THESIS

SPATIAL, DEMOGRAPHIC, AND PHYLOGENETIC PATTERNS OF *BARTONELLA* DIVERSITY IN BATS

Submitted by

Clifton Dyer McKee

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2015

Master's Committee

Advisor: Colleen T. Webb

Michael Y. Kosoy
W. Chris Funk
Tony Schountz
David T. S. Hayman

ABSTRACT

SPATIAL, DEMOGRAPHIC, AND PHYLOGENETIC PATTERNS OF *BARTONELLA* DIVERSITY IN BATS

Much recent attention has focused on bats as potentially exceptional reservoirs of pathogens. Bats are known to carry zoonotic viruses deadly to humans with no apparent signs of pathology, however the evolutionary and physiological processes that are behind this ability remain largely unknown. Despite this uncertainty, bats' long lifespans, deep evolutionary history, sociality, and migratory behavior make them a fascinating system in which to study patterns of diversity in viruses, bacteria, and other infectious organisms. This thesis explores ecological and evolutionary processes that structure the diversity of infectious bacteria in bats. I focus on *Bartonella*, a genus of vector-borne intracellular bacteria, because of its high prevalence and genetic diversity within bats. I examined the structure of *Bartonella* species assemblages in *Eidolon* spp. fruit bats across Africa and Madagascar using newly developed molecular and statistical tools. The results from this examination indicate that fruit bats from distant geographic locations host similar communities of *Bartonella*; I attribute this to widespread dispersal and communal roosting behavior in *Eidolon* spp. bats. To understand how *Bartonella* diversity has evolved and is structured geographically, I assembled a global dataset of *Bartonella* genotypes from bats and their ectoparasites. Using this dataset, I analyzed the contributions of cospeciation and sympatry among host species to the diversity of *Bartonella* in bats. Continued development of this research could provide a model system for the study of ecological and evolutionary processes contributing to pathogen diversification and infection dynamics in natural systems.

ACKNOWLEDGMENTS

I am deeply indebted to a number of people for their guidance and support during my work on this thesis. I am grateful for the training and mentorship that Michael Kosoy, Ying Bai, and Lynn Osikowicz have given me in the lab at the Centers for Disease Control (CDC) Division of Vector-Borne Diseases (DVBD). This project would have been impossible without their patience and direction. I sincerely thank David Hayman for his continued insights on bats, bacterial evolution, and everything in between. I have been fortunate to work with such a dedicated and passionate scientist. I appreciate the number of collaborators that have contributed samples to this work including Cara Brook, Ivan Kuzmin, Alison Peel, and Richard Suu-Ire. I owe enormous praise to my advisor Colleen Webb, who has helped me to advance as a scientist and academic, and has greatly influenced the scope and purpose of this thesis. I also thank additional members of my committee, the Webb and Kosoy labs, and other collaborators that have provided valuable input during the development of this work.

I must also give thanks to numerous people cheering for me along the way. To my family, I am forever grateful for your love and advice throughout my academic career; I would not be where I am now without you. To my friends, I am abundantly grateful for your encouragement and the perspective achieved from a shared glass of beer, a long run, and some time away from the computer screen. Finally, I am lucky to have the support of my partner, Ava Hoffman, who has been my councilor, editor, and exuberant cheerleader throughout this work.

Materials and laboratories facilities for experiments were provided by the CDC DVBD Bacterial Diseases Branch. This work was partially supported by the Wellcome Trust and the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and

Technology Directorate (US Department of Homeland Security) and the Fogarty International Center (NIH).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Generation and maintenance of *Bartonella* diversity in bats

Investigating the mechanisms that generate and maintain infection diversity can help to advance biodiversity theory in several ways. Microparasites like viruses and bacteria have rapid lifecycles that facilitate the measurement of ecological and evolutionary processes *in situ*. Furthermore, communities of parasites within host individuals are connected in a wider metacommunity framework through interactions and transmission among individuals. Diverse parasites can form communities inside host individuals that are subject to many of the same assembly processes as macro-scale ecosystems (Seabloom *et al.* 2015). Thus, reservoir hosts are true microcosms that offer excellent opportunities for testing the separate processes that generate and maintain biodiversity: ecological drift, selection, dispersal, and speciation (Vellend 2010).

Nevertheless, the measurement and characterization of microparasite communities in a host individual is largely concerned with unseen processes, with considerable uncertainty surrounding the density of parasites, interactions with the host organism, and interactions among parasites within the community (Telfer *et al*. 2010). To deal with this uncertainty, researchers must develop new methods to answer specific questions about how microparasite communities are structured. These methods will necessarily draw from many disciplines, from microbiology and molecular genetics to computer science and statistics. For some bacteria, the need for novel methods to measure coinfecting communities is pressing.

For *Bartonella* bacteria, there is some evidence that individual animals can carry a diverse assemblage of these parasites simultaneously (Kosoy *et al.* 2004; Abbot *et al.* 2007) and

that the resulting communities may drive the generation and maintainence of *Bartonella* diversity through interactions with the host immune system (Chan and Kosoy 2010), interactions among coinfecting bartonellae, and molecular evolution (Arvand *et al.* 2007; Berglund *et al.* 2010; Guy *et al.* 2012; Paziewska *et al.* 2011, 2012; Buffet *et al.* 2013). Yet the measurement of these phenomena are constrained by the fact that bartonella are fastidious bacteria that are notoriously difficult to detect or culture from an individual (Harms and Dehio 2010). Hence, I develop novel techniques to characterize the diversity of *Bartonella* species in samples from blood samples. I then compare the structure of these bacterial communities across geographically distant bat populations and distinct demographic groups. These tests allow me to make inference about the migration process of bacteria among individuals that can maintain *Bartonella* diversity patterns.

Additionally, *Bartonella* appear to have strong ecological associations with their hosts, with particular *Bartonella* species associated with restricted groups or species of mammals (Kosoy 2010). The diversity of *Bartonella* species may simply reflect random processes of isolation and divergence within host populations. Alternatively, the expansion of *Bartonella* diversity may be a directed process of cospeciation, wherein the bacteria develop specific adaptations to invade and persist in a new host species. When this latter process is extended over evolutionary time, the phylogenetic trees representing the diversity of host and parasite species may begin to mirror one another (Page 1994). From these patterns, one can infer the extent to which parasite and host speciation have been linked. However, testing this kind of relationship requires extensive sampling of host and parasite population diversity. To this end, *Bartonella* diversity has been especially well characterized in several taxa, particularly rodents and bats.

To understand the spatial, demographic, and phylogenetic mechanisms generating and maintaining *Bartonella* diversity, I focus on bats. Bats have several characteristics that make

them an ideal system for studying parasites communities. Bats have long lifespans and possibly have an exceptional ability to carry infections without signs of disease (Calisher *et al.* 2006; Brook and Dobson 2015). For some species of bats, long-distance dispersal and communal roosting provide a scenario that can support connectivity among metacommunities of parasites and would theoretically maintain *Bartonella* biodiversity across large geographic scales (Peel *et al.* 2013). In Chapter 2, I test this hypothesis using newly developed molecular methods that measure the relative abundance of coinfecting *Bartonella* species in a widely dispersed bat genus. Bats have been evolving for millions of years, producing the second-most diverse group of mammals on Earth next to rodents. If *Bartonella* have tracked the speciation of their bat reservoirs, then a pattern of cophylogeny should be apparent in molecular data. Furthermore, bats are globally distributed, which allows me to test the contribution of geographic isolation on divergence among bat species and their associated *Bartonella* bacteria. In Chapter 3, I analyze these cophylogenetic patterns in a global sampling of *Bartonella* from bats and their ectoparasites.

Although I test the spatial, demographic, and phylogenetic patterns of *Bartonella* diversity in bats, the methods from these studies are applicable to other systems. Characterization of bacterial communities may help to infer connectivity among populations of hosts, providing valuable insight about host and bacterial life history. Investigation of evolutionary relationships among host species and bacteria can reveal important insights on how microparasites adapt and cause disease in new mammal species. Finally, parasites and hosts represent microcosms that are invaluable systems to test fundamental ecological and evolutionary principles. Chapter 4 will comment on these connections and suggest some directions for future research.

REFERENCES

Abbot P, Aviles AE, Eller L, Durden LA (2007) Mixed infections, cryptic diversity, and vector-borne pathogens: evidence from *Polygenis* fleas and *Bartonella* species. Applied and Environmental Microbiology, 73, 6045–6052.

Arvand M, Feil EJ, Giladi M, Boulouis H-J, Viezens J (2007) Multi-locus sequence typing of *Bartonella henselae* isolates from three continents reveals hypervirulent and feline-associated clones. PLoS ONE, 2, e1346.

Berglund EC, Ellegaard K, Granberg F *et al.* (2010) Rapid diversification by recombination in *Bartonella grahamii* from wild rodents in Asia contrasts with low levels of genomic divergence in Northern Europe and America. Molecular Ecology, 19, 2241–2255.

Brook CE, Dobson AP (2015) Bats as "special" reservoirs for emerging zoonotic pathogens. Trends in Microbiology, 23, 172–180.

Buffet J-P, Pisanu B, Brisse S *et al.* (2013) Deciphering bartonella diversity, recombination, and host specificity in a rodent community. PLoS ONE, 8, e68956.

Calisher CH, Childs JE, Field HE, Holmes K V, Schountz T (2006) Bats: important reservoir hosts of emerging viruses. Clinical Microbiology Reviews, 19, 531–545.

Chan K-S, Kosoy MY (2010) Analysis of multi-strain *Bartonella* pathogens in natural host population – do they behave as species or minor genetic variants? Epidemics, 2, 165–72.

Guy L, Nystedt B, Sun Y *et al.* (2012) A genome-wide study of recombination rate variation in *Bartonella henselae*. BMC Evolutionary Biology, 12, 65.

Harms A, Dehio C (2012) Intruders below the radar: molecular pathogenesis of *Bartonella* spp. Clinical Microbiology Reviews, 25, 42–78.

Kosoy MY, Mandel E, Green D *et al.* (2004) Prospective studies of *Bartonella* of rodents. Part II. Diverse infections in a single rodent community. Vector-Borne and Zoonotic Diseases, 4, 296–305.

Page RD (1994) Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladistics, 10, 155–173.

Paziewska A, Harris PD, Zwolińska L, Bajer A, Siński E (2011) Recombination within and between species of the alpha proteobacterium *Bartonella* infecting rodents. Microbial Ecology, 61, 134–145.

Paziewska A, Siński E, Harris PD (2012) Recombination, diversity and allele sharing of infectivity proteins between *Bartonella* species from rodents. Microbial Ecology, 64, 525–536.

Peel AJ, Sargan DR, Baker KS *et al.* (2013) Continent-wide panmixia of an African fruit bat facilitates transmission of potentially zoonotic viruses. Nature Communications, 4, 2770.

Seabloom EW, Borer ET, Gross K *et al.* (2015) The community ecology of pathogens: coinfection, coexistence and community composition. Ecology Letters, 18, 401-415.

Telfer S, Lambin X, Birtles RJ *et al.* (2010) Species interactions in a parasite community drive infection risk in a wildlife population. Science, 330, 243–246.

Vellend M (2010) Conceptual synthesis in community ecology. The Quarterly Review of Biology, 85, 183–206.

CHAPTER 2

Phylogeography of *Bartonella* bacteria in *Eidolon* spp. fruit bats across Africa

Introduction

Detecting structure in well-mixed, migratory animal populations with molecular tools can be challenging, yet previous research has demonstrated that using genetic data from animal parasites can provide greater resolution for revealing cryptic population structure in hosts. Examples come from human ecology (Falush *et al.* 2003; Holmes 2004; Wirth *et al.* 2005) and several notable studies of wildlife populations (Nieberding *et al.* 2004; Biek *et al.* 2006; Criscione *et al.* 2006). In general, these studies have focused on genetic variants of microparasites (viruses and bacteria) found in host individuals and infer migrations based on the clustering of related parasite genotypes within geographic regions. One limitation to this approach is that not all parasites evolve rapidly enough to detect geographic clusters of isolates. Additionally, each individual host can carry multiple parasite genotypes or species. The presence and relative abundance of parasite species may not be independent of one another depending on transmission processes. However, the structure of parasite assemblages can also be informative on their own, potentially reflecting recent changes in transmission among host individuals linked to a shift in behavior. If these new transmission patterns persist, they can serve as a proxy for incipient population structure. I demonstrate an approach to detecting recent changes in transmission by analyzing the structure of species assemblages of cryptic and diverse infectious bacteria (*Bartonella* spp.) in a widely dispersed group of fruit bats (*Eidolon* spp.) in Africa and

Madagascar. For reasons outlined below, I believe bats and bartonella are an ideal system for this type of analysis.

Firstly, bartonellae are diverse and appear to cause persistent, nonpathological infections in host individuals (Harms and Dehio 2012) which may facilitate the formation of measurable communities of *Bartonella* species. Bartonellae are Gram-negative alpha-proteobacteria that primarily infect and persist in erythrocytes and endothelial cells of a wide variety of mammals globally, including rodents, bats, insectivores, carnivores, ungulates, and marine mammals (Kosoy 2010). Transmission of bartonellae between individuals is thought to occur primarily via blood-feeding fleas, ticks, lice, flies, and mites (Billeter *et al.* 2008). Over 30 species in the genus *Bartonella* have been formally described to date, and roughly half have been identified as emerging zoonotic pathogens in humans and domestic animals, causing a wide range of illnesses from mild, self-limiting fever to potentially fatal endocarditis (Chomel and Kasten 2010; Breitschwerdt *et al.* 2010). Furthermore, ongoing work has revealed interesting patterns of coinfection (Abbot *et al.* 2007; Chan and Kosoy 2010) and horizontal gene transfer (Arvand *et al.* 2007; Berglund *et al.* 2010; Guy *et al.* 2012; Paziewska *et al.* 2011, 2012; Buffet *et al.* 2013b) that likely play key roles in bartonella diversification.

Secondly, the abundance and diversity of bartonella can vary widely among mammal species, potentially reflecting differences in transmission and host specificity. Of the potential zoonotic reservoirs, rodents and bats appear to harbor the greatest diversity of *Bartonella* species (Bai and Kosoy 2012; Buffet *et al.* 2013a; Gutiérrez *et al.* 2015). Recent studies have identified more than 20 putative new *Bartonella* species in diverse bat communities from Europe, the Americas, Africa, and Southeast Asia, with several bat species hosting especially diverse and abundant bartonellae (Concannon *et al.* 2005; Kosoy *et al.* 2010a; Bai *et al.* 2011; Bai *et al.*

2012; Lin *et al* 2012; Veikkolainen *et al.* 2014; Kamani *et al.* 2014; Olival *et al.* 2014; Bai *et al.* 2015). Other studies have identified *Bartonella* species in a variety of bat ectoparasites (Loftis *et al.* 2005; Reeves *et al.* 2005; Reeves *et al* 2006; Billeter *et al.* 2012; Morse *et al.* 2012; Kamani *et al.* 2014; Reeves *et al.* 2007; Veikkolainen *et al.* 2014). Genetic evidence suggests that bats and their associated *Bartonella* species have co-diverged (Lei and Olival 2014), however sharing of *Bartonella* strains among bat species is likely, especially for species that share roosts or ectoparasites (Bai *et al.* 2011; Kamani *et al.* 2014). The diversity of bartonella in bats alone is interesting, but investigating the causes of variation in the diversity of bartonella among host individuals can make stronger connections to community ecology and bacterial evolution. Since infectious agents are dependent on their hosts and vectors for transmission, variation in parasite diversity among host populations may be linked to specific life history traits. Bats in particular show variation in traits such as sympatry, migration, longevity, litter size, body mass, and the use of torpor that can have strong effects on their viral diversity (Luis *et al.* 2013).

In this study, I focus on bartonella infections in widely distributed African fruit bats of the genus *Eidolon* (Pteropodidae). Compared to other bat species, bartonellae infecting *Eidolon* spp. fruit bats have been well characterized, providing a structured framework by which to measure communities of *Bartonella*. Furthermore, *Eidolon* spp. have interesting ecological traits that may influence their bacterial diversity. The straw-colored fruit bat (*E. helvum*) is the most widely distributed bat species in sub-Saharan Africa (Mickleburgh *et al.* 2008). In addition to the mainland population, there are additional, smaller populations on several islands in the Gulf of Guinea (Juste *et al.* 2000). Genetic and stable isotope evidence indicates that the straw-colored fruit bat (*Eidolon helvum*) is migratory and panmictic across continental Africa (Peel *et al.* 2010; Ossa *et al.* 2012; Peel *et al.* 2013); however, bats on the island of Annobón have been

determined to represent a genetically and morphologically distinct population (Juste *et al.* 2000). This level of panmixia is the highest recorded for any mammal species and may facilitate transmission of infections across the bat's geographic range (Peel *et al.* 2013). The closely related Madagascan fruit bat (*E. dupreanum*) is endemic and widespread on the island, although its distribution is patchy. Regular variation in the occupancy and abundance of roosts suggests that *E. dupreanum* is migratory, and recent molecular evidence shows that this species is also panmictic across Madagascar (MacKinnon *et al.* 2003; Shi *et al.* 2014). Both species frequently roost in large colonies which may facilitate transmission of infections (Sørensen and Halberg 2001; MacKinnon *et al*. 2003).

