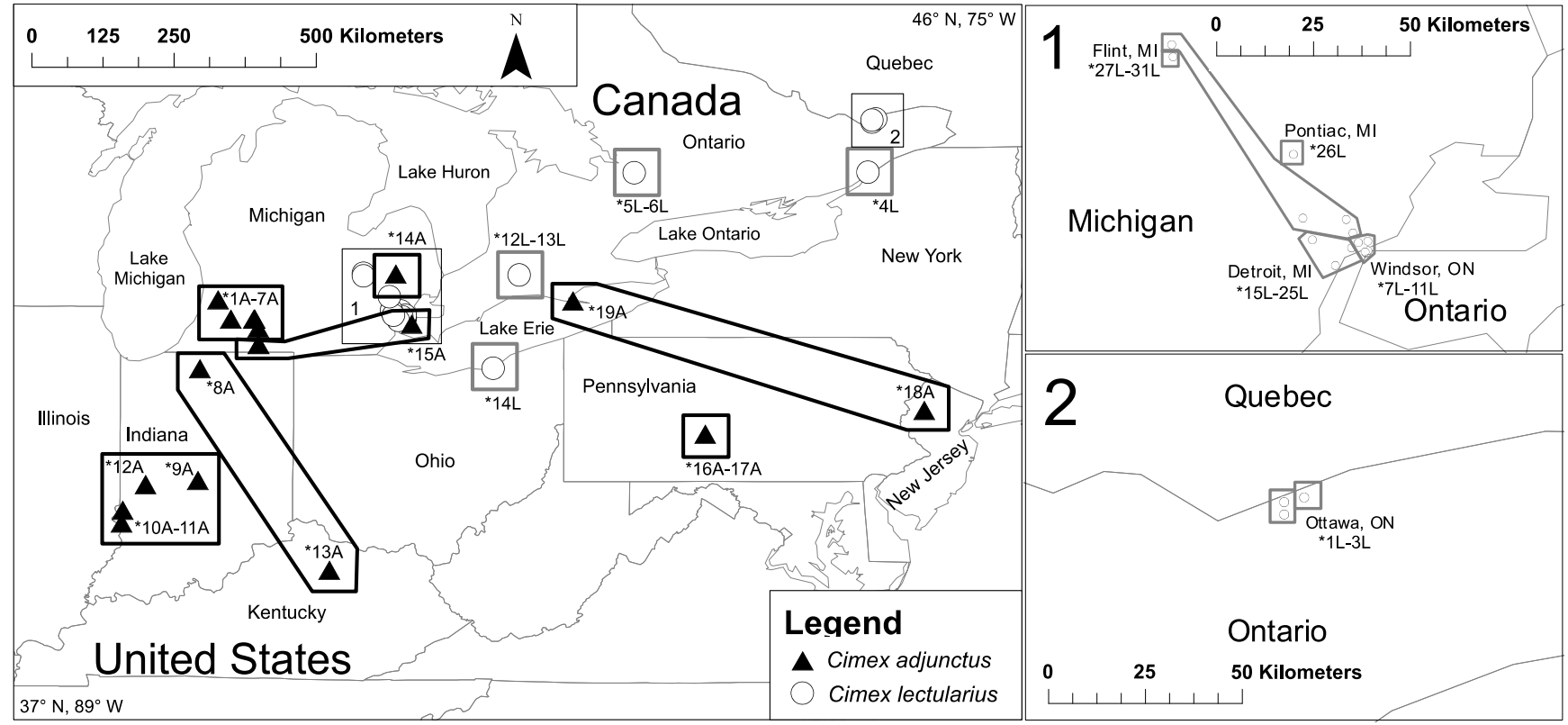

Figure 3.2 Study area showing locations of 75 *Cimex adjunctus* individuals collected from bats and bat roosts (black triangles), and 73 *Cimex lectularius* individuals collected in infested housing units (white circles), collected in the Great Lakes region of North America, with two close-ups around the region of Detroit, MI, and Ottawa, ON. Black buffers in *C. adjunctus* and grey buffers in *C. lectularius* represent genetic clusters for both species. Numbers preceded by an asterisk on the map correspond to site numbers in Table 3.1, Table B.1 and Table B.2, where numbers followed by an A correspond to *C. adjunctus* and numbers followed by an L to *C. lectularius*.

Canadian Postal Code Geocoded Database (Geocoder.ca, Geolytica Inc., Ottawa, Ontario, Canada). We sampled between one and four housing units per postal code, and we considered separate households in the same postal code as different sites. Upon collection, we immediately stored each sample individually in 95% ethanol until further analyses.

3.2.2 Genetic analyses

We used *C. adjunctus* genotypes at nine microsatellite loci, originally developed for *C. lectularius* (Cle002, Cle003, Cle013, Cle015 from Fountain et al. 2014; Clec15, Clec21, Clec48, Clec104 and BB28B, from Booth et al. 2012), from Talbot et al. (2016; Chapter 2). We extracted DNA from the whole insect for all *C. lectularius* samples using the DNeasy Blood and Tissue Kit (QIAGEN, Germantown, Maryland, United States). We amplified 20 microsatellite loci designed specifically for that species, and including the nine loci also used for *C. adjunctus* (Cle011 and Cle021, from Fountain et al. 2014; Clec11, Clec45, Clec96, Clec97, Clec99, BB21B, BB29B, BB31B and BB38B, from Booth et al. 2012). Amplifying a larger number of microsatellite loci in *C. lectularius* allowed us to compare the statistical resolution between using a smaller versus a larger set of markers (Table B.1). For all other analyses, we only used microsatellite markers that amplified for both *C. adjunctus* and *C. lectularius*.

We used a DNAEngine PTC-200 Thermal Cycler (Bio-Rad, Hercules, California, United States) to execute the polymerase chain reaction (PCR) amplification of *C. lectularius* samples. We performed PCR using the same protocols as in Booth et al. (2012) and Fountain et al. (2014) for the loci developed

by each study, respectively. We visualized PCR products by 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad, Hercules, California, United States) on a UV transilluminator to check the quality and size of amplified fragments. We then sized products on a 3730xl DNA Analyzer (Applied Biosystems, Foster City, California, United States; ABI). We called all microsatellite genotypes for each species using GeneMapper Software v.4.0 (ABI), and we checked all calls manually.

3.2.3 Statistical analyses

Hardy-Weinberg and linkage disequilibrium, genetic diversity, and pedigree analysis

For microsatellite loci, we tested for Hardy-Weinberg and linkage disequilibrium within each site that had data from more than one sampled individual, using Genepop 4.2. For each type of test, we corrected for multiple tests using Bonferroni correction, with a threshold α of 0.05. For sites with data from at least three sampled individuals, we calculated average number of alleles, expected and observed heterozygosity, and inbreeding coefficient. Sampling individuals from the same family group (*i.e.*, associated either by a parent-offspring or a full sibling relationship) can influence results of population genetic analyses (Goldberg and Waits 2010), and sampling of family groups has been reported in *C. lectularius* (Saenz et al. 2012). Therefore, we wanted to determine the relationship of each pair of individuals in the dataset, to ascertain that relatedness is not a source of bias in our comparison of genetic structure between the two species. To this end, we used ML-Relate, which provides

maximum likelihood estimates of relationship, to assess the proportion of all pairs of individuals, in each species, that are between individuals that are related as either parent-offspring or full-siblings and were collected from the same sampling site.

Individual-level analyses

We then tested our prediction of differing population structure between *C. adjunctus* and *C. lectularius*. To facilitate meaningful comparison between the two parasite species, we analyzed *C. lectularius* at the same seven markers as *C. adjunctus*. As a secondary analysis, we also examined *C. lectularius* at the whole panel of 20 markers, to ascertain whether using fewer markers had any effect on the results. We applied most analyses at the individual level, due to the fact that a large part of our dataset is composed of sites with only one individual sampled.

First, we looked for evidence of genetic clustering and isolation-by-distance (IBD). We conducted a Bayesian clustering analysis using Geneland 4.0.5, which takes into account geographic coordinates of samples. We used 10,000,000 iterations, thinned every 1,000th iteration, and a post-process burn-in of 2,000 after thinning, for *K* values between 1 and 20. We executed 10 runs, and kept the one with the higher posterior mean density, after burn-in. We attempted to identify the population to which each individual was assigned the most often, defined here as the population where the majority of Markov Chain Monte Carlo (MCMC) chains converged for any given individual. We also conducted a *K*-Means clustering analysis using GenoDive 2.0 (Meirmans 2012) on allele frequencies, for *K* values between 1 and 20, and using 50,000 simulation steps,

to validate results obtained with the Geneland method. We used Bayesian Information Criterion (BIC) values to determine the most likely *K* value.

Next, we conducted an individual-level analysis of IBD, using the estimate of genetic relatedness, r_W (Wang 2002), calculated with SpaGeDi 1.5. We calculated $1 - r_W$ for each pairwise relationship to obtain genetic distances. We calculated geographic distance (in km) between sample sites, corrected for sphericity of the earth, using the 'rdist.earth' function from the 'fields' package (Fields Development Team 2006) in R 3.1.3. We then fit pairwise genetic distance to geographic distance using multiple regression on distance matrices (MRM), in the 'MRM' function from the 'ecodist' package (Goslee and Urban 2007) in R 3.1.3, which uses a Mantel test derived linear regression model. We assessed significance through a permutation procedure (9,999 replicates) that takes into account non-independence of data points in distance matrices (Legendre et al. 1994; Lichstein 2007). An assumption of the r_W relatedness index is that individuals are in a large random mating population without population structure (Wang 2011). To correct for the population structure present in our dataset, we subsequently conditioned IBD models for genetic clustering; for each pair of individuals assigned to the same population in clustering analyses, we assigned a value of 0, and for each pair of individuals assigned to different sites, we assigned a value of 1, and then tested the effect of geographic distance, together with genetic clustering (defined as a 0/1 pairwise matrix), on genetic distance in an MRM model.

Site-level analyses

We also used some site-level analyses to complement results obtained at the individual level, using sites with at least three sampled individuals. A site is defined in our study as a single housing unit, for *C. lectularius*, or a single capture location or roost, for *C. adjunctus*. For all sites with more than one individual sampled, we conducted an Analysis of Molecular Variance (AMOVA), using GenoDive 2.0. We calculated expected heterozygosity averaged across sites and Hedrick's global G'_{ST} among sites (Hedrick 2005), also using GenoDive 2.0. G'_{ST} provides estimates of genetic differentiation that can be more meaningfully compared between species with different levels of genetic diversity (Meirmans and Hedrick 2011).

3.3 Results

We selected 75 individuals from the *C. adjunctus* dataset of Talbot et al. (2016; Chapter 2) from sites an average of 387 km apart (median = 278 km, between 0 and 1413 km); 54 of those were from the body of a bat, and 21 from the interior of a roost (Table B.1). We sampled between one and six *C. lectularius* individuals, the common bed bug, at infested housing units in the same region (Table B.2), leading to a collection of 73 individual *C. lectularius*, at an average of 373 km apart (median = 205 km, between 0 and 903 km). Genotype data for *C. adjunctus* is available in Talbot et al. (2016; Table A.4). Genotype data for *C. lectularius* for the whole panel of 20 microsatellite markers is available in Table B.3. Excluding sites with only one individual sampled (7 in *C. adjunctus*; 13 in *C.*

lectularius), average sample size per site was five in *C. adjunctus* and three in *C. lectularius* (over 12 sites in *C. adjunctus* and 18 sites in *C. lectularius*).

3.3.1 Genetic diversity, Hardy-Weinberg and linkage disequilibrium, and pedigree analysis

Two of the nine microsatellite markers used in *C. adjunctus* were monoallelic and we therefore excluded them from the analyses on both *C. adjunctus* and *C. lectularius*. Among the remaining seven microsatellite loci, in *C. adjunctus* and *C. lectularius* respectively, we observed between three and 16, and between two and 16 alleles. Respectively for *C. adjunctus* and *C. lectularius*, average number of alleles ranged from 1.7 to 3.6 and from 1.1 to 2.8, observed heterozygosity ranged from 0.11 to 0.36 and from 0.10 and 0.67, expected heterozygosity ranged from 0.33 to 0.50 and from 0.07 to 0.63, and the inbreeding coefficient varied between -0.01 and 0.75 and between -0.67 and 0.37 (Table 3.1). Inbreeding coefficients in *C. lectularius* were centered around zero, with a mean of -0.05. Inbreeding coefficients in *C. adjunctus* were higher, with mostly positive values and a mean of 0.41. We found three significant cases of deviation from Hardy-Weinberg equilibrium, characterized by homozygote excess, in *C. adjunctus* (one population at Clec104 and Cle015, and another population at Clec104). Since these incidences of deviation from Hardy-Weinberg equilibrium were not systematic across loci or sites, we retained these two markers and two sites for our analyses. We found no significant evidence, at a threshold of 0.05 after Bonferroni correction, of Hardy-Weinberg or linkage disequilibrium in any marker in *C. lectularius*. We also did not find any evidence of significant linkage

Table 3.1 Genetic diversity estimates for sites with at least three sampled individuals in *C. adjunctus* (numbers followed by A) and *C. lectularius* (numbers followed by L) in the Great Lakes region of North America, averaged across seven microsatellite markers. Site numbers correspond to those in Fig. 3.2, Table B.1 and Table B.2.

Site	Average number of alleles	Observed heterozygosity	Expected heterozygosity	Inbreeding coefficient G_{IS}
1A	2.286	0.357	0.354	-0.009
2A	1.857	0.190	0.423	0.549
3A	2.143	0.119	0.467	0.745
10A	1.800	0.200	0.500	0.600
14A	2.429	0.327	0.420	0.223
15A	3.571	0.248	0.393	0.369
16A	1.714	0.190	0.333	0.429
17A	1.714	0.286	0.429	0.333
C. adjunctus mean	2.189	0.240	0.415	0.405
1L	2.833	0.667	0.625	-0.067
2L	2.000	0.429	0.382	-0.121
3L	2.571	0.405	0.393	-0.030
4L	1.286	0.190	0.114	-0.667
5L	1.286	0.143	0.167	0.143
7L	2.000	0.381	0.500	0.238
9L	1.857	0.429	0.333	-0.286
11L	2.286	0.619	0.583	-0.061
14L	1.714	0.333	0.298	-0.120
15L	2.000	0.167	0.219	0.239
21L	2.286	0.429	0.440	0.027
23L	2.143	0.286	0.452	0.368
27L	1.143	0.095	0.071	-0.333
C. lectularius mean	1.954	0.352	0.352	-0.052

disequilibrium in any marker, in both species. Of all pairs of individuals in both species, a low proportion (2.8% in *C. adjunctus* and 2.9% in *C. lectularius*) showed a parent-offspring or full-sibling relationship and were from the same sampling site.

3.3.2 Spatial genetic structure

Bayesian clustering in Geneland revealed a coarser division of genetic structure in *C. adjunctus* than in *C. lectularius*, with seven genetic clusters in the former versus eleven for the latter (Table 3.2). Clusters were consistent with geographic sampling location in both species (Fig. 3.2), and only *C. adjunctus* had clusters containing individuals sampled moderately far away from each other. Clusters in *C. adjunctus* were also highly consistent with those in the larger-scale study of (Talbot et al. 2016; Chapter 2). *K*-means clustering in GenoDive also resulted in $K = 7$ in *C. adjunctus*. In contrast, the lowest BIC value was at $K = 10$ in *C. lectularius*, and the second lowest BIC value was at $K = 11$. We observed a non-significant IBD relationship in *C. adjunctus* ($P = 0.165$, $R^2 < 0.01$), but a significant relationship in *C. lectularius* ($P = 0.001$, $R^2 = 0.09$), when genetic relatedness was fitted only with geographic distance (Table 3.2). When IBD models were conditioned with genetic structure, we found significant IBD in both species ($P = 0.001$); the fit of the model was still very low in *C. adjunctus* ($R^2 = 0.04$; Table 3.2), but increased more than two-fold for *C. lectularius* ($R^2 = 0.22$).

Results from the AMOVA revealed a sharp difference between the species in the proportion of genetic variation among sites (Table 3.3), which was lower by almost 25% in *C. adjunctus* (33.3%) than in *C. lectularius* (57.9%). We

Table 3.2 Spatial genetic structure in *Cimex adjunctus* and in *Cimex lectularius* in the Great Lakes region of North America, estimated using microsatellite markers. Most likely *K* = the number of genetic clusters estimated using the Geneland method, IBD (r_W) = isolation-by-distance based on pairwise relatedness values, and IBD (r_W) + *K* = isolation-by-distance while correcting for population genetic structure.

Species		***Cimex adjunctus***	***Cimex lectularius***
Number of markers		7	7
Sample size		75	73
Most likely *K*		7	11
IBD (r_W)	*P*	0.165	0.001*
	R^2	< 0.01	0.09
IBD (r_W) + *K*	*P* (IBD)	0.358	0.001*
	P (*K*)	0.001*	0.001*
	R^2	0.04	0.22

*Statistically significant at α = 0.05

Table 3.3 Genetic diversity within and differentiation among sites with at least three sampled *Cimex adjunctus* and in *Cimex lectularius* in the Great Lakes region of North America, estimated using microsatellite markers. H_S = average expected heterozygosity, Hedrick's G'_{ST} = Hedrick's global G'_{ST} differentiation index (and 95% confidence intervals), and AMOVA = Analysis of Molecular Variance.

Species		***Cimex adjunctus***	***Cimex lectularius***
Number of markers		7	7
Sample size		60	50
H_S		0.389 (0.309 – 0.469)	0.351 (0.292 – 0.410)
Hedrick's G'_{ST}		0.233 (0.133 – 0.333)	0.750 (0.683 – 0.817)
AMOVA (% of variation)	Within sites	66.7	42.1
	Among sites	33.3	57.9

also found a significantly lower global differentiation index estimate in *C. adjunctus* (G'_{ST} = 0.233) than in *C. lectularius* (G'_{ST} = 0.750). These results indicate weaker genetic structuring in *C. adjunctus* than in *C. lectularius*. Average expected heterozygosity was slightly higher in *C. adjunctus* (H_S = 0.389) than *C. lectularius* (H_S = 0.351), although the confidence intervals for these estimates overlapped (Table 3.3). Finally, all analyses on *C. lectularius* did not show appreciable differences when applied on the dataset of seven microsatellite loci or on the whole panel of 20 markers, with the larger panel revealing only slightly stronger genetic differentiation and structuring (Table B.4).

3.4 Discussion

We confirmed our prediction of different population structure in *C. adjunctus* and *C. lectularius*, two congeneric ectoparasitic insects associated with different host species. Both species showed classic metapopulation characteristics, but each tended towards either a more patchy population or non-equilibrium metapopulation structure. In *C. adjunctus* we found moderate genetic differentiation, a small number of genetic clusters and no IBD, which suggest a mixture of classic metapopulation and patchy population attributes. In *C. lectularius* we found high genetic differentiation, a larger number of genetic clusters, and weak IBD, which suggest a mixture of classic and non-equilibrium metapopulation attributes. The high levels of genetic differentiation and weak pattern of IBD we observed in *C. lectularius* are consistent with existing literature on the species that points to a very strong genetic structure between housing units (Booth et al. 2012; Fountain et al. 2014), and the presence of weak IBD in

the eastern United States in a study area of comparable size (Saenz et al. 2012). Our results therefore support previous work in *C. lectularius* while providing new insight into the biology of the less well studied *C. adjunctus*. An interesting observation is that the degree of spatial genetic structuring in *C. lectularius* at a scale of a few hundred kilometers, observed here, is similar to what is observed for *C. adjunctus* across its entire range of several thousands of kilometers (Talbot et al. 2016; Chapter 2).

Two possible caveats of our between-species comparison are a higher mean number of samples per site and sampling over a longer time frame (increasing the possibility of genetic differentiation among samples due to drift), in *C. adjunctus* than in *C. lectularius*. However, both these factors would have a tendency to bias results towards higher genetic structuring in *C. adjunctus*, which we clearly did not observe in our results. We also found a weaker effect of geographic distance in *C. adjunctus* than in *C. lectularius*, despite the fact that geographic coordinates in *C. lectularius* were much less precise than *C. adjunctus* (we only had the postal code for each unit), which could lead to increased noise and a dampening of any underlying spatial pattern. Finally, we found similar patterns of genetic structure in *C. lectularius* when using a panel of twenty available markers versus the seven that overlapped with *C. adjunctus*, indicating that the seven markers used for our study provide reliable and meaningful estimates of genetic structure. Higher inbreeding coefficients in *C. adjunctus* reflect a generally lower observed than expected heterozygosity, which could be due to some pooling of individuals from separate populations (i.e., a Wahlund effect) arising from a lower certainty of roost assignment.

Host movement behaviour can be an important determinant of connectivity and population structure in hitchhiking parasites. For example, two tick host races parasitizing blacklegged kittiwake and Atlantic puffin respectively showed disparate genetic differentiation patterns at some spatial scales, likely caused by different dispersal patterns of the two host species (McCoy et al. 2003). Movement and parasite transport by humans is complex and makes use of a number of different modalities. Nonetheless, it is most likely that the dispersal kernels for parasites transported by humans versus flying bats are very different. The more patchy population structure in *C. adjunctus* than in *C. lectularius* observed here could be due to more frequent transport of parasites over a scale of tens to hundreds of kilometers by bats (Roberts et al. 2012; Norquay et al. 2013; Weller et al. 2016) than by humans.

Predicted genetic consequences of the three main models of population structure result largely from differences in connectivity among sub-populations (Mayer et al. 2009). Among different species that display classic metapopulation structure, variation in the degree of genetic differentiation of subpopulations will depend on both differences in connectivity as well as local dynamics that affect genetic drift. Higher genetic differentiation in *C. lectularius* versus *C. adjunctus* could also therefore be due to stronger effects of genetic drift, mediated by smaller effective population size (Garant et al. 2007; Gandon and Nuismer 2009). Differences in effective population size between the two parasite species might, in turn, be due in part to a difference in population dynamics mediated by pesticide use. Pesticides have been used extensively on household *C. lectularius* infestations (Davies et al. 2012), potentially causing local bottlenecks and

extinctions, whereas they may be only rarely used in known bat roosts. This could possibly explain why *C. lectularius* associated with bats are genetically more diverse than *C. lectularius* associated with humans (Booth et al. 2015). Local population crashes or extinctions may also occur in bat-associated *Cimex* parasites, due to roost-switching in bats (Bartonička and Růžičková 2013). However, pesticide use is arguably much more efficient at reducing or eliminating parasite populations than is roost-switching by hosts. Therefore populations of *C. lectularius* may experience higher rates of local bottlenecks and extinctions, and genetic drift may play a more important role in driving divergence among populations of human-associated versus bat-associated parasites.

To conclude, our results show appreciable differences in population structure and genetic differentiation between the two parasite species. Population structure, in turn, can have important implications for the transmission of vector-borne diseases and the eco-evolutionary dynamics between these parasites and their hosts (Vander Wal et al. 2014a, 2014b).

3.5 Literature Cited

Adelman ZN, Kilcullen KA, Koganemaru R et al. (2011) Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a single population. *Plos One*, **6**, e26228.

Andreakis N, Kooistra WHCF, Procaccini G (2009) High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Molecular Ecology*, **18**, 212–226.

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Bartonička T, Růžičková L (2013) Recolonization of bat roost by bat bugs (*Cimex pipistrelli*): could parasite load be a cause of bat roost switching? *Parasitology Research*, **112**, 1615–1622.
Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JTH, Darling SR, Gargas A, Niver R, Okoniewski JC, Rudd RJ, Stone WB (2009) Bat White-Nose Syndrome: an emerging fungal pathogen? *Science*, **323**, 227–227.
Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.
Booth W, Balvín O, Vargo EL, Vilímová J, Schal C (2015) Host association drives genetic divergence in the bed bug, *Cimex lectularius*. *Molecular Ecology*, **24**, 980–992.
Booth W, Saenz VL, Santangelo RG et al. (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.
Cooper R, Wang C, Singh N (2015) Mark-release-recapture reveals extensive movement of bed bugs (*Cimex lectularius* L.) within and between apartments. *Plos One***, 10**, e0136462.
Criado PR, Belda Junior W, Criado RFJ, Vasconcelos e Silva R, Vasconcellos C (2011) Bedbugs (Cimicidae infestation): the worldwide renaissance of an old partner of human kind. *The Brazilian journal of infectious diseases*, **15**, 74–80.
Davies TGE, Field LM, Williamson MS (2012) The re-emergence of the bed bug as a nuisance pest: implications of resistance to the pyrethroid insecticides. *Medical and Veterinary Entomology*, **26**, 241–254.
Ellison LE, O'Shea TJ, Neubaum DJ, Bowen RA (2007) Factors influencing movement probabilities of big brown bats (*Eptesicus fuscus*) in buildings. *Ecological applications: a publication of the Ecological Society of America*, **17**, 620–627.
Fields Development Team (2006) *Fields: tools for spatial data*. National Center for Atmospheric Research, Boulder, CO, USA.
Fountain T, Duvaux L, Horsburgh G, Reinhardt K, Butlin RK (2014) Human-facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*. *Molecular Ecology*, **23**, 1071–1084.
Furlonger CL, Dewar HJ, Fenton MB (1987) Habitat use by foraging insectivorous bats. *Canadian Journal of Zoology*, **65**, 284–288.
Gandon S, Nuismer SL (2009) Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *The American Naturalist*, **173**, 212–224.
Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, **21**, 434–443.
Goddard J (2009) Bed bugs (*Cimex lectularius*) and clinical consequences of their bites. *The Journal of the American Medical Association*, **301**, 1358–1366.
Goddard J, Baker GT, Ferrari FG, Ferrari C (2012) Bed bugs (*Cimex Lectularius*) and bat bugs (several *Cimex* species): a confusing issue. *Outlooks on Pest Management*, **23**, 125–127.

Goldberg CS, Waits LP (2010) Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. *Molecular Ecology Resources*, **10**, 304–313.
Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**.
Hanski I (1998) Metapopulation dynamics. *Nature*, **396**, 41–49.
Harrison S (1991) Local extinction in a metapopulation context: an empirical evaluation. *Biological Journal of the Linnean Society*, **42**, 73–88.
Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898.
Legendre P, Lapointe F-J, Casgrain P (1994) Modeling brain evolution from behavior: a permutational regression approach. *Evolution*, **48**, 1487.
Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology*, **188**, 117–131.
Mayer C, Schiegg K, Pasinelli G (2009) Patchy population structure in a short-distance migrant: evidence from genetic and demographic data. *Molecular Ecology*, **18**, 2353–2364.
McCoy KD, Boulinier T, Tirard C, Michalakis Y (2003) Host-dependant genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution*, **57**, 288–296.
Meirmans PG (2012) AMOVA-based clustering of population genetic data. *Journal of Heredity*, **103**, 744–750.
Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.
Mimet A, Houet T, Julliard R, Simon L (2013) Assessing functional connectivity: a landscape approach for handling multiple ecological requirements. *Methods in Ecology and Evolution*, **4**, 453–463.
Neel MC (2008) Patch connectivity and genetic diversity conservation in the federally endangered and narrowly endemic plant species *Astragalus albens* (Fabaceae). *Biological Conservation*, **141**, 938–955.
Norquay KJO, Martinez-Nuñez F, Dubois JE, Monson KM, Willis CKR (2013) Long-distance movements of little brown bats (*Myotis lucifugus*). *Journal of Mammalogy*, **94**, 506–515
Pearce RD, O'Shea TJ (2007) Ectoparasites in an urban population of big brown bats (*Eptesicus fuscus*) in Colorado. *The Journal of parasitology*, **93**, 518–530.
Penczykowski RM, Laine A-L, Koskella B (2016) Understanding the ecology and evolution of host-parasite interactions across scales. *Evolutionary Applications*, **9**, 37–52.
Rasic G, Keyghobadi N (2012) The pitcher plant flesh fly exhibits a mixture of patchy and metapopulation attributes. *Journal of Heredity*, **103**, 703–710.
Roberts BJ, Catterall CP, Eby P, Kanowski J (2012) Long-distance and frequent movements of the flying-fox *Pteropus poliocephalus*: implications for management. *Plos One*, **7**, e42532.
Saenz VL, Booth W, Schal C, Vargo EL (2012) Genetic analysis of bed bug populations reveals small propagule size within individual infestations but

high genetic diversity across infestations from the eastern United States. *Journal of Medical Entomology*, **49**, 865–875.
Storfer A, Murphy MA, Evans JS et al. (2007) Putting the “landscape” in landscape genetics. *Heredity*, **98**, 128–142.
Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.
Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Vander Wal E, Garant D, Calmé S et al. (2014a) Applying evolutionary concepts to wildlife disease ecology and management. *Evolutionary Applications*, **7**, 856–868.
Vander Wal E, Garant D, Pelletier F (2014b) Evolutionary perspectives on wildlife disease: concepts and applications. *Evolutionary Applications*, **7**, 715–722.
Wang J (2002) An estimator for pairwise relatedness using molecular markers. *Genetics*, **160**, 1203–1215.
Wang J (2011) Unbiased relatedness estimation in structured populations. *Genetics*, **187**, 887–901.
Weller TJ, Castle KT, Liechti F, Hein CD, Schirmacher MR, Cryan PM (2016) First direct evidence of long-distance seasonal movements and hibernation in a migratory bat. *Scientific Reports*, **6**.

4 Comparative analysis of landscape effects on spatial genetic structure of the big brown bat and one of its cimicid ectoparasites

4.1 Introduction

Landscape elements, and the composition and configuration of the surrounding landscape, affect dispersal and gene flow in a broad range of organisms (Manel et al. 2003; Storfer et al. 2010; Manel and Holderegger 2013). Gene flow in turn affects genetic structure, such that less gene flow is associated with increased spatial structure and differentiation (Bohonak 1999). The association between landscape variables and genetic structure or differentiation is now commonly used to infer which landscape elements may act to facilitate or impede gene flow (Storfer et al. 2007). Some studies have compared effects of the landscape on genetic structure of different species (Goldberg and Waits 2010; James et al. 2011; Rioux Paquette et al. 2014). Comparison of the effects of the landscape on ecologically interacting species has also received some attention (James et al. 2011), although comparative landscape genetic analysis of hosts and parasites is so far limited. While it is often assumed that genetic structure in parasites is correlated with dispersal patterns of their hosts, the strength of this correlation varies with several factors such as difference in generation time, degree of generalism of the parasite, and proportion of the life cycle of the parasite spent free from the host (Mazé-Guilmo et al. 2016). Even if a parasite depends entirely on the host for dispersal, specific details of how transmission and movement between host individuals occurs can lead to differences between parasite and host in genetic structure and dispersal patterns. For example, there is

a discrepancy between patterns of relatively strong genetic structure in a human roundworm parasite, which transmits through human faeces, and extensive movement in their human host. This discrepancy may be explained by the fact that the parasites transmit between host individuals during defecation, which primarily occurs within human households, resulting in parasite gene flow that is spatially restricted (Criscione et al. 2010). If transmission of parasites among host individuals occurs in environments that are not also the most conducive to host dispersal and gene flow, then the effects of land cover on genetic structure may differ between the parasite and its hosts. Therefore, potentially contrasting effects of the landscape on genetic structure of parasites and hosts have however not been described. Here, we analyze and compare the effect of landscape composition on the genetic structure of an ectoparasite and one of its host species.

Big brown bats (*Eptesicus fuscus*) are native to most of North America, being absent only in northern and eastern regions of Canada. They overwinter in underground openings (caves or mines) or buildings (Whitaker and Gummer 1992), and roost in attics of buildings (Ellison et al. 2007) or in trees (Willis et al. 2003; Arnett and Hayes 2009) in the summer. They forage widely over a range of land cover types with foraging activity occurring mainly in wetlands and developed areas (Furlonger et al. 1987; Lookingbill et al. 2010), although males show lesser foraging site fidelity than females (Wilkinson and Barclay 1997). While foraging, they often pause in structures, including under bridges, with other individuals and other species before resuming foraging activity (Adam and Hayes 2000). Generation time in big brown bats is between one and two years, depending on location and sex (Kurta and Baker 1990). In early fall, bats from many summer

roosts congregate at the entrance of winter hibernacula and copulate before hibernation; a process known as autumnal swarming (Kurta 1995). Therefore, gene flow in big brown bats occurs partly in the fall. Gene flow may also occur in the spring, when a small proportion of individuals return to a different summer roost than the one they occupied in the previous year, and during the summer, when some individuals switch summer roosts (Willis and Brigham 2004; Ellison et al. 2007). Males are thought to disperse among roosts during the summer more frequently than females (Vonhof et al. 2008). Gene flow in big brown bats may be relatively high, as suggested by low genetic differentiation across North America observed in two studies (Nadin-Davis et al. 2010; Turmelle et al. 2011). Nonetheless, gene flow also appears to be limited at larger distances. In a study in eastern Illinois and western Indiana (Vonhof et al. 2008), a significant isolation-by-distance (IBD) pattern was observed using microsatellite markers among six big brown bat summer maternity colonies, at an average distance of 54 km from each other. In addition to geographic distance, landscape features such as land cover composition could affect gene flow that results from big brown bat movements among summer roosts, and also between summer roots and hibernacula. Big brown bats are known to avoid field interiors and preferentially move along edges created by either forests or man-made structures, as do several other bat species including the little brown myotis, the northern myotis, the silver-haired bat, the hoary bat, the pipistrelle and the serotine (Verboom and Huitema 1997; Jantzen and Fenton 2013). Analysis of the associations between land cover and genetic structure may reveal additional effects the landscape on gene flow of big brown bats.

Big brown bats are also one of the key hosts of *Cimex adjunctus*, a widespread blood-feeding insect (Family Cimicidae) that is an ectoparasite of bats in North America. This insect occurs from the eastern seaboard to the Rocky Mountains, and from Labrador and the Northwest Territories south to Texas (Usinger 1966). *C. adjunctus* is an ectoparasite of warm-blooded animals, almost exclusively associated with bats, and is known to be a weak generalist, meaning that it associates with host species that are phylogenetically closely related to each other (Mazé-Guilmo et al. 2016). *C. adjunctus* parasitizes several other bat species in central and eastern North America, and although the full breadth of potential host species is not known, it includes the little brown myotis (*Myotis lucifugus*) and northern myotis (*Myotis septentrionalis*) (Usinger 1966; Talbot et al. 2016; Chapter 2). According to Usinger (1966), cimicid ectoparasites associated with bats may display between one and two generations per year, depending on the location. This parasite typically remains in the hosts' roosts, emerging from cracks in the walls to obtain blood meals (Usinger 1966). It is hypothesized to have limited inherent capacity for movement outside of roosts such that dispersal occurs primarily via individuals being carried by the host (Usinger 1966). Mist-net captures of bats transporting *C. adjunctus* (Talbot et al. 2016; Chapter 2) confirm this mode of dispersal. Therefore, gene flow in *C. adjunctus* is likely mediated by its bat hosts.