Previous work has demonstrated that *E. helvum* populations in several locations across Africa and *E. dupreanum* on Madagascar can host six *Bartonella* lineages, formally identified as distinct species by characterizing sequence divergence at eight neutral genetic loci (Kosoy *et al.* 2010a; Kamani *et al.* 2014; Bai *et al.* 2015; Brook *et al.* 2015). Bat flies (*Cyclopodia greeffi*) parasitizing these bats appear to carry these same strains and a variety of other sequence variants that may represent yet uncultured or fully characterized *Bartonella* species (Billeter *et al.* 2012; Kamani *et al.* 2014; Brook *et al.* 2015). Furthermore, evidence of homologous recombination among *Bartonella* species has been documented in cultured isolates from bats, suggesting that these *Bartonella* species may coexist at some point during the infection cycle (Bai *et al.* 2015). In a scenario where coexistence of *Bartonella* species is possible, traditional methods of molecular detection and sequencing are not sufficient.

Thus, I develop a multi-gene PCR detection system for the identification of diverse *Bartonella* infections along with statistical models that can quantify the relative abundance of potentially coinfecting *Bartonella* species in animal blood. I utilize these techniques to study the

phylogeography of *Bartonella* species in *Eidolon* fruit bats across Africa. I hypothesize that high population connectivity and massive roosting behavior in *Eidolon* spp. fruit bats will contribute to the transmission and maintenance of diversity in *Bartonella* species across Africa, with all locations showing similar *Bartonella* species assemblages. However, transmission of bacteria and genetic differentiation of host populations occur on separate ecological and evolutionary timescales, respectively, which may alter the distribution of *Bartonella* species and reveal population structure not seen in genetic data from *Eidolon* spp. This pattern may be especially strong on isolated islands such as Annobón and Madagascar, where there are genetically distinct populations (a separate species in the case of Madagascar) of *Eidolon* bats. I also test the hypothesis that demographic factors like age class, sex, and pregnancy status may predict prevalence of infection and *Bartonella* species abundances based on previous research on differences in Hendra virus antibody prevalence in pregnant and lactating females and different age classes of flying foxes (Plowright *et al.* 2008; Breed *et al.* 2011). The molecular and statistical techniques used in this study provide new tools for future studies of *Bartonella* infections in natural populations and contribute to our knowledge of the relationship between reservoirs, particularly bats, and the ecology and evolution of their infections.

Materials and methods

Study system

A source of complication for studying the distribution of *Bartonella* species in wild mammals is that bartonella infections are often difficult to detect. Sterile culturing has been the gold standard for bartonella detection and characterization in the past. However, bartonellae are extremely difficult to culture due to their strict nutritional requirements, slow growth rate, and tendency to be overgrown by other bacteria and fungi. Furthermore, culturing without serial

dilutions frequently isolates only one *Bartonella* strain. As a result, researchers have turned to molecular methods using PCR and sequencing to identify bartonella DNA in blood and tissues rapidly. With the ability to detect low levels of bacteremia, researchers can make better measurements of bartonella prevalence over time and in different populations.

However, molecular methods bring with them their own complications. Repeated testing can reveal conflicts in sequences from the same individual that represent potential coinfection by multiple strains (Kosoy *et al.* 2004b; Abbot *et al.* 2007). Furthermore, studies using multi-locus sequence typing and genome sequencing have shown that homologous recombination of protein-coding loci is common among co-circulating *Bartonella* species (Paziewska *et al.* 2011, 2012; Buffet *et al.* 2013b; Bai *et al.* 2015) and that detection using only one PCR marker is insufficient for accurately identifying an infection. Oligonucleotide primers used in PCR can have differing sensitivities for particular *Bartonella* species and may isolate a separate species in independent amplification and sequencing replicates. Finally, one set of primers may not be capable of detecting all species in a diverse assemblage, so the use of multiple primers may reveal cryptic *Bartonella* species. Hence, I attempt to address an unmet need for extended molecular methods that use replicate testing and multiple primers to measure the diversity of coinfecting *Bartonella* species in a sample, along with statistical methods to integrate these data into relative species abundances and compare species distributions across geographic and demographic factors.

Sample collection

Blood samples were taken from bats in different regions by separate teams for studying a variety of infectious agents. As such, the collection methods and accessory data about sampled bats varied among teams. Generally, sampling sites were chosen based on known bat roosts and nocturnal foraging sites. Sampling sites are mapped in Figure 2.1. Bats were captured using mist

nets and hand nets around caves or nocturnal foraging sites, then under manual restraint about 0.2-1 mL of blood was collected in one of two ways – from the propatagial vein using a citrated 1 mL syringe or via exsanguination by cardiac puncture following anesthesia by injection of ketamine chloride (0.05-0.1 m/g body weight) – then placed in a sterile collection tube.

Whole blood was immediately frozen at -80 ºC or blood clots were separated from serum then frozen at -80 ºC. Frozen samples were shipped to the Centers for Disease Control and Prevention Division of Vector-Borne Diseases in Fort Collins, Colorado on dry ice where they were kept at -20 ºC or below until DNA extraction. Further sampling details can be found in Kuzmin *et al.* (2008), Hayman *et al.* (2008), Kosoy *et al.* (2010a), Peel *et al.* (2012), and Brook *et al.* (2015).

For bats collected in Ghana, Madagascar, and islands in the Gulf Guinea (Bioko, São Tomé, Príncipe, and Annobón), weight, wing length, age class, sex, and pregnancy status of each bat was determined. For samples from Kenya and Nigeria, only weight, wing length, and sex of bats were recorded. Ageing bats beyond rough demographic classes is only possible by examining tooth cementum annuli (Hayman *et al.* 2012). Total sample sizes tested for each location were 22, 21, 29, 29, 47, 63, 22, and 23 for Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé, respectively.

DNA extraction and PCR amplification

Genomic DNA was extracted from each blood sample using the QIAamp DNA Blood Mini Kit and protocol (Qiagen, Valencia, CA). The concentration and quality of DNA in each sample was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE). Molecular detection of *Bartonella* spp. in blood samples was performed by amplifying partial fragments of six genetic loci via PCR: cell division protein gene (*ftsZ*), citrate

synthase gene (*gltA*), 16S-23S intergenic spacer region (ITS), NADH dehydrogenase gamma-subunit gene (*nuoG*), RNA polymerase beta-subunit gene (*rpoB*), and transfer messenger RNA gene (*ssrA*). Each of these loci has been used previously for the differentiation of *Bartonella* species (Table A1.1). PCR amplifications were performed in 25 μL reaction mixtures containing 12.5 μL of 2x GoTaq Green Master Mix (Promega, Madison, WI) including reaction buffer and 400 μM of each dNTP, 1 μL (0.4 μM) of each oligonucleotide primer, 5.5 μL of nuclease free water, and 5 μL (150 ng) of sample DNA template.

Single-round PCR amplification was sufficient for the isolation of ITS, *nuoG*, and *ssrA* sequences from positive wells. However, amplification of *ftsZ*, *gltA*, and *rpoB* sequences was low from one round of PCR, frequently yielding negative results for wells that were positive by ITS, *nuoG*, and *ssrA*. Thus, I utilized primers internal to the first round primers and performed nested PCR reactions to amplify *Bartonella* DNA further. Nested PCR amplifications were performed using similar reaction mixtures to the first round, but including 5 μL of the first round PCR product as the template for the nested reaction. Wells containing positive (*Bartonella doshiae*) and negative (nuclease-free water only) controls were included in all amplifications. Reaction mixtures were placed in a PTC 200 Peltier thermocyler (MJ Research, Inc., Waltham, MA) and exposed to a thermal program specific to the set of primers. Primers, thermal programs, and associated references are listed in Tables A1.1 and A1.2. Amplification products were separated and visualized by 2% agarose gel electrophoresis with ethidium bromide staining. PCR products were purified using Qiaquick PCR purification or gel extraction kits (Qiagen, Valencia, CA).

Sequencing and Bartonella *species identification*

Sequencing reactions were prepared in 20 μL mixtures with 8 μL 96x Big Dye buffer (Life Technologies, Carlsbad, CA), 2 μL (0.3 μM) of each primer (same as first round primers for ITS, *nuoG*, and *ssrA*; nested primers were used for *ftsZ*, *gltA*, and *rpoB*), 0.5-7 μL (3-20 ng) of sample PCR products (depending on fragment length), and 3-9.5 μL nuclease-free water. Reaction mixtures were exposed to a thermal program of 96 ºC for 5 minutes, followed by 25 cycles of 96 ºC for 10 seconds, 50 ºC for 5 seconds, and 60 ºC for 4 minutes. Reaction mixtures were sequenced in both directions using an Applied Biosystems Model 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA). DNA extraction, PCR reaction mixture preparation, PCR amplification, and sequencing were performed in separate dedicated laboratory rooms to reduce cross-contamination. Sequence traces were assembled and edited in Lasergene (DNASTAR, Madison, WI) and SnapGene (GSL Biotech LLC, Chicago, IL). Assembled sequences were identified via alignment and comparison with reference *Bartonella* species obtained from the GenBank database for each particular locus using ClustalX version 2.1 (Larkin *et al.* 2007) and the BLAST program (NCBI, Bethesda, MD). Sequences with no match on GenBank or were not identified as bartonella were considered negative samples and were excluded from the dataset. New sequences identified as bartonella but not identifiable to a species by at least 95% sequence similarity (La Scola *et al.* 2003) were considered putative novel species.

PCR amplification and sequencing were repeated three times for each of the six loci, for eighteen separate test runs for each blood sample. This accounts for the possibility of false negative runs and provides a quantitative measure of relative *Bartonella* species abundance for each sample, as explained below. Bats were considered positively infected if multiple runs of one locus yielded bartonella sequences and at least one other locus yielded a bartonella sequence.

This conservative requirement for the assessment of prevalence acts as a check among loci for the potential presence of false positives.

Modelling Bartonella *species abundance from multiple primers*

In order to quantify the relative abundance of *Bartonella* species among sampled bats, I adapt a multiple primer model developed by Johnson *et al.* (in preparation) for methanotrophic bacteria in the Great Plains. With perfect amplification of sequences from a positively infected sample, the number of positive runs y_{ijk} measured for *Bartonella* species k in sample i, amplified by primer set j would be proportional to the amount of DNA $e^{\phi_{ik}}$ present for species k in sample i. However, primer misamplification due to incomplete annealing or other errors in the PCR process can bias the expected number of positive runs. Thus, the amount of DNA will be proportional to the number of positive runs for species k in sample i multiplied by a term representing the misamplification bias $e^{\alpha_{jk}}$ unique to each primer set j and species k and constant across all samples. Therefore, I model *Bartonella* species abundances with the equation $Y_{ikl} \propto e^{\phi_{ik}+\alpha_{jk}}$.

Lacking strong prior information about the amount of bias for each *Bartonella* species from the six primer sets, I assume that primer sets are accurate on average, and biases among primer sets cancel out; formally, $\sum_{j=1}^{J} \alpha_{jk} = 0$ when primer sets are capable of amplifying species k. Thus, if a species has positive runs from only a single primer set ($J_k = 1$), then I assume that this primer set is accurate ($\alpha_{jk} = 0$), and that species is present in the sample. If none of the primer sets has positive runs for a species k in sample i ($y_{ijk} = 0$ for all primers j) then that species is absent from sample i.

A hierarchical Bayesian (Gelman & Hill 2007) multinomial model is proposed to account for the primer bias proportional to $e^{\alpha_{jk}}$ used to estimate the relative abundance of *Bartonella* species in each sample, proportional to k:

$$w_{ijk} \times y_{ijk} | \psi_{ij} \sim \text{Multinomial}\left(\sum_{k=1}^{S} w_{ijk} y_{ijk}, \omega_{ij}\right),$$

$$\omega_{ijk} \sim \frac{e^{\psi_{ijk}}}{\sum_{k=1}^{S} e^{\psi_{ijk}}},$$

$$\psi_{ijk} | \phi_{ik}, \alpha_{jk}, {\sigma^2}_{\psi,k} \sim \text{Normal}(\phi_{ik} + \alpha_{jk}, {\sigma^2}_{\psi,k})$$

where $e^{\psi_{ijk}}$ is proportional to the relative abundance ω_{ijk} of species k amplified by primer set j in sample i.

Effects associated with the location within each species are centered using the following conjugate priors: $\phi_{ik} \sim \text{Normal}(\lambda_{ik}, {\sigma^2}_{\phi,ik})$, $\lambda_{ik} \sim \text{Normal}(\mu_k, {\sigma^2}_{\lambda,k}{\sigma^2}_{\phi,ik})$, and $\mu_k \sim \text{Normal}(m = 0, {\sigma^2}_0{\sigma^2}_{\lambda,k})$, representing replicates within locations and overall. The prior on each primer effect α_{kl} is set to be $\text{Normal}(g = 0, {\sigma^2}_{\alpha,k})$. The prior distribution for all variance terms was chosen to be inverse-gamma for conjugacy, with shape and rate parameters set at 0.5 to generate a Student's T distribution as the marginal prior on all community effects.

Other models may have enough replications of μ_k and α_{jk} for both m and g to be identifiable, yet they are not identified in this multinomial model because ψ_{ijk}, ϕ_{ik}, λ_{ik}, μ_k, and α_{jk} are overparameterized and can only be identified up to some additive constant that cancels out when relative abundance is inferred (Gelman *et al.* 2003). This is acceptable since this study makes inference about differences in species assemblages, not about the abundance of individual

species. Further details on implementation using Markov chain Monte Carlo can be found in Johnson *et al.* (in preparation) and Appendix I, Section B.

Modelling Bartonella *species phylogeography*

Bartonella species assemblages were analyzed for phylogeographic and other patterns using a Bayesian multinomial logistic model adapted from Agresti (1990). Data on relative abundances of *Bartonella* species for each sampled bat were estimated from the multiple primer model. Accessory information on location, age class, and sex was compiled for each individual bat. Each combination of covariates is assumed to give rise to a multinomial response with a logistic link function. For each for location q, age class r, and sex s, the observed vector of species abundances $Z_{qrs\cdot} = Z_{qrs1}, \ldots, Z_{qrs7}$ has the distribution:

$$\mathrm{Z}_{qrs} \sim \mathrm{Multinomial}\left(p_{qrs}, n_{qrs}\right),$$

$$p_{qrst} = \frac{\pi_{qrst}}{\sum_t \pi_{qrst}},$$

$$\pi_{qrst} = e^{\beta_t + \gamma_{qt} + \delta_{rt} + \zeta_{st}}$$

where $n_{qrs} = \sum_t Z_{qrst}$ and γ_{qt}, δ_{rt}, and ζ_{st} represent the probability of species t occurring for each location q, age class r, or sex s, respectively. A corner point species is chosen for the intercept β_t for comparison with other species. Values for β_1, γ_{q1}, γ_{1q}, δ_{r1}, δ_{1r}, ζ_{s1}, and ζ_{1s} are all set to zero for identifiability. For computational efficiency, a multinomial-Poisson transformation was used based on the relationship between these two distributions:

$$Z_{qrst} \sim \text{Poisson}(\nu_{qrst}),$$

$$\nu_{qrst} = e^{\tau_{qrs}+\beta_t+\gamma_{qt}+\delta_{rt}+\zeta_{st}}$$

where ν_{qrst} is analogous to the *Bartonella* species relative abundances π_{qrst} in the untransformed equation and a dummy parameter τ_{qrs}. Unknown values of τ, β, γ, δ, and ζ are initially assigned independent diffuse priors of $\text{Normal}(0, {\sigma^2}_Z)$ where ${\sigma^2}_Z$ is the binomial variance in the species abundance data from the multiple primer model.

An important assumption of this model is that adjusted abundances within each predictor class (location, age class, and sex) can be averaged across individuals before analysis in the multinomial model. To check the validity of this assumption, I calculated species richness and evenness (Shannon-Wiener and Simpson indices) for all individuals using the "vegan" package (Oksanen *et al.* 2015) in R 3.0.3 (R Core Team 2015). I tested for a significant effect of location, age class, and sex using ANOVA; a lack of significant effects of any predictor indicates that averaging species abundances across individuals for these predictors is justified. Model selection for the multinomial analysis was performed by calculating the deviance information criterion (DIC) from the analysis of averaged datasets, selecting the top model with the lowest DIC value. More information on implementation of the Bayesian analysis using Markov chain Monte Carlo can be found in Appendix I, Section B.

Statistical analysis of bartonella prevalence

Logistic regression was used to assess patterns of bartonella prevalence among sampled bats. As mentioned above, bats were considered positively infected if multiple runs of one locus yielded bartonella sequences and at least one other locus yielded a bartonella sequence. All other

bats were considered bartonella-negative. Logistic regression analysis was implemented in SAS University Edition (SAS Institute, Cary, NC) using the "logistic" procedure with location, age class, sex, and pregnancy status of females as predictors in the regression model. Model selection was implemented using the Akaike information criterion, adjusted for finite sample sizes (AICc). Goodness-of-fit for all models was calculated by the area under the receiver operating characteristic curve (AUC). I consider an AUC greater than 0.7 to be a good fit (Hosmer and Lemeshow 2000).