Roost switching by bats in the summer is one very possible mechanism by which gene flow in both *C. adjunctus* and the hosts would occur. Whether *C. adjunctus* gene flow can occur during movements between summer roosts and winter hibernacula of bats is less clear because the extent to which *C. adjunctus*

overwinters in hibernacula is not known. Gene flow in *C. adjunctus* may also occur during bat foraging; movement of parasites between host individuals could occur at temporary night roosting areas, where bats from different summer day roosts congregate between bouts of feeding (Adam and Hayes 2000). Therefore, foraging movements of bats, although they do not result in bat gene flow, may affect gene flow in *C. adjunctus*. This is one possible mechanism by which discrepancies in gene flow patterns between bats and *C. adjunctus* could arise. While gene flow in *C. adjunctus* is potentially mediated by multiple bat species, the big brown bat is one of the most common and widespread hosts. Furthermore, key aspects of bat ecology that may contribute to ectoparasite gene flow are shared among several of *C. adjunctus*' hosts. For example, the use of edges at forests and developed areas for movement is common to many bat species (Verboom and Huitema 1997; Jantzen and Fenton 2013), as is the use of temporary roosting sites during foraging (Adam and Hayes 2000). Wetlands are also important sites of foraging activity for several other bat species including the eastern red bat (*Lasiurus borealis*), tri-coloured bat (*Perimyotis subflavus*) and little brown myotis (*M. lucifugus*) (Lookingbill et al. 2010).

In our study, we compared the effects of landscape composition on genetic differentiation in big brown bats, and in its parasite *C. adjunctus*. We hypothesized that gene flow of big brown bats preferentially occurs through land cover types that are known to facilitate movement, such as developed or forested areas. We therefore predicted a negative effect of these land covers types on bat genetic differentiation. We also hypothesized that bat gene flow is not associated with open land covers that are either avoided, such as open areas, or used

primarily for foraging, such as wetlands, and predicted a neutral or positive effect of these land covers types on bat genetic differentiation. For *C. adjunctus* we hypothesized that some portion of gene flow occurs during bat foraging, which does not result in gene flow in the bat itself. We therefore predicted that genetic differentiation of the two species could be affected differently by land cover, with a potentially significant negative effect of bat foraging areas, such as wetlands, on genetic differentiation of *C. adjunctus*.

4.2 Materials and Methods

4.2.1 Sample collection

We collected 2-mm wing biopsies from 142 big brown bats caught in mist nets or harp traps in the southern Great Lakes region (Fig. 4.1) between 1997 and 2010. Some of these samples were also used in Vonhof et al. (2008). Upon collection, samples were immediately stored in a 95% ethanol solution until further analysis.

We also collected 55 samples of *C. adjunctus* in the southern Great Lakes region (Fig. 4.1), from 2005 to 2014, that represent a portion of the samples used in (Talbot et al. 2016; Chapter 2). We removed all but six samples directly from mist-netted *E. fuscus* host individuals. Mist net capture locations were adjacent to a known summer roost (house, barn, church or school) of *E. fuscus*, or within forested national, provincial, state or territorial lands (Talbot et al. 2016; Chapter 2). Most mist-netted bats and the *C. adjunctus* individuals they harboured likely came from the adjacent known roost, although it is possible that a small

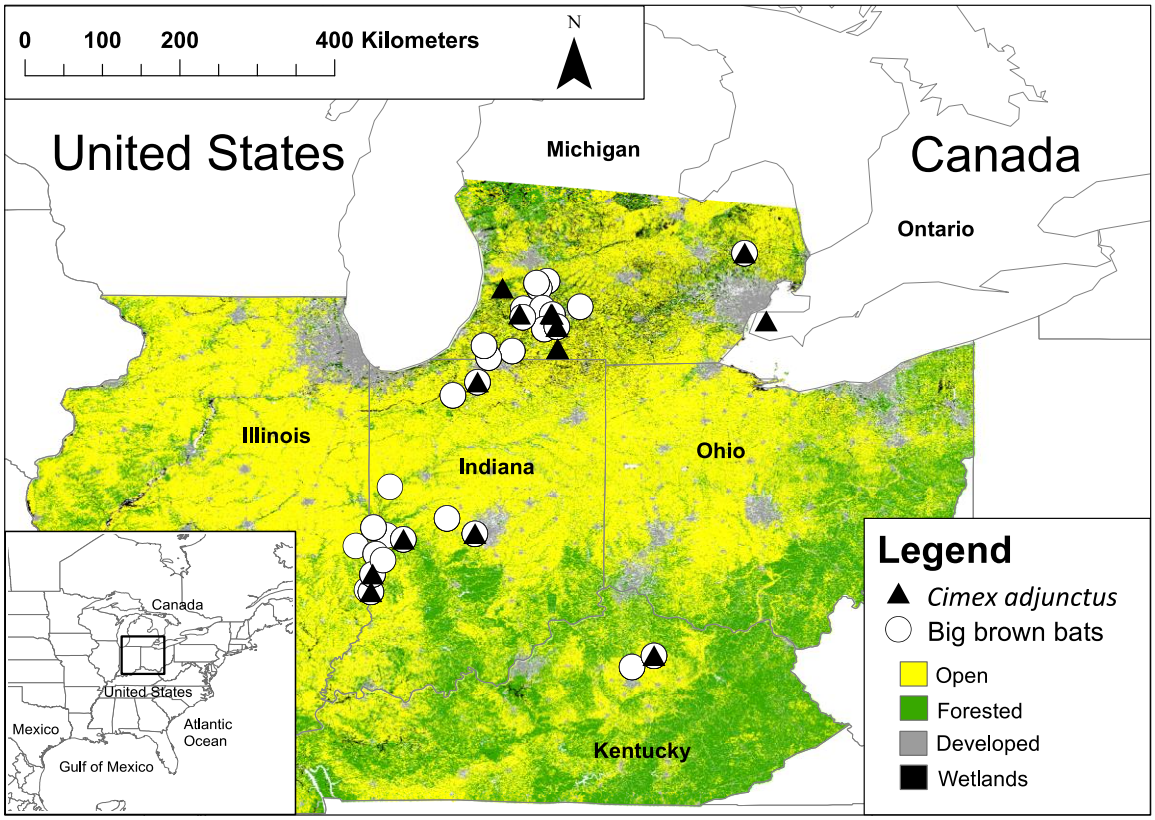

Figure 4.1 Study area, in the southern Great Lakes region of North America. White circles show sampling locations for the big brown bat, *Eptesicus fuscus,* and black triangles show sampling locations for its cimicid ectoparasite, *Cimex adjunctus*. Each of the four land cover types analyzed in our study is shown by a different shade.

proportion came from different roosts in the area. Overall, between 3 and 15% of mist-netted bats harboured a parasite, depending on the location. We also sampled six *C. adjunctus* individuals from the interior of a summer roost, in a house attic inhabited by *E. fuscus* (Talbot et al. 2016; Chapter 2). Because we could be certain of the roost site in this case, we considered this sampling location as distinct from its adjacent mist-netting capture location.

4.2.2 Genetic analyses

We genotyped big brown bats at eight microsatellite loci, originally developed for a range of bat species (MMG9 and MM25, from Castella and Ruedi 2000; TT20 from Vonhof et al. 2001; EF1, EF6, EF14, EF15 and EF20 from Vonhof et al. 2002). For samples that were also analysed by Vonhof et al. (2008), we used the genotype data reported in that paper. For all additional samples, we extracted DNA from wing biopsies using the DNeasy Blood and Tissue Kit (QIAGEN, Germantown, Maryland, United States) and genotyped each sample at the eight microsatellite loci using PCR chemistry and cycling conditions as in Vonhof et al. (2002). We used a DNAEngine Premium Thermal Cycler 200 (Bio-Rad, Hercules, California, United States) to execute the polymerase chain reaction (PCR) amplification. We visualized PCR products with 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad, Hercules, California, United States) on a UV transluminator to check the quality and size of amplified fragments. We then sized products on a 3730xl DNA Analyzer (Applied Biosystems, Foster City, California, United States).

Samples of *C. adjunctus* were previously genotyped at seven microsatellite loci, originally developed for *C. lectularius* (Cle002, Cle003, Cle013, Cle015 from Fountain et al. 2014; Clec15, Clec104 and BB28B from Booth et al. 2012), as described in Talbot et al. (2016; Chapter 2). We called microsatellite genotypes for each species using ABI's GeneMapper 4.0, and we checked all genotype calls manually.

4.2.3 Statistical analyses

Hardy-Weinberg, linkage disequilibrium and genetic diversity

For each species separately, we used Genepop 4.2 to test for Hardy-Weinberg and linkage disequilibrium in all sites with more than one individual sampled. We corrected *P* values for multiple hypothesis testing with Bonferroni correction, and used a threshold α of 0.05. Also, we calculated genetic diversity indices (total number of alleles, and average observed and expected heterozygosity and inbreeding coefficient G_{IS} across sites with more than one individual sampled) for each locus and averaged across all loci.

Comparative effect of geographic distance and land cover

We tested for isolation-by-distance (IBD) and effects of landscape composition on genetic differentiation, separately for *C. adjunctus* and the big brown bat, using an individual-based approach. We used r_W (Wang 2002), calculated with SpaGeDi 1.5, as a genetic relatedness index. We calculated 1 – r_W for each pair of individuals of each species to obtain genetic distances. We calculated geographic distance (in km) between sampling locations of individuals,

corrected for sphericity of the earth, using the 'rdist.earth' function from the 'fields' package (Fields Development Team 2006) in R 3.1.3. Next, to characterize land cover (Table 4.1) in the southern Great Lakes region of the United States, we used the National Land Cover Database (United States Geological Survey's Land Cover Institute, Sioux Falls, North Dakota, United States). We chose four types of land cover that may affect movements and behaviours of bats: wetland (two types combined: woody and emergent herbaceous), developed, forested (three types combined: deciduous, evergreen and mixed), and open (four types combined: hay and pasture, cultivated crops, barren land and grassland). Using ArcGIS 10.3 (ESRI, Redlands, United States), we created a buffer around a straight line between the capture location for each pair of individuals, for both species (Murphy et al. 2010; Rioux Paquette et al. 2014). We set the buffer's width to 54 km (27 km on either side of the line), the average distance between sampled big brown bat colonies in a previous study in which significant IBD was observed (Vonhof et al. 2008). Using Spatial Analyst (ArcGIS 10.3, ESRI), we calculated the proportion of each land cover type in each linear buffer, corresponding to each pair individuals.

To compare landscape composition between the parasite and the host, we simultaneously fit pairwise genetic distance ($1 - r_W$) to geographic distance and proportion of each type of land cover using multiple regression on distance matrices with the 'MRM' function from the 'ecodist' package (Goslee and Urban 2007) in R 3.1.3. This function determines significance of predictors through permutation (9,999 replicates) of distance matrices (Legendre et al. 1994; Lichstein 2007). We compared models for big brown bats and *C. adjunctus* to

Table 4.1 Description of each land cover type, from the United States Geological Survey's National Land Cover Database, used in the study, in the southern Great Lakes of North America. The mean proportion (and standard deviation) of each land cover type across all 54 km-wide buffers connecting pairs of samples sites is provided, separately for the big brown bat (*Eptesicus* fuscus) and one of its cimicid ectoparasites (*Cimex adjunctus*).

Land cover type	**Description**	**Average proportion**	
		Cimex adjunctus	***Eptesicus fuscus***
Developed	Areas with a mixture of constructed materials and vegetation, where constructed materials account for 30 to 100% of the cover, and vegetation accounts for 0 to 70% of the cover.	0.09 (0.12)	0.04 (0.06)
Forested	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover.	0.12 (0.09)	0.05 (0.09)
Open	Areas of cultivated crops, hay or pasture, dominated by gramanoid or herbaceous vegetation, or barren of any structure or vegetation	0.44 (0.25)	0.68 (0.29)
Wetlands	Areas where the soil or substrate is periodically saturated with or covered with water.	0.06 (0.05)	0.02 (0.04)

determine which land cover types have a significant positive or negative relationship with genetic distance in each of the host and the parasite.

We used an approach based on quantifying land cover composition in broad, linear buffers (Murphy et al. 2010; Rioux Paquette et al. 2014), as opposed to a resistance matrix approach (Spear et al. 2005, McRae and Beier 2007), for two reasons. First, our approach is arguably more appropriate for animals that are transported by such flying animals. Flying animals such as bats may easily move over smaller areas that are unsuitable or could otherwise represent high resistance (e.g., Amos et al. 2012). As a result, they are likely to respond to the composition of the landscape at a coarser scale rather than to detailed configuration of the landscape, and the paradigm of the resistance surface may not apply as well to such highly mobile, volant animals as it does to less mobile and non-volant animals. Second, our approach is less dependent on a priori knowledge or hypotheses of which landscape elements affect gene flow (Spear et al. 2010), which is particularly important for *C. adjunctus*, a species for which very little is known regarding basic aspects of ecology and movement.

4.3 Results

We obtained genotypes of 142 big brown bat individuals (49 males and 93 females; 114 adultsand 28 juveniles), from 32 roosts in the lower Great Lakes region of North America (Table C.1, Table C.2). We also obtained genotypes of 55 *C. adjunctus* from 15 roosts (Table C.3; microsatellite data available in Talbot et al. (2016; Table A.4). The average distance between roosts for big brown bat

samples was 141 km (range of 0.001 – 502 km). The average distance between roosts for *C. adjunctus* samples was 181 km (range of 0.012 – 511 km).

4.3.1 Hardy-Weinberg, linkage disequilibrium and genetic diversity

We found no significant evidence, after Bonferroni correction, of Hardy-Weinberg disequilibrium in big brown bats, nor linkage disequilibrium in either species. We found three significant cases of deviation from Hardy-Weinberg equilibrium in *C. adjunctus* (one population at Clec104 and Cle015, and another population at Clec104). These incidences of deviation from Hardy-Weinberg equilibrium were not systematic across loci, which would have suggested presence of null alleles, or across populations. Therefore, we retained these two markers and two populations for our analyses. Genetic diversity indices were overall higher in big brown bats than in *C. adjunctus* across microsatellite markers (Table 4.2), and values in C. adjunctus were very similar to those found in a study spanning a slightly larger study area in the same region (Chapter 3). Total number of alleles averaged at 28.9 in big brown bats and 5.6 in *C. adjunctus*, across microsatellite markers. Mean observed and expected heterozygosities, averaged across sites and across loci, were 0.815 and 0.861, respectively, in big brown bats and 0.256 and 0.434 in *C. adjunctus*. The mean inbreeding coefficient, averaged across sites and across loci, was 0.053 in big brown bats and 0.433 in *C. adjunctus*. Finally, pairwise genetic distances between individuals ($1 - r_W$) across the whole dataset were, on average, lower for big brown bats than for *C. adjunctus* (Big brown bat: 1.01 ± 0.11 (SD); *C. adjunctus*: 1.28 ± 0.61 (SD)).

Table 4.2 Genetic diversity indices (total number of alleles, N_{AL}, observed and expected heterozygosity, H_O and H_S, and inbreeding coefficient, G_{IS}) per microsatellite locus, and averaged across loci (Average). Values of H_O, H_E and G_{IS} are averaged across sites with more than one individual sampled, for big brown bats (*Eptesicus* fuscus; 141 individuals from 31 sites) and one of its cimicid ectoparasites (*Cimex adjunctus*; 50 individuals from 10 sites).

Species	Locus	N_{AL}	H_O	H_S	G_{IS}
Eptesicus fuscus	EF1	23	0.90	0.89	-0.01
	EF6	30	0.93	0.93	< 0.01
	EF14	31	0.87	0.89	0.02
	EF15	38	0.73	0.92	0.20
	EF20	29	0.79	0.90	0.12
	MMG9	46	0.87	0.96	0.09
	MMG25	19	0.63	0.66	0.05
	TT20	15	0.81	0.75	-0.07
	Average	28.9	0.815	0.861	0.053
Cimex adjunctus	Clec104	4	0.25	0.45	0.45
	Clec15	3	0.11	0.06	-0.81
	BB28B	4	0.53	0.40	-0.32
	Cle002	5	0.11	0.29	0.61
	Cle013	13	0.31	0.68	0.54
	Cle003	6	0.34	0.60	0.43
	Cle015	4	0.06	0.56	0.89
	Average	5.6	0.256	0.434	0.433

4.3.2 Comparative effect of geographic distance and land cover

In big brown bats, geographic distance, proportion of open land cover and proportion of developed land cover had significant relationships (Table 4.3) with genetic distance (final model $R^2 = 0.04$; Table 4.3). Genetic distance showed a positive relationship with both geographic distance ($P < 0.01$) and proportion of open land cover ($P < 0.01$), but a negative relationship with developed land cover ($P = 0.034$). These results suggest that geographic distance and open land cover may act to limit gene flow in big brown bats, while developed lands may facilitate gene flow.

In *C. adjunctus*, proportion of forested land cover and proportion of wetlands both had a marginally significant relationship with genetic distance (final model $R^2 = 0.06$; Table 4.3). The effect of forested land cover on genetic distance was positive ($P = 0.021$), while the effect of wetlands was negative ($P = 0.04$). These results suggest that forests may act to limit gene flow in *C. adjunctus* while wetlands may facilitate gene flow.

4.4 Discussion

4.4.1 Effect of land cover on genetic structure of the big brown bat and its ectoparasite

First, our results support earlier findings by Vonhof et al. (2008) of a significant positive relationship between geographic distance and genetic distance in big brown bats. Concordant with our predictions, we also found a significant effect of two land cover types on genetic structure in big brown bats. It has been suggested that bats preferentially move close to tall structures, either

Table 4.3 Effects of geographic distance and four different land cover types (Developed areas, Forested areas, Open areas, and Wetlands) on genetic distance ($1 - r_W$, where r_w is the related coefficient of Wang 2002) between individuals in the big brown bat (*Eptesicus* fuscus) and one of its cimicid ectoparasites (*Cimex adjunctus*), in the southern Great Lakes region of North America. Proportion of different land cover types were measured in 54 km-wide buffers between each pair of individuals, for each species separately. Models were fit using multiple regression on distance matrices (MRM). *P* values for significant effects are bolded.

Species		***Cimex adjunctus***	***Eptesicus fuscus***
Number of microsatellite markers		**7**	**8**
Sample size		**55**	**142**
Geographic distance	Slope	0.0005	0.0002
	SE	0.0005	< 0.0001
	P	0.111	**< 0.001**
Developed	Slope	0.1970	-0.0738
	SE	0.4089	0.0453
	P	0.567	**0.034**
Forested	Slope	0.9527	-0.0044
	SE	0.8036	0.0316
	P	**0.021**	0.859
Open	Slope	0.1808	0.0460
	SE	0.3177	0.0095
	P	0.404	**< 0.001**
Wetlands	Slope	-2.2797	-0.0183
	SE	1.9225	0.0588
	P	**0.040**	0.644
Final Model	R^2	0.06	0.04

trees or man-made structures, to avoid energy expenditures associated with moving against high winds (Jantzen and Fenton 2013). Therefore, open land cover, which represented a very large proportion of our study area, may be avoided. Consistent with this expectation, our results suggest that open land cover may act to limit gene flow in this species. Additionally, our results suggest that developed land cover may facilitate gene flow and support the hypothesis that big brown bats move preferentially along leeward edges of structural features (Jantzen and Fenton 2013).

Concordant with our predictions, we also found a significant effect of two land cover types, forested and wetlands, on genetic distance in *C. adjunctus*. These were different than the types of land cover found to affect big brown bat genetic distance, even though *C. adjunctus* almost entirely depends on its hosts to move outside of roosts (Usinger 1966). Furthermore, in contrast to our results on the big brown bat, we did not find IBD in *C. adjunctus*. Overall, our results suggest that a parasite and a host, while linked in their movements, may show differences in gene flow patterns. These differences may at least be partially explained by differences between the two species in the environments and types of land cover in which gene flow occurs. (Lookingbill et al. 2010) found the activity of several bat species, including the big brown bat, to be correlated with wetland cover. Our result of a negative effect of wetland cover on *C. adjunctus* genetic distance supports the hypothesis that gene flow in the ectoparasite may occur during foraging by bats in wetlands, possibly via transfer between individuals in temporary, communal roosts.

Our results suggest that forested areas impede on gene flow in the ectoparasite *C. adjunctus*. While several bat species are known to move along forest edges, they also show reduced activity in forest interiors and densely vegetated areas (Loeb and O'Keefe 2006; Jantzen and Fenton 2013). This restricting effect of contiguous or dense forest cover on bats could explain the positive effect of forest cover on *C. adjunctus* genetic distance. In addition, even when bats do forage in forested areas, it is possible that these environments provide few opportunities for *C. adjunctus* gene flow via transfer between host individuals, if there are few temporary, communal roosting sites for bats. While foraging in these environments, bats may be more likely to temporarily roost by themselves in trees. Finally, it is also possible that *C. adjunctus* experiences higher mortality or removal when bats travel through forested areas, although the exact mechanism by which this might occur is not clear.

Samples sizes in our study are larger for the big brown bat than its parasite. This is a function of the parasite being present on only a subset of sampled host individuals. While our sample size of *C. adjunctus* is relatively small, we used an individual-based analysis, which has been shown to allow for robust landscape genetic inference given small sample sizes (Prunier et al. 2013). Several other studies have used an individual-based approach with sample sizes similar to ours in drawing population genetic and landscape genetic inferences (Broquet et al. 2006; Laurence et al. 2013).

Finally, more information is needed on the effects of land cover on gene flow in males versus females, and in different age groups, in big brown bats. Sex-biased dispersal and sex-biased and age-biased parasitism, both suggested for

big brown bats (Pearce and O'Shea 2007; Vonhof et al. 2008), are important factors to take into account when comparing gene flow patterns between a host and a parasite.

4.4.2 Correlation between genetic differentiation of a host and a parasite

Although there are many examples in which host and parasite movement or gene flow are correlated (Nieberding et al. 2004, 2008, Bruyndonckx et al. 2009, Levin and Parker 2013), parasites often show higher levels of genetic differentiation than their hosts, possibly because of lower effective population size and shorter generation time in the parasite than the host (Mazé-Guilmo et al. 2016). For example, higher genetic structure in the trematode parasite (*Pagioporus shawi*) compared to its host, the steelhead trout (*Oncorhynchus mykiss*), led to parasite genotypes providing more accurate population assignments in the host than could be obtained by examining genotypes of the host itself (Criscione et al. 2006). Higher genetic differentiation in a host, although less common, is however also possible. For example, genetic structure among colonies was weaker for fleas than for their prairie dog hosts (Jones and Britten 2010). In addition to effective population size and generation time, additional factors that may uncouple the genetic structure of parasites from that of their hosts include the degree of generalism of the parasite, and the proportion of time spent in free-living stages by the parasite (Mazé-Guilmo et al. 2016).

Our results support the pattern of higher differentiation in the parasite, with higher pairwise genetic distances in *C. adjunctus* than in the big brown bat.

Two other studies on *C. adjunctus* conducted at two different spatial scales also found a much higher degree of genetic differentiation in the parasite (Talbot et al. 2016, 2017) than has previously been reported in two of its main hosts, the big brown bat (Vonhof et al. 2008, Nadin-Davis et al. 2010) and little brown myotis (Johnson et al. 2015). This difference was attributed to the fact that *C. adjunctus* is a weak generalist ectoparasite of highly mobile hosts, with a generation time that is likely much shorter than that of its hosts. Results from our landscape analyses suggest that there may be additional differences between *C. adjunctus* and its bat hosts in the location and timing of gene flow that contribute to their different genetic structure.

Although all parasite samples used in this study came from the body of big brown bats or in a roost inhabited by big brown bats, *C. adjunctus* can use several different bat species as hosts. In a range-wide study of the genetic structure of *C. adjunctus*, Talbot et al. (2016; Chapter 2) noted moderate differentiation among parasite samples from different host species at microsatellite markers, and very little differentiation at mitochondrial DNA. Therefore, individuals of *C. adjunctus* may switch host species somewhat regularly, a situation expected for a generalist ectoparasite. It is possible that the different responses of big brown bats and *C. adjunctus* to landscape composition partly reflect the fact that other bat species, such as *M. lucifugus* and *M. septentrionalis*, are also contributing to *C. adjunctus* gene flow. However, several key aspects of the ecology of big brown bats, including the use of wetlands for foraging, the use of temporary roosts while foraging, and seasonal patterns of gene flow are shared with other bat species that are potential hosts of *C.*

adjunctus (Adam and Hayes 2000; Furlonger et al. 1987; Lookingbill et al. 2010). As a result, our predictions regarding effects of land cover on *C. adjunctus* genetic differentiation arise not just from the behaviour of big brown bats, but also from the behaviour of multiple potential host species. Furthermore, because big brown bats are among the more widely dispersing of *C. adjunctus*' potential hosts, this bat species is likely to determine the upper limit of gene flow, and hence patterns of genetic differentiation, in the parasite.

While the effects of the landscape on gene flow and genetic structure of many animal species have been described (Storfer et al. 2010; Manel and Holderegger 2013), not much is known about how species that are dependent on the movements of other species, as is the case with many parasites, interact with the landscape (Sprehn et al. 2015). Our study has revealed a difference in the types of land cover that correlate with genetic differentiation of a generalist ectoparasite versus one of its potential bat host species. Our results suggest that in addition to factors such as host mobility and the generalist nature of the parasite (Mazé-Guilmo et al. 2016), differences between hosts and parasites in the nature, timing and location of gene flow events can also lead to discordant patterns of genetic structure.

4.5 Literature Cited

Adam MD, Hayes JP (2000) Use of bridges as night roosts by bats in the Oregon Coast Range. *Journal of Mammalogy*, **81**, 402–407.

Amos JN, Bennett AF, Mac Nally R et al. (2012) Predicting landscape-genetic consequences of habitat loss, fragmentation and mobility for multiple species of woodland birds. *Plos One*, **7**, e30888.

Arnett EB, Hayes JP (2009) Use of conifer snags as roosts by female bats in western Oregon. *Journal of Wildlife Management*, **73**, 214–225.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.
Booth W, Saenz VL, Santangelo RG et al. (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.
Broquet T, Johnson CA, Petit E et al. (2006) Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular Ecology*, **15**, 1689–1697.
Bruyndonckx N, Henry I, Christe P, Kerth G (2009) Spatio-temporal population genetic structure of the parasitic mite *Spinturnix bechsteini* is shaped by its own demography and the social system of its bat host. *Molecular Ecology*, **18**, 3581–3592.
Castella V, Ruedi M (2000) Characterization of highly variable microsatellite loci in the bat Myotis myotis (Chiroptera: Vespertilionidae). *Molecular Ecology*, **9**, 1000–1002.
Criscione CD, Anderson JD, Sudimack D et al. (2010) Landscape genetics reveals focal transmission of a human macroparasite. *Plos Neglected Tropical Diseases*, **4**, e665.
Criscione CD, Cooper B, Blouin MS (2006) Parasite genotypes identify source populations of migratory fish more accurately than fish genotypes. *Ecology*, **87**, 823–828.
Ellison LE, O'Shea TJ, Neubaum DJ, Bowen RA (2007) Factors influencing movement probabilities of big brown bats (*Eptesicus fuscus*) in buildings. *Ecological applications: a publication of the Ecological Society of America*, **17**, 620–627.
Fields Development Team (2006) *Fields: tools for spatial data*. National Center for Atmospheric Research, Boulder, CO, USA.
Furlonger CL, Dewar HJ, Fenton MB (1987) Habitat use by foraging insectivorous bats. *Canadian Journal of Zoology*, **65**, 284–288.
Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology*, **19**, 3650–3663.
Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**.
James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH (2011) Spatial genetic structure of a symbiotic beetle-fungal system: toward multi-taxa integrated landscape genetics. *Plos One*, **6**, e25359.
Jantzen MK, Fenton MB (2013) The depth of edge influence among insectivorous bats at forest–field interfaces. *Canadian Journal of Zoology*, **91**, 287–292.
Johnson LNL, McLeod BA, Burns LE, Arseneault K, Frasier TR, Broders HG (2015) Population genetic structure within and among seasonal site types in the little brown bat (*Myotis lucifugus*) and the northern long-eared bat (*M. septentrionalis*). *Plos One*, **10**, e0126309.
Jones PH, Britten HB (2010) The absence of concordant population genetic structure in the black-tailed prairie dog and the flea, *Oropsylla hirsuta*, with implications for the spread of *Yersinia pestis*. *Molecular Ecology*, **19**, 2038–2049.

Kurta A (1995) *Mammals of the Great Lakes region*. University of Michigan Press, Ann Arbor, MI, USA.
Kurta A, Baker RH (1990) *Eptesicus fuscus*. *Mammalian Species*, **356**, 1–10.
Laurence S, Smith MJ, Schulte-Hostedde AI (2013) Effects of structural connectivity on fine scale population genetic structure of muskrat, *Ondatra zibethicus*. *Ecology and Evolution*, **3**, 3524–3535.
Legendre P, Lapointe F-J, Casgrain P (1994) Modeling brain evolution from behavior: a permutational regression approach. *Evolution*, **48**, 1487.
Levin II, Parker PG (2013) Comparative host–parasite population genetic structures: obligate fly ectoparasites on Galapagos seabirds. *Parasitology*, **140**, 1061–1069.
Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology*, **188**, 117–131.
Loeb SC, O'Keefe JM (2006) Habitat use by forest bats in South Carolina in relation to local, stand, and landscape characteristics. *Journal of Wildlife Management*, **70**, 1210–1218.
Lookingbill TR, Elmore AJ, Engelhardt KAM et al. (2010) Influence of wetland networks on bat activity in mixed-use landscapes. *Biological Conservation*, **143**, 974–983.
Manel S, Holderegger R (2013) Ten years of landscape genetics. *Trends in Ecology and Evolution*, **28**, 614–621.
Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189–197.
Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G (2016) Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, **19**, 336–347.
Murphy MA, Dezzani R, Pilliod DS, Storfer A (2010) Landscape genetics of high mountain frog metapopulations. *Molecular Ecology*, **19**, 3634–3649.
Nadin-Davis SA, Feng Y, Mousse D, Wandeler AI, Aris-Brosou S (2010) Spatial and temporal dynamics of rabies virus variants in big brown bat populations across Canada: footprints of an emerging zoonosis. *Molecular Ecology*, **19**, 2120–2136.
Nieberding CM, Durette-Desset M-C, Vanderpoorten A et al. (2008) Geography and host biogeography matter for understanding the phylogeography of a parasite. *Molecular Phylogenetics and Evolution*, **47**, 538–554.
Nieberding C, Morand S, Libois R, Michaux JR (2004) A parasite reveals cryptic phylogeographic history of its host. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2559–2568.
Pearce RD, O'Shea TJ (2007) Ectoparasites in an urban population of big brown bats (*Eptesicus fuscus*) in Colorado. *Journal of Parasitology*, **93**, 518–530.
Prunier JG, Kaufmann B, Fenet S et al. (2013) Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: towards a better assessment of functional connectivity using an individual-based sampling scheme. *Molecular Ecology*, **22**, 5516–5530.
Rioux Paquette S, Talbot B, Garant D, Mainguy J, Pelletier F (2014) Modelling the dispersal of the two main hosts of the raccoon rabies variant in

heterogeneous environments with landscape genetics. *Evolutionary Applications*, **7**, 734–749.
Spear SF, Balkenhol N, Fortin M-J, Mcrae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, **19**, 3576–3591.
Sprehn CG, Blum MJ, Quinn TP, Heins DC (2015) Landscape genetics of *Schistocephalus solidus* parasites in threespine stickleback (*Gasterosteus aculeatus*) from Alaska. *Plos One*, **10**, e0122307.
Storfer A, Murphy MA, Evans JS et al. (2007) Putting the "landscape" in landscape genetics. *Heredity*, **98**, 128–142.
Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Molecular Ecology*, **19**, 3496–3514.
Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.
Turmelle AS, Kunz TH, Sorenson MD (2011) A tale of two genomes: contrasting patterns of phylogeographic structure in a widely distributed bat. *Molecular Ecology*, **20**, 357–375.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Verboom B, Huitema H (1997) The importance of linear landscape elements for the pipistrelle *Pipistrellus pipistrellus* and the serotine bat *Eptesicus serotinus*. *Landscape Ecology*, **12**, 117–125.
Vonhof MJ, Davis CS, Fenton MB, Strobeck C (2002) Characterization of dinucleotide microsatellite loci in big brown bats (*Eptesicus fuscus*), and their use in other North American vespertilionid bats. *Molecular Ecology Notes*, **2**, 167–169.
Vonhof MJ, Davis CS, Strobeck C, Fenton MB (2001) Characterization of microsatellite loci in Spix's disk-winged bats (*Thyroptera tricolor*). *Molecular Ecology Notes*, **1**, 73–75.
Vonhof MJ, Strobeck C, Fenton MB (2008) Genetic variation and population structure in big brown bats (*Eptesicus fuscus*): is female dispersal important? *Journal of Mammalogy*, **89**, 1411–1419.
Wang J (2002) An estimator for pairwise relatedness using molecular markers. *Genetics*, **160**, 1203–1215.
Whitaker JO, Gummer SL (1992) Hibernation of the big brown bat, *Eptesicus fuscus*, in buildings. *Journal of Mammalogy*, **73**, 312–316.
Wilkinson LC, Barclay RMR (1997) Differences in the foraging behaviour of male and female big brown bats (*Eptesicus fuscus*) during the reproductive period. *Écoscience*, **4**, 279–285.
Willis CKR, Brigham RM (2004) Roost switching, roost sharing and social cohesion: forest-dwelling big brown bats, *Eptesicus fuscus*, conform to the fission–fusion model. *Animal Behaviour*, **68**, 495–505.
Willis CKR, Kolar KA, Karst AL, Kalcounis-Rueppell MC, Brigham RM (2003) Medium- and long-term reuse of trembling aspen cavities as roosts by big brown bats (*Eptesicus fuscus*). *Acta Chiropterologica*, **5**, 85–90.

5 Host association and selection on salivary protein genes in bed bugs and related blood-feeding ectoparasites

5.1 Introduction

Selection pressures imposed by antagonistic interactions between species, such as predators and prey or parasites and hosts, can produce evolutionary 'arms races' and are important drivers of adaptation and diversification (Vermeij 1987; Hall et al. 2011; Pinheiro et al. 2013; Jacobs et al. 2014). Among parasites and their hosts, parasite species may evolve phenotypes that are increasingly efficient at using host resources. Hosts evolve phenotypes that are increasingly efficient at guarding against the loss of such resources. Such reciprocal evolutionary interactions may lead to parasite species becoming increasingly specialized to their host species, and to rapid evolution in genes involved in mediating the conflict (Decaestecker et al. 2007; Paterson et al. 2010).

Among parasite species, there is considerable variation in the extent to which the parasite individual is dependent upon and tied to a host individual. For example, a parasite may spend only part (temporary parasites), as opposed to all (permanent parasites), of its life cycle associated with its host (Balashov 2006, 2011). Also, parasites may occur within the host's body (endoparasite) or on the outside (ectoparasite) (Balashov 2006, 2011). Evolutionary arms races should be especially intense in permanent parasites and endoparasites, which are also often highly specialized (Hall et al. 2011; Leggett et al. 2013). Although temporary parasites and ectoparasites are likely to be more generalist and associate with more than one host species (Sponchiado et al. 2015), possibly due to being more

likely to encounter an alternative host, specialization and adaptation to a host can still occur in such parasites. For example, bat flies are temporary blood-feeding ectoparasites of bats, and yet they show very narrow host ranges (Dick and Patterson 2007). Genetic adaptation to the host has been best studied in species that are tightly linked to their host, such as permanent parasites or endoparasites (Hall et al. 2011; Abrams et al. 2013; Pinheiro et al. 2013), but has not been extensively examined in temporary parasites or ectoparasites (but see Mans et al. 2002, and Mahamdallie and Ready 2012).