Results

Sequencing and Bartonella *species identification*

Sequences from all six *Bartonella* species previously described from *Eidolon helvum* by Kosoy *et al.* (2010a) and Bai *et al.* (2015), currently named *E1-E5* and *Ew*, were successfully isolated via PCR. For a majority of *Bartonella*-positive bats, conflicting sequences representing multiple, potentially coinfecting *Bartonella* species were isolated from separate loci or repeated sequencing of a single locus. In addition to these 6 known species, *gltA* sequences were isolated that cluster with *Bartonella* sequences previously isolated from bat flies parasitizing *Eidolon helvum* but not *E. helvum* itself (Billeter *et al.* 2012). The phylogenetic tree for the *gltA* sequences shows that two clusters are distinct from *E1-E5* and *Ew* (Figure A1.1) with genetic distances greater than 5%, as estimated using Jukes-Cantor model (Jukes and Cantor 1969) indicating that these sequences likely represent novel species (La Scola *et al.* 2003). One *gltA* cluster had greatest similarity to "*Bartonella* sp. clone Cg 374" from Annobón and the second cluster had greatest similarity to "*Bartonella* sp. clone Cg 713-2" from Bioko (Billeter *et al.* 2012). Separate clusters of sequences were also found for each of the other five genes (*ftsZ*, ITS, *nuoG*, *rpoB*, and *ssrA*) that may represent these new *Bartonella* species (Figures A1.1-A1.6).

Nevertheless, these species have not been cultured and fully described, and I do not attempt to do so in this study. The PCR and sequencing platform I use may isolate sequences of coinfecting bartonellae or sequences that represent recombinant strains. Thus, attempting to classify a new species based on the identity of the bat, relative phylogenetic position, or another measure may be incorrect. Given this uncertainty, all sequences that cluster into one of these groups are simply identified as "new."

All sequences for each species were counted and summed for each location, as summarized in Figure 2.2. Species *E3*, *E5*, and *Ew* were common in many locations, similar to findings by Kamani *et al.* (2014) and Bai *et al.* (2015). Species *E1* was the most uncommon type, with sequences of this species only isolated from three out of the eight locations. *Eidolon dupreanum* appears to host the same assemblage of *Bartonella* species as its sister species *E. helvum*, despite the species' isolation on Madagascar. Total prevalence of bartonella DNA was based on the conservative criterion that bats were considered positively infected if multiple runs of one locus yielded bartonella sequences and at least one other locus yielded a bartonella sequence (for any *Bartonella* species). Even with this conservative measure, bartonella prevalence was high for all sampled locations: 90.9, 52.4, 75.9, 51.7, 57.4, 63.5, 59.1, and 52.1% for Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé, respectively (Table 2.1). Compared to previous estimates (26.1%) of bartonella prevalence in *E. helvum* from Kenya using culturing (Kosoy *et al.* 2010a), PCR and sequencing results in higher detection for all locations ($\chi^2 = 192.14$, DF = 8, $P < 2.2x10^{16}$; Figure A1.7).

Multiple primer model

The relative *Bartonella* species abundances as estimated from the three MCMC chains from the multiple primer model were averaged together since all of the chains appeared to

converge to a similar value for each sample and each *Bartonella* species. The resulting distribution of species across locations was similar to the distribution generated from unadjusted counts of species, with some small changes to particular species abundances based on bias estimates for each of the six PCR primers (Figure 2.3).

The overall bias of each primer set was quantified by extracting the highest probability density (HPD) interval for the α_{jk} parameters, representing the bias of primer set j for species k, across the three chains. Primer sets for a particular chain having a HPD interval that contained zero received a score of 0. Primers sets with a HPD interval that did not contain zero were assigned a score of +1 or -1 depending on whether the HPD interval was above or below zero, respectively. These scores were summed across each of the three chains for all six primer sets (Figure 2.4). Primer set *rpoB* was unable to amplify species *E1*, *E2*, and *E4*, so a score of -3 was automatically assigned to *rpoB* for those species to illustrate this bias. Based on this scoring scheme, *ftsZ* appears to be the least biased primer set for the evaluation of *Bartonella* species abundance in these samples.

Phylogeography model

An important assumption when using the multinomial model is that adjusted abundances within each predictor class (location, age class, and sex) can be averaged across individuals before analysis. Individual variation in species richness, Shannon-Wiener diversity, and Simpson diversity across locations, age classes, and sexes are shown in Figures A1.8-A1.10. Species richness did not differ across locations ($F = 1.83$, $DF = 7$, $P = 0.086$), age classes ($F = 2.38$, $DF = 2$, $P = 0.098$) or sexes ($F = 2.46$, $DF = 1$, $P = 0.12$). Species evenness did not differ across locations ($F = 1.45$, $DF = 7$, $P = 0.19$), age classes ($F = 2.63$, $DF = 2$, $P = 0.077$), or sexes ($F = 2.74$, $DF = 1$, $P = 0.1$) using the Shannon-Wiener diversity index. Finally, species evenness did

not differ across locations (F = 1.32, DF = 7, P = 0.25), age classes (F = 2.67, DF = 2, P = 0.074), or sexes (F = 1.84, DF = 1, P = 0.18) using the Simpson diversity index. Therefore, the assumption that *Bartonella* relative species abundances can be averaged across individuals for the multinomial model is justified.

Model selection using DIC from models in OpenBUGS chose the intercept as the best model (Table 2.2). Therefore, relative *Bartonella* species abundance distributions were not well explained by location, age class, or sex for the six locations where all covariates were recorded (excluding Kenya and Nigeria). To ensure that this result was not affected by the prior distribution of τ, β, γ, δ, and ζ parameters, additional analyses were run with higher and lower values of binomial variance ${\sigma^2}_Z$. Strong effects of covariates were only seen for very high variances that assign inappropriate weights on the ends of the binomial distribution (Table A1.3). Additional analyses were performed for abundance distributions based on all eight locations where only sex was recorded. Again, the intercept was chosen as the best model by DIC (Table 2.2) and this result was robust across variance values (Table A1.3). Species *E5* was used as the corner point comparison for the multinomial model in all comparisons. Additional analyses using species *E3* as the corner point produced nearly identical results.

Logistic models of bartonella prevalence

Table 2.1 summarizes the total number of positive samples for each location, age class, sex, and pregnancy status of females. Analysis of the six locations where age class and sex were reported using model selection chose several top models with nearly equivalent AICc and AUC values (Table A1.4). Age + sex was the top model by AICc, but the location + age + sex and location + age models had ΔAICc values less than two and AUC values greater than 0.7. Hence, I consider these three models as the top set for analysis. Age class was a strongly significant

predictor in all three models: age + sex (χ^2 = 24.2, DF = 2, P < 0.0001), location + age + sex (χ^2 = 19.56, DF = 2, P < 0.0001), and location + age (χ^2 = 18.49, DF =2, P < 0.0001). Odds ratio (OR) estimates for age classes show that juvenile bats are significantly less likely to carry bartonella than sexually immature bats (Figure 2.5), although the estimates varied for the three top models: age + sex (OR = 0.028), location + age + sex (OR = 0.011), and location + age (OR = 0.014). All of the 95% confidence intervals for odds ratio estimates of adults versus sexually immature bats included one, so bartonella prevalence does not significantly differ in these two age classes. The effects of sex and location were not present in all three models and their effects were not consistently significant. Only in the age + sex model was sex found to be a significant predictor of bartonella prevalence (χ^2 = 4.06, DF = 1, P = 0.044). Because the significant effect of sex disappears in the location + age + sex model, I do not consider this strong evidence for an effect of sex on bartonella prevalence. The effect of location was not significant in either model where it was included.

Model selection using data on female bats from the six locations where age class and pregnancy status were reported chose the age only model based on AICc, with no other models with ΔAICc values less than two (Table A1.5). Age class was a significant predictor of bartonella prevalence (χ^2 = 8.71, DF = 2, = 0.013), consistent with the models that included both males and females. Juvenile females were 0.027 times less likely to carry bartonella than sexually immature bats. Adult females were 0.483 times less likely to carry bartonella than sexually immature bats, but the 95% confidence interval for the odds ratio included one, so prevalence between adult and sexually immature females do not significantly differ. It should be noted that the age only model for females had an AUC value less than 0.7, so it is considered to have weak predictive power for analysis of bartonella prevalence (Hosmer and Lemeshow 2000).

Finally, analysis on all eight locations where only sex was reported found three top models by AICc: location + sex, location only, and sex only (Table A1.6). However, none of these models have AUC values greater than 0.7, so they all have low predictive power. No significant main effects of location or sex were found; this is illustrated by the fact that confidence intervals for the majority of locations, sexes, and pregnancy statuses of females overlap (Table 2.1). Overall, the results indicate that age class is a strong predictor of bartonella prevalence in bats, with juvenile bats less likely to be infected than sexually immature or adult bats; sex and location are not strong predictors of bartonella prevalence.

Discussion

The present study demonstrates apparent maintenance of diverse *Bartonella* species among widely dispersed populations of *Eidolon* spp. fruit bats. Using extended molecular and statistical techniques, I detect the presence and quantify the abundance of potentially coinfecting *Bartonella* species in individual blood samples from bats. Total bartonella DNA prevalence was high (>50%) across all sampled locations. Previously classified *Bartonella* species *E1-E5* and *Ew* from *Eidolon helvum* were found in newly sampled continental populations (Ghana, Nigeria) and isolated offshore islands (Annobón, Bioko, Príncipe, and São Tomé), as well as from the related species *E. dupreanum* endemic to Madagascar. In addition to the six described species, I report the presence in bats of putative novel *Bartonella* species previously sequenced only from bat flies from Ghana and islands in the Gulf of Guinea (Billeter *et al.* 2012). Comparison of locations, age classes, and sexes found no detectable difference in the distribution of *Bartonella* species abundances.

Contrasting with previous studies that used parasites to infer population structure in a reservoir host, no cryptic population structure could be detected in *Eidolon* fruit bats from

Bartonella species distributions. It should be noted however, that the multiple primer model assumes that locations differ in species abundances *a priori*. Thus, I find that abundances of species *E1* do differ across locations where it is present versus absent. However, these differences may not reflect actual absence in some of these locations and instead may be caused by a detection bias in the sampling if species *E1* is truly uncommon. For example, *E1* was found in Nigeria by culturing (Bai *et al.* 2015) although it was not present in the Nigerian samples used in this study.

This point aside, this analysis shows that populations of *Eidolon* spp. bats across large geographic distances share similar distributions of *Bartonella* species (Figure 2.3). Some salient life history features of *E. helvum* may explain this homogeneity. The distribution of *E. helvum* is not continuous across its range, but rather forms a connected network of populations that aggregate seasonally. The species is thought to migrate annually between the equatorial forests and the savannahs to the north and south, following shifts in the Intertropical Convergence Zone weather system and changes in seasonal food availability (Thomas 1983). This results in large fluctuations in the size of permanent colonies, typically 8-100 individuals (DeFrees and Wilson 1988), and the formation of large colonies up to 1.5 million individuals that persist for a short time (Sørensen and Halberg 2001; Hayman *et al.* 2012). Satellite telemetry has demonstrated that individual bats can migrate 370 km in one night and 2500 km over five months (Richter and Cumming 2008). This connectivity is expected to facilitate gene flow and the persistence of infections (Hess 1996). Plowright *et al.* asserted that the presence of large, weakly coupled and asynchronous metapopulations could explain the persistence of Hendra virus in Australian flying foxes (Plowright *et al.* 2011). Peel *et al.* detected antibodies to Lagos bat virus and henipaviruses in *E. helvum* roosts across continental Africa and islands in the Gulf of Guinea (Peel *et al.* 2012,

2013), which was linked to widespread panmixia in this species. Given that continental metapopulations of *E. helvum* are so well mixed and seasonal, then transmission of bartonella could be relatively consistent across geographic locations, resulting in the similar distribution of *Bartonella* species in Ghana, Nigeria, and Kenya.

Based on island biogeography theory, I expected that isolated islands might carry a restricted or potentially unique set of *Bartonella* species. By extension, genetically distinct bats like the population on Annobón (Juste *et al.* 2000; Peel *et al.* 2013) and the separate species *E. dupreanum* would be expected to carry even more restricted or distinct sets of *Bartonella* species due to their geographic and phylogenetic distance from mainland populations. The similarity found in *Bartonella* species assemblages between the Gulf of Guinea islands, Madagascar, and the mainland requires some speculation. The Gulf of Guinea islands are separated from the mainland and each other by only 30-350 km which are well within the longest overland flight distance (370 km) recorded for *E. helvum* (Richter and Cumming 2008). The distance across the Mozambique Channel to Madagascar is 460 km, however vagrant *E. helvum* have been found on the Cape Verde islands 570 km offshore (Jiménez and Hazevoet 2010). Low levels of gene flow among the Gulf of Guinea island populations (Juste *et al.* 2000) and between *E. helvum* and *E. dupreanum* (Shi *et al.* 2014) indicate that this is probably a rare occurrence; however, mating is not necessary for bartonella transmission. Aggressive encounters through biting and scratching or through sharing of ectoparasites among locals and vagrants could plausibly explain the transmission of diverse *Bartonella* species.

Alternatively, additional host species could facilitate transmission among *Eidolon* spp. in distant locations. Species *E3*, *E5*, and *Ew* have been found in other bat populations in Nigeria, specifically *Epomophorus*, *Micropteropus*, and *Rhinolophus* spp., which share the same bat fly

species (*Cyclopodia greeffi*) with *E. helvum* (Kamani *et al.* 2014). Although there is little overlap in species between Madagascar, the Gulf of Guinea islands, and the mainland, sampling of bartonella from African bats is far from exhaustive and more research may identify other "bridge" hosts.

Finally, host-parasite relationships between *Eidolon* spp. fruit bats and their associated *Bartonella* bacteria may be very old, preceding the divergence of *E. helvum* and *E. dupreanum* during the middle to late Miocene (Juste *et al.* 1999; Shi *et al.* 2014). Evidence for cospeciation has been found between bats and their bartonellae (Lei and Olival 2014), so limited divergence between *E. helvum* and *E. dupreanum* at a cellular and physiological level may reduce selection on the set of *Bartonella* species passed down from the common ancestor to the two *Eidolon* spp. Furthermore, E. *dupreanum* shares some similar demographic characteristics with *E. helvum* that could facilitate persistence of diverse *Bartonella* species on Madagascar. *E. dupreanum* is widespread but patchily distributed on the island. The bats roost in trees and caves with a typical colony size of 10-500 individuals. Regular variation in the occupancy and abundance of roosts suggests that *E. dupreanum* is migratory (MacKinnon *et al.* 2003) and genetic evidence indicates that the population is panmictic. Hence, ongoing transmission from mainland *E. helvum* populations to Madagascar may not be necessary if the size and connectivity of *E. dupreanum* populations are sufficient for endemic maintenance of a diverse assemblage of *Bartonella* species. Transmission by vagrant *Eidolon helvum* or other host species from the mainland or endemic maintenance in a large population are both plausible, non-mutually exclusive mechanisms that could explain the similarity of *Bartonella* species distributions across island populations. Further research is warranted to estimate the relative contributions of these mechanisms.

The detection of *Bartonella* sequences from bat blood that were first isolated from bat flies is not unprecedented. *Bartonella* species *E4* and *E5* were detected from *gltA* sequences in bat flies from Ghana and islands in the Gulf of Guinea (Billeter *et al.* 2012) before they were cultured from bat blood and fully classified as separate species (Bai *et al.* 2015). Kamani *et al.* (2014) found that *Eidolon helvum* and *Cyclopodia greeffi* bat flies in the same community share *Bartonella* species. Brook *et al.* (2015) detected *Bartonella* species *E4*, *E5*, and *Ew* in *Cyclopodia dubia* bat flies parasitizing *Eidolon dupreanum*. Judson *et al.* (2015) isolated identical *Bartonella* variants in Costa Rican bats and the bat flies directly parasitizing them. *E. helvum* is already exceptional among reservoir hosts for carrying six distinct *Bartonella* species; the identification of additional species is likely a function of sampling effort in *E. helvum* and improved detection tools.

Overall, the multi-gene PCR platform using DNA from blood does provide improved detection of bartonella compared to culturing. Additionally, I have determined that some primer sets have significant biases for some *Bartonella* species (Figure 2.4), but that the combination of multiple primer sets can capture the abundance of *Bartonella* species in a sample. I expect that this platform will be applicable to other animal systems and tissue types, and may assist in the identification of new and co-occurring *Bartonella* species in reservoir species.

Isolation of conflicting sequences representing separate *Bartonella* species across the six loci or among replicates at a single locus occurred frequently using the PCR platform. These conflicts may represent coinfecting *Bartonella* species, recombinant strains with loci reflecting mixed ancestry, or a combination of both scenarios (Abbot *et al.* 2007; Chan and Kosoy 2010; Arvand *et al.* 2007; Berglund *et al.* 2010; Guy *et al.* 2012; Paziewska *et al.* 2011, 2012; Buffet *et al.* 2013b). Given this uncertainty, I chose to model each sequence isolate as representing the

abundance of that species individually. Distinguishing between recombinant and coinfecting *Bartonella* strains would require culturing, isolation of separate colonies, and sequencing of multiple loci. However, bartonellae are fastidious and notoriously difficult to culture, so a culturing approach is expected to underestimate bartonella prevalence and may be unable to capture the true diversity of coinfecting and recombinant strains.

Integrating the multi-gene PCR platform and the multiple primer model developed by Johnson *et al.* is an attempt to capture and quantify the diversity of bartonellae in blood samples, but I recognize that these techniques are limited in some aspects. Differences in *Bartonella* species abundances may actually exist between distant fruit bat populations, but the number of sampled bats and sequencing replications in the PCR platform were both small, which in combination with misamplification biases of the primer sets may have reduced the power of this approach. Quantitative PCR (qPCR) and next-generation sequencing (NGS) techniques may prove to be more effective for capturing bartonella diversity and identification of coinfections or recombinant strains. In fact, the multiple primer model was originally designed for NGS data (Johnson *et al.*, in preparation). Unfortunately, these advanced molecular tools have not been validated for coinfecting *Bartonella* species to our knowledge and more development in this area is needed. Another improvement in this approach would be to account for missing species in the data explicitly. The multiple primer model assumes that if no sequences were isolated for a species, then that species is absent from the sample. An alternative statistical model may be able to account for negative detection error. Furthermore, population structure may exist within each of the *Bartonella* species across the sampled geographic range, but might not have been detected using the short fragments of genes amplified from the PCR based approach. More extensive

sampling and characterization of genetic diversity with *Bartonella* species using multi-locus sequence typing or whole genome sequencing may reveal cryptic bacterial population structure.