Adaptation of parasites to their specific hosts may be reflected in patterns of variation in parasite genes that are involved in mediating the host-parasite conflict, as suggested by Mans et al. (2002) and Arcà et al. (2014). Specifically, host adaptation is suggested when non-synonymous substitutions at such genes, among parasites associating with different hosts, are more frequent than expected under neutral evolution (*i.e.*, positive selection). In contrast, non-synonymous substitutions that are less frequent than expected under neutral evolution indicate selection for the conservation of gene sequences and function (i.e,, negative selection). Here, we determined whether there is evidence of positive selection and host-adaptation in a group of temporary ectoparasite species (genus *Cimex*, Order Hemiptera) that either associate with bats, humans or swallows, and that include a widespread human pest, the bed bug (*Cimex lectularis*).

Insects in the genus *Cimex* are temporary hematophagous ectoparasites of birds and mammals. They do not remain on their host at all times but rather stay in nests or roosts between blood meals (Usinger 1966). Most *Cimex* species are

associated exclusively with bats while a few associate mainly with humans and a few others with swallows (Usinger 1966; Goddard 2009; Criado et al. 2011). Bats are hypothesized as ancestral hosts of the genus, and a move to human or bird hosts may have occurred when these cohabited in the same environments as bats, e.g. in caves (Usinger 1966; Benoit 2011). The genus is traditionally divided into four species groups (Usinger 1966) whose identity was recently confirmed based on DNA analyses (Balvín et al. 2015). Members of the *Pilosellus* species group (represented by *C. adjunctus*, *C. brevis* and *C. latipennis* in the present study), associate mainly with bats and occur in North America (Usinger 1966), while members of the *Pipistrelli* species group (e.g. *Cimex pipistrelli* or *C. japonicus*) associate mainly with bats in the Palaearctic region. Among the species that associate with swallows represented in the study, which are phylogenetically related to the *C. pipistrelli* species group (Balvín et al. 2015), *Cimex vicarius* occurs in North America, whereas *Cimex hirundinis* occurs in Europe (Usinger 1966), and a third currently unnamed species (*Cimex sp.*) occurs in Japan (Balvín et al. 2015). Finally, members of the *Lectularius* and the *Hemipterus* species groups formerly associated mainly with bats, are represented by the cosmopolitan bed bug (*C. lectularius*) and the tropical bed bug (*C. hemipterus*), which both have created specific host lineages associated with humans (Usinger 1966).

Studies of mRNA and proteins expressed in the salivary glands of *C. lectularius* (Valenzuela et al. 1996; Valenzuela and Ribeiro 1998; Francischetti et al. 2010) provide insights into how certain salivary proteins in this species act to suppress host defences (coagulation and vasoconstriction) at the site of rupture of a blood vessel, where the ectoparasite feeds. The anti-platelet property of

apyrases results from catabolizing ADP released from damaged tissues. Apyrases in the saliva of *C. lectularius* could be used by the parasite to prevent clotting (Valenzuela et al. 1996; Francischetti et al. 2010). Nitrophorins, in the saliva of *C. lectularius*, have vasodilatory and anti-platelet property by transporting nitric oxide from the salivary glands to the feeding site (Valenzuela and Ribeiro 1998; Francischetti et al. 2010). Similar salivary proteins have been observed to be under positive selection in other blood-feeding arthropods, such as in mosquitoes (Arcà et al. 2014), and potentially also in soft ticks (Mans et al. 2002) (but not in others, such as sandflies; Mahamdallie and Ready 2012), and therefore may play an important role in parasite adaptation to feeding on different host taxa.

We looked for evidence of positive selection acting on two salivary protein gene fragments, one coding for an apyrase and one coding for a nitrophorin, among *Cimex* specimens associated with either bats, humans or swallows. We hypothesized that positive selection acts on those candidate genes, due to adaptation of specimens to blood feeding on phylogenetically different hosts. We therefore predicted codons of the two candidate genes would show significant signals of positive selection in most specimens. We also predicted that since association with humans and birds is hypothesized to have appeared after association with bats, specimens associated with humans or birds would show significant signals of positive selection at more codons that specimens associated with bats. We therefore predicted a significant difference in the number of codons showing signals of significant positive selection between specimens that are associated with different types of hosts (bats, humans or swallows).

5.2 Material and Methods

5.2.1 Specimen collection

We processed whole body samples, stored in 95% EtOH, of cimicid ectoparasites, and also used sequence data from previous studies (Balvín et al. 2015; Talbot et al. 2016; Table A.4). All cimicid samples were collected from either the body of one of various bat and swallow species, from a roost mainly inhabited by one of various bat or swallow species, or from human dwellings (Table D.1). We analysed specimens from species from the four main *Cimex* clades (*Lectularius*, *Pilosellus*, *Hemipterus* and *Pipistrelli* species groups; Balvín et al. 2015), as well as species associated with birds related to species in the *Pipistrelli* group. We identified individual samples to species using a combination of morphology (Usinger 1966) and DNA barcoding (Hebert et al. 2003), where we compared the Cytochrome Oxidase 1 (*CO1*) sequence from each sample to known *CO1* sequences for *Cimex* species from published sources (Balvín et al. 2015).

5.2.2 Genetic analyses

We extracted DNA from whole body samples using the DNeasy Blood and Tissue Kit (QIAGEN, Germantown, Maryland, United States). We then amplified fragments of the mitochondrial *CO1* gene and the nuclear Elongation Factor 1α (*EF1α*) gene using primers listed in Table D.2 (Balvín et al. 2012, 2015; Talbot et al. 2016). We used these loci to construct an independent phylogeny of the specimens used in our study, to account for expected phylogenetic variances

in our analyses of salivary protein genes. We also included published sequences (Table D.1) of 11 *Cimex* specimens for one or both of the *CO1* and *EF1α* genes.

We designed primers to amplify fragments of two salivary protein genes (apyrase and nitrophorin) that have a known function that is directly linked to efficiency of blood-feeding in *C. lectularius*, based on published mRNA sequences initially obtained from salivary glands of *C. lectularius* (Valenzuela et al. 1996; Valenzuela and Ribeiro 1998; Francischetti et al. 2010). We designed our primers to maximize the size of the resulting fragment while also maximizing the number of specimens for which we obtain successful amplification. Our apyrase primers amplified a genomic fragment more than a third (371 bp) of the entire coding sequence (969 bp; Dai et al. 2004), which does not contain introns. Our nitrophorin primers also amplify a fragment more than a third (300 bp) of the entire coding sequence (840 bp; Protein Data Bank identifier: 1NTF; Berman 2000; Weichsel et al. 2005), which also does not contain introns. Resulting fragments code for a diversity of structural elements in both proteins, and encompass areas in the interior and exterior of the 3D structure of the proteins (Dai et al. 2004; Weichsel et al. 2005).

We used a DNAEngine PTC-200 Thermal Cycler (Bio-Rad, Hercules, California, United States) for polymerase chain reaction (PCR) amplification for *CO1*, *EF1α*, and the apyrase and nitrophorin genes. For all gene fragments, we performed PCR in 25 µL final volume containing: 1X Taq polymerase buffer excluding $MgCl_2$ (Applied Biosystems, Foster City, California, United States), 1.5 mM of $MgCl_2$, 0.2 mM of each type of dNTP, 0.3 µM of each primer, 1 U of Taq polymerase (ABI), and 2 µL of DNA extraction product. We used the following

PCR cycling: an initial denaturation of 1 min at 94°C, followed by 35 cycles of 30 sec of denaturation at 94°C, 45 sec of annealing at a locus-specific temperature (42 – 57 °C; Table D.2) and 45 sec of extension at 72°C, finished by a final extension step of 5 min at 72°C. We visualized PCR products by 1.5% agarose gel electrophoresis using SYBR Green stain (Bio-Rad) on a UV transluminator to check the quality and size of amplified fragments. Then, we sequenced the amplified gene fragment for every sample using Sanger sequencing with BigDye terminator chemistry (ABI) and analyzed the fragments on a 3730xl DNA Analyzer (ABI). We aligned all sequences using the MUSCLE function in MEGA 6.06, and recorded heterozygous sites as ambiguous (N, which can be either of the four nucleotides).

5.2.3 Statistical analyses

We first constructed a hypothesized species tree, using information from both the *CO1* and the *EF1α* gene fragments, using the *BEAST framework in BEAST 2.4.2 (Heled and Drummond 2010). The purpose was to obtain a phylogenetic tree showing the tree topology (a proxy of species history) and the branch lengths (a proxy of branch-specific substitution rates) for all specimens. We used this phylogenetic tree in our analyses of selection (see following paragraphs) as an internal negative control (Pond et al. 2005). This procedure maximizes the robustness of inferences about selection in candidate gene sequences (here, apyrase and nitrophorin genes) of specimens from a variety of species. To build a species tree, we used parameters recommended by the authors (Heled et al. 2013) to create 10,000 species trees using data from both

loci: HKY substitution model with empirical allele frequencies, linear population size with constant root, 10,000,000 chains and storing every 5,000th chain. We used a relaxed log-normal clock because it was shown to perform better than strict clock when substitution rates are expected to vary among lineages (Brown and Yang 2011). We then computed the maximum credibility tree (Heled and Bouckaert 2013) out of all the species trees produced by *BEAST, using TreeAnnotator in BEAST 2.4.2. We discarded the first 20% of Markov Chain Monte Carlo as burn-in, set no posterior probability limit, and set node heights at common ancestors.

Second, we found the substitution model best representing each candidate gene, using a function implemented in the web interface of the HyPhy 2.2.1 package. We also performed a recombination detection analysis, using the single breakpoint recombination (SBP; Pond et al. 2006) analysis in the web interface of the HyPhy 2.2.1 package, to see if our alignments showed any sign of recombination, which could bias our analyses of signals of selection. We used Beta-Gamma site-to-site rate variation and three rate classes, as these options are the most general with the fewest number of parameters, and should be realistic in most situations (Pond et al. 2005).

Third, as a prior assessment of whether any of the candidate genes shows a whole-sequence signal of positive selection, we performed a partitioning approach for robust inference of selection (PARRIS) analysis (Scheffler et al. 2006). This analysis compares a null model (without positive selection) and an alternative model (with positive selection) on the candidate gene sequences, using a likelihood ratio test in the web interface of the HyPhy 2.2.1 package.

Fourth, we used a suite of analyses to test if any codon in the candidate gene sequences is, or has been, affected by selection at any point in time. We used more than one analysis, and considered only signals simultaneously obtained using every single analysis, to reduce the likelihood of false-positive results. We used three different likelihood-based approaches, implemented in the web interface of the HyPhy 2.2.1 package. Random-effects likelihood (REL; Pond and Frost 2005) computes the likelihood that nonsynonymous and synonymous substitution rates at each codon site fit with one of two predefined distributions, representing positive and negative selection, respectively, constructed using information from the hypothesized species tree. Fixed-effects likelihood (FEL; Pond and Frost 2005) compares nonsynonymous and synonymous substitution rates at each codon site with the global expected substitution rates calculated using information from the hypothesized species tree. Mixed-effects model of evolution (MEME; Murrell et al. 2012) compares nonsynonymous and synonymous substitution rates at each codon site with expected substitution rates specific to each node of the hypothesized species tree. All three methods can be used to detect codons under positive selection, but only REL and FEL are also aimed at detecting codons under negative selection. To reduce the possibility of false discovery, we considered only codons that were detected to be under positive selection by all three analyses ($\alpha < 0.05$ for the MEME and the FEL approaches, or Bayes factor >100 for the REL approach), or under negative selection by both REL and FEL. Additionally, we used the empirical Bayes procedure implemented in the MEME approach (Bayes factor >100) to identify nodes of the hypothesized species tree where a signal of significant positive

selection is observed, at each codon previously identified as being under positive selection simultaneously by the REL, FEL and MEME approaches. The purpose was to determine which specimens, or groups of specimens, are characterized by a signal of positive selection at each of those codons.

Finally, we analysed the effect of association to one of three types of host (bat, human and swallow), by each specimen, on the number of codons showing signals of positive selection with all three approaches, in any one specimen, using a standard analysis of variance function in R 3.2.2. We corrected the response and the predictor variables for phylogenetic independent contrasts, to account for shared history in each pair of specimens of the study. We built independent contrasts using our hypothesized species tree (Purvis and Rambaut 1995), with the 'pic' function from the 'ape' package (Paradis et al. 2004) in R 3.2.2. This procedure maximizes the robustness of our correlative inferences, because it takes into account phylogenetic distances between specimens.

5.3 Results

We amplified and sequenced the target fragments for both salivary protein candidate genes, *CO1*, and *EF1α* for 26 specimens of ten congeneric species (although three specimens did not amplify at the nitrophorin gene fragment; Table D.1). Sequence lengths in the apyrase fragment varied by 66 bp, possibly due to deletion or insertion events. Although these indels possibly represent adaptive change, there is currently no way to test selection associated to them. We trimmed the start and the end of each sequence for both salivary protein genes, so that they only included whole codons, resulting in an alignment

of 297 bp for the nitrophorin fragment and an alignment of 369 bp for the apyrase fragment. Up to 1% and 4% of the apyrase and nitrophorin gene sequences, respectively, in any individual, were heterozygous. For all analyses, we treated gaps and heterozygous loci as missing data, which was the most conservative option.

5.3.1 Species tree

As in Balvín et al. (2015), our species tree, based on *CO1* and *EF1α*, shows clear distinctions among most species (Fig. 5.1; except for *C. pipistrelli*, as in Balvín et al. (2015). One clade representing the *Pilosellus* group contains three bat-associated species from North America (*C. adjunctus*, *C. brevis* and *C. latipennis*), one clade representing the *Lectularius* group contains only the cosmopolitan *C. lectularius*, one clade representing the *Hemipterus* group contains only the tropical human-associated *C. hemipterus*, and a final clade contains species mainly associated with bats or swallows, in North America, Europe and Asia (*C. pipistrelli*, *C. japonicus*, *C. vicarius*, *C. hirundinis* and *C. sp*).

5.3.2 Signals of selection on the apyrase gene

The best substitution model for the apyrase gene fragment specified a similar rate of evolution between both types of transitions and A to C (or C to A) transversions, and a different rate of evolution for all other types of transversions. Using the SBP approach, all three criteria (AIC, AICc, and BIC) gave best support for no recombination in the dataset.

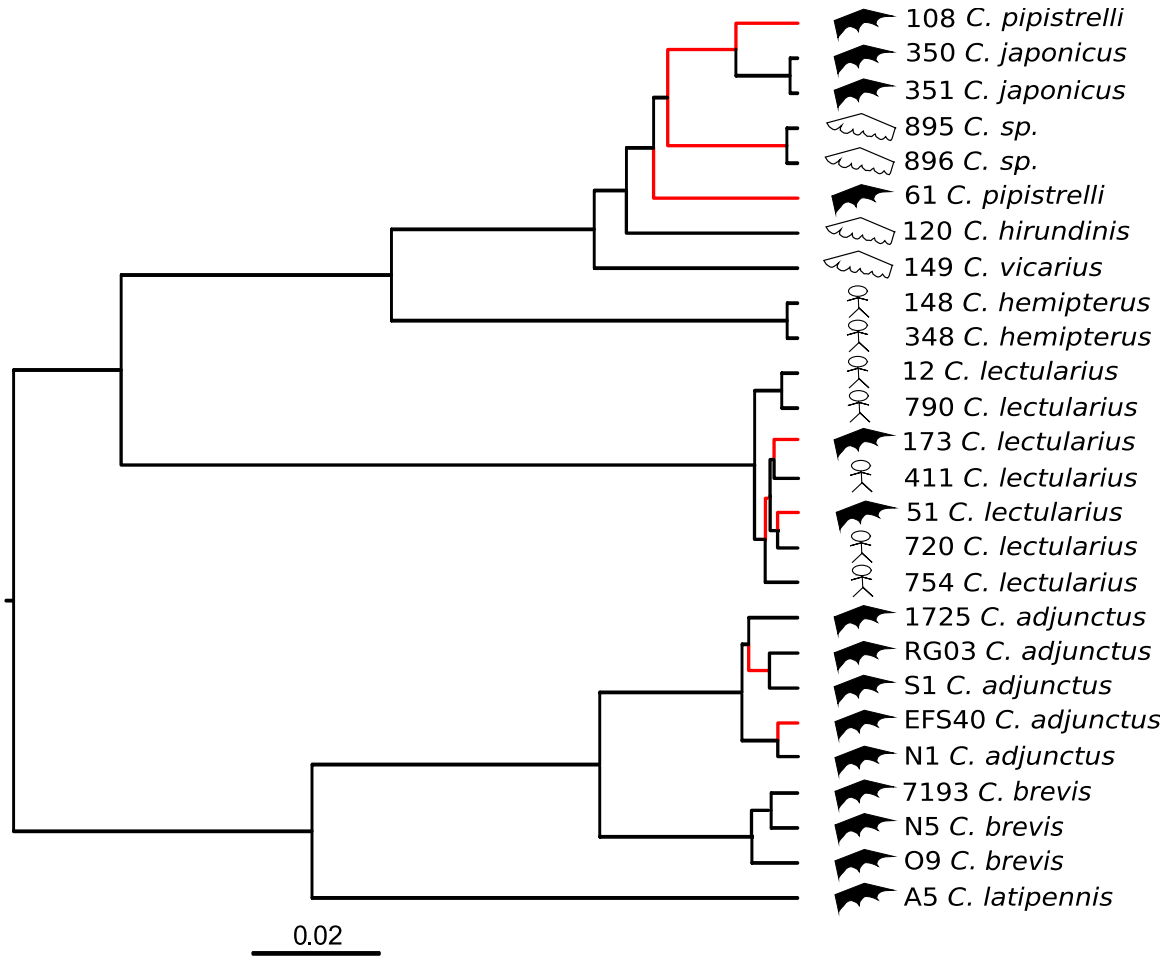

Figure 5.1 Hypothesized species tree of *Cimex spp.* specimens, based on *CO1* and *EF1α*, constructed with *BEAST 2.4.2. Scale represents substitutions per site. The specimen number (as in Table D.1), the species name and the host to which it was associated (stick figure for human, black pointed wing for bat and white rounded wing for swallow) of each sample are shown. Lighter-coloured branches indicate that significant positive selection was detected in the apyrase gene sequence of the corresponding specimen(s) at the terminal node, or at an internal node downstream from it (Bayes factor >100; calculated using the MEME approach), at any of five codons (37, 43, 63, 83, 93; previously identified as being under positive selection using the MEME, FEL and REL approaches).

Using the PARRIS approach to detect sequence-wide positive selection, we found a slightly higher log-likelihood value for the alternative model of positive selection (Log (L) = -1718.5 for the null model; Log (L) = -1714. 3 for the alternative model), leading to a likelihood ratio test value of 8.331 and a *P* of 0.016. The significant *P* value suggests evidence of positive selection in the dataset. The FEL approach identified six codons under positive selection, whereas the MEME and REL approaches identified eight and 15, respectively. Five codons were identified by all three approaches and therefore have strong support for being under positive selection (Table 5.1). The MEME approach suggests almost all evidence of positive selection in the five codons is in bat-associated lineages, and only one signal is in a swallow-associated lineage (Fig. 5.1; Fig. D.3 for results specific to each of the five codons). Furthermore, we found a significant effect (*df* = 1, *F* = 4.678, *P* = 0.041) of association with a particular type of host on the number of codons showing signals of positive selection. Specimens associated with bats had on average more codons showing signals of positive selection than specimens associated with humans or swallows (mean ± variance: bat = 1.7 ± 3.8; human = 0.3 ± 0.2; swallow = 0.5 ± 0.3). Observed substitutions at the five identified codons are likely to cause major structural changes to the apyrase protein, because they are all characterized by switches between charged, or polar, amino acids and nonpolar amino acids (Fig. D.4). Also, observed substitutions at three of the five identified codons (43, 63 and 93) represent changes between small and large amino acids. Amino acid substitutions at the five codons inferred to be under positive selection are however different among specimens showing signals of positive selection.

Table 5.1 Significant tests of selection on fragments of two genes coding for salivary proteins, apyrase and nitrophorin, in blood-feeding cimicid ectoparasites. The test value (*P* or Bayes factor) is given for all three codon-based analyses of positive and negative selection (MEME, FEL and REL), for each codon identified by all relevant analyses as showing a significant signal of selection. Codon numbers correspond to those used in Fig. D.3, Fig. D.4 and Fig. D.5.

Gene fragment	Type of selection	Codon number	MEME (*P*)	FEL (*P*)	REL (Bayes factor)
Apyrase	Positive	37	0.027	0.035	751.9
		43	0.036	0.014	1860.5
		63	0.013	0.011	169,142.0
		83	0.034	0.022	1162.7
		93	0.037	0.040	198.4
	Negative	32	N/A	0.001	432.7
		54		0.007	145.4
		104		< 0.001	479.0
		106		0.007	283.6
		112		0.005	248.3
		118		< 0.001	4,908.9
		123		< 0.001	3,001.6
Nitrophorin	Negative	1	N/A	0.007	875.2
		25		0.022	442.8
		32		< 0.001	675,630.0
		37		0.021	2,605.2
		40		0.020	484.9
		55		0.007	394.1
		69		0.005	13,861.0
		85		0.002	968.9
		88		0.003	212.1
		95		0.003	18,721.4
		99		0.005	12,754.0

The REL approach identified seven codons as potentially under negative selection, whereas the FEL approach identified the same seven codons, plus four additional ones. Therefore, seven codons have strong support for being under negative selection (Table 5.1; MEME cannot detect loci under negative selection). All seven codons show no variation in amino acid, and two of the identified codons (54 and 112) are binding sites for calcium and nucleotides, respectively (Dai et al. 2004).

5.3.3 Signals of selection on the nitrophorin gene

The best substitution model for the nitrophorin gene fragment specified a similar rate of evolution in all types of transitions and transversions, a model best known as F81 (Felsenstein 1981). Using the SBP approach, all three criteria (AIC, cAIC, and BIC) gave best support for no recombination in the dataset.

Using the PARRIS approach to detect sequence-wide positive selection, we found similar log-likelihood values for the two models (Log (L) = -1073.0 for the null model; Log (L) = -1072.8 for the alternative model), leading to a Likelihood Ratio Test value of 0.545 and a *P* of 0.761. The non-significant *P* value suggests no evidence of positive selection in the dataset.

The MEME approach identified three codons under positive selection, as opposed to one each in the FEL and the REL approaches. However, no single codon was identified by all three analyses. The REL approach identified 11 codons potentially under negative selection and the FEL approach identified the same 11 codons, plus three additional ones (Table 5.1). Therefore, no codon has strong support for being under positive selection and 11 codons have strong

support for being under negative selection. Ten of the 11 codons inferred as being under negative selection show no variation in amino acid (Codon 88 shows variation between a polar and a charged amino acids across specimens, Fig. D.5).

5.4 Discussion

5.4.1 Signals of selection on the apyrase and nitrophorin genes

We found signals of positive selection in the gene sequence coding for the salivary protein apyrase, among lineages of blood-feeding ectoparasites of the genus *Cimex*. A second candidate gene for the salivary protein nitrophorin did not show any signals of positive selection. We also found that the type of host (bats, humans or swallows) was correlated with the number of codons showing signals of positive selection at apyrase. However, contrary to our expectation, signals of positive selection were more frequent in bat-associated specimens than in swallow- or human-associated specimens. Our results indicate that association to bats has resulted in selective pressure on the gene coding for a salivary apyrase, a gene that is involved in preventing hemostasis at the feeding site in the host.

Host body temperature may be one possible factor that could result in divergent selection pressures on ectoparasites feeding on blood of bats, versus humans or swallows. We collected bat-associated individuals mostly on bats of Vespertilionidae, a diverse clade of insectivorous bats. Heterothermy is very common in these bats, and individuals frequently go into short bouts of torpor during which body temperature may drop rapidly (Stawski et al. 2014). Blood temperature of humans, on the other hand, is very stable, staying at around 37 degrees Celsius (Schmidt and Thews 1989). Blood temperature in barn swallows

is also stable at around 41 degrees Celsius (Møller 2010). DeVries et al. (2016) suggested host body temperature (simulated with a temperature-controlled feeder) may affect the blood-feeding behaviour of *C. lectularius*, where optimal temperatures ranged between 38-43°C. Therefore, apyrase in bat-associated taxa may be, or have been, affected by selective pressure from the relatively large daily temperature variation undergone in blood vessels of their bat hosts. However, this remains to be tested, by measuring the effect of amino acid substitutions on variation in 3D structure of cimicid salivary apyrase, and by measuring the efficiency of several salivary protein variants at different temperatures.

On average, the ratio of the volume of red blood cells on the total volume of blood (*i.e.* haematocrit) in birds and bats is higher than in humans on average (59% on average in bat species (Neuweiler 2000) and 55.9% in barn swallows in Europe (Saino et al. 1997); versus 42% in humans (Neuweiler 2000). Blood cells contain the ATP-rich heme complex, and ATP was shown to be the most effective phagostimulant in *C. lectularius* (Romero and Schal 2014), suggesting it is the main source of energy sought by *C. lectularius* while feeding. Ectoparasites feeding on humans may therefore need blood to flow for longer, compared to ectoparasites feeding on bats or swallows, to gain a similar amount of energy. As a result, one might expect selection on human-associated cimicids to prevent hemostasis more effectively. However we did not see strong evidence for this as we obtained only one significant signal of positive selection in one lineage of human-associated cimicids in the apyrase gene. However, this lineage also includes two specimens from bat-associated *C. lectularius* populations, which were both found to possess an additional putatively advantageous substitution.

Our results suggest lower haematocrit may not pose an important selective pressure on salivary apyrase or nitrophorin in cimicids feeding on humans, although it may pose a selective pressure on other genes not looked at in this study, or that the effects of lower haematocrit may be weak compared to those of variation in body temperature. Additionally, a decrease in overall genetic diversity in human-associated cimicid populations potentially caused by recent founding events, as suggested by Booth et al. (2015).

We observed significant signals of negative selection on fragments of both the apyrase and nitrophorin genes. Less than a third of codons identified as under negative selection in the apyrase fragment are active sites binding to calcium (a promoter) and to nucleotides (where ADP is hydrolysed; Dai et al. 2004). Codons displaying signals of negative selection in both proteins are located in the interior and exterior of the 3D structure of the proteins, suggesting selection affects the whole protein structure, and is not limited to the active sites (Dai et al. 2004; Protein Data Bank, RCSB). As expected, variation is very low at these specific codons in the two salivary protein genes, as changes in amino acid at those codons are presumably unfavourable. However, it is currently unclear how specific changes in amino acid at those codons would affect the function of the enzymes.

5.4.2 Species tree

Bats have been hypothesized to be the original host of species of *Cimex* and association to birds and humans putatively appeared subsequently (Usinger 1966). One interesting finding from our study is that most bat-associated

specimens show one or more signals of selection in a salivary protein gene, which is associated with feeding for these organisms, and almost none of the human-associated specimens show signals of positive selection. Interestingly, if bats are the original host of the group, one might expect selection on salivary proteins to operate more strongly on specimens associated with different hosts than bats. For example, persistence of a strain of the rabies virus in a new host (for example, from a bat species, which is thought to be the original host of rabies, to a canine species) requires multiple adaptive changes pertaining to replication and transmission (Mollentze et al. 2014). It is important to note that our study was performed on only two genes, and analysing genes coding for other functional traits related to host use are essential to fully understand adaptation of cimicids to their hosts. Our study however is an important first step in an investigation of signals of positive selection in salivary protein gene sequences in this diverse and cosmopolitan group of ectoparasites. Our study also provides evidence of positive selection linked with association to a specific type of host in a group of temporary ectoparasites associating with a range of mammals and birds.

5.5 Literature Cited

Abrams AJ, Cannatella DC, Hillis DM, Sawyer SL (2013) Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. *Journal of General Virology*, **94**, 2082–2093.

Arcà B, Struchiner CJ, Pham VM, Sferra G, Lombardo F, Pombi M, Ribeiro JMC (2014) Positive selection drives accelerated evolution of mosquito salivary genes associated with blood-feeding. *Insect Molecular Biology*, **23**, 122–131.

Balashov YS (2006) Types of parasitism of acarines and insects on terrestrial vertebrates. *Entomological Review*, **86**, 957–971.

Balashov YS (2011) Parasitism and ecological parasitology. *Entomological Review*, **91**, 1216–1223.

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Balvín O, Roth S, Vilímová J (2015) Molecular evidence places the swallow bug genus *Oeciacus* Stål within the bat and bed bug genus *Cimex* Linnaeus (Heteroptera: Cimicidae). *Systematic Entomology*, **40**, 652–665.

Benoit J (2011) Stress tolerance of bed bugs: a review of factors that cause trauma to *Cimex lectularius* and *C. hemipterus*. *Insects*, **2**, 151–172.

Berman HM (2000) The Protein Data Bank. *Nucleic Acids Research*, **28**, 235–242.

Brown RP, Yang Z (2011) Rate variation and estimation of divergence times using strict and relaxed clocks. *BMC Evolutionary Biolology*, **11**, 271.

Criado PR, Belda Junior W, Criado RFJ, Vasconcelos e Silva R, Vasconcellos C (2011) Bedbugs (Cimicidae infestation): the worldwide renaissance of an old partner of human kind. *The Brazilian journal of infectious diseases*, **15**, 74–80.

Dai J, Liu J, Deng Y, Smith TM, Lu M (2004) Structure and protein design of a human platelet function inhibitor. *Cell*, **116**, 649–659.

Decaestecker E, Gaba S, Raeymaekers JAM et al. (2007) Host–parasite “Red Queen” dynamics archived in pond sediment. *Nature*, **450**, 870–873.

DeVries ZC, Mick R, Schal C (2016) Feel the heat: activation, orientation and feeding responses of bed bugs to targets at different temperatures. *The Journal of Experimental Biology*, **219**, 3773–3780.

Dick CW, Patterson BD (2007) Against all odds: explaining high host specificity in dispersal-prone parasites. *International Journal for Parasitology*, **37**, 871–876.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.

Francischetti IMB, Calvo E, Andersen JF et al. (2010) Insight into the sialome of the bed bug, *Cimex lectularius*. *Journal of Proteome Research*, **9**, 3820–3831.

Goddard J (2009) Bed bugs (*Cimex lectularius*) and clinical consequences of their bites. *The Journal of the American Medical Association*, **301**, 1358–1366.

Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host-parasite coevolutionary arms races give way to fluctuating selection: bacteria-phage coevolutionary dynamics. *Ecology Letters*, **14**, 635–642.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.

Heled J, Bouckaert RR (2013) Looking for trees in the forest: summary tree from posterior samples. *BMC Evolutionary Biology*, **13**, 221.

Heled J, Bouckaert R, Xie W, Drummond AJ (2013) *BEAST in BEAST 2.2.x: estimating species trees from multilocus data. *GitHub*. See http://beast2-dev.github.io/beast-docs/beast2/STARBEAST/StarBEAST_tutorial.html (accessed on 10 November 2016).

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.
Jacobs FMJ, Greenberg D, Nguyen N et al. (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, **516**, 242–245.
Leggett HC, Buckling A, Long GH, Boots M (2013) Generalism and the evolution of parasite virulence. *Trends in Ecology and Evolution*, **28**, 592–596.
Mahamdallie SS, Ready PD (2012) No recent adaptive selection on the apyrase of Mediterranean *Phlebotomus*: implications for using salivary peptides to vaccinate against canine leishmaniasis: adaptive evolution of sandfly apyrase. *Evolutionary Applications*, **5**, 293–305.
Mans BJ, Louw AI, Neitz AWH (2002) Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. *Molecular biology and evolution*, **19**, 1695–1705.
Mollentze N, Biek R, Streicker DG (2014) The role of viral evolution in rabies host shifts and emergence. *Current Opinion in Virology*, **8**, 68–72.
Møller AP (2010) Body temperature and fever in a free-living bird. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **156**, 68–74.
Murrell B, Wertheim JO, Moola S et al. (2012) Detecting individual sites subject to episodic diversifying selection. *Plos Genetics*, **8**, e1002764.
Neuweiler G (2000) *The biology of bats*. Oxford University Press, New York, NY, USA.
Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
Paterson S, Vogwill T, Buckling A et al. (2010) Antagonistic coevolution accelerates molecular evolution. *Nature*, **464**, 275–278.
Pinheiro A, Woof JM, Abi-Rached L, Parham P, Esteves PJ (2013) Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens. *Plos One*, **8**, e73934.
Pond SLK, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, **22**, 1208–1222.
Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, **21**, 676–679.
Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW (2006) GARD: a genetic algorithm for recombination detection. *Bioinformatics*, **22**, 3096–3098.
Purvis A, Rambaut A (1995) Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Bioinformatics*, **11**, 247–251.
Romero A, Schal C (2014) Blood constituents as phagostimulants for the bed bug *Cimex lectularius* L. *Journal of Experimental Biology*, **217**, 552–557.

Saino N, Cuervo JJ, Ninni P, De Lope F, Moller AP (1997) Haematocrit correlates with tail ornament size in three populations of the Barn Swallow (*Hirundo rustica*). *Functional Ecology*, **11**, 604–610.
Scheffler K, Martin DP, Seoighe C (2006) Robust inference of positive selection from recombining coding sequences. *Bioinformatics*, **22**, 2493–2499.
Schmidt RF, Thews G (1989) *Human physiology*. Springer Berlin Heidelberg, Berlin, Germany.
Sponchiado J, Melo GL, Landulfo GA et al. (2015) Interaction of ectoparasites (Mesostigmata, Phthiraptera and Siphonaptera) with small mammals in Cerrado fragments, western Brazil. *Experimental and Applied Acarology*, **66**, 369–381.
Stawski C, Willis CKR, Geiser F (2014) The importance of temporal heterothermy in bats. *Journal of Zoology*, **292**, 86–100.
Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Valenzuela JG, Chuffe OM, Ribeiro JC (1996) Apyrase and anti-platelet activities from the salivary glands of the bed bug *Cimex lectularius*. *Insect Biochemistry and Molecular Biology*, **26**, 557–562.
Valenzuela JG, Ribeiro JM (1998) Purification and cloning of the salivary nitrophorin from the hemipteran *Cimex lectularius*. *The Journal of Experimental Biology*, **201**, 2659–2664.
Vermeij GJ (1987) *Evolution and escalation: an ecological history of life.* Princeton University Press, Princeton, NJ, USA.
Weichsel A, Maes EM, Andersen JF, Valenzuela JG, Shokhireva TK, Walker FA, Montfort WR (2005) Heme-assisted S-nitrosation of a proximal thiolate in a nitric oxide transport protein. *Proceedings of the National Academy of Science*, **102**, 594–599.

6 Host association influences variation at salivary protein genes in the bat ectoparasite *Cimex adjunctus*

6.1 Introduction

Antagonistic coevolution is a process whereby parasite or predator species develop phenotypic traits that are useful in extracting resources from their hosts or prey, while the hosts or prey develop reciprocal phenotypic traits that are useful at defending against resource loss to the parasite or predator (Hall et al. 2011; Pinheiro et al. 2013; Jacobs et al. 2014). When gene flow is not sufficient to counter the selection pressures presented by different hosts, even populations of a given parasite species may exhibit local adaptation to their sympatric hosts (Lively and Dybdahl 2000; Morgan et al. 2005; Garant et al. 2007). For example, the tick *Ixodes uriae*, a parasite of seabirds, shows adaptation to local kittiwake host populations (McCoy et al. 2002).