Demographic factors were found to influence bartonella prevalence. Specifically, I find that juveniles are less likely to be infected than other age classes (Figure 2.5). The slightly higher prevalence found in bats from Annobón and Ghana (Figure A1.7) may simply be due to sampling bias for sexually immature and adult bats. The presence of maternal antibodies may provide some protection for the youngest age class (Kallio *et al.* 2010; Garnier *et al.* 2012). Recent work has demonstrated that maternal antibodies against canine distemper and Hendra viruses in *Pteropus* spp. flying foxes (Epstein *et al.* 2013) and henipaviruses in *Eidolon helvum* (Baker *et al.* 2014) can persist in young bats for up to six months. Transfer of immunoglobulin G (IgG), the antibodies responsible for the majority of immunity against blood-borne microparasites, in *Pteropus alecto* from dam to pup was shown to occur primarily through milk (Wynne *et al.* 2013). Previous studies in six rodent species and black-tailed prairie dogs (*Cynomys ludovicianus*) found that bartonella prevalence was highest in younger, smaller rodents and declined in older individuals (Kosoy *et al.* 2004; Bai *et al.* 2008). Maternal antibodies may decay more rapidly in these species compared to bats. Alternatively, these differences in juvenile prevalence may reflect a large divide in the life history and demographic organization in rodents versus bats. Generally, rodents do not maintain high antibody titers to bartonella and may rely more on cell-mediated immunity than humoral immunity (Kosoy *et al.* 2004a). No studies have detected antibodies against bartonella in bats to date, and the study of bat immunology is still in its infancy (Baker *et al.* 2013). Measuring the demographic patterns of prevalence, possibly with better estimates of age from tooth cementum annuli (Hayman *et al.* 2012), will help to

understand transmission dynamics of bartonella in bat populations over time (Plowright *et al.* 2008; Dietrich *et al.* 2015).

Low juvenile prevalence may also stem from feeding preference of bat flies. Christe *et al.* (200) found that *Spinturnix myoti* mites feed more on female *Myotis myotis* bats in the early stages of pregnancy and less on bats in later pregnancy stages. This preference was inversely proportional to females' immunocompetence. The authors found no difference in ectoparasite load between lactating mothers and attached pups, but did find greater ectoparasite loads on young bats that had begun to roost independently than on lactating mothers or their attached newborns (Christe *et al.* 2000). Despite expectations that reduced immunocompetence in pregnant females may increase susceptibility to bartonella infection (Plowright *et al.* 2008; Breed *et al.* 2011; Baker *et al.* 2013), I found no difference in prevalence among non-pregnant, pregnant, and lactating females. The difference in bartonella prevalence seen among juvenile, sexually immature, and adult bats may reflect a shift in bat fly feeding behavior to young and newly independent bats, possibly in concert with the decline in maternal antibodies. The juveniles that are infected may have been horizontally infected due to sharing of bacteria or viruses with their infected dam, or possibly vertically infected as a fetus. Viable *Bartonella* bacteria have been detected in the blood of cotton rat (*Sigmodon hispidus*) embryos and neonates (Kosoy *et al.* 1998), suggesting that vertical transmission is possible. Data on bat fly infestation and identifications of dams and their pups were not recorded from all bats, so the correlation between ectoparasite numbers or mother-offspring transmission and bartonella prevalence could not be measured in this study. Further research on bat fly feeding behavior may reveal other interesting dynamics, such as bartonella-related mortality in bat flies and associated preference for uninfected host individuals (Witsenburg *et al.* 2014).

One important point that should be stated is that no study has empirically estimated the relative contribution of various possible transmission processes, either direct transmission, vector-borne transmission, or vertical transmission to bartonella prevalence and diversity in bats or any other mammal. If I assume that transmission is predominately vector-driven, then the inference I make about bat migration and the transmission of diverse *Bartonella* species across geographic regions is reliant on bat flies leaving their host and mixing randomly with other bats. Bat flies are known to leave their hosts to deposit pupae and to transfer to new individuals when bats are in close contact, as would be seen in massive roosts (Judson *et al.* 2015). Moreover, *Cyclopodia horsfieldi* bat flies were shown to lack population structure across Southeast Asia, owing to the movement and exchange of ectoparasites among *Pteropus* spp. bats (Olival *et al.* 2013), which likely has strong parallels to *Eidolon* spp. bats. Nevertheless, more research is needed to understand the dynamics of bat fly feeding behavior and bartonella transmission in bats.

Although there is no evidence that any of the *Bartonella* species detected in this study appear to cause pathology in their bat hosts, other animals, or humans, it should be noted that a recent study detected evidence of *Bartonella mayotimonensis*, the reported etiological agent of endocarditis in a human patient, in two species of insectivorous bats from Europe (Veikkolainen *et al.* 2014). Both *E. helvum* and *E. dupreanum* can roost in close proximity to humans, potentially facilitating spillover of bartonellae to humans. Transmission via direct handling and butchering by local hunters and bushmeat vendors is plausible, considering that both *Eidolon* species are frequently hunted (Kamins *et al.* 2011; Mickleburgh *et al.* 2009; MacKinnon *et al.* 2003; Jenkins and Racey 2008) and that participants in the bushmeat market may not understand the risk of infection posed by handling and consuming bats (Kamins *et al.* 2014). The etiologies

of acute, non-malarial febrile illness and endocarditis frequently go unexplained in developing countries (Crump *et al.* 2013) and zoonotic illnesses caused by bartonella may be more widespread than is currently appreciated (Kosoy *et al.* 2010b; Laudisoit *et al.* 2011; Rattanavong *et al.* 2014). Clearly, more research is needed to understand the extent to which zoonotic transmission of bartonella from bats to humans is occurring in Africa and elsewhere.

From a broad ecological and evolutionary perspective, this system of *Bartonella* species and their bat hosts can be used as a model for demonstrating the processes that structure parasite diversity in host populations. Vellend (2010) states there are four key processes that govern community dynamics – ecological drift, selection, dispersal, and speciation – which are applicable to parasite communities in a metapopulation framework (Seabloom 2015). For *Eidolon helvum*, its widespread movement across continental Africa provides unrestricted dispersal of *Bartonella* species into host individuals. If I assume that *Bartonella* species have predominately weak or neutral interactions, the process of community assembly in an individual host will be random, with some individuals carrying a more diverse set of *Bartonella* species than other individuals do. The contribution of ecological drift comes in the form of parasite population bottlenecks during the transmission process, wherein small parasite populations are more subject to stochastic processes than large populations, and may not persist in an infected individual. For a vector-borne infection like bartonella, the effect of ecological drift on diversity will be amplified. Although drift may be an important process driving within-host diversity, when averaged among host individuals dispersal will homogenize parasite communities, evidenced by similarity in *Bartonella* species assemblages across geographic locations. Ecological drift may also have a role in reducing diversity on isolated islands where dispersal is limited, however even limited dispersal may be able to maintain *Bartonella* species on islands,

especially if bats are chronically infected. Finally, speciation appears to be a weak process in this system, since identical *Bartonella* species have been found in *Eidolon helvum* and *E. dupreanum*, with no evidence of co-divergence in the bacterial species over the millions of years that these bat species have been separated (Shi *et al.* 2014).

This system is expected to be fruitful for continued research, both fundamental and applied. Future studies should explicitly test the contributions of bat immunity, parasite interactions (Telfer *et al.* 2010), and ecological drift to within-host bartonella dynamics, which could expand our understanding of disease ecology and evolution generally. Finally, spillover dynamics of infections at the animal-human interface are poorly understood, not least of which for bartonella, and more research on the maintenance and transmission of pathogens in reservoir species may help to prevent human disease.

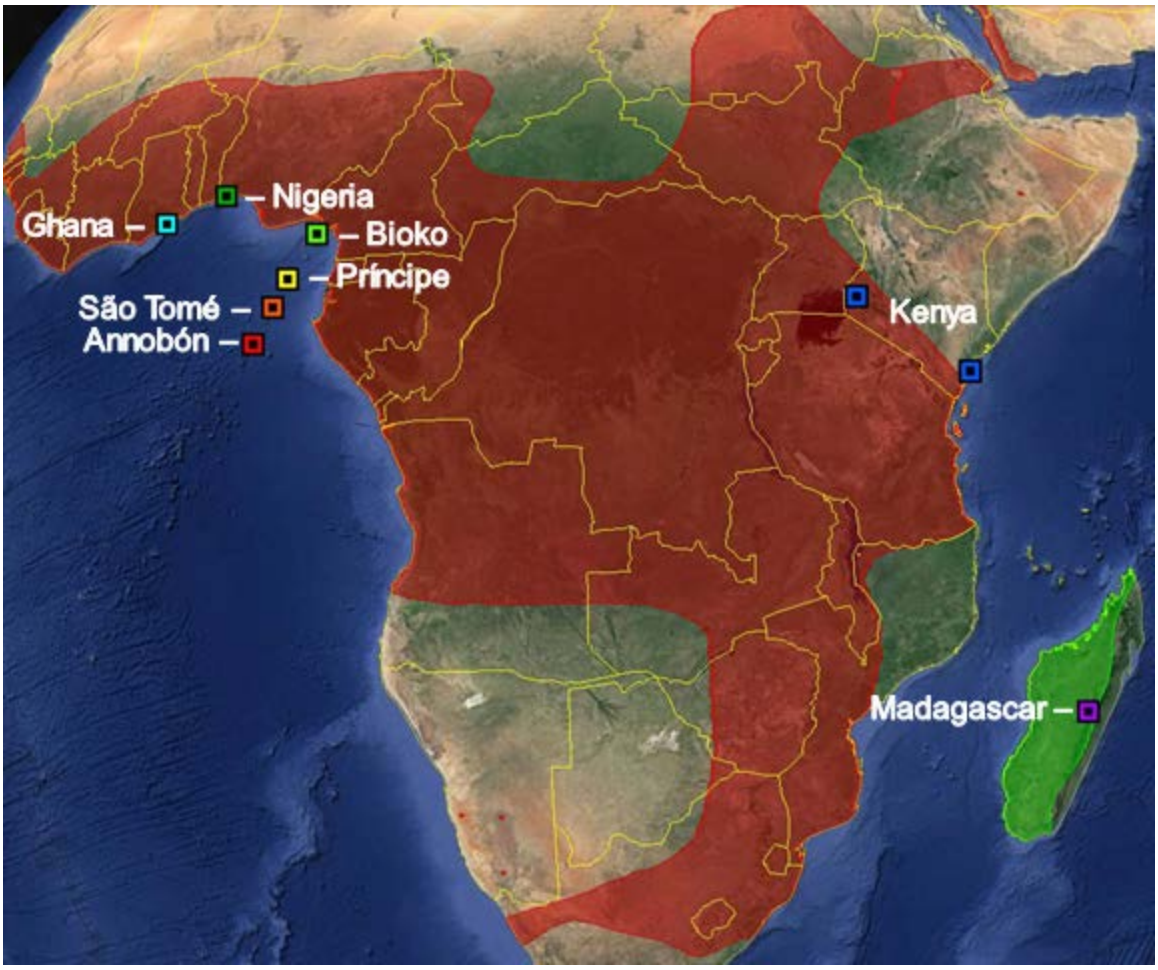

Figure 2.1 Map of sampling sites for *Eidolon helvum* across Africa and *Eidolon dupreanum* on Madagascar. The red shaded region and green shaded region represent the distributions of *Eidolon helvum* and *Eidolon dupreanum*, respectively. Adapted from Mickleburgh *et al.* (2008) and Andriafidison *et al.* (2008). Total sample sizes for each location were 22, 21, 29, 29, 47, 63, 22, and 23 for Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé, respectively.

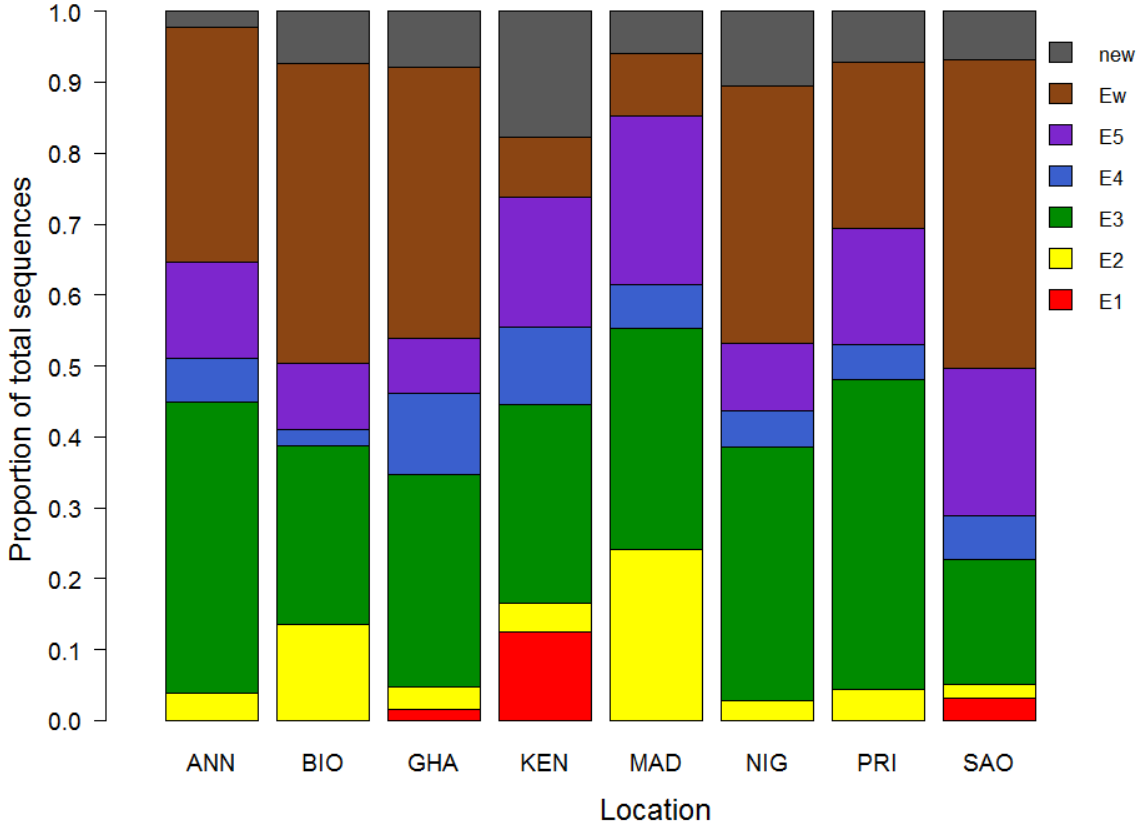

Figure 2.2 Counts of *Bartonella* species from PCR and sequencing. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé and total bartonella-positive samples were 20, 11, 22, 15, 27, 40, 13, and 12, respectively. Colors of stacked bars represent each of the *Bartonella* species found in *Eidolon* spp. bats.

Table 2.1 Summary counts of all bartonella positive samples for each location, age class, sex, and pregnancy status of females. n is total number of tested samples, p is the number of samples positive for any *Bartonella* spp. Approximate confidence intervals are calculated using the "add 2 successes and 2 failures" method from Agresti and Coull (1998).

	n	p	Mean (%)	Lower CI (%)	Upper CI (%)
Location					
Annobón	22	20	90.9	71.0	98.7
Bioko	21	11	52.4	32.4	71.7
Ghana	29	22	75.9	57.6	88.0
Kenya	29	15	51.7	34.4	68.6
Madagascar	47	27	57.4	43.3	70.5
Nigeria	63	40	63.5	51.1	74.3
Principe	22	13	59.1	38.7	76.8
São Tomé	23	12	52.2	33.0	70.8
Age class					
juvenile	25	4	16.0	5.8	35.3
sexually immature	27	23	85.2	66.9	94.7
adult	111	78	70.3	61.2	78.0
Sex					
female	118	67	56.8	47.8	65.4
male	135	91	67.4	59.1	74.8
Pregnancy status					
not pregnant	49	25	51.0	37.5	64.4
pregnant	16	12	75.0	50.0	90.3
lactating	21	14	66.7	45.2	83.0

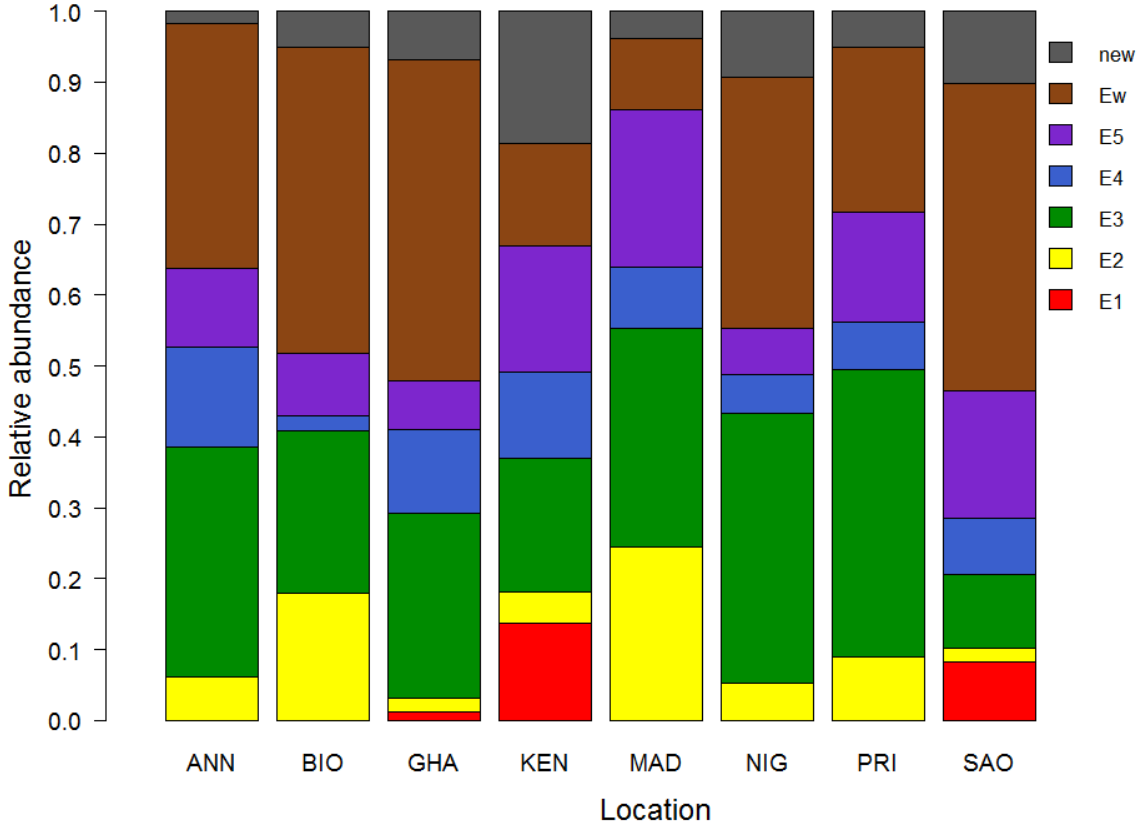

Figure 2.3 Relative *Bartonella* species abundances, adjusted from the multiple primer model. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé. Colors of stacked bars represent each of the *Bartonella* species found in *Eidolon* spp. bats.