In generalist parasites, individuals may parasitize either one of a variety of host species, and occasionally even switch between two host species (Balashov 2006, 2011). Host species sometimes live in different habitats or display different spatial distribution, so gene flow between parasite populations associated with different host species sometimes is lower than between populations associated with the same host species. Differentiation among parasite populations that results from restricted gene flow, rather than from selection pressures presented by different hosts, should be reflected only in neutral genetic markers. For example, two races of the strong generalist common bed bug, *Cimex lectularius*, associate with bats and humans, respectively, in Europe, and show little contemporary gene flow between them, leading to genetic divergence (Balvín

et al. 2012; Booth et al. 2015). Also, the tick *Ixodes uriae* shows evidence of genetically distinct host-specialized races (McCoy et al. 2003, 2005; Kempf et al. 2009). Parasite populations may also show adaptation to local host species, which would be reflected in adaptive genetic markers, although hardly any documented example of this process currently exists. Lower gene flow among parasite populations associated with different host species could in turn be caused by the adaptation of parasite genotypes to a local host species. Thus, adaptation to local hosts may result in divergence among parasite populations at both adaptive and neutral markers. *Cimex adjunctus* ectoparasites show higher neutral genetic variation among populations associated with different, versus the same, bat host species in North America (Talbot et al. 2016b; Chapter 2). We wanted to investigate if higher genetic variation between populations associated with different host species is driven by adaptation to local host species, in addition to different habitats and distribution of the host species. To that end, we looked for signals of host adaptation among populations of *C. adjunctus* known to associate with at least three bat species in North America.

Insects in the genus *Cimex* (Order: Hemiptera) are temporary ectoparasites of warm-blooded animals. They do not remain on their host at all times but rather remain in nests or roosts between blood meals (Usinger 1966). Most *Cimex* species are associated exclusively with bats, while a few associate with a more diverse range of hosts (Usinger 1966; Goddard 2009; Criado et al. 2011). *Cimex adjunctus* is a widespread ectoparasite of bats in North America, occurring from the eastern seaboard to the Rocky Mountains, and from Labrador and the Northwest Territories south to Texas, and parasitizes a number of

insectivorous bat species, including the big brown bat (*Eptesicus fuscus*), and several species of *Myotis* (Usinger 1966). It typically remains in the hosts' roosts, emerging from cracks in the walls only to obtain blood meals (Usinger 1966). Talbot et al. (2016b; Chapter 2) found that neutral genetic variation between *C. adjunctus* individuals associated with different host species was considerable, likely due to infrequent gene flow between them (although mitochondrial genetic data showed a much smaller value; Talbot et al. 2016b; Chapter 2). They suggested that gene flow between individuals associated with different host species may occur during short intervals through the night between feeding bouts, when many species, including several myotis species and big brown bats roost together. They also found a relatively high proportion of genetic variation among bat roosts, suggestive of low continent-scale gene flow (Talbot et al. 2016b; Chapter 2).

Certain proteins expressed in salivary glands of the common beg bug, *C. lectularius,* are hypothesized to contribute to suppressing host defensive factors that promote coagulation and vasoconstriction at the site of rupture of a blood vessel, where the ectoparasite feeds (Francischetti et al. 2010). Apyrases possess an anti-platelet activity by catabolizing ADP released from damaged tissues, and a version of those is present in the saliva of *C. lectularius*, and may be used to prevent clotting (Valenzuela et al. 1996). Nitrophorins, also present in the saliva of *C. lectularius*, transport nitric oxide, which has a vasodilatation and anti-platelet activity, from the insect's salivary glands to the feeding site (Valenzuela and Ribeiro 1998). In Chapter 5, I found signals of positive selection in salivary apyrase gene sequences, but not in salivary nitrophorin gene sequences, across

Cimex specimens representing multiple species that typically associate with bats. Their results suggest adaptation of *Cimex* individuals at the apyrase gene to the bat species with which they associate. In this study, we wanted to examine signals of adaptive variation at salivary protein genes within a single *Cimex* species, among individuals associated with different bat host species.

Three of the potential hosts of *C. adjunctus* (the big brown bat, *E. fuscus*, the little brown myotis, *Myotis lucifugus*, and the northern myotis, *Myotis septentrionalis*) are widespread in North America. Although all of them span the whole continent coast to coast, big brown bats do not go as far north as little brown bat, and northern myotis is much more common in the northeastern portion of North America: New England states, Quebec, Ontario, and the Maritimes provinces (Fenton and Barclay 1980; Kurta and Baker 1990; Caceres and Barclay 2000). There are several life history and habitat differences among the three bat species. First, while the big brown bat and the little brown myotis often roost in buildings (Furlonger et al. 1987; Ellison et al. 2007; Pearce and O'Shea 2007), the northern myotis usually roosts in trees (Czenze and Broders 2011). The northern myotis changes roost from day to day, whereas the little brown myotis and the big brown bat usually show more roost fidelity (Lewis 1995; Czenze and Broders 2011). Females of the two latter species congregate in larger numbers than female northern myotis, although males of all three species typically roost individually or in small numbers (Vonhof et al. 2008; Czenze and Broders 2011). *Myotis* species tend to roost in warm, humid, covered and edge habitats, whereas big brown bats are less restricted in their roosting locations, and tend to roost in lighted areas in urban or suburban environments (Furlonger et al. 1987).

Differences in roost preference and roosting numbers may lead to differences in the range of temperatures typically encountered by bats and their ectoparasites. Ambient and host body temperature can, in turn, affect the blood-feeding behavior of cimicid ectoparasites. *C. lectularius,* for example, feeds preferentially at a relatively narrow range of ambient temperatures (Johnson 1941; DeVries et al. 2016). Variation in ambient and host body temperature could be one factor that underlying adaptation of *C. adjunctus* populations to different host species.

Host body size and blood volume is a second potential factor influencing adaptation of *C. adjunctus* populations to different host species. Big brown bats are larger than the two *Myotis* species. Forearm length is generally representative of whole body size in bats, and is 34-38 mm in *M. septentrionalis* (Caceres and Barclay 2000), 33-41 mm in *M. lucifugus* (Fenton and Barclay 1980), and 39-54 mm in *E. fuscus*, (Kurta and Baker 1990). A larger body is usually associated with a larger total volume of blood (Dewey et al. 2008). Stress experienced as a result of blood feeding by a given parasite is potentially weaker in host animals that possess a larger total volume of blood. For that reason, defensive responses to blood-feeding parasites, and the selection pressures exerted on those parasites, may also be weaker in larger hosts.

We hypothesized that populations of *C. adjunctus* associated with host species that differ in roosting behaviour and body size may be differentiated at candidate genes linked to blood feeding. We investigated if there is any signal of local adaptation in *C. adjunctus* to sympatric host species at two salivary protein genes, an apyrase and a nitrophorin. We predicted more genetic variation among

populations associated with different host species at genes coding for the salivary protein genes as compared to neutral genetic markers. We also investigated the influence of association with ecologically different bat host species on allelic divergence in genotypes of individual *C. adjunctus* at the same two genes. We predicted that association with one of the three potential host species would be an important predictor of allelic divergence at genes coding for the apyrase and nitrophorin genes.

6.2 Material and Methods

6.2.1 Data collection

We primarily used *Cimex adjunctus* samples from a previous phylogeographic study of the species across its range in North America (a total 118 specimens previously used in Talbot et al. 2016b; Chapter 2; Fig. 6.1; Table E.1), along with a small number of additional samples (a total of 5 specimens; Table E.1). Most specimens were removed from mist-netted host individuals of *Eptesicus fuscus*, *Myotis lucifugus* or *Myotis septentrionalis*. We used mist nets to catch bats near summer roosts (house, barn, cabin, church, school or abandoned mine) or within forested national, provincial, state or territorial lands (Talbot et al. 2016b; Chapter 2). Most mist-netted bats and the *C. adjunctus* individuals they harboured likely came from the adjacent known roost, although it is possible that a small proportion came from different roosts in the area. Overall, between 3 and 15% of mist-netted bats harboured a parasite, depending on the location. We also used *C. adjunctus* individuals from the interior of two summer roosts. One roost was in a house attic inhabited by *E. fuscus* (Talbot et al. 2016b; Chapter 2). Upon

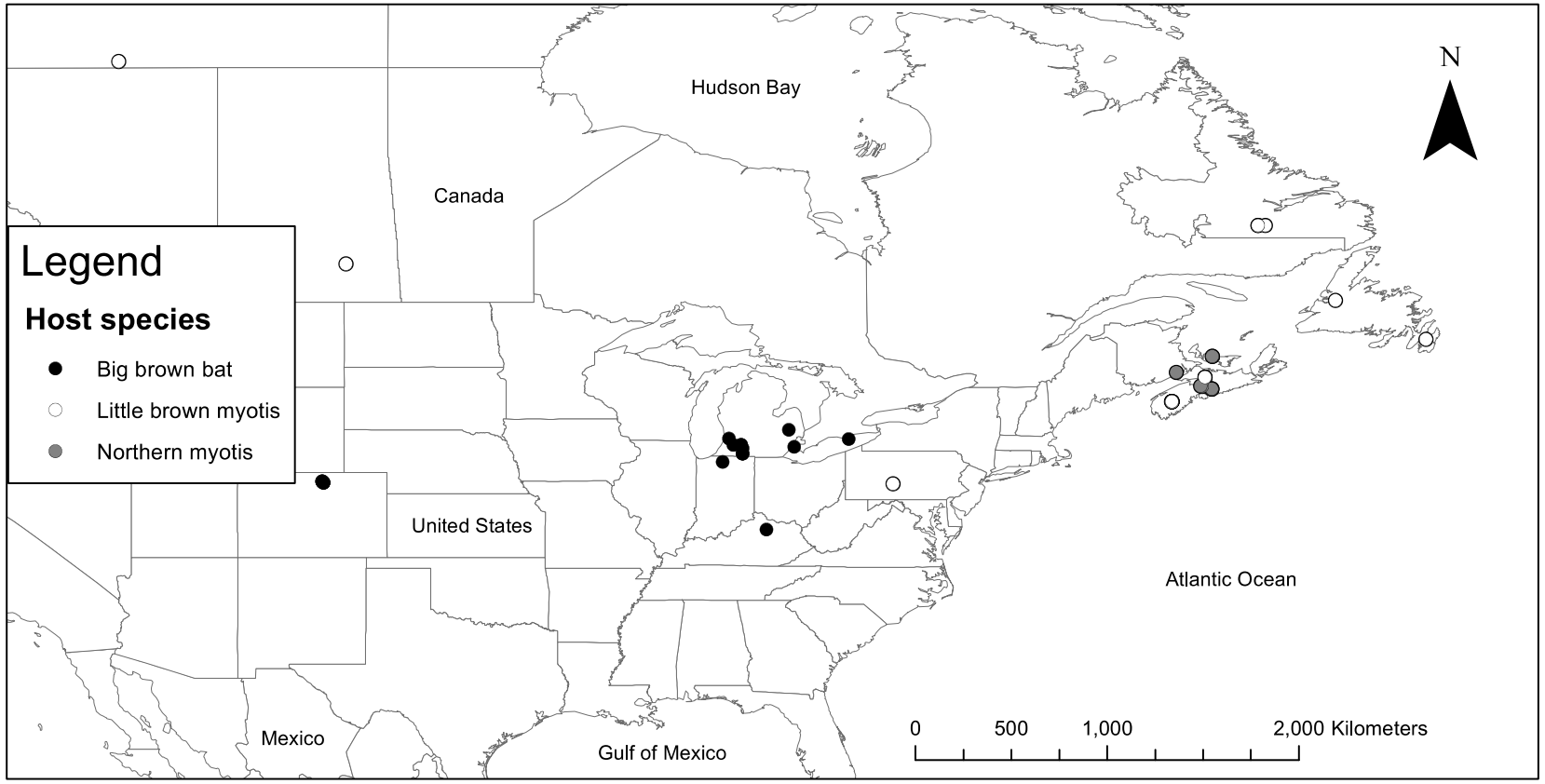

Figure 6.1 Sampling locations of *Cimex adjunctus* in North America. Created with ArcGIS v10.3 (ESRI, Redlands, USA). Sampling locations are color-coded to represent the host to which they were associated, as in the Legend.

collection, we immediately stored whole bodies, individually, in a 95% ethanol solution (0.95/1 mL) until further analyses. All samples included in this study were confirmed as being *C. adjunctus* using a DNA barcoding approach (Hebert et al. 2003). We compared the *CO1* sequence of each sample (see Talbot et al. 2016b; Chapter 2 for protocol) to known *CO1* sequences for *Cimex* species from published sources (Balvín et al. 2015). For each *C. adjunctus*, we recorded the exact GPS coordinates, the bat species on which it was collected (or the species inhabiting the roost where it was collected), and the year in which it was collected.

6.2.2 Genetic analyses

We extracted DNA from stored whole body samples using the DNeasy Blood and Tissue Kit (QIAGEN, Germantown, Maryland, United States). We then used primers designed by our group to amplify fragments of two genes coding for salivary proteins: an apyrase and a nitrophorin (Chapter 5). To the 5' end of these primers we added the Common Sequence 1 (CS1; on forward primers) and Common Sequence 2 (CS2; on reverse primers) universal tag sequences (Table E.2) from Fluidigm Corporation (South San Francisco, California, United States). These tags are needed for next generation sequencing library preparation using the Access Array system (Fluidigm Corp., South San Francisco, California, United States).

We used a DNAEngine PTC-200 Thermal Cycler (BIO-RAD, Hercules, California, United States) to execute the Polymerase Chain Reaction (PCR) amplification. We performed PCR in 25 µL final volume using the following recipe: 1X Taq Polymerase Buffer excluding $MgCl_2$ (Applied Biosystems, Foster City,

California, United States), 1.5 mM of $MgCl_2$, 0.2 mM of each type of dNTP, 0.3 µM of each primer, 1 U of Taq polymerase (ABI), and 2 µL of DNA extraction product. We used the following PCR program: an initial denaturation step of 1 min at 94°C, followed by 35 cycles of 30 sec of denaturation at 94°C, 45 sec of annealing at 53 – 55 °C (Table E.2) and 45 sec of extension at 72°C, finished by a final extension step of 5 min at 72°C. We visualized PCR products by 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad) on a UV transluminator to check the quality and size of amplified fragments. We then pooled the two salivary gene fragments together, for each individual.

We performed a second PCR step to prepare next generation sequencing libraries (for complete protocol, see Kennedy et al. 2014), wherein sequences from each individual sample are given a unique identifier. We visualized PCR products by 1.5% agarose gel electrophoresis using SYBR Green (Bio-Rad) on a UV transluminator to check the quality and size of amplified fragments. We then pooled all individuals together. We used the Illumina MiSeq paired-ends 250bp technology (Illumina Inc., San Diego, California, United States) to sequence all fragments, using a total of 100,000 reads.

We used Galaxy Tools to sort and align families of reads for each individual, and to produce single-strand consensus sequences (SSCS, using the Du Novo pipeline to detect mutations with increased accuracy; Schmitt et al. 2012). We produced SSCSs for families of reads containing two reads or more and determined a consensus nucleotide for sites with a phred score (Ewing et al. 1998) of 35 or more. We aligned all SSCSs using MEGA 6.06, and kept one copy of each allele (one allele in homozygotes and two in heterozygotes) for each of the

two salivary protein gene fragments, for each individual. We did not consider SSCSs that were present in only one copy, that were ambiguous for more than 5% of the sequence, that did not match with the reverse strand of the same read family, or that showed errors in the primer sequence, because these cases are indicative of amplification or sequencing errors. We converted DNA sequences into amino acid sequences, to remove all synonymous variation, using the universal genetic code, and attributed a sequence number to each unique allele based on the amino acid sequence. We calculated the divergence of each amino acid allele from all other observed alleles by calculating the sum of pairwise substitutions between each allele and all other alleles. Allelic divergence is a relative measure of how different each allele is compared to all other ones, such as if it underwent relatively more episodes of positive selection. In cases of insertions or deletions, we treated gaps as nucleotides, as these types of mutations may also contain signals of selection.

Finally, we collected microsatellite genotype data at nine loci (Clec15, Clec21, Clec48, Clec104, BB28B, Cle002, Clec003, Cle013, Cle015), for 118 samples (96% of this study's dataset; Table E.1), that were initially used in Talbot et al. (2016b; Chapter 2). Talbot et al. (2016b; Chapter 2) found a total of ten genetic clusters in a slightly larger dataset of 150 individuals, using allele frequencies at the same microsatellite loci, and genetic clusters were generally concordant with geographic location across North America. They attributed a genetic cluster number to each individual of their study, and we also used this cluster membership information in this study (Table E.1).

6.2.3 Statistical analyses

Local host adaptation

First, we tested our prediction of greater adaptive than neutral genetic variation among populations of *C. adjunctus* associated with different host species. We compared the results of analysis of molecular variance (AMOVA) on genotypes at the salivary protein genes (based on DNA sequence) to that on microsatellite loci. We used all sites, defined as unique sampling locations, with at least two sampled individuals, and with data for microsatellite loci and both salivary protein genes (a total of 86 individuals, or 70% of the dataset; Table E.1). Separately for each of the two adaptive (salivary protein) genes and the neutral dataset (nine microsatellite loci), we used hierarchical AMOVA to examine the proportion of genetic variation among populations associated with different host species, among populations from different sites but associated with the same host species (nested by host species), and within sites. We executed the AMOVA with GenoDive 2.0 for microsatellite data, and Arlequin 3.5 for salivary protein gene data.

Effect of host species on allelic divergence at salivary protein genes

Next, we tested our prediction that the host species with which individual *C. adjunctus* were associated would influence allelic divergence in genotypes at salivary protein genes. We used a generalized linear mixed-effects modelling approach, to account for both fixed- and random-effects variables. We used allelic divergence as the response variable (Table E.3). We accounted for non-independence of alleles from the same individual, and for neutral patterns of

genetic structuring, by treating the individual identifier and genetic cluster number as random factors. We nested the individual identifier within its genetic cluster, to account for the hierarchical organization of these variables. To test for any temporal or spatial component of variation at the salivary protein genes, we included the year of collection, and the latitude and longitude of the collection site as fixed factors. Finally, to test for an effect of the host species with which each individual was associated, we included host species as a fixed factor. We standardized numerical variables (latitude, longitude, year and allelic divergence) to reduce the risk of collinearity between variables in the models, using the 'scale' function in R 3.3.2. We compared models with every possible combination of fixed factors, as well as a null model that included only the random factors. We performed the analysis for each gene fragment separately, using the 'glmer' function from the 'lme4' package (Bates et al. 2015) in R 3.3.2, using the gaussian family. We used model selection to find all models that have a delta Akaike Information Criterion (ΔAICc) value ≤ 4, using the 'model.avg' function in the 'MuMIn' package (Barton 2013) in R 3.3.2. Using the same function, we calculated model-averaged coefficients for the four fixed variables. This approach allowed us to calculate the relative importance of year, latitude, longitude and host species on alleles at salivary protein genes in *C. adjunctus*.

6.3 Results

We obtained a genotype for at least one salivary protein gene fragment in 123 individuals, and for both fragments in 105 individuals (Table E.1). We obtained apyrase genotypes for 113 individuals, 78 being homozygotes and 35

being heterozygotes. We obtained nitrophorin genotypes for 115 individuals, 74 being homozygotes and 41 being heterozygotes. We obtained 26 unique apyrase DNA sequences, and 29 unique nitrophorin DNA sequences (with each sequence varying from all other ones on average by 18 and 11 mutations, respectively), which correspond to 17 unique apyrase amino acid sequences, and 21 unique nitrophorin amino acid sequences (each sequence varying from all other ones on average by 7 and 6 codons, respectively; Table E.3). Length of alleles in the apyrase fragment varied between 353 and 371 bp, due to at least two events of deletion and/or insertion. After alignment, we trimmed the start and end of all sequences, so that they included only whole codons, resulting in a 369 bp alignment for the apyrase fragment and a 297 bp alignment for the nitrophorin fragment.

6.3.1 Local host adaptation

AMOVA results were different among the three types of markers. For microsatellites, less than half of the total genetic variation was among populations associated with different host species (45.5%; Table 6.1), and more than a third was among populations from different sites but associated with the same host species (34.4%). For both candidate salivary protein genes, the proportion of genetic variation among populations associated with different host species was higher than for microsatellites (50.4% for nitrophorin and 68.8% for apyrase; Table 6.1). Similarly, for both candidate genes, the proportion of genetic variation among populations from different sites associated with the same host species was lower than for microsatellites (29.1% for nitrophorin and 24.5% for apyrase; Table 6.1).

Table 6.1 Results of analysis of molecular variance (AMOVA) on *Cimex adjunctus*, using microsatellite and salivary protein gene (apyrase and nitrophorin) data. Percentage of total genetic variation among host species, among sample sites, and within sample sites are shown.

Source of Variation	microsatellite	apyrase	nitrophorin
Among host species	45.5	68.8	50.4
Among sites	34.4	24.5	29.1
Within sites	20.1	6.7	20.5

These results suggest genetic variation in *C. adjunctus* at both candidate genes, but particularly apyrase, may be driven by adaptation related to the local host species.

6.3.2 Effect of host species on allelic divergence at salivary protein genes

The top model with the lowest AICc value for the apyrase fragment included only latitude and host species (Table 6.2). All other models had ΔAICc > 4. Model averaging revealed an Akaike weight of 0.82 for the top ranked model for apyrase (Table 6.2). There was no single, best model for the nitrophorin fragment. Four models had ΔAICc value less than 4 for the nitrophorin fragment, and all of them included host species. Three of those models also included either or a combination of latitude, longitude and year of collection (Table 6.2). Model averaging revealed an Akaike weight of 0.52 for the top model and 0.18, 0.14 and 0.09 for the second, third and fourth ranked models for nitrophorin (Table 6.2).

Model averaging revealed host species to be a highly important predictor in models of both apyrase and nitrophorin allelic divergence (relative variable importance = 1.00; Table 6.3). Models for both genes showed a larger effect of association with northern myotis, than with either big brown bat or little brown myotis, on allelic divergence. That is, the estimates of model coefficients were similar for big brown bat and little brown myotis, but both were very different from the estimate for northern myotis. Thus, association with one or the other of big brown bat and little brown myotis did not have a large effect on allelic divergence (Table 6.3). Latitude was also highly important in models of apyrase and moderately important in models of nitrophorin (relative variable importance = 1.00

Table 6.2 Model selection parameters (degrees of freedom, log-likelihood and ΔAICc) and model-averaged Akaike weights (Weight) of various generalized linear mixed-effects models explaining geographic distribution of apyrase and nitrophorin alleles of *C. adjunctus* in North America.

Gene fragment	Models	df	Log-likelihood	AICc	ΔAICc	Weight
Apyrase	Host + Latitude	7	-53.32	121.2	0.00	0.82
Nitrophorin	Host + Latitude	7	-130.3	275.1	0.00	0.52
	Host	6	-132.4	277.1	2.13	0.18
	Host + Longitude	7	-131.6	277.7	2.60	0.14
	Host + Latitude + Year	8	-131.0	278.6	3.52	0.09

Table 6.3 Model-averaged coefficients in models including all possible combinations of four variables (host species, longitude, latitude and year of collection) explaining the distribution of salivary protein genotypes of *Cimex adjunctus* in North America. Intercept (Host A) represents the northern myotis, Host B represents the little brown bat and Host C represents the big brown bat.

Gene fragment	Variable	Estimate	Standard error	Relative variable importance
Apyrase	Intercept (Host A)	1.154	0.213	1.00
	Host B	-1.546	0.156	
	Host C	-1.373	0.211	
	Longitude	-0.008	0.038	0.10
	Latitude	0.383	0.062	1.00
	Year	-0.007	0.027	0.09
Nitrophorin	Intercept (Host A)	1.318	0.210	1.00
	Host B	-1.246	0.193	
	Host C	-1.569	0.201	
	Longitude	0.114	0.088	0.20
	Latitude	0.165	0.059	0.66
	Year	-0.081	0.058	0.12

and 0.66 respectively; Table 6.3). Models for both genes showed a positive effect of latitude on allelic divergence. Longitude and year of collection were unimportant in models of both genes (relative variable importance < 0.3; Table 6.3). These results suggest the host species (northern myotis versus either big brown bats or little brown myotis) with which individual *C. adjunctus* were associated at the time of collection, and spatial location measured as latitude, are important predictors of allelic divergence in genotypes at both salivary protein genes.

For both salivary protein genes, specific amino acid alleles in *C. adjunctus* were associated with the different host species (Fig. 6.2). Alleles 1 and 4 of the apyrase gene fragment occurred at highest frequency (39 and 57%, respectively) in individuals associated with northern myotis, while alleles 3, 9 and 10 were the most common in individuals associated with big brown bats and little brown myotis (e.g. allele 3 occurred at 49 and 56% in individuals associated with big brown bat and little brown myotis, respectively; Fig. 6.2). Likewise, alleles 1, 19 and 20 of the nitrophorin gene fragment occurred at highest frequency (54, 15 and 27%, respectively) in individuals associated with northern myotis, alleles 1 and 3 were the most common in individuals associated with little brown myotis (13 and 74%, respectively), and alleles 3 and 4 were the most common in individuals associated with big brown bats (35 and 53%, respectively; Fig. 6.2). Overall, therefore, both genes varied most between populations of *C. adjunctus* associated with northern myotis versus the other two host species.

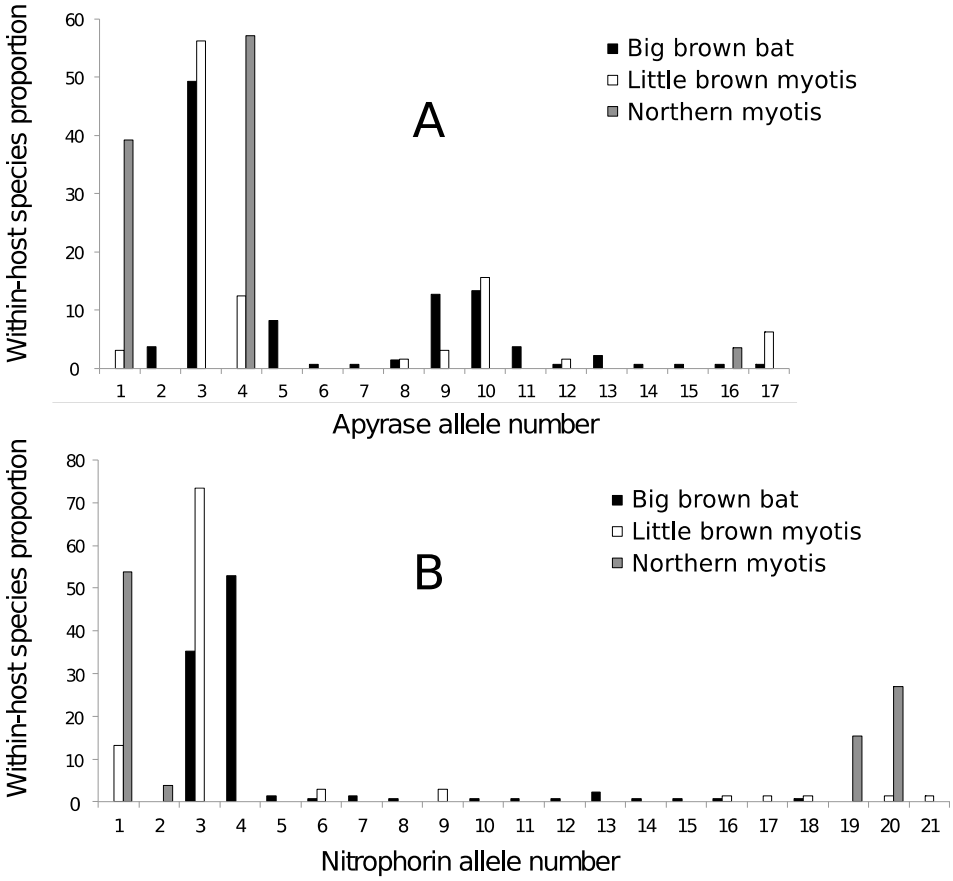

Figure 6.2 Bar plots representing the within-host species proportion of apyrase (A) or nitrophorin (B) alleles from individuals of *Cimex adjunctus* associated with one of three bat host species.

Alleles of the apyrase gene fragment associated with northern myotis (alleles 1 and 4) were different from those associated with the other hosts (alleles 3, 9 and 10) by a total of 18 amino acid substitutions, and one indel of five amino acids (Fig. 6.3). A total of 12 of these substitutions represent amino acids of different charge, and one is an extreme switch from a small, charged amino acid (alanine) to a large, non-polar amino acid (glutamic acid). One additional substitution represents a switch between a small and a large amino acid (glycine and valine; Fig. 6.3). By comparison, among those alleles associated with big brown bats and little brown myotis (alleles 3, 9 and 10), only five substitutions were found, although one represents a substantial change in both charge and size (valine and glutamic acid). The nitrophorin alleles associated with one or more host species (alleles 1, 3, 4, 19 and 20) varied by a total of 15 substitutions. Eight of these substitutions represent changes in amino acid charge, and one is an extreme switch from a small, nonpolar amino acid (glycine) to a large, charged amino acid (glutamic acid). Our results suggest variation in the three-dimensional structure of salivary apyrases and nitrophorins of *C. adjunctus* individuals associated with different bat host species, with more extensive variation between individuals associated with the northern myotis versus the other two host species, big brown bats and little brown myotis.

6.4 Discussion

6.4.1 Signal of local adaptation

Neutral genetic variation estimated with microsatellite loci was

Figure 6.3 Alignments of amino acid sequences of the five highest frequency alleles from each of two salivary protein gene fragments in *Cimex adjunctus* (alleles 1, 3, 4, 9 and 10 for the apyrase gene, and alleles 1, 3, 4, 19 and 20 for the nitrophorin gene). Amino acids that did not vary are shown only in the first allele. Amino acids that vary among alleles are color-coded to represent electrical characteristics: Light shade represents charged, medium shade represents polar uncharged and dark shade represents non-polar.

highest among host species, consistent with results of Talbot et al. (2016b; Chapter 2). Abundance of potential bat host species varies across the continent (Fenton and Barclay 1980; Kurta and Baker 1990; Caceres and Barclay 2000), and bats usually do not share summer roosts with other species. Also, northern myotis typically lives in different habitats than the other two species (Furlonger et al. 1987; Ellison et al. 2007; Pearce and O'Shea 2007; Czenze and Broders 2011). All those factors could reduce the extent of gene flow in *C. adjunctus* populations associated with different host species.

However, consistent with a scenario of local adaptation to their hosts, genetic variation among populations of blood-feeding *C. adjunctus* associated with different host species was higher at two salivary protein genes compared to neutral microsatellite loci. In addition to the homogenizing effect of gene flow and the differentiating effect of genetic drift, selection may influence genetic variation at certain genes that are important for survival and reproduction (Garant et al. 2007). In *C. adjunctus*, which eats blood, genes involved in the blood-feeding behaviour may experience strong selective pressure (Chapter 5; Mans et al. 2002; Mahamdallie and Ready 2012). Locally varying, or diversifying, selection pressure linked with the host species could drive populations of *C. adjunctus* on different host species to be more genetically different at genes involved in blood feeding than expected as a result of neutral processes alone (*i.e.*, genetic drift counteracted by gene flow). A potential source of diversifying selective pressure would be environmental conditions experienced by individuals living in different areas, which has been observed in a variety of different organisms (Nunes et al. 2011; Hancock et al. 2011; Talbot et al. 2016a; Pluess et al. 2016). Our results

therefore suggest that association of *C. adjunctus* with different host species that live in different habitats may result in diversifying selective pressure. Furthermore, it is possible that adaptation of parasite genotypes to a local host species may further reduce gene flow among parasite populations due to maladaptation of individuals that may try to colonize a new host species (Lively and Dybdahl 2000; Morgan et al. 2005).

Finally, genetic variation among populations of *C. adjunctus* associated with different host species was higher at apyrase than nitrophorin, which is consistent with the findings of Chapter 5, where signals of positive selection were found in the apyrase gene, but not in the nitrophorin gene. Interestingly, all five codons in which I previously (Chapter 5) detected signals of positive selection in the apyrase gene showed differences among alleles associated with northern myotis versus either with big brown bat or little brown myotis. In Chapter 5, I also found signals of negative selection at both apyrase and nitrophorin, suggestive of stabilizing selection on key portions of the sequence of the two salivary protein genes. Our result of greater adaptive than neutral variation associated with host species suggests that a negative selective pressure at the two candidate genes in *C. adjunctus* is potentially weaker than a diversifying selective pressure related to the host species. Finally, it is important to note that these analyses were performed on DNA sequences, and could thus incorporate a larger component of neutral genetic variation than would amino acid sequences.

6.4.2 Effect of host species on allelic divergence at salivary protein genes

As predicted, we found a significant and strong effect of the host species with which individual *C. adjunctus* associated on allelic divergence in genotypes at both salivary protein genes. This effect was mostly driven by association with northern myotis versus either big brown bat or little brown myotis. Additionally, allele frequencies at both genes, and allele chemistry at apyrase, seemed to vary more between individuals associated with the northern myotis than those associated with big brown bat or little brown myotis. The northern myotis differs from big brown bat and little brown myotis in that it typically does not roost in buildings (Furlonger et al. 1987; Ellison et al. 2007; Pearce and O'Shea 2007), but rather in trees (Czenze and Broders 2011). The microenvironment of tree roosts may be more variable than those of building roosts. Many bats that occur in North America, including big brown bats and several *Myotis* species, are heterothermic (Stawski et al. 2014), meaning they frequently go into torpor. In *Eptesicus fuscus*body temperatures may drop by more than 15°C in 30 minutes (Audet and Fenton 1988). *C. adjunctus* living in trees may therefore experience higher variance in ambient temperature and also higher variance in their host's body temperature. In *C. lectularius*, ambient temperature has an important effect on feeding activity (Johnson 1941, DeVries et al. 2016). Both ambient and host body temperatures could affect blood-feeding behaviour of *C. adjunctus* as well, although no relevant experimental data are available for this species. Differences in ambient or host body temperatures experienced by *C. adjunctus* exploiting tree-dwelling versus building-dwelling bats could result in different selection pressures on blood feeding and related salivary proteins.

Our results also point to a possible spatial component of selection pressure, related mainly to the latitude. This could be a result of a difference in spatial distribution of the three bat host species (Fenton and Barclay 1980; Kurta and Baker 1990; Caceres and Barclay 2000). Alternatively, it could be an effect of average temperature, which typically decreases as latitude increases in North America, and potentially brings an additional stress on blood feeding in *C. adjunctus*, which is likely to be dependent on ambient temperature (Johnson 1941; DeVries et al. 2016).

Higher variance in host body and ambient temperatures in tree roosts could mean less than optimal feeding temperatures. Therefore, to meet their energy needs, *C. adjunctus* living in tree roosts have to be more efficient at extracting blood than those roosting in buildings. Apyrase and nitrophorin suppress hemostasis by the host (Valenzuela et al. 1996; Valenzuela and Ribeiro 1998; Francischetti et al. 2010). Hemostasis is a combined physiological response of hosts against feeding hematophagous ectoparasites, including vasoconstriction and coagulation. Our results suggest most amino acid differences among apyrase and nitrophorin alleles are between alleles shared by individuals associated with either little brown myotis or big brown bats versus those associated with northern myotis. Amino acid differences potentially affect the three-dimensional structure of the proteins (Dai et al. 2004), because they change the electrical affinity or the size of the amino acid, and in the case of the apyrase, the length of the amino acid sequence. These differences in turn possibly affect the efficiency of the enzymes at preventing hemostasis in the host.

In conclusion, we observed more adaptive genetic variation among sites inhabited by different host species than expected under a neutral scenario of only genetic drift and gene flow, in two salivary protein genes, in *C. adjunctus*. We also observed an important effect of the host species with which individual *C. adjunctus* were associated on the parasite's genotype at both salivary protein genes. Our results suggest selective pressure exerted on two salivary protein genes in *C. adjunctus* is strongly dependent on the host species with which they associate.