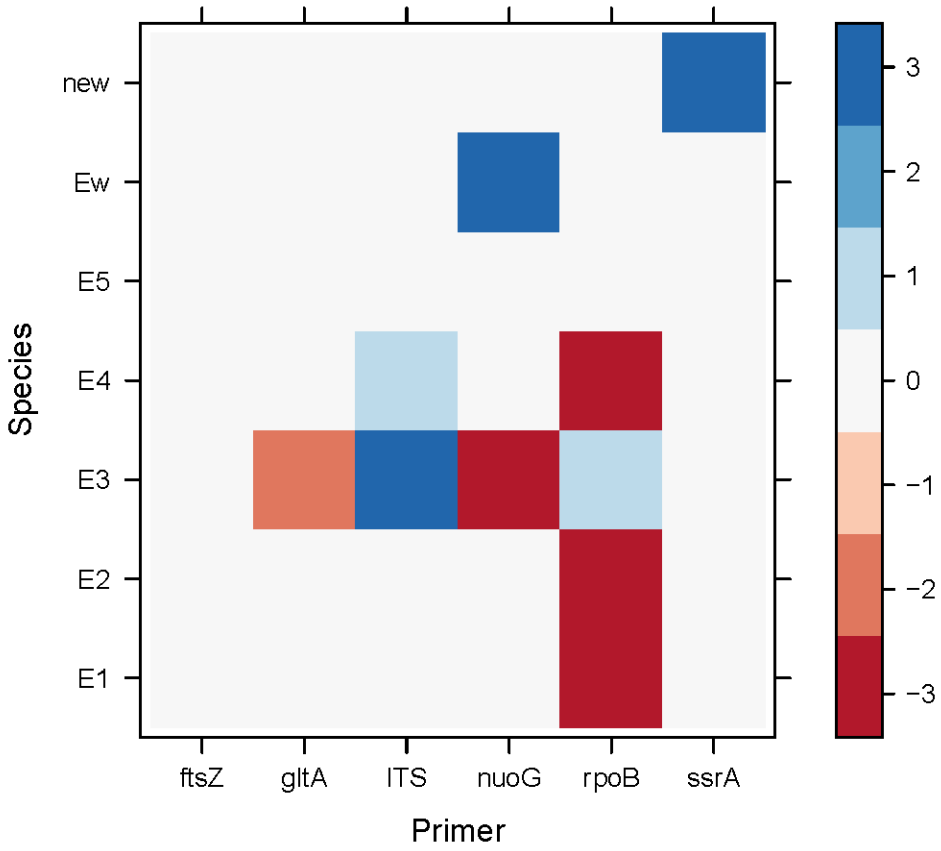

Figure 2.4 Primer biases assessed from the three MCMC chains of the multiple primer model. Blue cells indicate positive bias and red cells indicate negative bias for a species by a primer set.

Table 2.2 (A) Model selection using the deviance information criterion (DIC) from the phylogeography model of location, age class, and sex. This test includes six out of eight locations (excluding Kenya and Nigeria because age classes were not reported). Species *E5* was used as the corner point comparison for the multinomial model. (B) Model selection using DIC from the phylogeography model of location and sex. This test includes all locations (Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé). Species *E5* was used as the corner point comparison for the multinomial model.

Model	**(A) DIC**	**(A) ΔDIC**	**(B) DIC**	**(B) ΔDIC**
(Intercept)	1.25	0	1.22	0
Location	7.73	6.48	10.24	9.02
Location + Sex	17.50	16.25	22.69	21.47
Location + Age	28.36	27.11	-	-
Location + Age + Sex	50.31	49.06	-	-

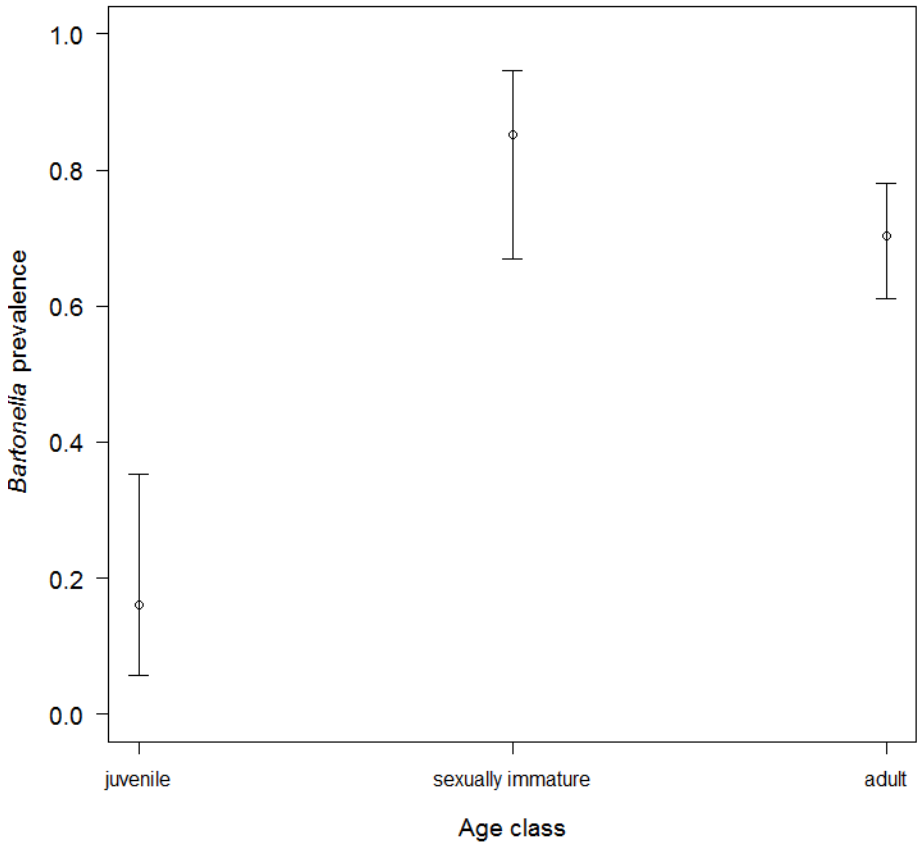

Figure 2.5 Comparison of *Bartonella* spp. prevalence in *Eidolon* spp. fruit bats across sampled age classes. Sample sizes were 25, 27, and 111 for juvenile, sexually immature, and adult bats, respectively. Bats were considered positively infected if multiple runs of one locus yielded *Bartonella* sequences and at least one other locus yielded a *Bartonella* sequence. Point estimates represent total bartonella abundance for all bats in that age class. Binomial confidence intervals are estimated using the "add two successes and two failures" method from Agresti and Coull (1998). Juveniles were found to have significantly lower prevalence than sexually immature or adult bats.

REFERENCES

Abbot P, Aviles AE, Eller L, Durden LA (2007) Mixed infections, cryptic diversity, and vector-borne pathogens: evidence from *Polygenis* fleas and *Bartonella* species. Applied and Environmental Microbiology, 73, 6045–6052.

Andriafidison D, Cardiff SG, Goodman SM, *et al.* (2008) *Eidolon dupreanum*. In: The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 11 February 2015.

Agresti A (1990) Categorical data analysis. New York: Wiley and Sons, Inc.

Agresti A, Coull BA (1998) Approximate is better than 'exact' for interval estimation of binomial proportions. The American Statistician, 52, 119–126.

Arvand M, Feil EJ, Giladi M, Boulouis H-J, Viezens J (2007) Multi-locus sequence typing of *Bartonella henselae* isolates from three continents reveals hypervirulent and feline-associated clones. PLoS ONE, 2, e1346.

Baker KS, Suu-Ire R, Barr J *et al.* (2014) Viral antibody dynamics in a chiropteran host. Journal of Animal Ecology, 83, 415–428.

Baker ML, Schountz T, Wang L-F (2013) Antiviral immune responses of bats: a review. Zoonoses and Public Health, 60, 104–116.

Bai Y, Kosoy M (2012) *Bartonella* infections in rodents and bats in tropics. In: Rodriguez-Morales A, editor. Current Topics in Tropical Medicine. Rijeka: InTech, 51–66.

Bai Y, Kosoy MY, Ray C, Brinkerhoff R, Collinge SK (2008) Temporal and spatial patterns of *Bartonella* infection in black-tailed prairie dogs (*Cynomys ludovicianus*). Microbial Ecology, 56, 373–382.

Bai Y, Kosoy MY, Recuenco S *et al.* (2011) *Bartonella* spp. in bats, Guatemala. Emerging Infectious Diseases, 17, 1269–1272.

Bai Y, Recuenco S, Gilbert AT *et al.* (2012) Prevalence and diversity of *Bartonella* spp. in bats in Peru. The American Journal of Tropical Medicine and Hygiene, 87, 518–523.

Bai Y, Hayman DTS, McKee CD, Kosoy MY (2015) Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLoS Neglected Tropical Diseases, 9, e0003478.

Berglund EC, Ellegaard K, Granberg F *et al.* (2010) Rapid diversification by recombination in *Bartonella grahamii* from wild rodents in Asia contrasts with low levels of genomic divergence in Northern Europe and America. Molecular Ecology, 19, 2241–2255.

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

Billeter SA, Hayman DTS, Peel AJ *et al.* (2012) *Bartonella* species in bat flies (Diptera: Nycteribiidae) from western Africa. Parasitology, 139, 324–329.

Billeter SA, Levy M, Chomel BB, Breitschwerdt EB (2008) Vector transmission of *Bartonella* species with emphasis on the potential for tick transmission. Medical and Veterinary Entomology, 22, 1–15.

Breed AC, Breed MF, Meers J, Field HE (2011) Evidence of endemic Hendra virus infection in flying-foxes (*Pteropus conspicillatus*) – implications for disease risk management. PLoS ONE, 6, 1–7.

Breitschwerdt EB, Maggi RG, Chomel BB, Lappin MR (2010) Bartonellosis: An emerging infectious disease of zoonotic importance to animals and human beings. Journal of Veterinary Emergency and Critical Care, 20, 8–30.

Brook CE, Bai Y, Dobson AP *et al.* (2015) *Bartonella* spp. in fruit bats and blood-feeding ectoparasites in Madagascar. PLoS Neglected Tropical Diseases. 9, e0003532.

Buffet J-P, Kosoy M, Vayssier-Taussat M (2013a) Natural history of *Bartonella*-infecting rodents in light of new knowledge on genomics, diversity and evolution. Future Microbiology, 8, 1117–1128.

Buffet J-P, Pisanu B, Brisse S *et al.* (2013b) Deciphering bartonella diversity, recombination, and host specificity in a rodent community. PLoS ONE, 8, e68956.

Chan K-S, Kosoy MY (2010) Analysis of multi-strain *Bartonella* pathogens in natural host population – do they behave as species or minor genetic variants? Epidemics, 2, 165–72.

Chomel BB, Kasten R (2010) Bartonellosis, an increasingly recognized zoonosis. Journal of Applied Microbiology, 109, 743–750.

Christe P, Arlettaz R, Vogel P (2000) Variation in intensity of a parasitic mite (*Spinturnix myoti*) in relation to the reproductive cycle and immunocompetence of its bat host (*Myotis myotis*). Ecology Letters, 3, 207–212.

Concannon R, Wynn-Owen K, Simpson V, Birtles RJ (2005) Molecular characterization of haemoparasites infecting bats (Microchiroptera) in Cornwall, UK. Parasitology, 131, 489–496.

Criscione CD, Cooper B, Blouin MS (2006) Parasite genotypes identify source populations of migratory fish more accurately than fish genotypes. Ecology, 87, 823-828.

Crump JA, Morrissey AB, Nicholson WL *et al.* (2013) Etiology of severe non-malaria febrile illness in northern Tanzania: a prospective cohort study. PLoS Neglected Tropical Diseases, 7, e2324.

DeFrees S, Wilson D (1988) *Eidolon helvum*. Mammalian Species, 312, 1–5.

Dietrich M, Wilkinson DA, Benlali A *et al.* (2015) *Leptospira* and paramyxovirus infection dynamics in a bat maternity enlightens maintenance in wildlife. Environmental Microbiology.

Epstein JH, Baker ML, Zambrana-Torrelio C *et al.* (2013) Duration of maternal antibodies against canine distemper virus and Hendra virus in pteropid bats. PLoS ONE, 8, e67584.

Falush D, Wirth T, Linz B, *et al.* (2003) Traces of human migrations in *Helicobacter pylori* populations. Science, 299, 1582-1585.

Field H, Young P, Yob JM *et al.* (2001) The natural history of Hendra and Nipah viruses. Microbes and Infection, 3, 307–314.

Garnier R, Ramos R, Staszewski V *et al.* (2012) Maternal antibody persistence: a neglected life-history trait with implications from albatross conservation to comparative immunology. Proceedings of the Royal Society of London. Series B, Biological Sciences, 279, 2033–2041.

Gelman A, Carlin JB, Stern HS, Rubin DB (2003) Bayesian data analysis. Boca Raton: CRC Press.

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press.

Gutiérrez R, Krasnov B, Morick D *et al.* (2015) *Bartonella* infection in rodents and their flea ectoparasites: an overview. Vector-Borne and Zoonotic Diseases, 15, 27–39.

Guy L, Nystedt B, Sun Y *et al.* (2012) A genome-wide study of recombination rate variation in *Bartonella henselae*. BMC Evolutionary Biology, 12, 65.
Harms A, Dehio C (2012) Intruders below the radar: molecular pathogenesis of *Bartonella* spp. Clinical Microbiology Reviews, 25, 42–78.
Hayman DTS, McCrea R, Restif O *et al.* (2012) Demography of straw-colored fruit bats in Ghana. Journal of Mammalogy, 93, 1393–1404.
Hess G (1996) Disease in metapopulation models: implications for conservation. Ecology, 77, 1617-1632.
Holmes EC (2004) The phylogeography of human viruses. Molecular Ecology, 13, 745-756.
Hosmer DW, Lemeshow S (2000) Applied logistic regression, second edition. New York: Wiley and Sons, Inc.
Iehlé C, Razafitrimo G, Razainirina J *et al.* (2007) Henipavirus and Tioman virus antibodies in pteropodid bats, Madagascar. Emerging Infectious Diseases, 13, 159–161.
Jenkins RKB, Racey PA (2008) Bats as bushmeat in Madagascar. Madagascar Conservation and Development, 3, 22-30.
Jiménez S, Hazevoet CJ (2010) First record of straw-coloured fruit bat *Eidolon helvum* (Kerr 1792) for the Cape Verde Islands. Zoologia Caboverdiana, 1, 116-118.
Johnson NG, Koyama A, Webb CT, von Fischer JC (in preparation) Diversity of methanotrophic bacterial communities across the Great Plains.
Judson S, Frank H, Hadly E (2015) Bartonellae are prevalent and diverse in Costa Rican bats and bat flies. Zoonoses and Public Health.
Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN, editor, Mammalian Protein Metabolism. New York: Academic Press, 21-132.
Juste J, Alvarez Y, Tabarés E *et al.* (1999) Phylogeography of African fruitbats (Megachiroptera). Molecular Phylogenetics and Evolution, 13, 596–604.
Juste J, Ibáñez C, Machordom A, Ibanez C (2000) Morphological and allozyme variation of *Eidolon helvum* (Mammalia: Megachiroptera) in the islands of the Gulf of Guinea. Biological Journal of the Linnean Society, 71, 359–378.
Kallio ER, Begon M, Henttonen H *et al.* (2010) Hantavirus infections in fluctuating host populations: the role of maternal antibodies. Proceedings of the Royal Society of London. Series B, Biological Sciences, 277, 3783–3791.
Kamani J, Baneth G, Mitchell M *et al.* (2014) *Bartonella* species in bats (Chiroptera) and bat flies (Nycteribiidae) from Nigeria, West Africa. Vector-Borne and Zoonotic Diseases, 14, 625–632.
Kamins AO, Restif O, Ntiamoa-Baidu Y *et al.* (2011) Uncovering the fruit bat bushmeat commodity chain and the true extent of bat hunting in Ghana, West Africa. Biological Conservation, 144, 3000–3008.
Kamins AO, Rowcliffe JM, Ntiamoa-Baidu Y *et al.* (2014) Characteristics and risk perceptions of Ghanaians potentially exposed to bat-borne zoonoses through bushmeat. EcoHealth, 12, 104-120.
Kosoy MY (2010) Ecological associations between bacteria of the genus *Bartonella* and mammals. Biology Bulletin, 37, 716–724.
Kosoy MY, Bai Y, Lynch T *et al.* (2010a) *Bartonella* spp. in bats, Kenya. Emerging Infectious Diseases, 16, 1875–1881.

Kosoy MY, Bai Y, Sheff K *et al.* (2010b) Identification of *Bartonella* infections in febrile human patients from Thailand and their potential animal reservoirs. The American Journal of Tropical Medicine and Hygiene, 82, 1140–1145.
Kosoy MY, Mandel E, Green D, Marston E, Childs JE (2004a) Prospective studies of *Bartonella* of rodents. Part I. Demographic and temporal patterns in population dynamics. Vector-Borne and Zoonotic Diseases, 4, 285–295.
Kosoy MY, Mandel E, Green D *et al.* (2004b) Prospective studies of *Bartonella* of rodents. Part II. Diverse infections in a single rodent community. Vector-Borne and Zoonotic Diseases, 4, 296–305.
Kosoy M, Regnery R, Kosaya O *et al.* (1998) Isolation of *Bartonella* spp. from embryos and neonates of naturally infected rodents. Journal of Wildlife Diseases, 34, 305–309.
Kuzmin IV, Niezgoda M, Franka R *et al.* (2008) Lagos bat virus in Kenya. Journal of Clinical Microbiology, 46, 1451–1461.
Larkin MA, Blackshields G, Brown NP, *et al.* (2007) Clustal W and Clustal X version 2.0. Bioinformatics, 23, 2947-2948.
La Scola B, Zeaiter Z, Khamis A, Raoult D (2003) Gene-sequence-based criteria for species definition in bacteriology: the *Bartonella* paradigm. Trends in Microbiology, 11, 318–321.
Laudisoit A, Iverson J, Neerinckx S *et al.* (2011) Human seroreactivity against *Bartonella* species in the Democratic Republic of Congo. Asian Pacific Journal of Tropical Medicine, 4, 320–322.
Lei BR, Olival KJ (2014) Contrasting patterns in mammal-bacteria coevolution: *Bartonella* and *Leptospira* in bats and rodents. PLoS Neglected Tropical Diseases, 8, e2738.
Leroy EM, Kumulungui B, Pourrut X *et al.* (2005) Fruit bats as reservoirs of Ebola virus. Nature, 438, 575–576.
Lin J-W, Hsu Y-M, Chomel BB *et al.* (2012) Identification of novel *Bartonella* spp. in bats and evidence of Asian gray shrew as a new potential reservoir of *Bartonella*. Veterinary Microbiology, 156, 119–126.
Loftis AD, Gill JS, Schriefer ME, *et al.* (2005) Detection of *Rickettsia*, *Borrelia*, and *Bartonella* in *Carios kelleyi* (Acari: Argasidae). Journal of Medical Entomology, 42, 473-480.
Luis AD, Hayman DT, O'Shea TJ *et al.* (2013) A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proceedings of the Royal Society of London. Series B, Biological Sciences, 280, e20122753.
MacKinnon J, Hawkins CE, Racey PA (2003) Pteropodidae. In: Goodman SM, Benstead JP, editors, The natural history of Madagascar. Chicago: The University of Chicago Press, 1299–1302.
Mickleburgh S, Hutson AM, Bergmans W, Fahr J, Racey PA (2008) *Eidolon helvum*. In: The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 11 February 2015.
Mickleburgh S, Waylen K, Racey PA (2009) Bats as bushmeat – a global review. Oryx, 43, 217–234.
Morse SF, Olival KJ, Kosoy MY *et al.* (2012) Global distribution and genetic diversity of *Bartonella* in bat flies (Diptera: Hippoboscoidea: Streblidae, Nycteribiidae). Infection, Genetics and Evolution, 12, 1717–1723.

Nieberding C, Morand S, Libois R, Michaux JR (2004) A parasite reveals cryptic phylogeographic history of its host. Proceedings of the Royal Society of London. Series B, Biological Sciences, 271, 2559-2568.
Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: community ecology package. R package version 2.2-1.
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 & Vectors, 6, 231.
Olival KJ, Dittmar K, Bai Y *et al.* (2015) *Bartonella* spp. in a Puerto Rican bat community. Journal of Wildlife Diseases, 51, 274-278.
Ossa G, Kramer-Schadt S, Peel AJ, Scharf AK, Voigt CC (2012) The movement ecology of the straw-colored fruit bat, *Eidolon helvum*, in sub-Saharan Africa assessed by stable isotope ratios. PLoS ONE, 7, e45729.
Paziewska A, Harris PD, Zwolińska L, Bajer A, Siński E (2011) Recombination within and between species of the alpha proteobacterium *Bartonella* infecting rodents. Microbial Ecology, 61, 134–145.
Paziewska A, Siński E, Harris PD (2012) Recombination, diversity and allele sharing of infectivity proteins between *Bartonella* species from rodents. Microbial Ecology, 64, 525–536.
Peel AJ, Rossiter SJ, Wood JLN, Cunningham AA, Sargan DR (2010) Characterization of microsatellite loci in the straw-colored fruit bat, *Eidolon helvum* (Pteropodidae). Conservation Genetics Resources, 2, 279–282.
Peel AJ, Sargan DR, Baker KS *et al.* (2013) Continent-wide panmixia of an African fruit bat facilitates transmission of potentially zoonotic viruses. Nature Communications, 4, 2770.
Plowright RK, Field HE, Smith C *et al.* (2008) Reproduction and nutritional stress are risk factors for Hendra virus infection in little red flying foxes (*Pteropus scapulatus*). Proceedings of the Royal Society of London. Series B, Biological Sciences, 275, 861–869.
Plowright RK, Foley P, Field HE *et al.* (2011) Urban habituation, ecological connectivity and epidemic dampening: the emergence of Hendra virus from flying foxes (*Pteropus* spp.). Proceedings of the Royal Society of London. Series B, Biological Sciences, 278, 3703–3712.
Rattanavong S, Fournier P-E, Chu V *et al.* (2014) *Bartonella henselae* endocarditis in Laos – "The unsought will go undetected." PLoS Neglected Tropical Diseases, 8, e3385.
Reeves WK, Loftis AD, Gore JA, Dasch GA (2005) Molecular evidence for novel bartonella species in *Trichobius major* (Diptera: Streblidae) and *Cimex adjunctus* (Hemiptera: Cimicidae) from two southeastern bat caves, USA. Journal of Vector Ecology, 30, 339-341.
Reeves WK, Dowling AP, Dasch GA (2006) Rickettsial agents from parasitic dermanyssoidea (Acari: Mesostigmata). Experimental and Applied Acarology, 38, 181-188.
Reeves WK, Rogers TE, Durden LA, Dasch GA (2007) Association of *Bartonella* with the fleas (Siphonaptera) of rodents and bats using molecular techniques. Journal of Vector Ecology, 32, 118-122.
Richter H, Cumming GS (2008) First application of satellite telemetry to track African straw-coloured fruit bat migration. Journal of Zoology, 275, 172–176.

Seabloom EW, Borer ET, Gross K *et al.* (2015) The community ecology of pathogens: coinfection, coexistence and community composition. Ecology Letters, 18, 401-415.
Shi JJ, Chan LM, Peel AJ *et al.* (2014) A deep divergence time between sister species of *Eidolon* (Pteropodidae) with evidence for widespread panmixia. Acta Chiropterologica, 16, 279–292.
Sørensen U, Halberg K (2001) Mammoth roost of nonbreeding straw-coloured fruit bat *Eidolon helvum* (Kerr, 1792) in Zambia. African Journal of Ecology, 39, 213-215.
Telfer S, Lambin X, Birtles RJ *et al.* (2010) Species interactions in a parasite community drive infection risk in a wildlife population. Science, 330, 243–246.
Thomas D (1983) The annual migrations of three species of West African fruit bats (Chiroptera: Pteropodidae). Canadian Journal of Zoology, 61, 2266-2272.
Veikkolainen V, Vesterinen EJ, Lilley TM, Pulliainen AT (2014) Bats as reservoir hosts of human bacterial pathogen, *Bartonella mayotimonensis*. Emerging Infectious Diseases, 20, 960–967.
Vellend M (2010) Conceptual synthesis in community ecology. The Quarterly Review of Biology, 85, 183–206.
Wirth T, Meyer A, Achtman M (2005) Deciphering host migrations and origins by means of their microbes. Molecular Ecology, 14, 3289-3306.
Witsenburg F, Schneider F, Christe P (2014) Signs of a vector's adaptive choice: on the evasion of infectious hosts and parasite-induced mortality. Oikos, 1–9.

CHAPTER 3

Phylogenetic and geographic constraints on *Bartonella* transmission among bat species

Introduction

Disease ecology, evolutionary biology, and community ecology share important conceptual roots and unifying questions (Seabloom *et al.* 2015). One of the key processes governing the diversity of organisms in any environment is the formation of new species (Vellend 2010). For microparasites like viruses and bacteria, which spend at least part of their life cycle inside of a animal host, persistence within a host is determined by a parasites's adaptation to a particular set of extracellular or intracellular environmental covariates (Parrish *et al.* 2008; Longdon *et al.* 2014). In the short term, parasite adaptation may take the form of specific changes in genes important to invasion and replication within host cells. However, over longer timescales, isolation of a microparasite within a single host species may result in subsequent changes in the rest of the genome.

As parasites switch between species, their phylogenetic similarity may begin to track that of their host species if this process is occurring synchronously with host speciation. Cross-species transmission may then become constrained by host species relatedness, such that parasites adapted to a particular host species will not successfully persist in a phylogenetic distant host species (Streicker *et al.* 2010; Longdon *et al.* 2011; Faria *et al.* 2013). These phylogenetic constraints may also be reinforced by geographic barriers, with distantly related host speices sharing very little geographic overlap and thereby reducing the probability of transmission among divergent host lineages. Thus, a pattern of host-parasite cospeciation may

emerge. For bats, there is evidence that rabies viruses (Hughes *et al.* 2005), coronaviruses (Cui *et al.* 2007), and malarial parasites (Schaer *et al.* 2013, 2015) have diverged along with their host species.

For *Bartonella*, a genus of facultative intracellular bacteria, there is a strong separation between lineages infecting different orders of mammals including rodents, bats, ungulates, carnivores, and marine mammals (Kosoy 2010). The process of adaptation to its primary cell niche (erythrocytes and endothelial cell) in its mammalian host and also adaptation to specific arthropod vectors (fleas, ticks, flies, and mites) are expected to drive the divergence and formation of new *Bartonella* species. Numerous studies have demonstrated that bats show a high prevalence and genetic diversity of *Bartonella* bacteria (Concannon *et al.* 2005; Kosoy *et al.* 2010; Bai *et al.* 2011; Bai *et al.* 2012; Lin *et al.* 2012; Veikkolainen *et al.* 2014; Olival *et al.* 2015; Kamani *et al.* 2014; Bai *et al.* 2015). Given that bats are an evolutionary ancient lineage of mammals (O'Leary *et al.* 2013), the accumulation of parasite diversity may not be surprising; however, there may be a more interesting pattern that reflects deep divergence of *Bartonella* lineages that track the radiation of bat species.

Therefore, I test the hypothesis that bat-*Bartonella* relationships show a strong pattern of cophylogeny. Lei and Olival (2014) found significant phylogenetic congruence between bat species and *Bartonella*. In the intervening time, however, there have been several other articles published identifying novel *Bartonella* genotypes in bat species (Veikkolainen *et al.* 2014; Olival *et al.* 2015; Kamani *et al.* 2014; Bai *et al.* 2015). Furthermore, the study by Lei and Olival did not use published data on *Bartonella* found in ectoparasitic bat flies (Morse *et al.* 2012; Billeter *et al.* 2012). Hence, I compiled a larger dataset to test for congruence between *Bartonella* and their associated bat host species. Like the previous study, I anticipate that there will be a

significant cophylogenetic signal, and with a more diverse sampling of bat families, I expect to find patterns of cophylogeny in *Bartonella* that reflect these higher divisions of bat taxa.

An important confounding factor in the analysis of host-parasite relationships, however, is that host phylogeny may be constrained by geography, reflecting a history of colonization and speciation. That is, closely related species may tend to share more of their geographic ranges than distantly related species. Using a dataset that includes a global selection of bat species, I test the degree to which host species relatedness and sympatry are correlated, an associated that was not explored by Lei and Olival. I predict that this correlation is strong and that sympatry among bat host species will show global congruence with *Bartonella* phylogeny, since bat species sharing geographic space are expected to share parasites more often than allopatric species.

Global tests of phylogenetic congruence are expected to detect that, on average, transmission events reflected in host-parasite relationships that cross large phylogenetic distances are infrequent. However, processes of parasite duplication, extinction, and host switching can distort the overall trend of cospeciation (Page 1994). Therefore, I seek to quantify the number of *Bartonella* host species transitions that cross bat families and suborders with Bayesian phylogenetic analyses. I hypothesize that transition events between bat species will occur more frequently within the same bat family and less frequently for cross-family and cross-order transitions.

Finally, I test additional hypotheses related to the fact that *Bartonella* are vector-borne parasites and that, generally, bacterial infections of bats are poorly studied. First, some *Bartonella* genotypes may be more frequently associated with ectoparasites, potentially as endosymbionts (Billeter *et al.* 2008, 2012; Morse *et al.* 2012; Zhu *et al.* 2014). Thus, their phylogenetic distance from other *Bartonella* genotypes may not reflect the evolutionary history

of bat species. I use tests of individual host-parasite linkages to examine whether *Bartonella* genotypes associated with ectoparasites of bats are less congruent with bat phylogeny than genotypes found directly in bats. Second, I examine the effects of sampling bias on the diversity of *Bartonella* genotypes found in bat species represented in this dataset. For example, the straw-colored fruit bat (*Eidolon helvum*) has been shown to carry six distinct *Bartonella* species, the highest diversity detected in any reservoir species (Bai *et al.* 2015), however this diversity may simply be a function of sampling intensity, as has been seen in viral diversity in bats (Luis *et al.* 2013).

Overall, I am interested in understanding the mechanisms that generate and maintain parasite diversity, a question that is central to disease ecology and is shared with community ecology and evolutionary biology (Kurtenbach *et al.* 2006; Seabloom *et al.* 2015). Specifically, this study will further our understanding regarding how bat-*Bartonella* relationships are formed and the phylogenetic and geographic determinants of cross-species transmission. Our analysis of *Bartonella* phylogeny in relation to host sympatry and evolutionary history may aid in the diagnosis of bartonellosis in humans or domestic animals and help determine the most probable reservoir species.

Materials and methods

Compiled sequence data

Sequence data for this study were first compiled from a previous analysis of bat-*Bartonella* cophylogeny by Lei and Olival (2014). These data include partial citrate synthase gene sequences (*gtlA*) for *Bartonella* genotypes isolated from bats from the UK, Kenya, Guatemala, Peru, and Taiwan (Concannon *et al.* 2005; Kosoy *et al.* 2010; Bai *et al.* 2011; Bai *et al.* 2012; Lin *et al.* 2012). The *gltA* gene has been shown to provide good phylogenetic resolution

among known *Bartonella* species and subspecies and is widely used for detection of *Bartonella* infections (Norman *et al.* 1995). I also included sequences from several recent studies that have isolated additional *Bartonella* sequences (*gltA*) from bats in Finland, Puerto Rico, Nigeria, and several other countries in Africa (Veikkolainen *et al.* 2014; Olival *et al.* 2015; Kamani *et al.* 2014; Bai *et al.* 2015). Other studies have identified *Bartonella* genotypes in ectoparasites from bats, particularly bat flies and fleas, using *gltA* sequences (Morse *et al.* 2012; Billeter *et al.* 2012; Veikkolainen *et al.* 2014). Finally, I searched for additional unpublished sequences on GenBank using the search terms "bat* bartonella" and found *gltA* sequences from *Bartonella* in bats and ectoparasites from Peru, Poland, and Vietnam. From each unique *Bartonella gltA* genotype found on GenBank, I extracted data on the genus and species of the bat host (Table A2.1). For *gltA* genotypes isolated from ectoparasites, I extracted the genus and species of the ectoparasite and the bat host (Table A2.2).