6.5 Literature Cited

Audet D, Fenton MB (1988) Heterothermy and the use of torpor by the bat *Eptesicus fuscus* (Chiroptera: Vespertilionidae): a field study. *Physiological Zoology*, **61**, 197–204.

Balashov YS (2006) Types of parasitism of acarines and insects on terrestrial vertebrates. *Entomological Review*, **86**, 957–971.

Balashov YS (2011) Parasitism and ecological parasitology. *Entomological Review*, **91**, 1216–1223.

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Balvín O, Roth S, Vilímová J (2015) Molecular evidence places the swallow bug genus *Oeciacus* Stål within the bat and bed bug genus *Cimex* Linnaeus (Heteroptera: Cimicidae). *Systematic Entomology*, **40**, 652–665.

Barton K (2013) *MuMIn: Multi-model inference*. Institute of Nature Conservation, Krakow, Poland.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**.

Booth W, Balvín O, Vargo EL, Vilímová J, Schal C (2015) Host association drives genetic divergence in the bed bug, *Cimex lectularius*. *Molecular Ecology*, **24**, 980–992.

Caceres MC, Barclay RMR (2000) *Myotis septentrionalis*. *Mammalian Species*, **34**, 1–4.

Criado PR, Belda Junior W, Criado RFJ, Vasconcelos e Silva R, Vasconcellos C (2011) Bedbugs (Cimicidae infestation): the worldwide renaissance of an old partner of human kind. *The Brazilian journal of infectious diseases*, **15**, 74–80.

Czenze ZJ, Broders HG (2011) Ectoparasite community structure of two bats (*Myotis lucifugus* and *M. septentrionalis*) from the Maritimes of Canada. *Journal of Parasitology Research*, **2011**, 1–9.
Dai J, Liu J, Deng Y, Smith TM, Lu M (2004) Structure and protein design of a human platelet function inhibitor. *Cell*, **116**, 649–659.
DeVries ZC, Mick R, Schal C (2016) Feel the heat: activation, orientation and feeding responses of bed bugs to targets at different temperatures. *The Journal of Experimental Biology*, **219**, 3773–3780.
Dewey FE, Rosenthal D, Murphy DJ, Froelicher VF, Ashley EA (2008) Does size matter?: Clinical applications of scaling cardiac size and function for body size. *Circulation*, **117**, 2279–2287.
Ellison LE, O'Shea TJ, Neubaum DJ, Bowen RA (2007) Factors influencing movement probabilities of big brown bats (*Eptesicus fuscus*) in buildings. *Ecological Applications*, **17**, 620–627.
Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. 1. Accuracy assessment *Genome Research*, **8**, 175–185.
Fenton MB, Barclay RMR (1980) *Myotis lucifugus*. *Mammalian Species*, **142**, 1–8.
Francischetti IMB, Calvo E, Andersen JF et al. (2010) Insight into the sialome of the bed bug, *Cimex lectularius*. *Journal of Proteome Research*, **9**, 3820–3831.
Furlonger CL, Dewar HJ, Fenton MB (1987) Habitat use by foraging insectivorous bats. *Canadian Journal of Zoology*, **65**, 284–288.
Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, **21**, 434–443.
Goddard J (2009) Bed bugs (*Cimex lectularius*) and clinical consequences of their bites. *The Journal of the American Medical Association*, **301**, 1358–1366.
Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host-parasite coevolutionary arms races give way to fluctuating selection: bacteria-phage coevolutionary dynamics. *Ecology Letters*, **14**, 635–642.
Hancock AM, Witonsky DB, Alkorta-Aranburu G et al. (2011) Adaptations to climate-mediated selective pressures in humans. *Plos Genetics*, **7**, e1001375.
Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.
Jacobs FMJ, Greenberg D, Nguyen N et al. (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature*, **516**, 242–245.
Johnson CG (1941) The ecology of the bed-bug, *Cimex lectularius* L., in Britain. *Journal of Hygiene*, **41**, 345–461.
Kempf F, Boulinier T, De MeeûS T, Arnathau C, McCoy KD (2009) Recent evolution of host-associated divergence in the seabird tick *Ixodes uriae*. *Molecular Ecology*, **18**, 4450–4462.
Kennedy SR, Schmitt MW, Fox EJ et al. (2014) Detecting ultralow-frequency mutations by Duplex Sequencing. *Nature Protocols*, **9**, 2586–2606.

Kurta A, Baker RH (1990) *Eptesicus fuscus*. *Mammalian Species*, **356**, 1–10.
Lewis SE (1995) Roost fidelity of bats: a review. *Journal of Mammalogy*, **76**, 481–496.
Lively CM, Dybdahl MF (2000) Parasite adaptation to locally common host genotypes. *Nature*, **405**, 679–681.
Mahamdallie SS, Ready PD (2012) No recent adaptive selection on the apyrase of Mediterranean *Phlebotomus*: implications for using salivary peptides to vaccinate against canine leishmaniasis: adaptive evolution of sandfly apyrase. *Evolutionary Applications*, **5**, 293–305.
Mans BJ, Louw AI, Neitz AWH (2002) Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. *Molecular biology and evolution*, **19**, 1695–1705.
McCoy KD, Boulinier T, Schjørring S, Michalakis Y (2002) Local adaptation of the ectoparasite *Ixodes uriae* to its seabird host. *Evolutionary Ecology Research*, **4**, 441–456.
McCoy KD, Boulinier T, Tirard C, Michalakis Y (2003) Host-dependant genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution*, **57**, 288–296.
McCoy KD, Chapuis E, Tirard C et al. (2005) Recurrent evolution of host-specialized races in a globally distributed parasite. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2389–2395.
Morgan AD, Gandon S, Buckling A (2005) The effect of migration on local adaptation in a coevolving host–parasite system. *Nature*, **437**, 253–256.
Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient: selection in occelated lizards. *Molecular Ecology*, **20**, 193–205.
Pearce RD, O'Shea TJ (2007) Ectoparasites in an urban population of big brown bats (*Eptesicus fuscus*) in Colorado. *The Journal of parasitology*, **93**, 518–530.
Pinheiro A, Woof JM, Abi-Rached L, Parham P, Esteves PJ (2013) Computational analyses of an evolutionary arms race between mammalian immunity mediated by immunoglobulin A and its subversion by bacterial pathogens. *Plos One*, **8**, e73934.
Pluess AR, Frank A, Heiri C et al. (2016) Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. *New Phytologist*, **210**, 589–601.
Schmitt MW, Kennedy SR, Salk JJ et al. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proceedings of the National Academy of Sciences*, **109**, 14508–14513.
Stawski C, Willis CKR, Geiser F (2014) The importance of temporal heterothermy in bats. *Journal of Zoology*, **292**, 86–100.
Talbot B, Chen T-W, Zimmerman S et al. (2016a) Combining genotype, phenotype, and environment to infer potential candidate genes. *Journal Of Heredity*, **108**, 207-216.

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016b) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Valenzuela JG, Chuffe OM, Ribeiro JC (1996) Apyrase and anti-platelet activities from the salivary glands of the bed bug *Cimex lectularius*. *Insect Biochemistry and Molecular Biology*, **26**, 557–562.
Valenzuela JG, Ribeiro JM (1998) Purification and cloning of the salivary nitrophorin from the hemipteran *Cimex lectularius*. *The Journal of Experimental Biology*, **201**, 2659–2664.
Vonhof MJ, Strobeck C, Fenton MB (2008) Genetic variation and population structure in big brown bats (*Eptesicus fuscus*): is female dispersal important? *Journal of Mammalogy*, **89**, 1411–1419.

7 General discussion

7.1 Use of hosts for movement by generalist ectoparasites

7.1.1 Continent-scale genetic structure and demographic history

In Chapter 2, I found a similar correlation between continent-scale geographic distances and genetic distances in *Cimex adjunctus*, a generalist ectoparasite of bats, and two of its key hosts, the big brown bat and the little brown myotis, in North America. However, I also found that, for the same geographical distances, levels of spatial genetic structure in *C. adjunctus* are generally much higher than in the two bat species. Similarly, demographic history of *C. adjunctus* is largely different than that of the two bat species.

Although *C. adjunctus* probably relies on its hosts for dispersal among roosts, it is interesting to see patterns suggestive of long-distance dispersal, where individuals from a single genetic cluster are spread over thousands of kilometers. On the other hand, there are as many as six genetic clusters in *C. adjunctus* in the relatively small area of the Great Lakes region. These patterns suggest a large degree of randomness in how cimicid insects disperse using their hosts. For example, they may inadvertedly find themselves on the body of a host individual at the moment it is leaving either to switch roosts or to forage. A succession of similar events could lead to occasional long-distance movements as suggested by my results. In an interesting parallel, bed bugs associated with humans typically disperse while within luggage, furniture or clothes (Davies et al. 2012). In some cases, they may disperse between continents while transported in human belongings.

My results on demographic history suggest that a small increase of effective population size of *C. adjunctus* occurred in the last few tens of thousands of years, possibly following Pleistocene climate warming (Swenson and Howard 2005). However, the increase of effective population size was of much higher magnitude in at least one of the hosts of *C. adjunctus*, the little brown bat. One possibility is that anti-parasitism strategies used by the host maintain low parasite numbers (ter Hofstede and Fenton 2005; Bartonička and Růžičková 2013). It seems as though hosts influence patterns of neutral genetic variation of their generalist ectoparasites, but the process may be much more complex than in more specialized parasite-host systems.

7.1.2 Spatial genetic structure of a weak versus a strong generalist

In Chapter 3, I found lower levels of contemporary population genetic structure in *C. adjunctus*, a species mostly found associated with highly mobile hosts, *i.e.* bats, than in *C. lectularius*, a geographically sympatric congener that is a much stronger generalist.

Although *C. lectularius* is known to associate with bats in Europe (Balvín et al. 2012), it was never documented on them in North America. Most typical hosts of *C. lectularius* in North America, other than humans, are heavily associated with humans, such as domestic animals (Usinger 1966). On the other hand, *C. adjunctus* associates with a number of bat species, and almost never on any other type of host (Usinger 1966). Bats are known to disperse quite frequently, for a number of factors, including bad environmental conditions in the roost, high numbers of parasites or the drive to associate with a large number of

conspecifics (Willis and Brigham 2004; Ellison et al. 2007; Bartonička and Růžičková 2013). Flight also allows them to be less impeded by landscape elements than terrestrial animals. My results are interesting because they suggest that a species associating with host species that are known to move regularly and over relatively long distances show lower levels of population genetic structure than a species associating with a much larger variety of animals.

Another important factor that may drive my results is the use of pyrethroid insecticides, which were heavily used on *C. lectularius* populations associated with humans and domestic animals for several tens of years (Davies et al. 2012; Zhu et al. 2013). Extirpations caused by the use of those pesticides may drive stronger genetic structure in *C. lectularius* than *C. adjunctus* in North America. Altogether, my results suggest the environmental conditions associated with a type of host may have an influence of neutral genetic variation, even for generalist ectoparasites.

7.1.3 Effect of landscape on spatial genetic structure of parasite and a host

In Chapter 4, I found that land cover affects spatial genetic structure in the big brown bat and *C. adjunctus* differently, potentially due to a difference in location and timing of gene flow between the two species.

While *C. adjunctus* likely relies on its hosts for dispersal, differences in life history result in differences in gene flow patterns between this ectoparasite and its hosts. In fact, bats have a complicated life cycle consisting of switching from multi-species hibernacula to usually mono-species summer roosts (Kurta 1995). Cimicid ectoparasites, on the other hand, have a relatively simpler life

cycle, consisting of reproduction within a roost during the summer, and occasional roost switching triggered by their bat hosts (Usinger 1966). It is interesting to see how, for parasites and hosts whose movements are linked, large differences in patterns of neutral genetic variation may be observed. These differences arise despite potentially linked life history characteristics that nonetheless lead to differences in the timing and nature of gene flow.

7.1.4 Conclusions

I conducted a continent-scale comparison of a cimicid species and two of its potential hosts (Chapter 2), as well as a smaller-scale comparison, exploring effects of landscape composition, of a cimicid species and one of its potential hosts (Chapter 4). I also conducted a comparison of spatial genetic structure of a weak generalist ectoparasite associated with highly mobile hosts and a closely related strong generalist ectoparasite associated with a wider range of animals (Chapter 3). Altogether, results of Chapter 2-4 support the emerging hypothesis that generalist ectoparasites and their highly mobile hosts possibly display weak, but positive, correlation in neutral genetic variation (Mazé-Guilmo et al. 2016).

7.2 Adaptation of generalist ectoparasites to their hosts

7.2.1 Adaptation to host type in cimicid ectoparasites

In Chapter 5, across the *Cimex* genus, I found positive selection on a gene coding for a protein that is useful in blood feeding, an apyrase. Furthermore, I found that host type (swallow, human or bat) had a significant effect on the

number of codons that show a significant signal of positive selection, in the apyrase gene.

The original hosts of cimicid insects are thought to be bats or birds (Usinger 1966). Bats and swallows can overlap in their choice of roosts (*i.e.* attics in man-made structures, hollow trees, etc.), but specialization by cimicids on humans was probably accompanied by radical changes in living conditions. First, humans are capable of heating their homes. Heating brings relatively constant temperature throughout the year, which may allow a larger number of generations per year for cimicid parasites (Usinger 1966), and also increase the frequency of feeding (Johnson 1941; DeVries et al. 2016). Second, human blood is energetically of much lower value than bat or bird blood, because of its lower hematocrit (Schmidt and Thews 1989; Saino et al. 1997; Neuweiler 2000). If cimicid parasites switched to humans from birds or bats, as hypothesized, the switch must have been accompanied by a trade-off between more constant environmental temperature and efficiency of energy intake while blood feeding on the host.

Interestingly, most *Cimex* specimens showing signals of positive selection at salivary protein genes were associated with bats. Although these results do not entirely rule out bats as the original hosts of cimicids, they suggest adaptation from a putative original bat host species to other bat host species may have been accompanied by molecular changes at least in the salivary apyrase. My results would therefore suggest that the complicated life cycle of bats may present complex, species-specific selective pressures on genes involved in blood feeding in their generalist ectoparasites.

7.2.2 Adaptation to host species in a bat-associated cimicid ectoparasite

In Chapter 6, I found a continent-scale signal of local adaptation by *C. adjunctus* to the respective bat host species with which they associate, in two genes coding for proteins involved in blood-feeding, an apyrase and a nitrophorin. I also found that the host species (big brown bat, little brown myotis or northern myotis) was the strongest driver of genotypic variation at the two genes. Furthermore, genotypic variation was mostly between the building-roosting species, the big brown bat and the little brown myotis, and the tree-roosting species, the northern myotis.

My results suggest specific ecological characteristics associated with a host species are a strong driver of selective pressure even in a generalist ectoparasite. Altogether, even in the face of gene flow, it appears possible for individuals of a generalist ectoparasite species to show some degree of specialization to a host species.

7.2.3 Conclusions

I observed signals of positive selection in the *Cimex* genus at a single gene that is directly involved in blood feeding (Chapter 5). I also showed that populations from a single ectoparasite species show local adaptation to a specific bat host species at two genes that are directly involved in blood feeding (Chapter 6). Altogether, results of Chapters 5-6 suggest that hosts have an influence on genes coding for salivary proteins of blood-feeding ectoparasites across the *Cimex* genus, and also within *C. adjunctus*. These results support the perhaps

counter-intuitive notion that blood-feeding ectoparasites that may associate with a range of different hosts also may show adaptation to specific hosts.

7.3 General conclusion of the dissertation

Overall, I found that hosts influence the neutral and the adaptive genetic variation of their generalist temporary ectoparasites, which represent a large portion of all parasites. However, these effects may be more complex than in other types of parasites. Blood-feeding ectoparasitic arthropods (such as bed bugs, ticks, kissing bugs and flies) often are vectors of pathogens or are otherwise a serious nuisance, and are a cause of huge socio-economic problems throughout the world (Bargues et al. 2006; Delaunay et al. 2010; Dharmarajan et al. 2011; Morens and Fauci 2013). Understanding more about gene flow patterns and blood-feeding adaptations in those arthropods may offer valuable clues to control and manage the propagation of parasites and pathogens (Morens and Fauci 2013).

7.4 Caveats, knowledge gaps and research needs

7.4.1 Sample sizes

The most important caveat of this thesis is undoubtedly the small sample size per site. A typical sampling design in population genetics involves collection of a large number of individuals from several populations across the study area (Prunier et al. 2013). Such a sampling design allows the application of various metrics and analyses, such as detection of migrants between sites, divergence of populations, likelihood of bottleneck events, etc. (Hanski 1998;

Luikart et al. 1998; Hedrick 2005; Meirmans and Hedrick 2011). However, many types of molecular ecological analyses developed in the last several years allow a sampling scheme consisting of a few individuals sampled in a large number of sites spread randomly across the study area. For example, analyses of past demography using coalescent models (Heller et al. 2013), or analyses of the effect of the land cover on spatial genetic structure (Prunier et al. 2013) can be performed on data from a pooled sampling scheme, consisting of a few individuals sampled in each of a large number of sites across the study area. A trade-off exists between the number of populations and the number of specimens per population that can be sampled for population genetic studies, particularly for species that are very difficult to sample due to their small size and their occurrence in locations that are difficult to access.

7.4.2 Number of markers

Another important caveat of this thesis is the relatively small number of markers used. For my analyses of neutral genetic variation, microsatellite markers are very useful because of their ease of access and because they mostly are not affected by selection (Meirmans and Hedrick 2011; Haasl and Payseur 2013). In my project, I used microsatellite markers originally developed for the well-known *C. lectularius* to study a closely related species, the bat-associated *C. adjunctus*. Only a small proportion of microsatellite markers developed in two studies on *C. lectularius* (totalling 45 markers; Booth et al. 2012; Fountain et al. 2014) could be used in *C. adjunctus*. In my analyses of adaptive genetic variation, I only analysed two candidate genes. Although these

represent all known salivary protein genes that have been directly linked to efficiency of blood feeding in cimicids (Valenzuela et al. 1996; Valenzuela and Ribeiro 1998; Francischetti et al. 2010), they only represent two small coding regions of the whole genome. Studying adaptation to the host using only two candidate genes has a potential of being biased by locus-specific processes, such as methylation or selective sweeps (Pritchard et al. 2010; Massicotte et al. 2011), and may not fully capture the effect of the host on its parasites.

Another type of genetic marker increasingly used in population genetic studies is the single nucleotide polymorphism (SNP). Now largely analyzed by high-throughput sequencing, this type of marker is usually less variable than microsatellites, but a far larger number of markers can be genotyped (Davey and Blaxter 2010). The advantage of a very large panel of SNPs is that they represent the whole genome, and not only specific locations where microsatellite markers or candidate genes would be situated. SNP-based studies have therefore a smaller chance of being biased by locus-specific processes (Beaumont and Balding 2004; Holderegger et al. 2008; de Villemereuil et al. 2014). Using association analyses, it is possible to identify SNP loci that are strongly linked to loci under selection, as well as loci that behave as neutral markers (*i.e.* in migration-drift equilibrium; Coop et al. 2010; Hancock et al. 2011; Frichot et al. 2013; Talbot et al. 2016). One can therefore study both neutral and adaptive genetic variation using genome-wide SNPs. The use of such markers would be a valuable tool to study the effect of the host of neutral and adaptive genetic variation in parasites.

7.4.3 Model species

The *Cimex* genus is interesting primarily because of the widespread distribution and the generalist nature of the species within it. Furthermore, we observe a range in the degree of generalism among *Cimex* species (Usinger 1966). For example, although many species are exclusive to bats or swallows, *C. lectularius* is known to parasitize humans, bats, chicken and other domestic animals. However, *C. lectularius* is probably the only strong generalist in the genus (Usinger 1966). Also, it seems no species in the genus is highly specialized on any one species of host. For a more precise measurement of the effect of generalism in a parasite on adaptation to its hosts, it would be interesting to study a group of parasites showing more variation in their degree of generalism. Bat flies of the Nycteribid family would be interesting to study, because they show either high host specificity or generalism (Dick and Patterson 2007). However, these insects exclusively parasitize bats in all cases; *C. lectularius* would be considered a much stronger generalist than Nycteribid bat flies. In conclusion, *Cimex* is an informative group for studies of the effect of association with different hosts on parasite genetic variation.

7.4.4 Limits of genetic data

As demonstrated in this dissertation, genetic data are useful to study a broad range of different processes in ecology and evolution. This is especially true when the processes of interest are very difficult to measure directly because the study organism is very small or hard to sample. One limitation of genetic data however is that studying gene flow and genetic structure as a proxy for

movement will always incorporate a degree of uncertainty (Bohonak 1999). For example, movement while foraging is probably highly uncorrelated with gene flow in bats. If one aims to study foraging movement in bats therefore, one would need to use other tools. Another limitation of genetic data is that although it is possible to detect signals of selection on sequences, there is no way to directly measure fitness and therefore, adaptation, using purely genetic data (Barrett and Hoekstra 2011). For example, the gold standard for demonstrating local adaptation is to transplant individuals between environments (for example different host species) and measure their resulting fitness, for example growth or number of offspring (Nuismer and Gandon 2008). Although genetic data can offer a great breadth of information about a diverse array of processes in ecology and evolution, they should ideally be used within a broad framework of different complementary approaches, to obtain a comprehensive perspective on the processes of interest (Hancock et al. 2011; Talbot et al. 2016).

7.5 Literature Cited

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Bargues MD, Klisiowicz DR, Panzera F et al. (2006) Origin and phylogeography of the Chagas disease main vector *Triatoma infestans* based on nuclear rDNA sequences and genome size. *Infection, Genetics and Evolution*, **6**, 46–62.

Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767–780.

Bartonička T, Růžičková L (2013) Recolonization of bat roost by bat bugs (*Cimex pipistrelli*): could parasite load be a cause of bat roost switching? *Parasitology Research*, **112**, 1615–1622.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21–45.
Booth W, Saenz VL, Santangelo RG et al. (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.
Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.
Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416–423.
Davies TGE, Field LM, Williamson MS (2012) The re-emergence of the bed bug as a nuisance pest: implications of resistance to the pyrethroid insecticides. *Medical and Veterinary Entomology*, **26**, 241–254.
Delaunay P, Blanc V, Del Giudice P et al. (2010) Bedbugs and infectious diseases. *Clinical Infectious Diseases*, **52**, 200–210.
DeVries ZC, Mick R, Schal C (2016) Feel the heat: activation, orientation and feeding responses of bed bugs to targets at different temperatures. *The Journal of Experimental Biology*, **219**, 3773–3780.
Dharmarajan G, Beasley JC, Rhodes OE (2011) Heterozygote deficiencies in parasite populations: an evaluation of interrelated hypotheses in the raccoon tick, Ixodes texanus. *Heredity*, **106**, 253–260.
Dick CW, Patterson BD (2007) Against all odds: explaining high host specificity in dispersal-prone parasites. *International Journal for Parasitology*, **37**, 871–876.
Ellison LE, O'Shea TJ, Neubaum DJ, Bowen RA (2007) Factors influencing movement probabilities of big brown bats (*Eptesicus fuscus*) in buildings. *Ecological applications: a publication of the Ecological Society of America*, **17**, 620–627.
Fountain T, Duvaux L, Horsburgh G, Reinhardt K, Butlin RK (2014) Human-facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*. *Molecular Ecology*, **23**, 1071–1084.
Francischetti IMB, Calvo E, Andersen JF et al. (2010) Insight into the sialome of the bed bug, *Cimex lectularius*. *Journal of Proteome Research*, **9**, 3820–3831.
Frichot E, Schoville SD, Bouchard G, Francois O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.
Haasl RJ, Payseur BA (2013) Microsatellites as targets of natural selection. *Molecular Biology and Evolution*, **30**, 285–298.
Hancock AM, Witonsky DB, Alkorta-Aranburu G et al. (2011) Adaptations to climate-mediated selective pressures in humans. *Plos Genetics*, **7**, e1001375.
Hanski I (1998) Metapopulation dynamics. *Nature*, **396**, 41–49.
Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution; International Journal of Organic Evolution*, **59**, 1633–1638.

Heller R, Chikhi L, Siegismund HR (2013) The confounding effect of population structure on Bayesian Skyline Plot inferences of demographic history. *Plos One*, **8**, e62992.
ter Hofstede HM, Fenton MB (2005) Relationships between roost preferences, ectoparasite density, and grooming behaviour of neotropical bats. *Journal of Zoology*, **266**, 333–340.
Holderegger R, Herrmann D, Poncet B et al. (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, **1**, 273–283.
Johnson CG (1941) The ecology of the bed-bug, *Cimex lectularius* L., in Britain. *Journal of Hygiene*, **41**, 345–461.
Kurta A (1995) *Mammals of the Great Lakes region*. University of Michigan Press, Ann Arbor, MI, USA.
Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963–974.
Massicotte R, Whitelaw E, Angers B (2011) DNA methylation: A source of random variation in natural populations. *Epigenetics*, **6**, 421–427.
Mazé-Guilmo E, Blanchet S, McCoy KD, Loot G (2016) Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, **19**, 336–347.
Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.
Morens DM, Fauci AS (2013) Emerging infectious diseases: threats to human health and global stability. *Plos Pathogens*, **9**, e1003467.
Neuweiler G (2000) *The biology of bats*. Oxford University Press, New York, NY, USA.
Nuismer SL, Gandon S (2008) Moving beyond common- garden and transplant designs: insight into the causes of local adaptation in species interactions. *The American Naturalist*, **171**, 658–668.
Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208–R215.
Prunier JG, Kaufmann B, Fenet S et al. (2013) Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: towards a better assessment of functional connectivity using an individual-based sampling scheme. *Molecular Ecology*, **22**, 5516–5530.
Saino N, Cuervo JJ, Ninni P, De Lope F, Moller AP (1997) Haematocrit correlates with tail ornament size in three populations of the Barn Swallow (*Hirundo rustica*). *Functional Ecology*, **11**, 604–610.
Schmidt RF, Thews G (1989) *Human physiology*. Springer Berlin Heidelberg, Berlin, Germany.
Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.

Talbot B, Chen T-W, Zimmerman S et al. (2016) Combining genotype, phenotype, and environment to infer potential candidate genes. *Journal of Heredity*, **108**, 207–216.
Usinger RL (1966) *Monograph of Cimicidae (Hemiptera, Heteroptera)*. Entomological Society of America, Annapolis, MD, USA.
Valenzuela JG, Chuffe OM, Ribeiro JC (1996) Apyrase and anti-platelet activities from the salivary glands of the bed bug *Cimex lectularius*. *Insect Biochemistry and Molecular Biology*, **26**, 557–562.
Valenzuela JG, Ribeiro JM (1998) Purification and cloning of the salivary nitrophorin from the hemipteran *Cimex lectularius*. *The Journal of Experimental Biology*, **201**, 2659–2664.
de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.
Willis CKR, Brigham RM (2004) Roost switching, roost sharing and social cohesion: forest-dwelling big brown bats, *Eptesicus fuscus*, conform to the fission–fusion model. *Animal Behaviour*, **68**, 495–505.
Zhu F, Gujar H, Gordon JR et al. (2013) Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Scientific Reports*, **3**.

Appendix A : Chapter 2 Supplementary Material

Table A.1 List of *Cimex adjunctus* specimens included in analyses. Site refers to each unique sampling location in the study, and correspond to those in Fig. 2.1, Fig. 2.2, Table 2.2 and Table A.2. Host species refers to the bat species from which the samples were collected, or which was inhabiting the roost from which the samples were collected (EPFU : *Eptesicus fuscus*, MYLU : *Myotis lucifugus*, MYSE : *Myotis septentrionalis*). Individuals with the same haplotype number share the same mitochondrial *CO1* haplotype sequence (we could not obtain *CO1* information for six individuals). We provide the Genbank accession number for each unique haplotype the first time it appears in the table. The genetic cluster (identified by Geneland 4.0.5 based on microsatellite data) to which each specimen was assigned is also given (we could not obtain microsatellite data for 10 individuals).

ID	Site	Province or state	Country	Host species	Year	Haplotype #	Genbank accession #	Genetic cluster #
380	1	New Brunswick	Canada	MYSE	2001	8	KU534903	7
431	2	Nova Scotia	Canada	MYLU	2003	6	KU534901	8
444	2	Nova Scotia	Canada	MYSE	2003	8		8
468	2	Nova Scotia	Canada	MYLU	2003	10	KU534905	Unavailable
941	2	Nova Scotia	Canada	MYLU	2004	12	KU534907	8
471-Cim1	3	Nova Scotia	Canada	MYSE	2003	8		8
471-Cim2	3	Nova Scotia	Canada	MYSE	2003	8		8
4823	4	Nova Scotia	Canada	MYLU	2008	6		8
4839	4	Nova Scotia	Canada	MYLU	2008	4	KU534899	8
1469	5	Prince Edward Island	Canada	MYSE	2005	10		7
1459	6	Prince Edward Island	Canada	MYSE	2005	8		7
1460	6	Prince Edward Island	Canada	MYSE	2005	8		7
1725-Cim1	7	Nova Scotia	Canada	MYSE	2005	11	KU534906	7
1725-Cim2	7	Nova Scotia	Canada	MYSE	2005	8		Unavailable
1727.4	7	Nova Scotia	Canada	MYSE	2006	3	KU534898	7
1730.2	7	Nova Scotia	Canada	MYSE	2006	8		7
5573	8	Nova Scotia	Canada	MYLU	2009	8		7
5574	8	Nova Scotia	Canada	MYSE	2009	7	KU534902	7
7532	8	Nova Scotia	Canada	MYSE	2010	6		7
6366	9	Nova Scotia	Canada	MYSE	2010	6		7
6476	10	Nova Scotia	Canada	MYSE	2010	9	KU534904	7
7045	11	Nova Scotia	Canada	MYSE	2010	9		7
7259	12	Labrador	Canada	MYLU	2011	4		7

11477	13	Labrador	Canada	MYLU	2012	5	KU534900	9
11494	13	Labrador	Canada	MYLU	2012	5		9
10186	14	Newfoundland	Canada	MYLU	2013	1	KU534896	9
10030	15	Newfoundland	Canada	MYLU	2012	2	KU534897	9
12471	16	Labrador	Canada	MYSE	2014	3		Unavailable
EFS40-10	17	Michigan	USA	EPFU	2005	23	KU534918	5
EFS40-12	17	Michigan	USA	EPFU	2005	23		5
EFS40-14	17	Michigan	USA	EPFU	2005	23		5
EFS40-2	17	Michigan	USA	EPFU	2005	23		5
EFS40-4	17	Michigan	USA	EPFU	2005	23		5
EFS40-6	17	Michigan	USA	EPFU	2005	23		5
EFS40-7	17	Michigan	USA	EPFU	2005	23		5
EFS42-35	18	Michigan	USA	EPFU	2005	24	KU534919	5
EFS42-36	18	Michigan	USA	EPFU	2005	13	KU534908	Unavailable
EFS42-39	18	Michigan	USA	EPFU	2005	17	KU534912	5
EFS42-43	18	Michigan	USA	EPFU	2005	17		5
EFS42-67	18	Michigan	USA	EPFU	2005	13		5
EFS42-Roost-Cim1	19	Michigan	USA	EPFU	2005	17		5
EFS42-Roost-Cim2	19	Michigan	USA	EPFU	2005	13		5
EFS42-Roost-Cim3	19	Michigan	USA	EPFU	2005	17		5
EFS42-Roost-Cim4	19	Michigan	USA	EPFU	2005	17		5
EFS42-Roost-Cim5	19	Michigan	USA	EPFU	2005	17		5
EFS42-Roost-Cim6	19	Michigan	USA	EPFU	2005	17		5
EF-293-Roost	20	Michigan	USA	EPFU	2005	22	KU534917	5
2006-MV-17	21	Michigan	USA	EPFU	2006	25	KU534920	5
LW-E01	22	Michigan	USA	EPFU	2005	26	KU534921	6
LW-E02	22	Michigan	USA	EPFU	2005	27	KU534922	6
LW-E03	23	Michigan	USA	EPFU	2005	28	KU534923	5
MV496-Cim1	24	Indiana	USA	EPFU	2010	31	KU534926	3
MV496-Cim2	24	Indiana	USA	EPFU	2010	31		3
EFAB-9-Cim1	25	Indiana	USA	EPFU	2005	29	KU534924	3
EFAB-9-Cim2	25	Indiana	USA	EPFU	2005	30	KU534925	3
EFWG-16	26	Indiana	USA	EPFU	2005	32	KU534927	3
EFWN-23	26	Indiana	USA	EPFU	2005	32		3
EFWN-28	26	Indiana	USA	EPFU	2005	17		3
EFGB-18	27	Indiana	USA	EPFU	2005	33	KU534928	1
EFMS-16	28	Indiana	USA	EPFU	2005	23		3
MV642	29	Kentucky	USA	EPFU	2010	19	KU534914	3
MV652-Cim1	29	Kentucky	USA	EPFU	2010	20	KU534915	3
MV652-Cim2	29	Kentucky	USA	EPFU	2010	21	KU534916	Unavailable
MV194-Cim1	30	Michigan	USA	EPFU	2008	13		2

MV194-Cim2	30	Michigan	USA	EPFU	2008	Unavailable		2
MV205	30	Michigan	USA	EPFU	2008	13		2
RH01	30	Michigan	USA	EPFU	2008	17		2
RH02	30	Michigan	USA	EPFU	2008	18	KU534913	2
RH03	30	Michigan	USA	EPFU	2008	18		2
RH04	30	Michigan	USA	EPFU	2008	13		2
RH05	30	Michigan	USA	EPFU	2008	13		2
RH06	30	Michigan	USA	EPFU	2008	13		2
RH07	30	Michigan	USA	EPFU	2008	17		Unavailable
RH09	30	Michigan	USA	EPFU	2008	13		2
RH11	30	Michigan	USA	EPFU	2008	18		2
RH12	30	Michigan	USA	EPFU	2008	13		2
RH13	30	Michigan	USA	EPFU	2008	Unavailable		2
RH14	30	Michigan	USA	EPFU	2008	Unavailable		2
RH15	30	Michigan	USA	EPFU	2008	18		2
O1	31	Ontario	Canada	EPFU	2014	8		6
O2	31	Ontario	Canada	EPFU	2014	8		6
O3	31	Ontario	Canada	EPFU	2014	8		6
O4	31	Ontario	Canada	EPFU	2014	8		6
O5	31	Ontario	Canada	EPFU	2014	8		6
O6	31	Ontario	Canada	EPFU	2014	8		6
O7	31	Ontario	Canada	EPFU	2014	8		6
CCC-Roost-Cim2	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-Roost-Cim3	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-Roost-Cim4	32	Pennsylvania	USA	MYLU	2006	15	KU534910	4
CCC-Roost-Cim5	32	Pennsylvania	USA	MYLU	2006	3		4
CCC-Roost-Cim6	32	Pennsylvania	USA	MYLU	2006	14	KU534909	4
CCC-Roost-Cim7	32	Pennsylvania	USA	MYLU	2006	15		4
CCC-Roost-Cim8	32	Pennsylvania	USA	MYLU	2006	15		4
CCC-Roost-Cim9	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-Roost-Cim10	32	Pennsylvania	USA	MYLU	2006	Unavailable		4
CCC-Roost-Cim11	32	Pennsylvania	USA	MYLU	2006	Unavailable		4
CCC-Roost-Cim12	32	Pennsylvania	USA	MYLU	2006	Unavailable		4
CCC-Roost-Cim13	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-Roost-Cim14	32	Pennsylvania	USA	MYLU	2006	6		4
CCC-Roost-Cim15	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-Roost-Cim16	32	Pennsylvania	USA	MYLU	2006	13		4
CCC-4	33	Pennsylvania	USA	MYLU	2006	13		4
CCC-45	33	Pennsylvania	USA	MYLU	2006	13		4
CCC-50	33	Pennsylvania	USA	MYLU	2006	3		4
RG-01	34	Colorado	USA	EPFU	2005	35	KU534930	10

RG-02	35	Colorado	USA	EPFU	2005	35		Unavailable
RG-03-Cim1	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim2	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim3	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim4	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim5	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim6	36	Colorado	USA	EPFU	2005	35		Unavailable
RG-03-Cim7	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim8	36	Colorado	USA	EPFU	2005	35		10
RG-03-Cim9	36	Colorado	USA	EPFU	2005	36	KU534931	10
RG-04	37	Colorado	USA	EPFU	2005	39	KU534934	10
RG-06-Cim1	38	Colorado	USA	EPFU	2005	35		10
RG-06-Cim2	38	Colorado	USA	EPFU	2005	32		10
RG-06-Cim3	38	Colorado	USA	EPFU	2005	35		10
RG-06-Cim4	38	Colorado	USA	EPFU	2005	35		10
RG-06-Cim5	38	Colorado	USA	EPFU	2005	35		Unavailable
RG-07-Cim1	39	Colorado	USA	EPFU	2005	39		10
RG-07-Cim2	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim3	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim4	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim5	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim6	39	Colorado	USA	EPFU	2005	40	KU534935	10
RG-07-Cim7	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim8	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim9	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim10	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim11	39	Colorado	USA	EPFU	2005	35		10
RG-07-Cim12	39	Colorado	USA	EPFU	2005	35		10
RG-08-Cim1	40	Colorado	USA	EPFU	2005	37	KU534932	10
RG-08-Cim2	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim3	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim4	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim5	40	Colorado	USA	EPFU	2005	35		10
RG-08-Cim6	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim7	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim8	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim9	40	Colorado	USA	EPFU	2005	37		10
RG-08-Cim11	40	Colorado	USA	EPFU	2005	35		10
RG-08-Cim14	40	Colorado	USA	EPFU	2005	35		10
RG-08-Cim15	40	Colorado	USA	EPFU	2005	37		10
RG-09-Cim1	41	Colorado	USA	EPFU	2005	38	KU534933	10

RG-09-Cim2	41	Colorado	USA	EPFU	2005	35		10
RG-09-Cim3	41	Colorado	USA	EPFU	2005	35		10
RG-09-Cim4	41	Colorado	USA	EPFU	2005	32		10
RG-09-Cim5	41	Colorado	USA	EPFU	2005	38		10
RG-09-Cim6	41	Colorado	USA	EPFU	2005	35		10
RG-09-Cim7	41	Colorado	USA	EPFU	2005	35		10
RG-09-Cim8	41	Colorado	USA	EPFU	2005	35		Unavailable
HC-4	42	New Jersey	USA	MYLU	2006	13		4
O8	43	Ontario	Canada	EPFU	2014	16	KU534911	4
S1-Cim1	44	Saskatchewan	Canada	MYLU	2014	34	KU534929	3
S1-Cim2	44	Saskatchewan	Canada	MYLU	2014	34		3
N1	45	Northwest Territories	Canada	MYLU	2014	13		7
N3	45	Northwest Territories	Canada	MYLU	2014	41	KU534936	7
N4	45	Northwest Territories	Canada	MYLU	2014	17		7

Table A.2 Details of sampling locations where we collected *Cimex adjunctus* specimens used in this study. Site refers to each unique sampling locationand correspond to those in Fig. 2.1, Fig. 2.2, Table 2.2 and Table A.1. Sample size refers to the number of specimens collected at each site. Capture method refers to the way *C. adjunctus* samples were collected: either from the body of a bat that was captured outside a roost with a mist net or a harp trap (Bat capture), or from inside of the roost itself (Roost visit). Site characteristics refers either to the type of roost (house, barn, cabin, church or abandoned mine) in, or next to which, we collected samples, or to the forested land jurisdiction where collection was performed.