Cytochrome b (*cytb*) gene sequences (Table A2.3) were collected from GenBank for each bat host species; this mitochondrial gene provides good phylogenetic resolution among mammalian species (Kocher *et al.* 1989; Bradley *et al.* 2001; Agnarsson *et al.* 2011). For bats identified only to the genus level or in cases where a suitable *cytb* sequence could not be found, representative or substitute species were chosen (as in Lei and Olival 2014). The criteria for representative and replacements species are discussed in detail in Appendix II, Section B. Sensitivity analysis using alternative suitable replacement bat species suggest that these host substitutions do not alter the observed cophylogenetic patterns. Host bat family and suborder were recorded based on IUCN Red List of Threatened Species (IUCN 2014), the Mammal Species of the World 3rd Edition (Wilson and Reeder 2005), and published articles (Teeling *et al.* 2002; Agnarsson *et al.* 2011) (Table A2.3).

In total, this dataset includes 155 unique *Bartonella* genotypes from 54 bat species, 37 genera, 10 families, and both recognized suborders, Yinpterochiroptera and Yangochiroptera (Teeling *et al.* 2002; Agnarsson *et al.* 2011). To check for evidence of sampling bias in measured diversity of *Bartonella* genotypes from each bat species, I counted the number of sampled bats of each species from the research studies included in the dataset (excluding unpublished sequences) and counted the number of articles published on each species by searching the binomial species name in Web of Science (Table A2.3). Log-transformed host-parasite links were tested for correlation with log-transformed values of sampling effort.

Compiled geographic range data

Shape files for geographic ranges of each bat species were downloaded from the International Union for Conservation of Nature (IUCN) Red List website (http://www.iucnredlist.org/technical-documents/spatial-data) (IUCN 2014). Using the command "over" from the R package "sp" and the commands "gIntersection" and "gArea" from the package "rgeos" for each species in the dataset, I calculated a) if each pair of bats' ranges overlapped, and if they overlapped, b) the area of the intersection between the two ranges (R Core Team 2015; Pebesma 2005; Bivand *et al.* 2013; Bivand and Rundel 2014). A matrix of area overlaps was obtained for these pairwise comparisons. Percent overlap of species ranges was calculated by dividing the area of intersection of each pair of species (*ij*) relative to the total area of each species' range. This creates an asymmetric matrix such that the percent range overlap of species *i* and species *j* is not equal to the percent range overlap of species *j* and species *i*.

Phylogenetic analysis of sequence data

Lengths of *gltA* sequence isolates varied considerably in the *Bartonella* dataset, so sequence lengths were trimmed to 334 base pairs covered by all of the isolates. The total length

of *cytb* sequences in the bat species dataset was 1140 base pairs. *Brucella melitensis* AM040264 was chosen as the outgroup for the *Bartonella* phylogeny and the duck-billed platypus, *Ornithorynchus anatinus* HQ379928, was chosen as the outgroup for the bat phylogeny (Lei and Olival 2014). Sequences were aligned with MAFFT using the G-INS-I method (Katoh and Standley 2013). Maximum likelihood (ML) phylogenetic trees were generated with MEGA6 (Tamura *et al.* 2013) using the generalized time reversible substitution model (Nei and Kumar 2000) with five gamma categories (GTR+G). Support for nodes in the tree was estimated from 1000 bootstrap replicates. To illustrate bat-*Bartonella* linkages, tanglegrams were drawn on ML trees using the "cophyloplot" command in the "ape" package in R (R Core Team 2015; Paradis *et al.* 2004).

Correlation between bat phylogeny and sympatry

A Mantel test (Mantel 1967) was used to find the correlation between the two matrices, bat phylogenetic distance and bat geographic range overlap. First, distances were calculated from branch lengths of the ML tree (patristic distances) of bat species using the "cophenetic" function in the "ape" package in R. (R Core Team 2015; Paradis *et al.* 2004). Second, the geographic range overlap matrix was transformed into a distance matrix with the "dist" function in the R "stats" package using the maximum distance method. This calculation takes the minimum percent overlap between each pair of species (by dividing the overlap area by the largest range size of the pair) and subtracts the percentage from one. This creates a symmetrical matrix that can used in the global fit tests. Thus, like phylogenetic distances where closely related species have low distance values, species with highly overlapping ranges have low geographic distance values. The "mantel" command in the "vegan" package in R (R Core Team 2015; Oksanen *et al.* 2015) was used calculate the correlation between the matrices using 10000 permutations.

Tests of cophylogeny

Cophylogenetic analyses were performed using several complementary approaches, specifically global fit tests and partitioned Bayesian phylogenetic trees. Global fit methods account for two confounding factors: some bat species host multiple *Bartonella* genotypes and some *Bartonella* genotypes are linked with multiple bat species. Bayesian phylogenetic analyses were used to reconstruct changes in host bat traits over the topology of the *Bartonella* and bat phylogenies.

Global fit tests

Global fit analyses were first performed on the ML trees of bat species and *Bartonella* genotypes. Two patristic distance matrices were calculated from bat and *Bartonella* trees using the "cophenetic" command in the "ape" package in R (R Core Team 2015; Paradis *et al.* 2004). A third matrix was generated for host-parasite links, which allows for multiple linkages among bat species and *Bartonella* genotypes. Two methods were used to measure the fit between bat and *Bartonella* tree topologies through the matrix of host-parasite linkages, the distance-based ParaFit (Legendre *et al.* 2002) and the Procrustean Approach to Cophylogeny (PACo) (Balbuena *et al.* 2013). Both tests were implemented using the "ape" and "vegan" packages in R (R Core Team 2015; Paradis *et al.* 2004; Oksanen *et al.* 2015) with 10000 permutations. ParaFit tests the overall congruence between host and parasite topologies using only the patristic distance matrices. PACo uses Procrustean superimposition, wherein the host and parasite distance matrices are converted into two-dimensional ordinations and the parasite ordination is rotated to fit the host ordination. In this way, PACo explicitly tests the degree to which parasite phylogeny depends on the host phylogeny and is considered a more conservative test than ParaFit (Balbuena *et al.* 2013). Residual values from PACo were saved to quantify the number of significant

linkages among bats and *Bartonella* genotypes. These values were used to test the hypothesis that genotypes found in ectoparasites do not significantly diverge from the overall cophylogenetic trend.

Based on my hypothesis that bat phylogeny and bat geographic range overlap are correlated, I test the degree to which *Bartonella* phylogeny is supported by host species range overlap. ParaFit and PACo tests were used on the patristic distance matrix of *Bartonella* genotypes, the distance matrix of geographic overlaps and the matrix of host-parasite linkages. Because patristic distance and geographic overlap matrices are calculated with different methods, global fit values from the separate analyses are not directly comparable. Hence, I only make inference on the significance of the global fit tests, not the magnitude of any global fit values.

Bayesian phylogenetic analysis and reconstruction of host switches

Following a previous study reconstructing host switching events among *Bartonella* genotypes in rats (Hayman *et al.* 2013), Bayesian Markov chain Monte Carlo (MCMC) analysis of *Bartonella* sequence data from bats was performed using BEAST 1.8.2 (Drummond *et al.* 2012). The GTR+G substitution model with 5 gamma categories (Nei and Kumar 2000) was used for the MAFFT alignments (Katoh and Standley 2013) of *Bartonella gltA* sequences. Base frequencies were estimated from the data and nodes of the tree were estimated using substitutions per site. The population sizes of *Bartonella* genotypes were assumed to be constant for the coalescent model. Sequences were assigned discrete traits based on the family of the host bat (10 families), the suborder of the host bat (Yangochiroptera and Yinpterochiroptera), and the region in which the host bat was captured (Africa, Europe, North America/Caribbean, South America, and Southeast Asia). BEAST independently estimates the rate of these discrete state transitions across the topology of trees generated from *Bartonella* sequence data. Starting with a

prior value of one, the clock rate for each discrete state is estimated from the average number of state transitions across all nodes in the phylogenetic tree. Individual family, suborder, and region transition rates were estimated, starting with a diffuse gamma prior distribution with shape and scale parameters set to one and an initial value of one. All other priors for nucleotide frequencies and substitution rates were kept at the default, diffuse settings.

A MCMC chain length of 120×10^6 iterations was chosen for the analysis, sampling every 12000 iterations to ensure that the effective sample sizes (ESS) of all parameters was >200. Tracer 1.6.1 (University of Edinburgh, UK) was used to assess the mixing and convergence of parameters. Following the completion of the MCMC analysis, the first 10% of maximum clade credibility (MCC) trees were discarded as burn-in using TreeAnnotator (available at http://tree.bio.ed.ac.uk/software/). A second, identical Bayesian MCMC analysis was performed on bat *cytb* sequences, using the discrete traits (bat host family, suborder, and capture region) to compare how the states transitioned across the bat tree topology versus the *Bartonella* tree topology.

Gamma-distributed discrete state transition rates were estimated from the posterior of the MCMC chains. I inspected the median and 95% highest probability density (HPD) interval of each rate to find families, suborders, and geographic regions that had a significant number of exchanges over the topology of the phylogeny. Clock rates, or the mean number of transitions across all nodes, for each state were also inspected to quantify the overall trend in exchanges among bat families, suborders, and regions. All posterior state transition rates and tree likelihoods were extracted using the program Tracer 1.6. Finally, to corroborate the results of the cophylogenetic tests performed using ML trees, global fit analyses (ParaFit, PACo) were

repeated using the Bayesian phylogenies of bat species and *Bartonella* genotypes, as well as the bat species geographic overlap matrix and the Bayesian tree of *Bartonella* genotypes.

Results

Phylogenetic analysis

The maximum likelihood (ML) tree of bat *cytb* sequences (Figure A2.1) matches well with previous phylogenies of bats, with good support (>50%) at the level of individual families (Teeling *et al.* 2002; Agnarsson *et al.* 2011). However, deeper nodes at the level of suborders (Yinpterochiroptera and Yangochiroptera) were poorly resolved. The ML tree of *Bartonella gltA* sequences (Figure A2.2) also had good support (>50%) for closely related genotypes and at the putative species level (<5% sequence divergence) (La Scola *et al.* 2003), but deeper nodes had low support. The tanglegram linking bat species and *Bartonella* genotypes (Figure A2.3) clearly illustrates the pattern of multiple host-parasite associations, which supports my decision to use global fit tests. Overall, there does not appear to be an obvious congruence between bat and *Bartonella* topologies, however manual rearrangement of branches was difficult and the global fit tests should be able to detect any overall association trends.

Correlation between bat phylogeny and sympatry

The map of species distributions (Figure 3.1) indicate that there is a high level of range overlap (indicated by darker shading) among bats in the dataset, particularly within North America and the Caribbean, South America, Europe, Africa, and Southeast Asia. The Mantel test (Mantel 1967) shows that matrices of bat phylogenetic distances and geographic range overlaps are significantly correlated (Pearson correlation coefficient = 0.34, P = 1E-5). Given this result, I performed global fit analyses on both host phylogenetic distance and host geographic overlap to test the degree to which these covariates predict *Bartonella* phylogeny.

Global fit tests

ParaFit and PACo analyses provided strong support for a cophylogenetic relationship between *Bartonella* and bats in the dataset (ParaFitGlobal = 16.2, P = 1E-5; m^2 global value = 11.3, P = 1E-4). The majority of bat-*Bartonella* links (90/158) had residual values below the overall mean, indicating that host species phylogeny strongly predicts the associate parasite phylogeny. Sixteen and eleven linkages showed residual values greater than the mean (μ) plus 1x and 1.5x the interquartile range (IQR) of residuals, respectively, indicating that these host-parasite relationships are outliers in the overall cophylogenetic trend.

Global fit analyses using bat host species geographic overlaps also showed a signficant trend (ParaFitGlobal = 74.6, P = 1E-5; m^2 global value = 45.0, P = 1E-4). Eighty-one out of the 158 overlap-*Bartonella* linkages had residual values below the overall mean. Fourteen linkages showed residuals greater than μ+IQR, and five linkages showed residuals greater than μ+1.5xIQR. These results indicate that the degree of overlap among bat species ranges is also a strong predictor of *Bartonella* phylogeny, suggesting that related *Bartonella* genotypes are more likely to be shared among sympatric bats. However, it is important to note that sympatry and bat host phylogeny are correlated, so the effects of these two covariates on *Bartonella* phylogeny cannot be completely separated.

Procrustes superimposition plots (Balbuena *et al.* 2013) of *Bartonella* genotype and bat host phylogenetic distance ordinations from PACo indicate that there is good separation among bat species at the level of families and suborders using the *cytb* gene (Figure A2.4). There is a large amount of phylogenetic overlap among bats from Southeast Asia and Africa and among bats from North America, the Caribbean, and South America. For ordinations of *Bartonella* genotypes using the *gltA* gene (Figure A2.5), the limited amount of sequence information in only

334 base pairs prevents clear separation of genotypes in ordination space based on host bat family, suborder, or geographic region. Finally, ordination of geographic range overlaps demonstrates poor separation of bat families and suborders, but provides better separation of bat species across regions (Figure A2.6). This separation is most obvious among bats from Europe, Africa, and Southeast Asia, yet there is still a large amount of overlap among bats from North America, the Caribbean, and South America.

Bayesian phylogenetic analysis and reconstruction of host switches

The Bayesian phylogenetic analyses of *Bartonella gltA* genotypes and bat *cytb* sequences yielded trees (Figures 3.2-3.5; Figures A2.7-A2.8) with good convergence and large effective sample sizes (ESS>200) for all parameters. There was strong posterior probability (PP > 80%) across the topology of these trees, even for deeper nodes (Figures 3.2-3.5; Figures A2.7-A2.8). Overall tree likelihoods (-ln) were 23198.7 (ESS = 6835) and 8192.8 (ESS = 4398) for bat and *Bartonella* phylogenies, respectively. Coloring the branches of the bat and *Bartonella* trees according to bat host family reveals a qualitative fit between the two trees in terms of their topology and the formation of distinct clades corresponding to bat superfamilies (Figures 3.2-3.3). However, there are five *Bartonella* genotypes (KP100353, KP100358, JN172066, KP100343, and KP100346) at the base of the tree that do not neatly fit into these groups (Figure 3.3). These taxonomic similarities extend to the level of bat suborders, with bacterial clades separated into the Yinpterochiroptera and Yangochiroptera (Figure 3.4; Figure A2.7). Both trees show that the Yangochiroptera clades are polyphyletic, which conflicts with other phylogenies (Agnarsson *et al.* 2011). The Bayesian analysis was also run using the Hasegawa, Kishino, and Tano (HKY) substition model and this polyphyly was not resolved (not shown), hence it is likely an issue attributable to insufficient data in the short sequences used to make the trees. Finally,

there appears to be a clear division between Old World and New World *Bartonella* genotypes (Figure 3.5; Figure A2.8) that is mirrored in the bat phylogeny. Nevertheless, there is some exchange between *Bartonella* genotypes in the Vespertilionoidea clade across Europe, North America, and South America.

Extracted posterior estimates of state transitions among families, suborders, and geographic regions were generally low; Table 3.1 shows only the transition rates with a median value greater than one. All of the family transition rates listed in Table 3.1 are between pairs of families within the same suborder and five out of six were in the same superfamily (Figure 3.3). However, only one pair of families, Pteropodidae and Rhinolophiae, showed exchange rates signficantly higher than one (median = 4.2) when compared to the prior gamma distribution with a starting value of one. This trend is reflected in the family clock rate, which estimates that only 1.6 cross-family transitions occur on average across the 155 *Bartonella* genotypes. The suborder clock rate is lower than the prior expectation of one, estimating only 0.3 cross-suborder transitions across the tree. Exchanges between the two suborders do occur, as can be seen in Figure 3.4. However, the median number of exchanges is 1.4, which is not significantly greater than the prior expectation of one transition. There is a significant amount of exchange between several geographic regions, particularly between Africa-Southeast Asia and North America/Caribbean-South America, which had median numbers of transitions (3.4 and 4.3, respectively) signficantly greater than one.

Finally, repeated global fit analyses yielded similar results to the tests using ML trees, with strong support for a cophylogenetic relationship between bats and *Bartonella* (ParaFitGlobal = 401.2, P = 1E-5; m^2 global value = 219.0, P = 1E-4). The tanglegram associating bat species to *Bartonella* genotypes using Bayesian analysis (Figure 3.6) better

illustrates the topological congruence between the two phylogenies. The strong correlation between bat phylogeny and geographic range overlap remained (Mantel test, Pearson correlation coefficient = 0.41, P = 1E-5) and the relationship between host species sympatry and *Bartonella* phylogeny was significant (ParaFitGlobal = 49.0, P = 1E-5; m^2 global value = 40.7, P = 1E-4). Eighty-nine of 158 PACo residual values from the phylogenetic comparison were below the mean of residual values, with 18 and 11 residuals above μ+IQR and μ+1.5xIQR, respectively. Comparison between residuals for bat-borne versus ectoparasite-borne *Bartonella* genotypes shows no difference in fit to bat phylogeny for these groups (Kolmogorov-Smirnov D = -0.0386, P = 0.97). Residual values for the geographic range overlap comparison comprised 99 values below the mean (χ^2 = 9.6266, P = 0.001), 17 above μ+IQR, and 8 above μ+1.5xIQR. Again, there was no difference in residuals for bat and ectoparasites for fit between geographic overlap to bat phylogeny (Kolmogorov-Smirnov D = 0.1541, P = 0.36). The outliers in the global fit analyses reflect either cross-family transitions or associations between bats and *Bartonella* genotypes that are very distant from other clades. For instance, *Hipposideros larvatus* from Kenya appears to be carrying a genotype (KP100355) related to *Bartonella* associated with the families Miniopteridae and Vespertilionidae. *Pteronotus davyi* appears to carry several genotypes (HM597202, HM597205, and KX416248) that are associated with phyllostomid bats. The five other outliers (KP100353, KP100358, JN172066, KP100343, and KP100346) are basal lineages associated with *Eidolon helvum* and *Rhinolophus* spp.