Site	Province/State	Country	Sample size	Capture method	Site characteristics
1	New Brunswick	Canada	1	Bat capture	Fundy National Park
2	Nova Scotia	Canada	4	Bat capture	Kejimkujik National Park
3	Nova Scotia	Canada	2	Bat capture	Kejimkujik National Park
4	Nova Scotia	Canada	2	Bat capture	Kejimkujik National Park
5	Prince Edward Island	Canada	1	Bat capture	Provincial land
6	Prince Edward Island	Canada	2	Bat capture	Provincial land
7	Nova Scotia	Canada	4	Bat capture	Dollar Lake Provincial Park
8	Nova Scotia	Canada	3	Bat capture	Abandoned mine
9	Nova Scotia	Canada	1	Bat capture	House
10	Nova Scotia	Canada	1	Bat capture	House
11	Nova Scotia	Canada	1	Bat capture	Abandoned mine
12	Newfoundland and Labrador	Canada	1	Bat capture	Cabin
13	Newfoundland and Labrador	Canada	2	Bat capture	Cabin
14	Newfoundland and Labrador	Canada	1	Bat capture	Cabin
15	Newfoundland and Labrador	Canada	1	Bat capture	House
16	Newfoundland and Labrador	Canada	1	Bat capture	Provincial land
17	Michigan	United States	7	Bat capture	House
18	Michigan	United States	5	Bat capture	House
19	Michigan	United States	6	Roost visit	House
20	Michigan	United States	1	Bat capture	House
21	Michigan	United States	1	Bat capture	State land

22	Michigan	United States	2	Bat capture	State land
23	Michigan	United States	1	Bat capture	State land
24	Michigan	United States	2	Bat capture	House
25	Indiana	United States	2	Bat capture	House
26	Indiana	United States	3	Bat capture	School
27	Indiana	United States	1	Bat capture	Barn
28	Indiana	United States	1	Bat capture	School
29	Kentucky	United States	3	Bat capture	State land
30	Michigan	United States	15	Bat capture	House
31	Ontario	Canada	7	Bat capture	House
32	Pennsylvania	United States	15	Bat capture	Church
33	Pennsylvania	United States	3	Roost visit	Church
34	Colorado	United States	1	Bat capture	School
35	Colorado	United States	1	Bat capture	State land
36	Colorado	United States	9	Bat capture	House
37	Colorado	United States	1	Bat capture	House
38	Colorado	United States	5	Bat capture	House
39	Colorado	United States	12	Bat capture	House
40	Colorado	United States	12	Bat capture	House
41	Colorado	United States	8	Bat capture	Church
42	New Jersey	United States	1	Bat capture	Church
43	Ontario	Canada	1	Bat capture	Long Point Provincial Park
44	Saskatchewan	Canada	2	Bat capture	Provincial land
45	Northwest Territories	Canada	3	Bat capture	Territorial land

Table A.3 Sequence, primer concentration ([Primer]), annealing temperature (T°), total number of alleles, and size range information for nine microsatellite markers previously designed for *Cimex lectularius* (by 1: Fountain et al. 2014 and 2: Booth et al. 2012), used here on *Cimex adjunctus* samples. For each marker, we also show the mean number of alleles (N_{AL}), expected heterozygosity (H_S), observed heterozygosity (H_O) and inbreeding coefficient (G_{IS}) across all sites with five or more sampled individuals.

Marker name	Sequence	[Primer]	T°	Total number of alleles	Size range	Source	N_{AL}	H_S	H_O	G_{IS}
Cle002	F: CATGAAATTGGGAGTTTCTAGTTTC R: TTACCGCCCATGTAAACGAG	0.25 µL	See text	11	201-235	1	5	0.241	0.140	0.418
Cle003	F: TTCGTTTGTGTAGAACCTTGG R: TACGTCCCTACAAGCTCACC	0.25 µL		9	154-213	1	8	0.504	0.325	0.354
Cle013	F: TTCACAGATTTAAGCCTAACTGGTC R: CAAATAACCTCGAATTCATACGC	0.25 µL		31	132-252	1	20	0.679	0.514	0.244
Cle015	F: TCATATGGGCGGATTAGAGC R: TAACAATCTGGAGGCGGAAC	0.25 µL		17	268-457	1	6	0.594	0.208	0.651
Clec15	F: GTTTGCAACCGGCACTGG R: AGTTAGCACGGTAAGCCCG	1.00 µL	59°C	7	196-234	2	4	0.125	0.114	0.086
Clec21	F: CCATTCCAACCCTGCTTACTG R: ACGACTAAACATGTCCAAACTC	0.60 µL	59°C	3	207-216	2	2	0.019	0.000	1.000
Clec48	F: CCGGTTCGCCCATTGAAAC R: TGCATGGCAACTGATGTAACTG	0.80 µL	59°C	2	227-228	2	1	0.000	0.000	N/A
Clec104	F: TCCCTACCAATCGGACATTC R: GCCCCCTTCCAGTTTTATGT	0.30 µL	59°C	7	240-249	2	4	0.164	0.061	0.627
BB28B	F: GGGCGAGAGAGAGTGATACC R: TTGTAGCGCCCTCTTCAACT	1.20 µL	61°C	6	92-212	2	4	0.328	0.426	-0.296

Table A.4 Genotypic data at 9 microsatellite markers for 150 *Cimex adjunctus* individuals (we could not obtain microsatellite data for 10 individuals). The first three digits in each string describe the length (in base pairs) of the first allele, and the last three describe the length of the second allele. Missing data is identified with "000000".

ID	Clec48	Clec104	Clec21	Clec15	BB28B	Cle002	Cle013	Cle003	Cle015
380	227227	240240	213213	208208	092212	229231	000000	202202	000000
431	228228	247247	207207	209209	189189	229229	157157	184184	270270
444	227227	240240	213213	208208	092092	220220	245247	202202	289442
941	228228	247247	207207	209209	189189	229229	176176	184184	270270
471-Cim1	227227	240240	213213	208208	092092	233233	137243	202202	289442
471-Cim2	227227	240240	213213	197208	092092	229229	216243	202202	442442
4823	228228	247247	207207	209209	189189	229229	176176	184202	268270
4839	228228	247247	207207	209209	189189	229229	176176	184202	270270
1469	227227	240240	000000	208208	092092	220220	202202	202202	457457
1459	227227	240240	213213	208208	092212	220220	202216	202202	435435
1460	227227	240240	213213	208208	092212	220220	202202	202202	435435
1725-Cim1	227227	240240	000000	208208	092092	220220	202206	202202	435453
1727.4	227227	240240	000000	197208	092092	233233	202210	202202	435435
1730.2	227227	240240	213213	197208	092092	220220	202206	202202	435435
5573	227227	240240	213213	196196	092092	231231	210252	200202	431435
5574	227227	240240	213213	196196	092092	220220	206210	202202	442457
7532	227227	240240	213213	196208	092092	220220	206206	202202	291435
6366	227227	240240	213213	196208	092092	220220	210254	202202	435453
6476	227227	238240	213213	208208	092092	220220	210210	202202	291291
7045	227227	240240	213213	000000	092212	220220	202202	202202	435435
7259	227227	240240	213213	196208	092212	220237	202202	202202	429429
11477	228228	247247	207207	209209	189189	229229	148176	184202	270278
11494	228228	247247	207207	209209	189189	229229	176176	184184	270270
10186	228228	247247	207213	209209	189189	227227	148176	184184	270270
10030	228228	240247	207207	209209	189189	227227	176176	184202	270270
EFS40-10	228228	247247	207207	209209	161189	229229	157163	184195	270270

EFS40-12	228228	247247	207207	209209	135189	229229	000000	184195	000000
EFS40-14	228228	247247	207207	209209	161189	229229	157163	154184	270270
EFS40-2	228228	247247	207207	209209	161189	229229	157163	184200	268268
EFS40-4	228228	247247	207207	209209	161189	229229	157157	184184	268268
EFS40-6	228228	247247	207207	000000	189189	224231	157157	000000	268268
EFS40-7	228228	247247	207207	209209	161189	229229	157163	184200	268268
EFS42-35	228228	247247	207207	209209	189189	229229	157157	184200	268270
EFS42-39	228228	247247	207207	209209	161189	229229	000000	200200	402402
EFS42-43	228228	247247	207207	209209	161189	229229	176176	184184	268402
EFS42-67	228228	247247	000000	000000	189189	229229	000000	000000	396396
EFS42-Roost-Cim1	228228	247247	207207	209209	189189	201229	157157	184184	000000
EFS42-Roost-Cim2	228228	247247	207207	209209	161189	229229	165165	184184	402402
EFS42-Roost-Cim3	228228	247247	207207	209209	161189	201201	000000	184184	268268
EFS42-Roost-Cim4	228228	247247	207207	209209	161189	229229	157157	200200	270270
EFS42-Roost-Cim5	228228	247247	207207	209209	189189	231231	181181	200200	268268
EFS42-Roost-Cim6	228228	247247	207207	209209	189189	229229	000000	184200	270270
EF-293-Roost	228228	247247	207207	209209	189189	229229	163163	184200	270270
2006-MV-17	228228	247247	207207	209212	189189	229229	163163	184184	268270
LW-E01	228228	247247	207207	209209	161189	229229	157178	184195	268268
LW-E02	228228	248248	207207	209209	161189	229229	168168	184184	268268
LW-E03	228228	247247	207207	209209	161189	229229	174174	184200	268268
MV496-Cim1	228228	246248	207207	209212	161189	229229	157206	184195	000000
MV496-Cim2	228228	247247	207207	209212	189189	229229	000000	195195	000000
EFAB-9-Cim1	228228	247248	207207	209209	189189	229229	157157	000000	000000
EFAB-9-Cim2	228228	246246	000000	000000	189189	233233	000000	000000	000000
EFWG-16	228228	248248	207207	000000	189189	229229	000000	000000	270270
EFWN-23	228228	247248	207207	209209	189189	229229	000000	184184	000000
EFWN-28	228228	247248	000000	000000	000000	229229	000000	000000	000000
EFGB-18	228228	248248	000000	234234	000000	229235	183183	202202	000000
EFMS-16	228228	247249	207207	000000	165165	229229	157157	195195	000000
MV642	228228	248248	207207	209209	161189	229229	000000	184184	270270

MV652-Cim1	228228	246247	207207	209209	189189	229229	152152	195195	000000
MV194-Cim1	228228	247247	207207	209209	161189	229231	000000	184195	270270
MV194-Cim2	228228	247247	207207	209209	161189	229231	157202	195195	268268
MV205	228228	247247	207207	209209	161189	229229	157176	195195	270270
RH01	228228	246246	207207	209209	161189	201229	157157	184184	270270
RH02	228228	247247	207207	209209	189189	229229	157157	195195	268268
RH03	228228	247247	207207	209209	161189	201201	157157	195195	268268
RH04	228228	247247	207207	209209	189189	229229	157157	184195	268268
RH05	228228	247247	207207	209209	189189	229231	157157	184184	268268
RH06	228228	247247	207207	209209	189189	229231	143143	184184	270270
RH09	228228	245245	207207	209209	189212	231231	157157	184195	268268
RH11	228228	247247	207207	209209	189189	229229	000000	000000	270270
RH12	228228	247247	207207	209209	212212	201231	157157	000000	270270
RH13	228228	247247	207207	209209	161189	229229	157157	184209	268268
RH14	228228	247248	207207	196209	161189	229231	132157	195195	270270
RH15	228228	247247	207207	209209	161212	201229	189189	195195	270270
O1	228228	247247	207207	209209	161189	229229	178178	184195	268268
O2	228228	248248	207207	209209	161189	229229	163178	184195	268268
O3	228228	247248	207207	209209	161189	229229	178178	184200	268268
O4	228228	248248	207207	209209	161189	201229	157178	195195	268268
O5	228228	248248	207207	209209	161189	229229	178178	195195	270270
O6	228228	248248	207207	209209	189189	229229	178206	195195	268268
O7	228228	247248	207207	209209	161189	201229	206206	213213	270270
CCC-Roost-Cim2	228228	246248	207207	209209	189189	229229	176176	184184	270270
CCC-Roost-Cim3	228228	247247	207207	209209	189189	229229	157157	184184	270270
CCC-Roost-Cim4	228228	248248	207207	209209	161189	201229	157176	195195	270270
CCC-Roost-Cim5	228228	247247	207207	209209	189189	229229	157176	184184	270270
CCC-Roost-Cim6	228228	247247	207207	209209	189189	229229	206206	184184	268268
CCC-Roost-Cim7	228228	247247	207207	209209	161189	229229	157172	184195	270270
CCC-Roost-Cim8	228228	246246	207207	209209	161189	201229	157157	184195	268268
CCC-Roost-Cim9	228228	246246	207207	209212	189189	229229	157157	184200	268270

CCCRoost-Cim10	228228	247247	207207	209209	189189	229229	157157	184192	270270
CCCRoost-Cim11	228228	248248	207207	209209	189189	229229	157157	184184	270270
CCCRoost-Cim12	228228	246246	207207	209209	189189	229229	183183	184184	270292
CCC-Roost-Cim13	228228	247248	207207	209209	189189	229229	157176	184184	268268
CCC-Roost-Cim14	228228	246247	207207	209209	189212	229229	157176	184184	268270
CCC-Roost-Cim15	228228	246246	207207	209209	189212	229229	168184	184184	270270
CCC-Roost-Cim16	228228	246246	207207	209209	189189	229229	163176	184195	270270
CCC-4	228228	246246	207207	209209	189189	229229	157157	184184	268270
CCC-45	228228	246248	207207	209209	135189	229229	176176	184202	270270
CCC-50	228228	248248	207207	209209	189189	229229	157157	184184	270270
RG01	228228	247247	207207	209212	189189	229229	187193	200200	402402
RG03-Cim1	228228	247247	207207	212212	189189	228228	000000	000000	000000
RG03-Cim2	228228	247247	207207	209212	189189	229229	184189	200200	268400
RG03-Cim3	228228	247247	207207	212212	189189	229229	184184	184184	000000
RG03-Cim4	228228	247247	207207	212212	161189	229229	187187	184195	270401
RG03-Cim5	228228	247247	207207	212212	189189	229229	174184	200200	270401
RG03-Cim7	228228	247247	207207	212212	161189	229229	184184	184200	402402
RG03-Cim8	228228	247247	207207	212212	161189	229229	183184	000000	270270
RG03-Cim9	228228	247247	207207	000000	189189	229229	000000	000000	000000
RG04	228228	247247	207207	209209	161189	229229	184189	184200	270402
RG06-Cim1	228228	247247	207207	212212	161189	229229	184202	184184	270270
RG06-Cim2	228228	247247	207207	212212	000000	229229	189189	000000	000000
RG06-Cim3	228228	247247	207207	212212	161189	229229	176176	184184	270270
RG06-Cim4	228228	247247	207207	212212	189189	229229	000000	184184	270402
RG07-Cim1	228228	247247	207207	212212	189189	229229	000000	188200	402402
RG07-Cim2	228228	247247	207207	212212	189189	229229	184189	200200	401401
RG07-Cim3	228228	247247	207207	209212	189189	229229	000000	200200	401401
RG07-Cim4	228228	247247	216216	212212	189189	229229	174183	184200	401401
RG07-Cim5	228228	247247	207207	212212	189189	229229	184189	200200	401401
RG07-Cim6	228228	247247	207207	209212	189189	229229	000000	184184	000000
RG07-Cim7	228228	247247	207207	212212	161189	229229	184184	184184	270270

RG07-Cim8	228228	247247	207207	212215	161189	229229	183187	184200	268401
RG07-Cim9	228228	247247	207207	209212	189189	229229	184189	184200	270401
RG07-Cim10	228228	247247	207207	212212	189189	229229	174184	200200	400400
RG07-Cim11	228228	247247	207207	212212	189189	229229	174184	200200	270401
RG07-Cim12	228228	247247	207207	209212	189189	229229	174184	200200	270402
RG08-Cim1	228228	247247	207207	212212	189189	229229	195202	000000	268402
RG08-Cim2	228228	247247	207207	212212	189189	229229	165195	000000	268402
RG08-Cim3	228228	247247	207207	212212	189189	229229	000000	184184	000000
RG08-Cim4	228228	247247	207207	212212	189189	229229	165189	200200	270270
RG08-Cim5	228228	247247	207207	212212	189189	229229	165208	184184	270402
RG08-Cim6	228228	247247	207207	212212	189189	229229	184189	184184	268402
RG08-Cim7	228228	247247	207207	212212	189189	229229	202202	200200	270402
RG08-Cim8	228228	247247	207207	212212	189189	229229	165189	200200	402402
RG08-Cim9	228228	247247	207207	212212	189189	229229	183189	184184	402402
RG08-Cim11	228228	247247	207207	212212	189189	229229	221221	000000	402402
RG08-Cim14	228228	247247	207207	212212	189189	229229	195195	184200	401401
RG08-Cim15	228228	247247	207207	212212	189189	229229	183189	000000	401401
RG09-Cim1	228228	247247	207207	000000	189189	229229	187189	200200	000000
RG09-Cim2	228228	247247	207207	212212	189189	229229	183187	200200	270402
RG09-Cim3	228228	247247	207207	209212	189189	229229	187187	200200	402402
RG09-Cim4	228228	247247	207207	209209	161189	229229	165189	200200	402402
RG09-Cim5	228228	247247	207207	209212	189189	229229	183189	200200	402402
RG09-Cim6	228228	247247	207207	209209	189189	229229	184189	200200	000000
RG09-Cim7	228228	247247	207207	209209	189189	229229	189189	000000	000000
HC-4	228228	247248	207207	209212	189189	201229	183206	163184	268270
O8	228228	247247	207207	209209	161189	229229	206206	192200	270270
S1-Cim1	228228	247247	000000	209212	189189	229229	000000	184192	431435
S1-Cim2	228228	247247	207207	209212	189189	229229	000000	000000	268270
N1	227227	240240	213213	197208	092092	230230	183206	202202	434444
N3	227227	240240	213213	197197	092092	230230	206210	202202	291431
N4	227227	240240	213213	197208	092092	228228	226226	202202	431435

Table A.5 Host species and sampling year from which *C. adjunctus* samples assigned to each genetic cluster were collected.

Cluster number	Host species	Sampling year
1	*E. fuscus*	2005
2	*E. fuscus*	2005, 2006
3	*E. fuscus*	2008
4	*M. lucifugus*, *M. septentrionalis*	2005, 2006, 2009, 2010, 2011
5	*M. lucifugus*	2012, 2013
6	*E. fuscus*, *M. lucifugus*	2006, 2014
7	*E. fuscus*	2005
8	*M. lucifugus*, *M. septentrionalis*	2003, 2004, 2008
9	*E. fuscus*, *M. lucifugus*	2005, 2010, 2014
10	*E. fuscus*	2005, 2014

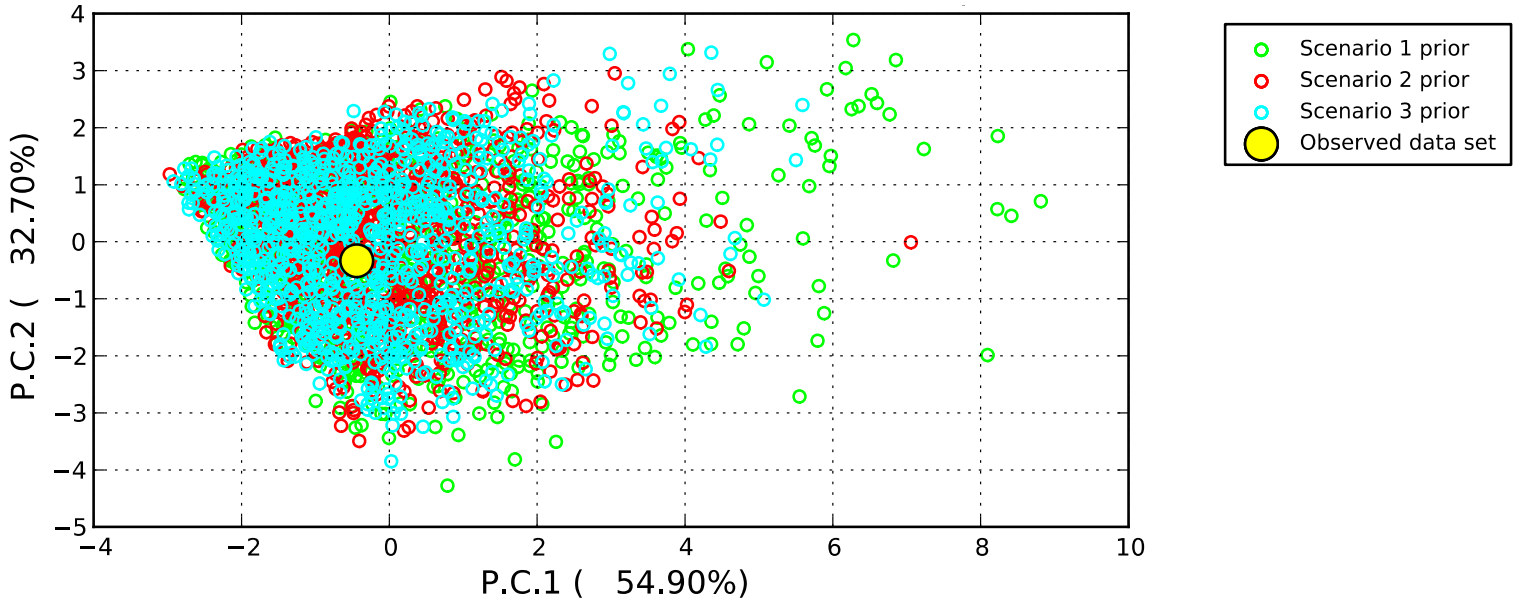

Figure A.6 First and second principal components (percentage of explained variation in parentheses) of simulated values under each putative demographic scenario for *Cimex adjunctus*, as determined by approximate Bayesian computation (ABC), and comparison with observed values.

Literature Cited

Booth W, Saenz VL, Santangelo RG et al. (2012) Molecular markers reveal infestation dynamics of the bed bug (Hemiptera: Cimicidae) within apartment buildings. *Journal of Medical Entomology*, **49**, 535–546.

Fountain T, Duvaux L, Horsburgh G, Reinhardt K, Butlin RK (2014) Human-facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*. *Molecular Ecology*, **23**, 1071–1084.

Appendix B : Chapter 3 Supplementary Material

Table B.1 List of *Cimex adjunctus* specimens included in analyses. Site refers to each unique sampling location in the study, and corresponds to those in Fig. 3.2 and Table 3.1. Host species refers to the bat species from which the samples were collected, or which was inhabiting the roost from which the samples were collected (EPFU : *Eptesicus fuscus*, MYLU : *Myotis lucifugus*). Capture method refers to the way *C. adjunctus* samples were collected: either from the body of a bat that was captured outside a roost with a mist net or a harp trap (Bat capture), or from inside of the roost itself (Roost visit). Site characteristics refers either to the type of roost (house, barn, school, church) in, or next to which, we collected samples, or to the forested land jurisdiction where collection was performed.

ID	Site	Capture Method	Site characteristics	Province /State	Country	Host species	Year
EFS40-10	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-12	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-14	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-2	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-4	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-6	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS40-7	1A	Bat capture	House	Michigan	USA	EPFU	2005
EFS42-35	2A	Bat capture	House	Michigan	USA	EPFU	2005
EFS42-39	2A	Bat capture	House	Michigan	USA	EPFU	2005
EFS42-43	2A	Bat capture	House	Michigan	USA	EPFU	2005
EFS42-67	2A	Bat capture	House	Michigan	USA	EPFU	2005
EFS42-Roost-Cim1	3A	Roost visit	House	Michigan	USA	EPFU	2005
EFS42-Roost-Cim2	3A	Roost visit	House	Michigan	USA	EPFU	2005
EFS42-Roost-Cim3	3A	Roost visit	House	Michigan	USA	EPFU	2005
EFS42-Roost-Cim4	3A	Roost visit	House	Michigan	USA	EPFU	2005

EFS42-Roost-Cim5	3A	Roost visit	House	Michigan	USA	EPFU	2005
EFS42-Roost-Cim6	3A	Roost visit	House	Michigan	USA	EPFU	2005
EF-293-Roost	4A	Bat capture	House	Michigan	USA	EPFU	2005
2006-MV-17	5A	Bat capture	State land	Michigan	USA	EPFU	2006
LW-E01	6A	Bat capture	State land	Michigan	USA	EPFU	2005
LW-E02	6A	Bat capture	State land	Michigan	USA	EPFU	2005
LW-E03	7A	Bat capture	State land	Michigan	USA	EPFU	2005
MV496-Cim1	8A	Bat capture	House	Indiana	USA	EPFU	2010
MV496-Cim2	8A	Bat capture	House	Indiana	USA	EPFU	2010
EFAB-9-Cim1	9A	Bat capture	House	Indiana	USA	EPFU	2005
EFAB-9-Cim2	9A	Bat capture	House	Indiana	USA	EPFU	2005
EFWG-16	10A	Bat capture	School	Indiana	USA	EPFU	2005
EFWN-23	10A	Bat capture	School	Indiana	USA	EPFU	2005
EFWN-28	10A	Bat capture	School	Indiana	USA	EPFU	2005
EFGB-18	11A	Bat capture	Barn	Indiana	USA	EPFU	2005
EFMS-16	12A	Bat capture	School	Indiana	USA	EPFU	2005
MV642	13A	Bat capture	State land	Kentucky	USA	EPFU	2010
MV652-Cim1	13A	Bat capture	State land	Kentucky	USA	EPFU	2010
MV194-Cim1	14A	Bat capture	House	Michigan	USA	EPFU	2008
MV194-Cim2	14A	Bat capture	House	Michigan	USA	EPFU	2008
MV205	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH01	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH02	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH03	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH04	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH05	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH06	14A	Bat capture	House	Michigan	USA	EPFU	2008

RH09	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH11	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH12	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH13	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH14	14A	Bat capture	House	Michigan	USA	EPFU	2008
RH15	14A	Bat capture	House	Michigan	USA	EPFU	2008
O1	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O2	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O3	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O4	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O5	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O6	15A	Bat capture	House	Ontario	Canada	EPFU	2014
O7	15A	Bat capture	House	Ontario	Canada	EPFU	2014
CCC-Roost-Cim2	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim3	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim4	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim5	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim6	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim7	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim8	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim9	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim10	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim11	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim12	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim13	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim14	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-Roost-Cim15	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006

CCC-Roost-Cim16	16A	Roost visit	Church	Pennsylvania	USA	MYLU	2006
CCC-4	17A	Bat capture	Church	Pennsylvania	USA	MYLU	2006
CCC-45	17A	Bat capture	Church	Pennsylvania	USA	MYLU	2006
CCC-50	17A	Bat capture	Church	Pennsylvania	USA	MYLU	2006
HC-4	18A	Bat capture	Church	New Jersey	USA	MYLU	2006
O8	19A	Bat capture	Long Point Provincial Park	Ontario	Canada	EPFU	2014

Table B.2 List of *Cimex lectularius* specimens included in analyses. Site refers to each unique sampling location in the study, and corresponds to those in Fig. 3.2 and Table 3.1.

ID	Site	Province/State	Country	Year
1-1	1L	Ontario	Canada	2014
1-2	1L	Ontario	Canada	2014
1-3	1L	Ontario	Canada	2014
2-1	2L	Ontario	Canada	2014
2-2	2L	Ontario	Canada	2014
2-3	2L	Ontario	Canada	2014
2-4	2L	Ontario	Canada	2014
2-5	2L	Ontario	Canada	2014
3-1	3L	Ontario	Canada	2014
3-3	3L	Ontario	Canada	2014
3-4	3L	Ontario	Canada	2014
3-5	3L	Ontario	Canada	2014
3-6	3L	Ontario	Canada	2014
3-7	3L	Ontario	Canada	2014
4-1	4L	Ontario	Canada	2014
4-2	4L	Ontario	Canada	2014
4-3	4L	Ontario	Canada	2014
4-4	4L	Ontario	Canada	2014
4-5	4L	Ontario	Canada	2014
4-6	4L	Ontario	Canada	2014
5-1	5L	Ontario	Canada	2014
5-2	5L	Ontario	Canada	2014
5-3	5L	Ontario	Canada	2014

6-1	6L	Ontario	Canada	2014
7-1	7L	Ontario	Canada	2014
7-2	7L	Ontario	Canada	2014
7-3	7L	Ontario	Canada	2014
8-1	8L	Ontario	Canada	2014
8-2	8L	Ontario	Canada	2014
9-1	9L	Ontario	Canada	2014
9-2	9L	Ontario	Canada	2014
9-3	9L	Ontario	Canada	2014
10-2	10L	Ontario	Canada	2014
11-1	11L	Ontario	Canada	2014
11-2	11L	Ontario	Canada	2014
11-3	11L	Ontario	Canada	2014
12-1	12L	Ontario	Canada	2014
13-1	13L	Ontario	Canada	2014
15-1	14L	Ohio	USA	2014
15-2	14L	Ohio	USA	2014
15-3	14L	Ohio	USA	2014
20-1	15L	Michigan	USA	2014
20-2	15L	Michigan	USA	2014
20-3	15L	Michigan	USA	2014
20-4	15L	Michigan	USA	2014
20-5	15L	Michigan	USA	2014
20-6	15L	Michigan	USA	2014
22-1	16L	Michigan	USA	2014
22-2	16L	Michigan	USA	2014
23-1	17L	Michigan	USA	2014

24-2	18L	Michigan	USA	2014
25-2	19L	Michigan	USA	2014
25-3	19L	Michigan	USA	2014
26-1	20L	Michigan	USA	2014
27-1	21L	Michigan	USA	2014
27-2	21L	Michigan	USA	2014
27-3	21L	Michigan	USA	2014
28-1	22L	Michigan	USA	2014
28-2	22L	Michigan	USA	2014
29-1	23L	Michigan	USA	2014
29-2	23L	Michigan	USA	2014
29-3	23L	Michigan	USA	2014
30-1	24L	Michigan	USA	2014
31-1	25L	Michigan	USA	2014
34-2	26L	Michigan	USA	2014
34-3	27L	Michigan	USA	2014
35-1	27L	Michigan	USA	2014
35-2	27L	Michigan	USA	2014
35-3	27L	Michigan	USA	2014
36-1	28L	Michigan	USA	2014
37-1	29L	Michigan	USA	2014
38-1	30L	Michigan	USA	2014
39-1	31L	Michigan	USA	2014

Table B.3 Genotypic data at 20 microsatellite markers for 73 *Cimex lectularius* individuals. The top three digits describe the length (in base pairs) of the first allele, and the lower three describe the length of the second allele. Missing data is identified with “000” both on the top and the bottom row.

ID	BB 38B	Clec 11	BB 31B	Clec 21	Clec 48	Clec 96	Clec 97	Clec 104	Clec 45	Clec 99	BB 21B	BB 28B	BB 29B	Clec 15	Cle 013	Cle 015	Cle 002	Cle 021	Cle 011	Cle 003
1-1	110	244	177	267	228	164	280	237	151	117	151	095	167	210	000	000	241	210	170	000
	110	244	177	267	228	164	280	237	151	126	157	116	251	210	000	000	270	229	170	000
1-2	110	247	177	267	228	164	280	231	151	117	151	101	169	210	239	292	270	229	172	253
	110	255	177	267	228	164	288	231	151	126	157	116	251	210	242	302	270	229	172	276
1-3	110	255	177	266	228	164	280	231	154	126	157	111	169	210	216	000	246	229	139	253
	112	255	177	266	228	164	280	237	154	126	169	116	169	210	248	000	263	229	182	276
2-1	112	255	180	266	228	164	280	231	151	130	157	114	192	210	216	288	246	229	139	276
	116	255	191	267	231	164	280	231	151	133	169	116	192	216	242	288	246	229	139	276
2-2	112	244	180	266	231	167	278	231	154	133	157	114	192	210	216	292	246	242	139	259
	116	255	180	267	231	167	280	231	154	133	169	116	192	216	216	292	246	242	139	276
2-3	112	255	191	266	231	164	280	231	151	130	157	114	171	210	000	292	246	242	142	000
	116	255	191	267	231	164	280	231	151	133	169	116	192	216	000	292	246	242	142	000
2-4	116	255	180	266	231	164	280	231	151	130	157	116	171	210	216	292	246	229	139	266
	116	255	191	266	231	164	280	231	151	133	157	116	171	210	242	292	267	242	160	276
2-5	112	244	180	266	228	167	278	231	151	133	157	114	171	210	216	288	246	242	139	276
	116	255	191	267	231	167	280	231	151	133	169	116	171	210	216	292	267	242	160	276
3-1	116	247	180	266	228	164	280	231	151	117	167	095	187	210	224	302	246	229	139	262
	116	255	180	266	228	164	288	231	154	120	167	099	187	210	275	302	246	239	139	266
3-3	114	247	180	266	228	164	288	231	151	117	151	095	187	210	224	302	246	229	139	262
	116	247	180	266	228	164	288	231	154	120	167	105	187	210	226	302	246	229	139	266
3-4	114	247	180	266	228	164	288	250	151	120	167	095	187	210	224	302	246	229	139	262
	116	247	180	266	228	164	288	250	154	120	167	099	187	210	226	302	246	229	139	276
3-5	114	253	180	266	228	164	278	231	151	117	151	099	187	210	224	302	246	239	139	205
	116	255	180	266	228	164	280	231	154	117	167	105	187	210	226	302	246	239	139	262
3-6	116	255	177	266	228	164	280	231	151	120	167	095	187	210	239	302	246	229	139	264
	116	255	180	266	228	164	280	231	151	133	167	105	187	210	242	302	246	229	139	276
3-7	116	253	180	266	228	164	278	250	151	120	151	095	187	210	224	302	246	229	139	266
	116	253	180	266	228	164	278	250	151	120	167	105	187	210	226	302	246	239	139	266
4-1	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	229	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	246	239	139	266
4-2	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	239	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	246	239	139	266
4-3	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	229	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	246	229	139	266

4-4	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	239	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	267	239	160	266
4-5	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	229	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	267	239	160	266
4-6	116	255	180	266	228	164	280	231	154	117	157	116	190	216	239	296	246	229	139	266
	116	255	180	266	228	164	280	231	154	117	157	116	190	216	242	296	246	229	139	266
5-1	112	255	177	266	228	164	280	231	151	117	000	095	251	216	000	000	000	242	160	266
	112	255	177	266	228	164	280	231	151	117	000	095	251	216	000	000	000	242	160	266
5-2	112	255	177	266	228	164	280	231	151	117	151	095	251	210	239	302	242	242	164	266
	112	255	177	266	228	164	280	231	151	117	151	095	251	210	242	302	242	242	164	266
5-3	112	255	177	000	000	164	280	231	151	117	151	095	251	210	239	302	242	242	164	266
	112	255	177	000	000	164	280	231	151	117	151	095	251	210	242	302	242	242	164	266
6-1	112	255	176	266	228	149	288	234	154	120	163	114	251	210	248	292	241	229	160	266
	116	255	177	266	228	149	288	234	154	133	163	114	251	210	250	302	267	229	166	266
6-2	114	255	176	266	228	164	280	231	154	117	151	095	251	210	248	302	241	239	160	244
	116	255	176	266	228	164	288	231	154	117	151	095	251	210	250	302	267	239	160	266
6-3	116	255	177	266	228	164	280	231	154	117	151	095	251	210	242	302	241	239	166	244
	116	255	177	266	228	164	288	231	154	120	151	114	251	210	242	302	241	239	166	266
7-1	114	255	176	266	228	167	280	231	151	117	149	097	171	210	270	292	261	239	160	266
	114	255	176	266	228	167	280	231	151	117	149	099	251	210	272	302	267	239	172	266
7-2	114	255	176	266	228	164	280	231	151	117	149	097	171	210	270	292	261	239	160	266
	114	255	176	266	228	164	280	231	151	117	149	097	251	210	272	302	267	239	172	266
8-1	114	255	177	266	228	164	280	231	151	133	151	095	251	210	213	292	241	239	164	216
	116	255	177	266	228	164	280	231	151	133	167	095	251	210	216	302	241	239	164	216
8-2	114	255	177	266	228	164	280	231	151	117	151	095	251	210	239	292	241	239	164	216
	116	255	177	266	228	164	280	231	154	133	167	116	251	210	242	302	241	239	164	266
8-3	114	255	177	266	228	164	280	231	151	117	151	095	251	210	213	292	241	239	164	216
	116	255	177	266	228	164	280	231	151	133	167	116	251	210	216	292	241	239	164	266
9-2	112	255	177	266	228	164	280	231	151	117	151	095	169	210	239	292	242	239	160	266
	112	255	177	266	228	164	280	231	151	117	169	111	169	210	242	292	267	239	180	266
10-1	114	255	177	253	228	164	280	231	154	117	149	095	169	210	239	302	241	239	160	266
	116	255	177	266	228	164	280	231	154	133	151	097	251	216	242	302	267	239	164	266
10-2	114	255	177	253	228	164	280	231	151	117	149	095	169	210	239	306	244	239	160	266
	116	255	177	266	228	164	288	231	154	117	153	097	251	216	242	306	267	239	160	273
10-3	114	255	177	266	228	164	280	234	154	117	149	097	251	210	239	302	239	239	160	266
	114	255	177	266	228	164	280	234	154	133	149	097	251	210	242	306	267	239	160	273
11-1	112	255	177	266	228	164	280	231	151	117	151	095	251	210	239	302	241	242	164	264
	112	255	177	266	228	164	280	231	151	117	151	095	251	210	242	302	241	242	164	266
12-1	116	247	184	266	228	164	278	231	151	120	159	105	167	210	219	000	261	237	178	244
	116	255	194	266	228	164	288	231	151	120	176	116	190	210	219	000	261	237	178	271
13-1	116	247	184	266	228	164	278	231	151	120	159	105	167	210	219	292	257	237	178	244

	116	255	191	266	228	164	288	231	154	120	176	116	190	210	219	302	261	237	178	244
15-1	110	247	176	266	228	164	278	231	151	117	151	101	167	210	211	292	246	237	168	244
	114	247	177	267	228	164	278	231	151	133	159	101	192	210	216	292	263	244	182	256
15-2	114	255	177	267	228	164	278	231	151	133	151	101	192	210	211	292	246	237	168	242
	114	255	177	267	228	164	278	231	151	133	159	101	251	210	216	292	246	237	168	256
15-3	110	255	177	267	228	164	291	231	151	117	151	101	167	210	216	292	246	237	168	244
	114	255	177	267	228	164	291	231	151	133	169	118	192	210	216	292	263	261	182	244
20-1	110	247	183	266	228	149	288	246	154	126	163	116	204	210	219	292	239	242	166	274
	110	247	183	267	228	149	288	246	154	126	163	118	204	210	219	292	239	242	166	274
20-2	110	255	183	266	228	167	278	231	154	117	163	122	152	210	219	292	244	239	170	274
	110	255	183	266	228	167	278	231	154	117	163	124	177	210	219	292	244	239	170	274
20-3	110	255	183	266	228	164	278	231	154	126	159	124	152	210	219	292	244	239	170	274
	110	255	183	266	228	167	278	231	154	133	163	126	177	210	219	292	244	242	170	274
20-4	110	255	183	266	228	164	278	231	154	117	163	122	177	210	219	292	244	242	170	274
	110	255	183	266	228	167	278	231	154	133	163	124	177	210	219	292	244	242	170	274
20-5	110	255	183	266	228	167	278	231	154	126	163	122	152	210	000	292	244	239	170	274
	110	255	183	266	228	167	278	231	154	133	163	124	177	210	000	292	244	242	170	274
20-6	110	255	182	266	228	167	278	231	154	126	163	122	152	000	219	292	244	242	168	274
	110	255	183	266	228	167	278	231	154	133	163	124	152	000	219	292	246	242	168	274
22-1	110	247	183	266	228	149	288	246	154	126	169	116	204	000	219	292	239	242	166	274
	110	247	183	266	228	149	288	246	154	126	169	118	204	000	219	292	239	242	166	274
22-2	110	247	183	266	228	149	288	246	154	126	169	116	204	210	219	292	239	242	166	273
	110	247	183	266	228	149	288	246	154	126	169	118	204	210	219	292	239	242	166	273
23-1	110	255	182	266	228	164	288	231	154	133	169	116	167	210	226	292	244	239	170	242
	112	255	182	267	228	167	288	231	154	133	169	118	220	210	229	292	244	242	170	244
24-2	110	247	183	266	228	149	288	246	154	126	163	118	204	210	219	292	239	242	166	274
	110	247	183	266	228	149	288	246	154	126	163	118	204	210	219	292	239	242	166	274
25-2	114	255	196	267	228	167	288	231	154	133	165	109	152	000	242	000	244	242	170	242
	114	255	196	267	228	167	288	231	154	133	169	118	220	000	242	000	244	242	170	251
25-3	110	255	196	266	228	167	288	231	154	133	165	109	220	210	226	292	244	261	170	251
	114	255	196	267	228	167	288	231	154	133	169	118	220	210	229	292	244	261	170	251
26-1	112	255	180	266	228	167	288	231	154	117	157	116	220	210	000	306	239	242	166	216
	114	255	196	267	228	167	288	231	154	133	169	118	233	210	000	306	263	242	182	274
27-1	110	247	193	266	228	164	288	231	154	117	169	116	167	210	239	292	244	242	170	216
	114	255	196	266	228	164	288	231	154	117	169	118	220	210	242	292	244	242	170	216
27-2	114	255	182	266	228	164	288	231	154	117	163	118	167	210	239	292	244	242	168	216
	114	255	196	266	228	164	288	231	154	117	167	124	220	210	242	292	246	242	168	274
27-3	114	255	183	266	228	164	278	231	154	130	000	116	204	000	219	000	263	242	182	244
	114	255	193	266	228	164	288	231	154	130	000	118	220	000	219	000	270	242	182	274
28-1	110	255	193	266	228	164	288	231	151	130	000	122	233	000	000	292	239	239	162	216
	144	255	196	267	228	164	288	231	154	133	000	124	233	000	000	292	267	239	162	244

28-2	110	255	183	266	228	176	288	231	151	126	169	116	233	210	226	292	242	242	170	242
	110	255	193	266	231	176	288	231	151	126	169	118	233	210	236	292	244	242	170	242
29-1	144	255	196	267	228	167	278	231	154	126	163	109	204	210	207	306	267	229	162	262
	144	255	196	267	228	167	278	231	154	130	165	124	204	210	207	306	267	242	162	262
29-2	110	255	193	266	228	167	278	231	151	130	163	109	204	210	229	292	244	229	162	262
	144	255	196	267	228	176	288	231	154	130	165	109	233	210	229	306	244	239	170	262
29-3	110	255	186	266	228	176	288	231	151	130	163	118	204	210	207	306	261	242	172	216
	114	255	196	267	228	176	288	231	151	133	169	124	204	210	207	314	267	242	172	262
30-1	110	255	180	266	228	164	278	231	154	130	163	116	220	210	219	306	244	242	170	216
	110	255	191	266	228	167	288	231	154	133	169	118	233	210	219	314	244	244	170	260
31-1	114	247	193	267	228	164	288	231	154	126	163	122	204	210	239	292	244	239	182	274
	114	247	193	267	228	164	288	231	154	126	163	124	204	210	242	292	263	239	182	274
34-2	116	247	180	000	000	167	278	231	154	130	000	101	220	000	219	000	241	242	168	264
	116	247	191	000	000	167	278	231	154	130	000	101	220	000	219	000	241	242	168	264
34-3	116	247	180	267	228	164	278	231	154	130	151	101	220	210	239	306	244	242	170	244
	116	247	191	267	228	164	278	231	154	133	151	101	220	210	242	306	244	244	170	244
35-1	110	247	183	266	228	149	288	246	154	126	163	118	204	210	219	000	239	242	166	274
	110	247	183	266	228	149	288	246	154	126	163	118	204	210	219	000	239	242	166	274
35-2	110	247	183	266	228	149	288	246	154	126	163	116	204	210	219	292	239	242	166	274
	110	247	183	267	228	149	288	246	154	126	163	118	204	210	219	292	239	242	166	274
35-3	110	247	183	266	228	149	288	246	154	129	163	116	204	210	219	292	239	242	166	274
	110	247	183	266	228	149	288	246	154	129	163	118	204	210	219	292	239	242	166	274
36-1	110	247	180	266	228	149	278	231	151	133	163	105	167	210	000	000	239	242	000	242
	114	255	183	266	228	164	278	246	154	133	169	118	204	210	000	000	239	242	000	244
37-1	114	255	180	266	228	164	280	231	154	133	163	105	167	210	219	292	234	226	166	274
	114	255	183	266	228	164	280	231	154	133	163	118	204	210	219	314	237	246	184	276
38-1	116	253	177	266	231	164	298	231	151	133	159	114	190	210	219	292	263	239	162	253
	116	253	191	266	231	164	298	231	151	133	159	114	220	210	219	292	263	239	162	266
39-1	110	255	191	266	228	164	288	231	151	126	153	109	194	2102	216	302	257	229	168	266
	114	255	191	267	228	164	288	231	154	133	153	109	251	10	216	302	257	229	168	266

Table B.4 Spatial genetic structure, diversity, and differentiation in *C. lectularius* samples, estimated using two different panels of microsatellite markers varying in the number of markers. Most likely number of genetic clusters (*K*) estimated using the Geneland method, isolation-by-distance (IBD) and IBD while correcting for population genetic structure (IBD + K), average expected heterozygosity (H_S), Hedrick's global G'_{ST} differentiation index (and 95% confidence intervals), and AMOVA = Analysis of Molecular Variance.

Number of markers		**20**	**7**
Most likely K		13	11
IBD (r_W)	*P*	0.001*	0.001*
	R^2	0.11	0.09
IBD (r_W) + *K*	*P* (IBD)	0.001*	0.001*
	P (*K*)	0.001*	0.001*
	R^2	0.26	0.22
H_S		0.310 (0.250 – 0.370)	0.353 (0.226 – 0.471)
Hedrick's G'_{ST}		0.768 (0.706 – 0.828)	0.739 (0.626 – 0.863)
AMOVA (% of variation)	Within sites	32.0	40.5
	Among sites	68.0	59.5

*Statistically significant at α = 0.05

Appendix C : Chapter 4 Supplementary Material

Table C.1 Big brown bat, *Eptesicus fuscus,* specimens included in our study in the southern Great Lakes region of North America. 'Genotype source' refers to whether the specimen was genotyped at eight microsatellite markers in this current study or in the study of Vonhof et al. 2008. Site refers to each unique sampling location in the study, and correspond to those in Table C.3. Site characteristics refers either to the type of roost (house, barn, school, church) in, or next to which, we collected samples, or to the forested land jurisdiction where collection was performed. Number of parasites refers to the number of *Cimex adjunctus* specimens collected from the same site and used in this study.

ID	Year	Genotyped source	Site	State	Site characteristics	Number of parasites
CLR-01	2005	This study	1	Michigan	House	0
CLR-02	2005	This study	1	Michigan	House	0
EF-293-1	2005	This study	2	Michigan	House	1
EF-293-10	2005	This study	2	Michigan	House	1
EF-293-11	2005	This study	2	Michigan	House	1
EF-293-12	2005	This study	2	Michigan	House	1
EF-293-13	2005	This study	2	Michigan	House	1
EF36-1	2005	This study	3	Michigan	State land	0
EF36-10	2005	This study	3	Michigan	State land	0
EF36-11	2005	This study	3	Michigan	State land	0
EF36-2	2005	This study	3	Michigan	State land	0
EF36-3	2005	This study	3	Michigan	State land	0
EFAB-1	2005	This study	4	Indiana	House	2
EFAB-10	2005	This study	4	Indiana	House	2
EFAB-11	2005	This study	4	Indiana	House	2
EFAB-2	2005	This study	4	Indiana	House	2
EFAB-3	2005	This study	4	Indiana	House	2

EFF-10	2005	This study	5	Michigan	State land	0
EFF-11	2005	This study	5	Michigan	State land	0
EFF-12	2005	This study	5	Michigan	State land	0
EFF-13	2005	This study	5	Michigan	State land	0
EFF-14	2005	This study	5	Michigan	State land	0
EFGB-10	2005	This study	6	Indiana	Barn	1
EFGB-11	2005	This study	6	Indiana	Barn	1
EFGB-12	2005	This study	6	Indiana	Barn	1
EFGB-13	2005	This study	6	Indiana	Barn	1
EFMC-1	2005	This study	7	Michigan	State land	0
EFMC-10	2005	This study	7	Michigan	State land	0
EFMC-11	2005	This study	7	Michigan	State land	0
EFMC-12	2005	This study	7	Michigan	State land	0
EFMC-13	2005	This study	7	Michigan	State land	0
EFMS-1	2005	This study	8	Indiana	School	1
EFMS-10	2005	This study	8	Indiana	School	1
EFMS-11	2005	This study	8	Indiana	School	1
EFMS-12	2005	This study	8	Indiana	School	1
EFMS-13	2005	This study	8	Indiana	School	1
EFS42-1	2005	This study	9	Michigan	House	4
EFS42-10	2005	This study	9	Michigan	House	4
EFS42-11	2005	This study	9	Michigan	House	4
EFS42-12	2005	This study	9	Michigan	House	4
EFS42-13	2005	This study	9	Michigan	House	4
EFWG-1	2005	This study	10	Indiana	School	3
EFWG-10	2005	This study	10	Indiana	School	3
EFWG-11	2005	This study	10	Indiana	School	3

EFWG-12	2005	This study	10	Indiana	School	3
EFWG-13	2005	This study	10	Indiana	School	3
GF-01	1997	This study	11	Indiana	State land	0
GF-02	1997	This study	11	Indiana	State land	0
GF-03	1997	This study	11	Indiana	State land	0
GF-04	1997	This study	11	Indiana	State land	0
GF-05	1997	This study	11	Indiana	State land	0
MA-01	1997	This study	12	Indiana	House	0
MA-02	1997	This study	12	Indiana	House	0
MV-150	2008	This study	13	Michigan	House	0
MV-151	2008	This study	13	Michigan	House	0
MV-152	2008	This study	13	Michigan	House	0
MV-153	2008	This study	13	Michigan	House	0
MV-154	2008	This study	13	Michigan	House	0
MV-162	2008	This study	14	Michigan	House	0
MV-163	2008	This study	14	Michigan	House	0
MV-164	2008	This study	14	Michigan	House	0
MV-165	2008	This study	14	Michigan	House	0
MV-166	2008	This study	14	Michigan	House	0
MV-188	2008	This study	15	Michigan	House	15
MV-189	2008	This study	15	Michigan	House	15
MV-190	2008	This study	15	Michigan	House	15
MV-191	2008	This study	15	Michigan	House	15
MV-192	2008	This study	15	Michigan	House	15
MV-250	2008	This study	16	Indiana	House	0
MV-251	2008	This study	16	Indiana	House	0
MV-252	2008	This study	16	Indiana	House	0

MV-253	2008	This study	16	Indiana	House	0
MV-254	2008	This study	16	Indiana	House	0
MV-265	2009	This study	17	Michigan	House	0
MV-266	2009	This study	17	Michigan	House	0
MV-267	2009	This study	17	Michigan	House	0
MV-268	2009	This study	17	Michigan	House	0
MV-269	2009	This study	17	Michigan	House	0
MV-282	2009	This study	18	Michigan	House	0
MV-288	2009	This study	19	Michigan	Barn	0
MV-289	2009	This study	19	Michigan	Barn	0
MV-323	2009	This study	20	Indiana	House	0
MV-324	2009	This study	20	Indiana	House	0
MV-325	2009	This study	20	Indiana	House	0
MV-327	2009	This study	20	Indiana	House	0
MV-419	2009	This study	21	Michigan	House	0
MV-420	2009	This study	21	Michigan	House	0
MV-421	2009	This study	21	Michigan	House	0
MV-422	2009	This study	21	Michigan	House	0
MV-491	2010	This study	22	Indiana	House	2
MV-492	2010	This study	22	Indiana	House	2
MV-494	2010	This study	22	Indiana	House	2
MV-495	2010	This study	22	Indiana	House	2
MV-516	2010	This study	23	Kentucky	Barn	0
MV-517	2010	This study	23	Kentucky	Barn	0
MV-518	2010	This study	23	Kentucky	Barn	0
MV-519	2010	This study	23	Kentucky	Barn	0
MV-520	2010	This study	23	Kentucky	Barn	0

MV-572	2010	This study	24	Michigan	House	0
MV-573	2010	This study	24	Michigan	House	0
MV-574	2010	This study	24	Michigan	House	0
MV-575	2010	This study	24	Michigan	House	0
MV-576	2010	This study	24	Michigan	House	0
MV-642	2010	This study	25	Kentucky	State land	2
MV-643	2010	This study	25	Kentucky	State land	2
MV-644	2010	This study	25	Kentucky	State land	2
MV-645	2010	This study	25	Kentucky	State land	2
MV-646	2010	This study	25	Kentucky	State land	2
MV-87	2006	This study	26	Michigan	State land	0
MV-88	2006	This study	26	Michigan	State land	0
MV-89	2006	This study	26	Michigan	State land	0
MV-90	2006	This study	26	Michigan	State land	0
LC01	1997	Vonhof et al. 2008	27	Indiana	State land	0
LC02	1997	Vonhof et al. 2008	27	Indiana	State land	0
LC03	1997	Vonhof et al. 2008	27	Indiana	State land	0
LC04	1997	Vonhof et al. 2008	27	Indiana	State land	0
LC05	1997	Vonhof et al. 2008	27	Indiana	State land	0
1	1997	Vonhof et al. 2008	28	Indiana	School	0
2	1997	Vonhof et al. 2008	28	Indiana	School	0
3	1997	Vonhof et al. 2008	28	Indiana	School	0
4	1997	Vonhof et al. 2008	28	Indiana	School	0
5	1997	Vonhof et al. 2008	28	Indiana	School	0
SM01	1997	Vonhof et al. 2008	29	Indiana	Church	0
SM02	1997	Vonhof et al. 2008	29	Indiana	Church	0
SM03	1997	Vonhof et al. 2008	29	Indiana	Church	0

SM04	1997	Vonhof et al. 2008	29	Indiana	Church	0
SM05	1997	Vonhof et al. 2008	29	Indiana	Church	0
HB01	1997	Vonhof et al. 2008	30	Indiana	Barn	0
HB02	1997	Vonhof et al. 2008	30	Indiana	Barn	0
HB03	1997	Vonhof et al. 2008	30	Indiana	Barn	0
HB04	1997	Vonhof et al. 2008	30	Indiana	Barn	0
HB05	1997	Vonhof et al. 2008	30	Indiana	Barn	0
PA19	1997	Vonhof et al. 2008	31	Illinois	House	0
PA20	1997	Vonhof et al. 2008	31	Illinois	House	0
PA21	1997	Vonhof et al. 2008	31	Illinois	House	0
PA29	1997	Vonhof et al. 2008	31	Illinois	House	0
PA02	1997	Vonhof et al. 2008	31	Illinois	House	0
WB01	1997	Vonhof et al. 2008	32	Indiana	House	0
WB02	1997	Vonhof et al. 2008	32	Indiana	House	0
WB03	1997	Vonhof et al. 2008	32	Indiana	House	0
WB04	1997	Vonhof et al. 2008	32	Indiana	House	0
WB05	1997	Vonhof et al. 2008	32	Indiana	House	0

Table C.2 Genotypic data at 8 microsatellite markers for 142 big brown bat (*Eptesicus fuscus*) individuals collected in the southern Great Lakes region of North America. The first three digits in each string describe the length (in base pairs) of the first allele, and the last three describe the length of the second allele. Missing data is identified with "000000".

ID	EF1	EF6	EF14	EF15	EF20	MMG9	MMG25	TT20
CLR-01	169177	179181	114118	127140	100111	140140	121133	178188
CLR-02	173177	175192	112116	139140	102110	127134	136136	180184
EF-293-1	163167	184184	118120	129129	93106	116159	121133	180192
EF-293-10	167171	175187	110114	127129	110110	151154	133139	180186
EF-293-11	171175	179188	108112	111135	78089	146170	133133	184190
EF-293-12	171177	184189	108120	129129	78106	164164	121133	180182
EF-293-13	179179	175182	98102	129133	87093	165172	124133	180186
EF36-1	179179	165187	116118	133133	87093	130160	133139	180184
EF36-10	165179	173187	116118	127127	87087	130174	133133	180184
EF36-11	171177	181187	98110	127129	78106	160160	113133	182186
EF36-2	167169	175192	110112	127133	104108	127176	113124	180182
EF36-3	177179	183183	102102	129131	110110	127140	124136	180182
EFAB-1	175175	192192	110120	127129	106106	138161	133136	180184
EFAB-10	179209	187190	110116	133135	100111	180180	133139	180184
EFAB-11	167177	177180	110114	127141	106106	141161	133133	180184
EFAB-2	171177	179189	95116	121129	84097	165172	133139	180184
EFAB-3	165177	173181	95118	129129	93100	151164	136136	184184
EFF-10	165173	171177	98100	127127	106106	153165	121121	180182
EFF-11	167175	173187	110118	118133	106106	157157	133136	182186
EFF-12	167169	179193	95137	131143	93100	161161	121133	180186
EFF-13	177204	183183	116118	121129	100106	140159	133136	180186
EFF-14	169171	175179	112118	141141	78095	164164	133133	182182

EFGB-10	177177	173177	95112	118118	100108	143143	133133	182186
EFGB-11	165215	171173	118137	127150	87108	182182	127133	180180
EFGB-12	169171	187189	100116	121158	104114	168170	133142	182184
EFGB-13	177177	173187	95112	118135	100110	143155	121121	180186
EFMC-1	173175	180187	110112	118123	104110	160164	133133	182182
EFMC-10	171177	165183	110118	133137	95106	116116	133133	180188
EFMC-11	169171	173179	110114	127141	95108	140149	121133	182184
EFMC-12	167177	177184	118118	138138	78106	153167	124133	186186
EFMC-13	169175	156167	104114	135141	97110	164173	121133	180180
EFMS-1	165217	175179	95120	129135	78104	150157	136136	182186
EFMS-10	167169	173193	108120	129152	89108	162165	121133	180184
EFMS-11	167177	181185	114116	119119	104110	150161	124133	182182
EFMS-12	171173	175179	110114	129148	87104	140159	133133	182182
EFMS-13	167177	177185	95114	118127	78108	158168	133133	180184
EFS42-1	175202	185189	108124	129129	108110	160169	121133	180180
EFS42-10	165219	187196	114118	129135	108110	153165	132136	184186
EFS42-11	171177	188196	118137	131142	102110	155155	121136	180184
EFS42-12	173179	177185	114116	139141	106106	151157	136142	182184
EFS42-13	167175	187190	112118	125133	108108	116156	133133	184188
EFWG-1	173181	181187	104108	119127	110111	170176	133136	180182
EFWG-10	175179	179187	110116	110135	82108	158161	124136	180186
EFWG-11	171173	173177	118118	123133	104110	151164	133133	180180
EFWG-12	171173	179181	104108	125141	108108	162167	133139	180184
EFWG-13	169198	171175	110112	116129	89108	146164	133133	180184
GF-01	175179	181186	102118	121127	108110	116165	133133	180188
GF-02	173177	175183	104112	104148	102110	160164	133133	180180
GF-03	171175	179185	93118	129131	100120	127167	133133	180180

GF-04	171175	184186	108116	118127	106124	151165	133133	180184
GF-05	167173	185187	108112	135135	108110	140143	113133	184188
MA-01	163177	175192	93114	123127	87104	165168	121133	180184
MA-02	167177	165183	112114	137137	78097	127154	121124	180188
MV-150	165167	170185	102102	133141	104104	140161	124124	180186
MV-151	177202	183185	114114	127141	100100	141161	133133	180186
MV-152	171175	184192	110114	104141	110110	155165	133133	179184
MV-153	167169	169184	114118	129133	95106	146160	121133	180182
MV-154	165169	165192	112114	123125	87106	165165	133133	180186
MV-162	171171	183188	116137	127158	102108	151165	121133	182184
MV-163	175177	175187	116122	135135	102106	158165	113133	182186
MV-164	167171	165187	95118	129133	95116	142151	133133	182186
MV-165	171171	156175	114122	146146	102108	151151	121133	182186
MV-166	171181	189189	110116	129143	104108	116116	113133	182184
MV-188	167175	184188	102112	135135	106110	141176	130130	182184
MV-189	169173	185192	112112	118118	102106	158176	133133	182184
MV-190	173175	173192	98118	104141	100108	154154	133136	182184
MV-191	171202	181190	120120	104104	104104	150161	133139	182184
MV-192	171177	181187	110118	104125	110110	127150	124133	182188
MV-250	163173	177188	102120	123133	78108	153162	133133	180186
MV-251	171177	177187	98110	131133	78106	151164	133133	178180
MV-252	169171	186192	108120	146146	78110	180183	132136	180184
MV-253	173177	183192	116118	129131	100100	167167	133139	180180
MV-254	171171	183189	104110	127127	95110	116165	133136	184188
MV-265	165169	169181	110112	104104	108110	140176	127133	180180
MV-266	173177	169192	116116	118137	84110	140182	121133	180184
MV-267	169171	185187	112116	131133	106108	158164	113127	180182

MV-268	177181	188192	112118	123131	78087	161164	139139	178182
MV-269	167202	185185	110114	127129	110110	116126	113133	180184
MV-282	175177	183189	112114	127129	106111	146158	113133	180188
MV-288	169211	175183	98098	127135	78097	155164	113133	180182
MV-289	179181	175192	116120	143143	104111	164170	121133	180184
MV-323	167167	163177	112114	127127	106106	162169	133133	182182
MV-324	171177	177196	110114	119119	89104	153164	121133	180184
MV-325	167179	179185	98100	118129	106110	127170	127136	178182
MV-327	177179	177181	110110	129137	104110	164180	113133	184184
MV-419	165173	185185	110120	131133	100106	154176	136136	186186
MV-420	171173	177177	120120	135135	87087	156164	136136	180182
MV-421	167171	173183	112118	118118	84102	150164	127133	180186
MV-422	169171	175194	116120	127129	100106	140151	124133	182186
MV-491	169169	169192	95120	104127	106106	140164	124133	180186
MV-492	169177	173188	104110	135135	108108	151155	124133	178180
MV-494	165179	183183	114118	123141	106106	116169	133136	180184
MV-495	167177	181185	110114	123137	106110	161180	113133	180184
MV-516	171179	181187	102130	129131	104108	143168	133133	184184
MV-517	171177	173185	108116	133158	106106	161167	133136	182184
MV-518	167169	185187	116116	133133	106106	153164	133136	180184
MV-519	167177	163169	102116	133147	93104	116155	124136	182186
MV-520	173177	173183	110114	143143	78104	116149	136136	180180
MV-572	167179	187189	116120	118135	89104	146146	121133	180184
MV-573	171171	175192	112114	118118	106106	146153	133133	184184
MV-574	173179	188189	95116	118127	78108	127176	121145	180184
MV-575	169173	175179	120120	118118	100106	161161	133133	180182
MV-576	158171	179190	112141	123123	106106	141176	133133	184184

MV-642	173173	173181	104114	118127	104106	153167	121133	186190
MV-643	171177	183189	95114	133133	111116	172185	121133	182184
MV-644	165202	175183	112120	116116	95110	142165	133139	180184
MV-645	183202	175185	112112	129129	89106	158160	133133	180180
MV-646	158179	185192	98112	129137	104120	127156	133133	180182
MV-87	169173	156187	112114	129129	102102	146167	133133	182184
MV-88	169173	185187	116116	104148	82104	140154	133142	180182
MV-89	167171	175192	114114	129135	78104	140154	121133	178182
MV-90	158171	189192	112112	121129	100104	127170	133136	180180
LC01	173175	173182	107115	131131	85089	147158	134134	185189
LC02	171173	177181	107109	129138	85091	147149	134137	187189
LC03	173217	183184	109113	105129	81087	161165	134134	185187
LC04	167171	182183	103105	127152	77077	139187	125143	187191
LC05	169177	182187	97109	111137	93093	126152	134134	183187
1	169171	155181	117119	141156	71077	126159	134134	183187
2	167171	173182	97107	127169	79089	147168	113128	185191
3	169177	167183	115115	127133	61075	145147	125134	183187
4	175177	181183	113115	131158	91091	155157	134134	183187
5	175177	179189	107115	125131	73089	144185	125134	183183
SM01	173175	187193	97119	127139	83091	126146	134137	183185
SM02	169171	181185	103117	131152	61091	142149	122134	185185
SM03	159183	163179	109111	127152	85089	116152	134143	187187
SM04	169179	155186	105107	125129	61089	116149	134143	183185
SM05	171177	173185	109115	131157	89095	145162	134134	185187
HB01	177177	183186	111115	129131	83085	153166	134134	183189
HB02	163171	175183	111115	135139	61089	158164	134134	187187
HB03	167169	178191	101113	133144	61085	152160	134134	183187

HB04	167203	191194	111113	127140	85087	162183	113134	185187
HB05	169177	183183	103115	111119	77091	142161	134143	183183
PA19	173175	175177	111113	129154	81089	126150	125140	185191
PA20	169177	185187	95107	125133	61093	149164	134134	183187
PA21	169181	173175	95109	156162	75087	155165	134134	187193
PA29	177197	185195	113115	131161	81091	147168	137143	185189
PA02	171197	185195	111115	147153	77087	162168	113134	185189
WB01	169177	179185	107113	125127	77087	157187	128134	187189
WB02	171171	163185	107113	131158	61089	148150	134137	183185
WB03	171171	173185	109113	149175	83085	153165	131134	183189
WB04	173177	177189	99115	127135	85087	150156	134134	183189
WB05	165169	173187	113121	129135	83083	140164	134137	185185

Table C.3 *Cimex adjunctus* specimens included in our study in the southern Great Lakes region of North America. Site refers to each unique sampling location in the study, and corresponds to those in Table C.1. Site numbers with an asterisk correspond to sites where no big brown bat data were collected. Capture method refers to the way *C. adjunctus* samples were collected: either from the body of a bat that was captured outside a roost with a mist net or a harp trap ('Bat capture'), or from inside of the roost itself ('Roost visit'). Site characteristics refers either to the type of roost (house, barn, school, church) in, or next to which, we collected samples, or to the forested land jurisdiction where collection was performed.

ID	Year	Site	State/Province	Capture Method	Site characteristics
EF-293-Roost	2005	2	Michigan	Bat capture	House
EFAB-9-Cim1	2005	4	Indiana	Bat capture	House
EFAB-9-Cim2	2005	4	Indiana	Bat capture	House
EFGB-18	2005	6	Indiana	Bat capture	Barn
EFMS-16	2005	8	Indiana	Bat capture	School
EFS42-35	2005	9	Michigan	Bat capture	House
EFS42-39	2005	9	Michigan	Bat capture	House
EFS42-43	2005	9	Michigan	Bat capture	House
EFS42-67	2005	9	Michigan	Bat capture	House
EFWG-16	2005	10	Indiana	Bat capture	School
EFWN-23	2005	10	Indiana	Bat capture	School
EFWN-28	2005	10	Indiana	Bat capture	School
MV194-Cim1	2008	15	Michigan	Bat capture	House
MV194-Cim2	2008	15	Michigan	Bat capture	House
MV205	2008	15	Michigan	Bat capture	House
RH01	2008	15	Michigan	Bat capture	House
RH02	2008	15	Michigan	Bat capture	House
RH03	2008	15	Michigan	Bat capture	House
RH04	2008	15	Michigan	Bat capture	House
RH05	2008	15	Michigan	Bat capture	House

RH06	2008	15	Michigan	Bat capture	House
RH09	2008	15	Michigan	Bat capture	House
RH11	2008	15	Michigan	Bat capture	House
RH12	2008	15	Michigan	Bat capture	House
RH13	2008	15	Michigan	Bat capture	House
RH14	2008	15	Michigan	Bat capture	House
RH15	2008	15	Michigan	Bat capture	House
MV496-Cim1	2010	22	Indiana	Bat capture	House
MV496-Cim2	2010	22	Indiana	Bat capture	House
MV642	2010	25	Kentucky	Bat capture	State land
MV652-Cim1	2010	25	Kentucky	Bat capture	State land
EFS40-10	2005	33*	Michigan	Bat capture	House
EFS40-12	2005	33*	Michigan	Bat capture	House
EFS40-14	2005	33*	Michigan	Bat capture	House
EFS40-2	2005	33*	Michigan	Bat capture	House
EFS40-4	2005	33*	Michigan	Bat capture	House
EFS40-6	2005	33*	Michigan	Bat capture	House
EFS40-7	2005	33*	Michigan	Bat capture	House
EFS42-Roost-Cim1	2005	34*	Michigan	Roost visit	House
EFS42-Roost-Cim2	2005	34*	Michigan	Roost visit	House
EFS42-Roost-Cim3	2005	34*	Michigan	Roost visit	House
EFS42-Roost-Cim4	2005	34*	Michigan	Roost visit	House
EFS42-Roost-Cim5	2005	34*	Michigan	Roost visit	House
EFS42-Roost-Cim6	2005	34*	Michigan	Roost visit	House
2006-MV-17	2006	35*	Michigan	Bat capture	State land
LW-E01	2005	36*	Michigan	Bat capture	State land
LW-E02	2005	36*	Michigan	Bat capture	State land

LW-E03	2005	37*	Michigan	Bat capture	State land
O1	2014	38*	Ontario	Bat capture	House
O2	2014	38*	Ontario	Bat capture	House
O3	2014	38*	Ontario	Bat capture	House
O4	2014	38*	Ontario	Bat capture	House
O5	2014	38*	Ontario	Bat capture	House
O6	2014	38*	Ontario	Bat capture	House
O7	2014	38*	Ontario	Bat capture	House

Literature Cited

Vonhof MJ, Strobeck C, Fenton MB (2008) Genetic variation and population structure in big brown bats (*Eptesicus fuscus*): is female dispersal important? *Journal of Mammalogy*, **89**, 1411–1419.

Appendix D : Chapter 5 Supplementary Material

Table D.1 List of *Cimex* specimens included in analyses. Host refers to the species with which the ectoparasite was found associated, either being on its body or in a roost or housing unit inhabited mainly by the species. Accession numbers for each of the four genes analysed in the study are listed, except when no amplification could be obtained (indicated with "No amplification").

Sample name	Species	Host	Country	*CO1* accession number	*EF1α* accession number	Apyrase accession number	Nitrophorin accession number
12	*Cimex lectularius*	*Homo sapiens* (Human)	Canada	KY561671	KY561686	KY561709	KY561735
411	*Cimex lectularius*	*Homo sapiens* (Human)	Canada	KY561672	KY561687	KY561710	KY561736
720	*Cimex lectularius*	*Homo sapiens* (Human)	Austria	KY561673	KY561688	KY561711	KY561737
754	*Cimex lectularius*	*Homo sapiens* (Human)	Italy	KY561674	KY561689	KY561712	KY561738
790	*Cimex lectularius*	*Homo sapiens* (Human)	Finland	KY561675	KY561690	KY561713	KY561739
51	*Cimex lectularius*	*Myotis myotis* (Bat)	Czech Republic	KY561676	KY561691	KY561714	KY561740
173	*Cimex lectularius*	*Myotis myotis* (Bat)	Czech Republic	KJ937980	KY561692	KY561715	KY561741
148	*Cimex hemipterus*	*Homo sapiens* (Human)	Indonesia	KY561677	KY561693	KY561716	KY561742
348	*Cimex hemipterus*	*Homo sapiens* (Human)	Malaysia	KY561678	KY561694	KY561717	KY561743
RG03	*Cimex adjunctus*	*Eptesicus fuscus* (Bat)	USA	KU534930	KY561695	KY561718	KY561744

EFS40	*Cimex adjunctus*	*Eptesicus fuscus* (Bat)	USA	KU534918	KY561696	KY561719	KY561745
S1	*Cimex adjunctus*	*Myotis lucifugus* (Bat)	Canada	KU534929	KY561697	KY561720	KY561746
N1	*Cimex adjunctus*	*Myotis lucifugus* (Bat)	Canada	KU534908	KY561698	KY561721	KY561747
1725	*Cimex adjunctus*	*Myotis septentrionalis* (Bat)	Canada	KU534906	KY561699	KY561722	KY561748
O9	*Cimex brevis*	*Lasionycteris noctivagans* (Bat)	Canada	KY561679	KY561700	KY561723	KY561749
N5	*Cimex brevis*	*Myotis lucifugus* (Bat)	Canada	KY561680	KY561701	KY561724	KY561750
7193	*Cimex brevis*	*Myotis lucifugus* (Bat)	Canada	KY561681	KY561702	KY561725	KY561751
A5	*Cimex latipennis*	*Myotis ciliolabrum* (Bat)	Canada	KY561682	KY561703	KY561726	KY561752
61	*Cimex pipistrelli*	*Myotis myotis* (Bat)	Czech Republic	GU985529	KY561704	KY561727	No amplification
108	*Cimex pipistrelli*	*Myotis daubentoni* (Bat)	Czech Republic	KY561683	KY561705	KY561728	KY561753
350	*Cimex japonicus*	*Verpertilio superans* (Bat)	Japan	KC503541	KF018744	KY561729	KY561754
351	*Cimex japonicus*	*Verpertilio superans* (Bat)	Japan	KY561684	KY561706	KY561730	KY561755
149	*Cimex vicarius*	*Petrochelidon pyrrhonota* (Swallow)	USA	GU985541	KF018738	KY561731	No amplification
120	*Cimex hirundinis*	*Delichon urbica* (Swallow)	Czech Republic	GU985543	KF018736	KY561732	KY561756
895	*Cimex sp.*	*Delichon*	Japan	GU985542	KY561707	KY561733	No amplification

896	*Cimex sp.*	*dasypus* (Swallow) *Delichon dasypus* (Swallow)	Japan	KY561685	KY561708	KY561734	KY561757

Table D.2 Information on primers used for amplification of target genes in our study. Length of target fragment, authors who first introduced them, and optimal polymerase chain reaction annealing temperature are shown.

Gene	Length of fragment	Name of primer	Sequence from 5' to 3'	Authors	Annealing temperature
Cytochrome Oxidase 1 (*CO1*)	576 bp	Lep1Fdeg	F: ATTCAACCAATCATAAAGATATNGG	Balvín et al. 2012	42°C
		Lep1Rdeg	R: TAWACTTCWGGRTGTCCRAARAATCA		
		CO1adj1F	F: TATGAGCAGGCATGTTAGGG	Talbot et al. 2016	49°C
		CO1adf1R	R: ATAGATGTTGATAAAGAATTGGG		
Elongation Factor 1α (*EF1α*)	509 bp	Shirley	F: GCYTCGTGGTGCATYTCSAC	Balvín et al. 2015	57°C
		Prowler	R: CAGGCTGATTGYGCTGTACTTATYCTTGC		
Apyrase	369 bp	CadjA1F	F: GTCCAGCCGACGAATGTAGT	Current study	55°C
		CadjA1R	R: CCAGTCCTGTTCTGTCATCG		
		A2F	F: CCAGCCGACGAATGTAGTTG	Current study	56°C
		A3R	R: TTACCATCTCCGTCAGCCAG		
Nitrophorin	297 bp	N2F	F: CGATCAGAAGAAATCAGGCGG	Current study	53°C
		N2R3R	R: AGGTAGGCTTGAAGGTGACC		

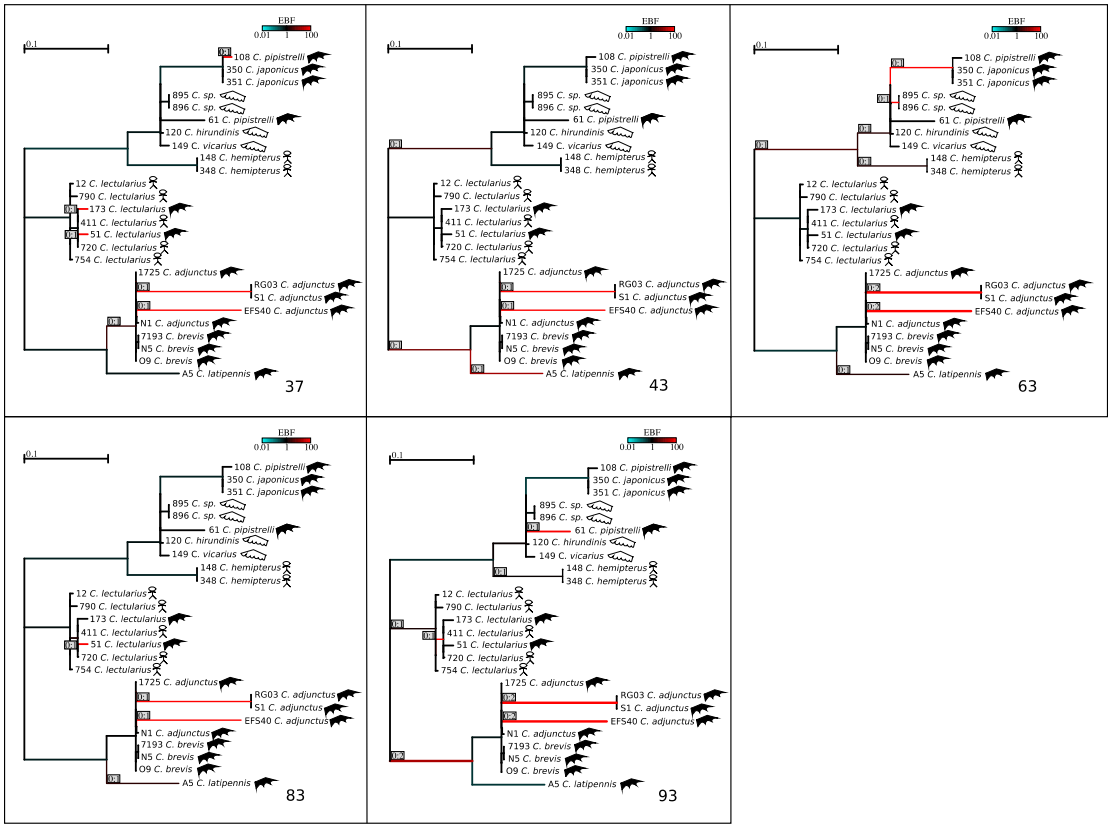

Figure D.3 Hypothesized species tree of *Cimex spp.* specimens, based on *CO1* and *EF1α*, constructed with *BEAST 2.4.2. Scale represents substitutions per site. Positive selection is inferred at specific nodes in five codons of an apyrase gene (37, 43, 63, 83 and 93; identifier is on the bottom right corner of each box). The shading of the lines indicates empirical Bayes factor (EBF), calculated using the MEME approach, where bright red represents strong inference of positive selection at that codon, teal represents strong inference of negative selection, and black represents no divergence from neutral expectation. Grey boxes at the base of branches show the number of synonymous versus nonsynonymous mutations in the codon for the respective node.

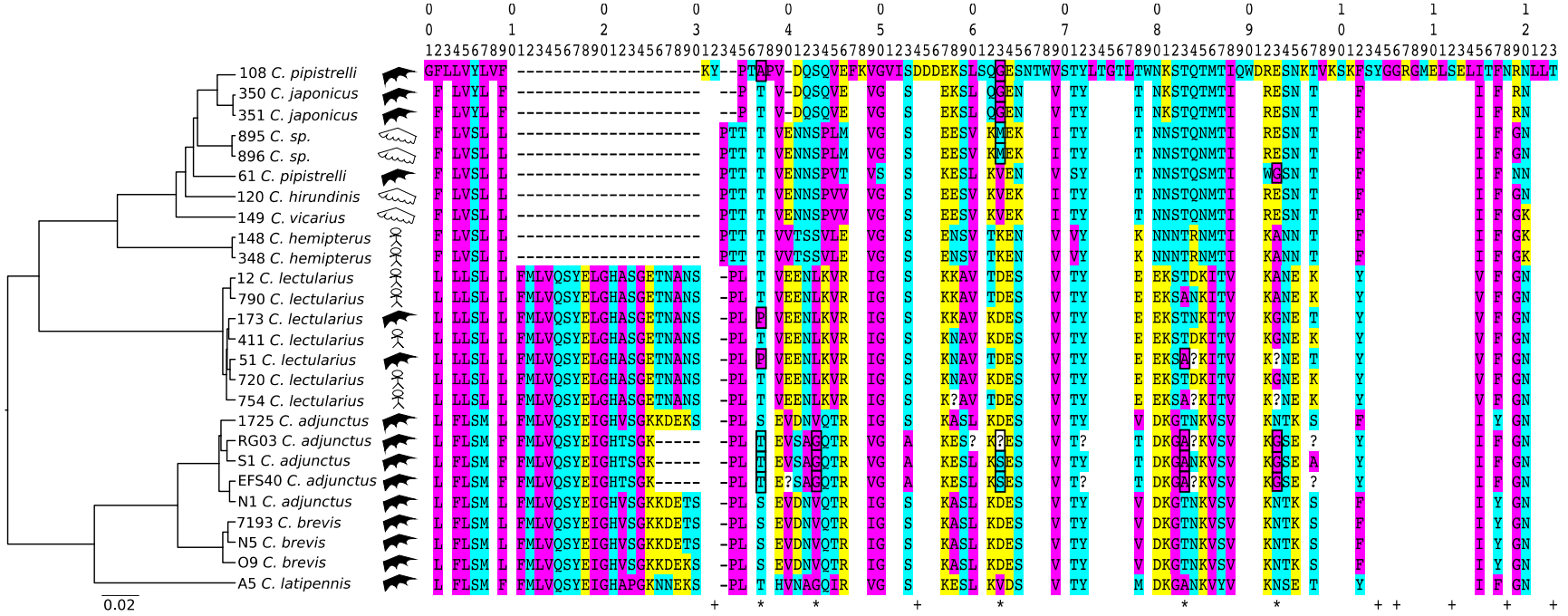

Figure D.4 Translated (from DNA sequence) amino acid sequence of an apyrase gene for each cimicid ectoparasite specimen in this study. Codon numbers (top rows) correspond to those in Table 5.1 and Fig. D.3. Color represents the chemical properties of amino acids (purple: nonpolar, blue: polar uncharged, yellow: charged). Blanks refer to codon positions with no variation, hyphens refer to gaps in the alignments, question marks refer to codons containing heterozygous sites, asterisks refer to codons inferred to be under positive selection, and pluses refer to codons inferred to be under negative selection. A hypothesized species tree, based on *CO1* and *EF1α*, constructed with *BEAST 2.4.2 (scale represents substitutions per site), the specimen number (as in Table D.1), the species name and the host to which it was associated (stick figure for human, black pointed wing for bat and white rounded wing for swallow) of each sample is shown. Black rectangles around codon positions refer to detection of a significant signal of positive selection at the specific terminal node or at an internal node downstream from it, in the hypothesized species tree, using the MEME approach.

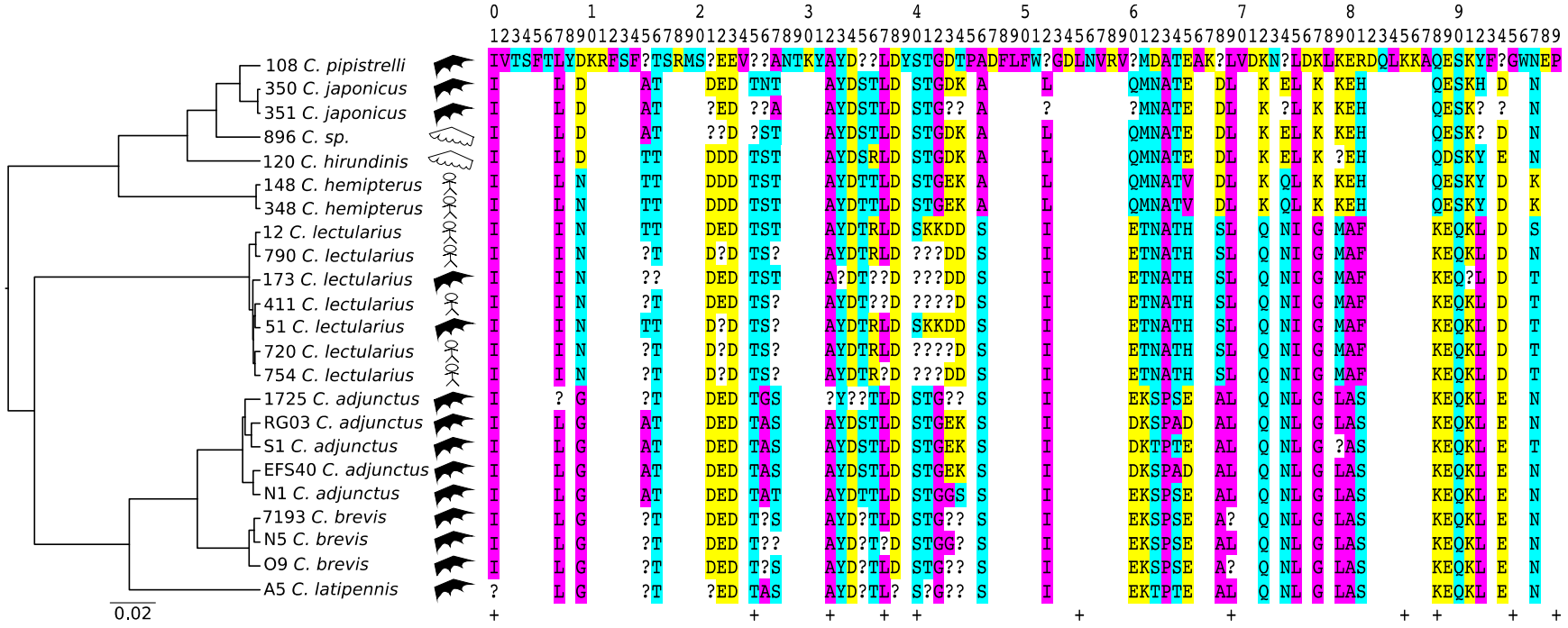

Figure D.5 Translated (from DNA sequence) amino acid sequence of a nitrophorin gene for each cimicid ectoparasite specimen in this study. Codon numbers (top rows) correspond to those in Table 5.1. Color represents the chemical properties of amino acids (purple: nonpolar, blue: polar uncharged, yellow: charged). Blanks refer to codon positions with no variation, hyphens refer to gaps in the alignments, question marks refer to codons containing heterozygous sites, and pluses refer to codons inferred to be under negative selection. A hypothesized species tree, based on *CO1* and *EF1α*, constructed with *BEAST 2.4.2 (scale represents substitutions per site), the specimen number (as in Table D.1), the species name and the host to which it was associated (stick figure for human, black pointed wing for bat and white rounded wing for swallow) of each sample is shown.

Literature Cited

Balvín O, Munclinger P, Kratochvíl L, Vilímová J (2012) Mitochondrial DNA and morphology show independent evolutionary histories of bedbug *Cimex lectularius* (Heteroptera: Cimicidae) on bats and humans. *Parasitology Research*, **111**, 457–469.

Balvín O, Roth S, Vilímová J (2015) Molecular evidence places the swallow bug genus *Oeciacus* Stål within the bat and bed bug genus *Cimex* Linnaeus (Heteroptera: Cimicidae). *Systematic Entomology*, **40**, 652–665.

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.

Appendix E : Chapter 6 Supplementary Material

Table E.1 Information on individual *C. adjunctus* specimens included in our study. ID refers to the individual identifier of each individual. Apyrase and nitrophorin genotypes refer to the amino acid allele number, as in Table E.3, of the two alleles at each individual. Genetic cluster refers to the cluster number identified in Talbot et al. (2016; Chapter 2), based on microsatellite genotypes. Year refers to when the specimen was collected. Latitude and longitude refer to the geographic coordinates of collection of each specimen. Host species refers to the species with which each specimen was found associated (1: *E. fuscus*; 2: *M. lucifugus*; 3: *M. septentrionalis*). Site refers to each unique sampling location. Prior use refers to: 1. if a specimen was used previously in Talbot et al. (2016; Chapter 2) or 2. was collected during this study.

ID	Apyrase genotype	Nitrophorin genotype	Genetic cluster	Year	Latitude	Longitude	Host species	Site	Prior use
431	3/3	3/3	8	2003	44.33346	-65.20493	2	1	1
444	4/4	Not applicable	8	2003	44.33346	-65.20493	3	1	1
941	3/3	3/3	8	2004	44.33346	-65.20493	2	1	1
1459	1/4	20/1	4	2005	46.46317	-63.31047	3	2	1
1460	4/4	1/20	4	2005	46.46317	-63.31047	3	2	1
1725-Cim1	1/1	20/1	4	2005	44.93087	-63.32218	3	2	1
1727.4	4/4	1/19	4	2006	44.93087	-63.32218	3	3	1
1730.2	4/4	1/19	4	2006	44.93087	-63.32218	3	3	1
4823	3/3	21/3	8	2008	44.33346	-65.20493	2	3	1
4839	3/3	3/3	8	2008	44.33346	-65.20493	2	4	1
5573	1/1	20/1	4	2009	45.48071	-63.65522	2	4	1
5574	1/16	20/1	4	2009	45.48071	-63.65522	3	5	1
6366	1/1	1/20	4	2010	45.04187	-63.78792	3	5	1
7532	1/1	1/1	4	2010	45.48071	-63.65522	3	5	1
6476	1/1	1/20	4	2010	45.0627	-63.8592	3	6	1
7045	4/4	1/2	4	2010	45.25271	-62.11658	3	7	1
7259	4/4	1/1	4	2011	52.60055	-60.82865	2	8	1
380-Cim2	1/4	19/1	4	2001	45.6957	-65.00849	3	9	1
471-Cim1	4/4	1/20	8	2003	44.33016	-65.23428	3	10	1
471-Cim2	4/4	1/19	8	2003	44.33016	-65.23428	3	10	1
CCC-45	10/10	6/3	6	2006	40.47685	-78.28799	2	11	1
CCC-4	10/3	3/3	6	2006	40.47685	-78.28799	2	11	1
CCC-50	3/3	3/3	6	2006	40.47685	-78.28799	2	11	1
CCC-Roost-Cim10	10/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim11	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1

CCC-Roost-Cim12	8/9	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim13	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim14	10/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim15	3/10	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim16	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim2	10/10	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim3	9/12	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim4	Not applicable	18/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim5	3/10	6/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim6	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim7	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim8	3/3	3/3	6	2006	40.47685	-78.28799	2	12	1
CCC-Roost-Cim9	3/10	9/3	6	2006	40.47685	-78.28799	2	12	1
EF-293-Roost	13/14	3/3	2	2005	42.28908	-85.41092	1	13	1
EFS40-10	11/11	3/3	2	2005	42.29483	-85.42367	1	14	1
EFS40-12	Not applicable	4/4	2	2005	42.29483	-85.42367	1	14	1
EFS40-14	Not applicable	3/4	2	2005	42.29483	-85.42367	1	14	1
EFS40-2	11/11	10/3	2	2005	42.29483	-85.42367	1	14	1
EFS40-4	5/11	11/3	2	2005	42.29483	-85.42367	1	14	1
EFS40-7	13/10	4/4	2	2005	42.29483	-85.42367	1	14	1
EFS42-39	3/10	4/4	2	2005	42.14528	-85.35503	1	15	1
EFS42-43	10/10	4/4	2	2005	42.14528	-85.35503	1	15	1
EFS42-Roost-Cim2	10/10	4/4	2	2005	42.14528	-85.35503	1	16	1
EFS42-Roost-Cim3	10/10	7/12	2	2005	42.14528	-85.35503	1	16	1
EFS42-Roost-Cim4	2/2	Not applicable	2	2005	42.14528	-85.35503	1	16	1
EFS42-Roost-Cim6	Not applicable	4/4	2	2005	42.14528	-85.35503	1	16	1
LW-E01	10/10	3/3	10	2005	41.89061	-85.34915	1	17	1
LW-E02	10/10	3/3	10	2005	41.89061	-85.34915	1	17	1
LW-E03	13/16	13/13	2	2005	42.29012	-85.79005	1	18	1
2006-MV-17	10/10	4/4	2	2006	42.59428	-85.98747	1	19	1
MV496-Cim1	3/10	14/4	9	2010	41.50544	-86.2773	1	20	1
MV496-Cim2	10/10	15/16	9	2010	41.50544	-86.2773	1	20	1
MV642	10/3	3/13	9	2010	38.32762	-84.2291	1	21	1
MV652-Cim1	2/3	Not applicable	9	2010	38.32762	-84.2291	1	21	1
N1	4/4	1/1	4	2014	60.09333	-112.24732	2	22	1
N3	4/4	1/1	4	2014	60.09333	-112.24732	2	22	1
N4	4/4	1/1	4	2014	60.09333	-112.24732	2	22	1
O1	5/5	Not applicable	10	2014	42.211091	-82.9265	1	23	1
O2	5/3	3/3	10	2014	42.211091	-82.9265	1	23	1
O3	3/7	3/3	10	2014	42.211091	-82.9265	1	23	1
O4	5/5	5/3	10	2014	42.211091	-82.9265	1	23	1

O5	5/3	5/6	10	2014	42.211091	-82.9265	1	23	1
O6	6/5	4/3	10	2014	42.211091	-82.9265	1	23	1
O7	5/5	3/7	10	2014	42.21109	-82.9265	1	23	1
O8	3/2	4/3	6	2014	42.577544	-80.377328	1	24	1
O13-Cim1	3/3	3/3	Not applicable	2015	42.6551	-80.46961	1	25	2
O13-Cim2	3/3	3/3	Not applicable	2015	42.6551	-80.46961	1	25	2
O13-Cim3	3/3	3/3	Not applicable	2015	42.6551	-80.46961	1	25	2
RG-01Cim1	Not applicable	4/4	7	2005	40.53241	-105.01655	1	26	1
RG-03Cim1	9/3	Not applicable	7	2005	40.53241	-105.01655	1	27	1
RG-03Cim2	3/8	Not applicable	7	2005	40.53241	-105.01655	1	27	1
RG-03Cim5	3/9	4/4	7	2005	40.53241	-105.01655	1	27	1
RG-03Cim7	3/3	4/4	7	2005	40.53241	-105.01655	1	27	1
RG-03Cim8	3/3	4/4	7	2005	40.53241	-105.01655	1	27	1
RG-04Cim1	9/12	4/4	7	2005	40.54548	-105.0586	1	28	1
RG-06Cim1	3/3	4/4	7	2005	40.52294	-105.01799	1	29	1
RG-06Cim3	Not applicable	4/4	7	2005	40.52294	-105.01799	1	29	1
RG-07Cim10	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim11	Not applicable	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim2	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim4	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim5	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim7	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-07Cim9	3/3	4/4	7	2005	40.5287	-105.00829	1	30	1
RG-08Cim1	Not applicable	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-08Cim11	9/9	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-08Cim14	9/9	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-08Cim15	9/9	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-08Cim2	9/9	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-08Cim6	9/9	4/4	7	2005	40.57805	-105.07017	1	31	1
RG-09Cim2	9/3	4/4	7	2005	40.59204	-105.08571	1	31	1
RG-09Cim3	3/3	4/4	7	2005	40.59204	-105.08571	1	32	1
RG-09Cim4	3/9	4/4	7	2005	40.59204	-105.08571	1	32	1
RG-09Cim5	3/15	4/4	7	2005	40.59204	-105.08571	1	32	1
RG-09Cim7	9/9	Not applicable	7	2005	40.59204	-105.08571	1	32	1
RG-09Cim8	3/8	4/4	Not applicable	2005	40.59204	-105.08571	1	32	1
MV194-Cim1	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
MV194-Cim2	2/3	3/3	3	2008	42.99915	-83.1777	1	33	1
MV205-Cim1	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
RH01	3/3	4/4	3	2008	42.99915	-83.1777	1	33	1
RH02	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
RH03	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1

RH04	3/3	4/3	3	2008	42.99915	-83.1777	1	33	1
RH05	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
RH09	3/17	18/3	3	2008	42.99915	-83.1777	1	33	1
RH11	5/3	3/3	3	2008	42.99915	-83.1777	1	33	1
RH12	3/3	4/3	3	2008	42.99915	-83.1777	1	33	1
RH13	3/3	8/3	3	2008	42.99915	-83.1777	1	33	1
RH14	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
RH15	3/3	3/3	3	2008	42.99915	-83.1777	1	33	1
S1-Cim1	Not applicable	3/9	9	2014	50.79072	-103.97047	2	34	1
S1-Cim2	Not applicable	3/3	9	2014	50.79072	-103.97047	2	34	1
S2	17/17	Not applicable	Not applicable	2015	49.5718	-109.8787	2	35	2
S7	17/17	16/17	Not applicable	2015	49.5718	-109.8787	2	35	2
11477	3/3	3/3	5	2012	52.60568	-61.16388	2	36	1
11494	3/3	3/3	5	2012	52.60568	-61.16388	2	36	1
10186	3/3	3/3	5	2013	47.26388	-53.28183	2	37	1
10030	3/3	3/3	5	2012	49.09374	-57.53579	2	38	1

Table E.2 Information on the amplification primers used in our study.

Gene	Length of fragment	Name of primer	Sequence from 5' to 3'	Annealing Temperature
Apyrase	369 bp	CadjA1F	F: [ACACTGACGACATGGTTCTACA][1]GTCCAGCCGACGAATGTAGT	55°
		CadjA1R	R: [TACGGTAGCAGAGACTTGGTCT][2]CCAGTCCTGTTCTGTCATCG	
Nitrophorin	297 bp	N2F	F: [ACACTGACGACATGGTTCTACA][1]CGATCAGAAGAAATCAGGCGG	53°
		N2R3R	R: [TACGGTAGCAGAGACTTGGTCT][2]AGGTAGGCTTGAAGGTGACC	

[1]CS1 Fluidigm universal tag (Fluidigm Corporation, South San Francisco, California, United States)
[2]CS2 Fluidigm universal tag (Fluidigm Corporation, South San Francisco, California, United States)

Table E.3 Unique apyrase and nitrophorin sequences obtained in our study. DNA sequence number refers to the number attributed to each unique DNA sequence. We provide the Genbank accession number for each unique DNA sequence. Amino acid sequence number refers to the number attributed to each unique converted amino acid sequence using the Universal Genetic Code. Allelic divergence represents the sum of pairwise substitutions between each allele (the first time it appears in the table) is and all other alleles.

Gene	DNA sequence number	Genbank accession number	Amino acid sequence number	Allelic divergence
Apyrase	1	KY561758	1	297
	2	KY561759	2	100
	3	KY561760	3	77
	4	KY561761	4	297
	5	KY561762	5	90
	6	KY561763	6	87
	7	KY561764	7	79
	8	KY561765	8	83
	9	KY561766	9	91
	10	KY561767	10	87
	11	KY561768	3	
	12	KY561769	11	90
	13	KY561770	12	78
	14	KY561771	13	88
	15	KY561772	10	
	16	KY561773	13	
	17	KY561774	14	105
	18	KY561775	15	80
	19	KY561776	10	
	20	KY561777	13	
	21	KY561778	16	88
	22	KY561779	4	

	23	KY561780	1	
	24	KY561781	16	
	25	KY561782	17	89
	26	KY561783	17	
	1	KY561784	1	161
	2	KY561785	2	151
	3	KY561786	1	
	4	KY561787	3	108
	5	KY561788	4	93
	6	KY561789	1	
	7	KY561790	1	
	8	KY561791	3	
	9	KY561792	5	102
	10	KY561793	3	
	11	KY561794	6	97
	12	KY561795	4	
	13	KY561796	7	88
Nitrophorin	14	KY561797	8	96
	15	KY561798	9	99
	16	KY561799	10	111
	17	KY561800	11	126
	18	KY561801	12	113
	19	KY561802	13	97
	20	KY561803	14	133
	21	KY561804	15	99
	22	KY561805	16	114
	23	KY561806	16	
	24	KY561807	17	179
	25	KY561808	18	127
	26	KY561809	6	
	27	KY561810	19	192

	28	KY561811	20	202
	29	KY561812	21	124

Literature Cited

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.

VITA

Name:	Benoit Talbot
Post-secondary Education and Degrees:	Université de Montréal Montréal, Québec, Canada 2007-2010 B.Sc. Université de Sherbrooke Sherbrooke, Québec, Canada 2010-2012 M.Sc. University of Western Ontario London, Ontario, Canada 2013-2017 Ph.D.
Honours and Awards:	Schmeelk Canada Fellowship 2016-2017 Ontario Graduate Scholarship 2016-2017 Ruth Horner Arnold Fellowship 2015 National Science and Engineering Research Council of Canada scholarships 2008, 2009, 2010-2011, 2013-2016 Fonds du Québec de Recherche en Nature et Technologies scholarship 2011-2012
Related Work Experience	Teaching Assistant University of Western Ontario 2014-2016 Teaching Assistant Université de Sherbrooke 2011-2012

Publications:

Talbot B, Balvin O, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2017) Host association and selection on salivary protein genes in bed bugs and related blood-feeding ectoparasites. *Royal Society Open Science*, **In Press**.

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2017) Comparative analysis of landscape effects on spatial genetic structure of the big brown bat and one of its cimicid ectoparasites. *Ecology and Evolution*, **In Press**.

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2017) Population structure in two geographically sympatric and congeneric ectoparasites: *Cimex adjunctus* and *Cimex lectularius*, in the North American Great Lakes region. *Canadian Journal of Zoology*, **In Press**.

Talbot B, Chen TW, Zimmerman S, Joost S, Eckert AJ, Crow TM, Semizer-Cuming D, Manel S (2016) Combining genotype, phenotype and environment to infer potential candidate genes. *Journal of Heredity*, **108**, 207-216.

Talbot B, Vonhof MJ, Broders HG, Fenton B, Keyghobadi N (2016) Range-wide genetic structure and demographic history in the bat ectoparasite *Cimex adjunctus*. *BMC Evolutionary Biology*, **16**.

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

Talbot B, Garant D, Rioux-Paquette S, Mainguy J, Pelletier F (2014) Genetic structure and diversity of rabid raccoons (*Procyon lotor*). *Écoscience*, **20**, 345-351.

Talbot B, Garant D, Rioux-Paquette S, Mainguy J, Pelletier F (2012) Lack of genetic structure and female-specific effect of dispersal barriers in a rabies vector, the striped skunk (*Mephitis mephitis*). *Plos One*, **7**, 1-7.