Sampling bias

I looked for the presence of bias in the dataset by testing the correlation between the number of host-parasite links and sampling effort. Specifically, I used two measures of sampling effort: the number of published articles on each bat species found on Web of Science and the

total sample size of each bat species tested in the individual studies which contribute to the dataset. There was a significant log-log correlation between the number of host-parasite links and Web of Science articles (Pearson correlation coefficient = 0.32, P = 0.019), even when the bat with the highest number of articles (*Myotis myotis*, 2751 articles) was removed (Pearson correlation coefficient = 0.31, P = 0.026). There was a significantly positive log-log correlation between the number of links and individual species sample sizes (Pearson correlation coefficient = 0.66, P = 5.4E-7). This correlation was still significant when the species with the highest number of host-parasite links (*Eidolon helvum*, 53 links) was removed (Pearson correlation coefficient = 0.58, P = 2.5E-5). This suggests that high levels of *Bartonella* diversity found in several bat species are probably due in part to sampling bias, although this does not completely exclude the effects of ecological and evolutionary processes that may increase *Bartonella* diversity in some bats.

Discussion

Our analysis of *Bartonella* genotypes infecting bats supports previous work by Lei and Olival (2014) that found significant congruence between bat and *Bartonella* phylogenies, indicating a general cophylogenetic trend. The dataset I used included a larger number of *Bartonella gltA* sequences from bats and bat ectoparasites representing a greater number of families and from more regions than this previous study, including *Bartonella* sequences from ectoparasitic bat flies, fleas, and mites (Billeter *et al.* 2012; Morse *et al.* 2012; Veikkolainen *et al.* 2014). The inclusion of more sequences could have easily diluted the cophylogenetic signal observed previously, especially if many of the host-parasite associations arose from apparent host-switching events over large phylogenetic distances. Yet my analysis shows that this overall congruence between bats and *Bartonella* is robust to the new sequences and perhaps even

enhanced. Figure 3.3 shows a clear visual congruence between bat species and *Bartonella* genotypes when the branches are colored by the host families. The formation of distinct and well-supported clades of *Bartonella* genotypes linking families, superfamilies, and suborders of bats (Figure 3.4; Figure A2.7) suggest that *Bartonella* have been co-diverging with bats over significant evolutionary time. Lei and Olival did not explore this pattern previously, which may have been due to low representation of bat families in their dataset.

Despite the clear overall trend, just over half of all host-parasite linkages were found to be significant, which is the same result seen by Lei and Olival. It should be noted however that I use the more conservative tests of host-parasite linkages developed by Balbuena *et al.* (2013). Trees generated from Bayesian inference clearly showed that the most significant outliers from the global fit analyses using ParaFit (Legendre *et al.* 2002) and PACo (Balbuena *et al.* 2013) were produced by links between bats and *Bartonella* genotypes associated with a different family of bats than the apparent host species, or by *Bartonella* genotypes at the base of the phylogenetic tree with uncertain relationships with other sequence types. In fact, these outliers branch from the outgroup (*Brucella melitensis*) deeper than any of the *Bartonella* genotypes associated with particular bat families. One possibility is that these are symbiotic *Bartonella* genotypes primarily adapted to the ectoparasite vector (Morse *et al.* 2012; Zhu *et al.* 2014), and the presence of the bacteria in the bat speces is accidental and/or transient. Lei and Olival did not include bat ectoparasites in their analysis, which prevented them from separating bat-adapted and ectoparasite-adapted genotypes. However, the test of residuals for *Bartonella* genotypes isolated from bats versus ectoparasites showed that, on average, there was no difference in the fit between these two groups to host phylogeny. Thus, if there are *Bartonella* genotypes strictly

endosymbiotic in their arthropod hosts in this dataset, they are infrequent enough that they do not significantly skew the overall trend of co-divergence among bats and *Bartonella*.

An important confounding factor in the study of host-parasite relationships is the influence of geography, specifically the correlation between host species relatedness and sympatry. If two bat species are closely related and also have a high degree of range overlap, it might be expected that these two bats would share similar parasites. High amounts of sympatry and interaction at common roosts may be able to facilitate cross-species transmissions despite phylogenetic barriers. The previous analysis by Lei and Olival did not explore the interaction of bat sympatry and phylogeny. I find that bats in the dataset showing a large amount of geographic overlap are more likely to be related to each other. Therefore, I repeated the global fit analyses using bat sympatry to match with *Bartonella* phylogeny and find a high degree of congruence. I conclude that *Bartonella* phylogeny is strongly predicted by both host phylogeny and degree of range overlap among species. Unfortunately, the different dimensions of the host and parasite matrices prevented me from directly testing the effect of an interaction between host phylogeny and sympatry. A follow-up analysis using a pruned dataset or other methods to test for spatial autocorrelation would help to clarify how host-parasite relationships are structured.

I also explicitly tested my hypothesis that host-switching events would be constrained by phylogenetic distance by modeling transitions between bat families and suborders, and across geographic regions. Again, this was not directly explored by Lei and Olival's study. Our Bayesian trees clearly show that the vast majority of host switching and parasite duplication events occur within the same bat family and that transitions between families and suborders happen infrequently. Specifically, my analysis estimates that only 1.6 cross-family transitions and 0.3 cross-order transitions occur across all nodes of the tree. These rates are in strong

contrast to the number of within-species duplications or cross-species transitions that occur within the same family, with 144 total occurrences in the consensus tree. These results support the expectation that transitions of *Bartonella* between bat host species would be constrained by host relatedness, as has been demonstrated for bat rabies (Streicker *et al.* 2010; Faria *et al.* 2013). Nevertheless, the dataset of *Bartonella* sequences is still limited, so my estimations of transmission rates across phylogenetic scales may be inaccurate. Transitions between geographic regions happen more frequently, with an average of 2.1 region transitions across all nodes of the tree (Figure 3.5). The regions contributing most to this rate are exchanges between North America, the Caribbean Islands, and South America, as well as Africa and Southeast Asia. There is a high level of sympatry among phyllostomid bats represented in the dataset from the Americas, with numerous species having ranges that span from the entire region. Hence, the interaction between many closely related host species in sympatry would be expected to facilitate transmission of *Bartonella* across species boundaries. For Africa and Southeast Asia, the apparent exchanges of *Bartonella* across species is difficult to explain geographically. There are no obvious bridge species that would connect these two regions in the dataset (Figure 3.1), however it is possible that a bridge species exists and has not yet been sampled.

Our analyses captured some very general trends in the evolution of bats and *Bartonella*, but there are still substantial gaps in our understanding of the mechanisms that contribute to this pattern. These gaps may begin to be closed with the acquisition of new *Bartonella* sequences from other bat species and other regions. The 54 species used in this study represent less than 5% of the ~1240 species of bats worldwide, with sampling from only 24 (12%) of 196 countries. Figure 3.1 highlights some of these geographic deficiencies, particularly Australia and the Pacific Islands, Central and East Asia, the Middle East, and North America. Our test of bias in

research effort indicates that we have probably only scratched the surface of *Bartonella* diversity in bats, even within individual species.

Another important gap in the study of bat-*Bartonella* relationships is the limited amount of information contained within the citrate synthase gene (*gltA*), the most popular marker used for the detection of *Bartonella*. The short sequence length prevents me from resolving the position of many branches across the *Bartonella* phylogenetic tree or measuring mutation rates for the estimation of divergence times (Hayman *et al.* 2013). Estimated divergence times for clades of *Bartonella* genotypes would have been especially useful in my analysis to compare with published bat phylogenies, to see if host species and parasite genotypes began to radiate at the same time. However, to estimate divergence times we would need more sequence information, perhaps in the form of multi-locus sequence typing (MLST) or whole-genome sequencing (WGS) datasets. MLST or WGS datasets could also measure the frequency of lateral gene transfer (LGT) and recombination events which could confound patterns of cophylogeny. For example, some of the apparent host-switching events may not represent invasion by an entirely separate genotype of *Bartonella*, but rather just the *gltA* gene that has undergone homologous recombination into a separate genome after coinfection of two genotypes within an individual mammalian or arthropod host. Recent studies have shown that rates of LGT and recombination in *Bartonella* are higher than previously expected given its intracellular lifestyle (Vos and Didelot 2009; Berglund *et al.* 2009, 2010; Paziewska *et al.* 2011, 2012; Buffet *et al.* 2013b; Bai *et al.* 2015). Therefore, sequencing of multiple genomic regions or genes related to the host cell invasion process may be more informative for showing fine-scale differences among *Bartonella* that better reflect their transmission history.

The study of bat-*Bartonella* evolutionary relationships, and by extension host-parasite relationships generally, is not only interesting from a biological perspective, but can also aid in the identification of zoonoses in humans and domestic animals. For instance, Lin *et al.* (2012) saw that *gltA* isolates from *Miniopterus schreibersii* bats were 96% similar to isolates found in stray dogs in Thailand by Bai *et al.* (2010), suggesting potential spillover. In 2014, Veikkolainen *et al.* found sequences in vespertilionid bats that were very similar to *Bartonella mayotimonensis*, a novel agent of endocarditis in a human patient from the United States (Lin *et al.* 2010). Numerous other cases of human and animal bartonellosis have been ultimately attributed to zoonotic origin. Studying how these *Bartonella* evolve and persist in their reservoir species may help to understand the mechanisms that facilitate emergence in novel host species and cause disease. The specific methods used in this study are particularly useful for diverse and rapidly evolving microparasites like bacteria and viruses. Application to other systems could reveal general mechanisms of host-parasite evolution and discover deep relationships at the root of some of our most destructive infectious diseases.

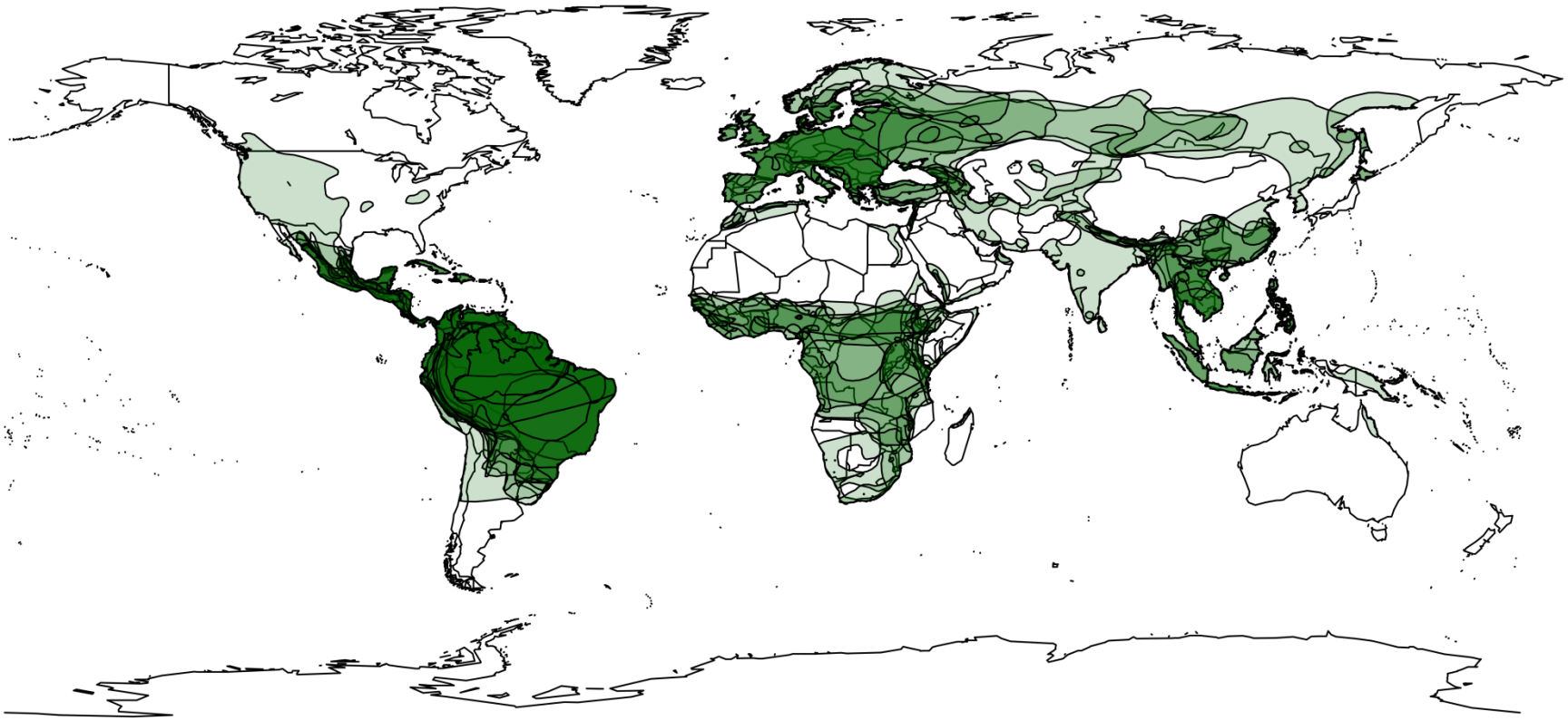

Figure 3.1 Geographic distributions of bat species represented in the study. Darker regions show high levels of range overlap among sampled species, particularly within North and South America, Europe, Africa, and Southeast Asia.

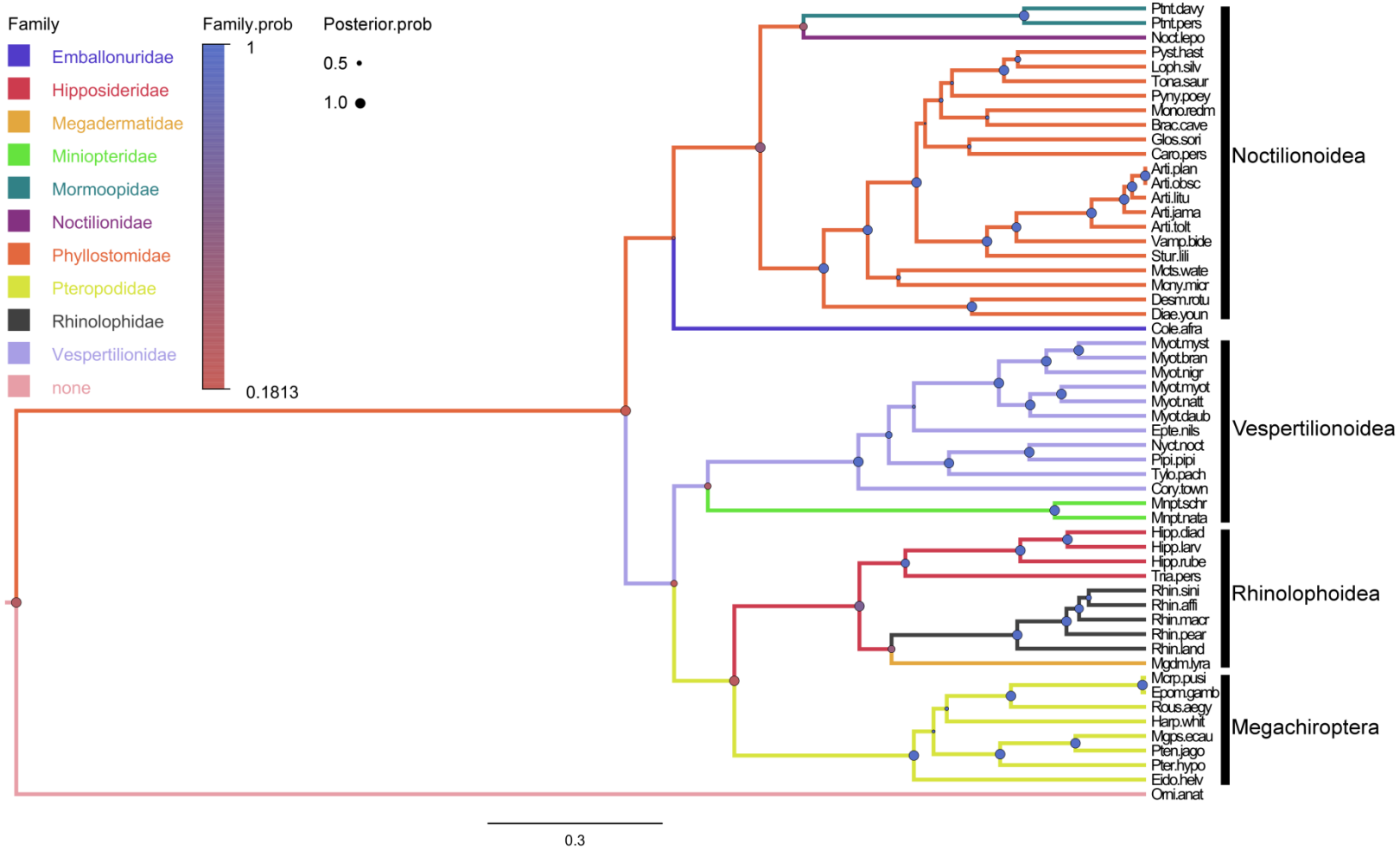

Figure 3.2 Bayesian phylogeny of bat host species reconstructing bat families, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of bat cytochrome *b* (*cytb*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the family at that node (state.prob). Clades of bat families are grouped by recognized superfamilies (Noctilionoidea, Vespertilionoidea, and Rhinolophoidea) and the megabats (Megachiroptera). Mean tree likelihood (-ln) = 23077.7, ESS = 7073; mean family tree likelihood (-ln) = 47.7, ESS = 8054. Details on tip labels for bat species are listed in Table A2.3.

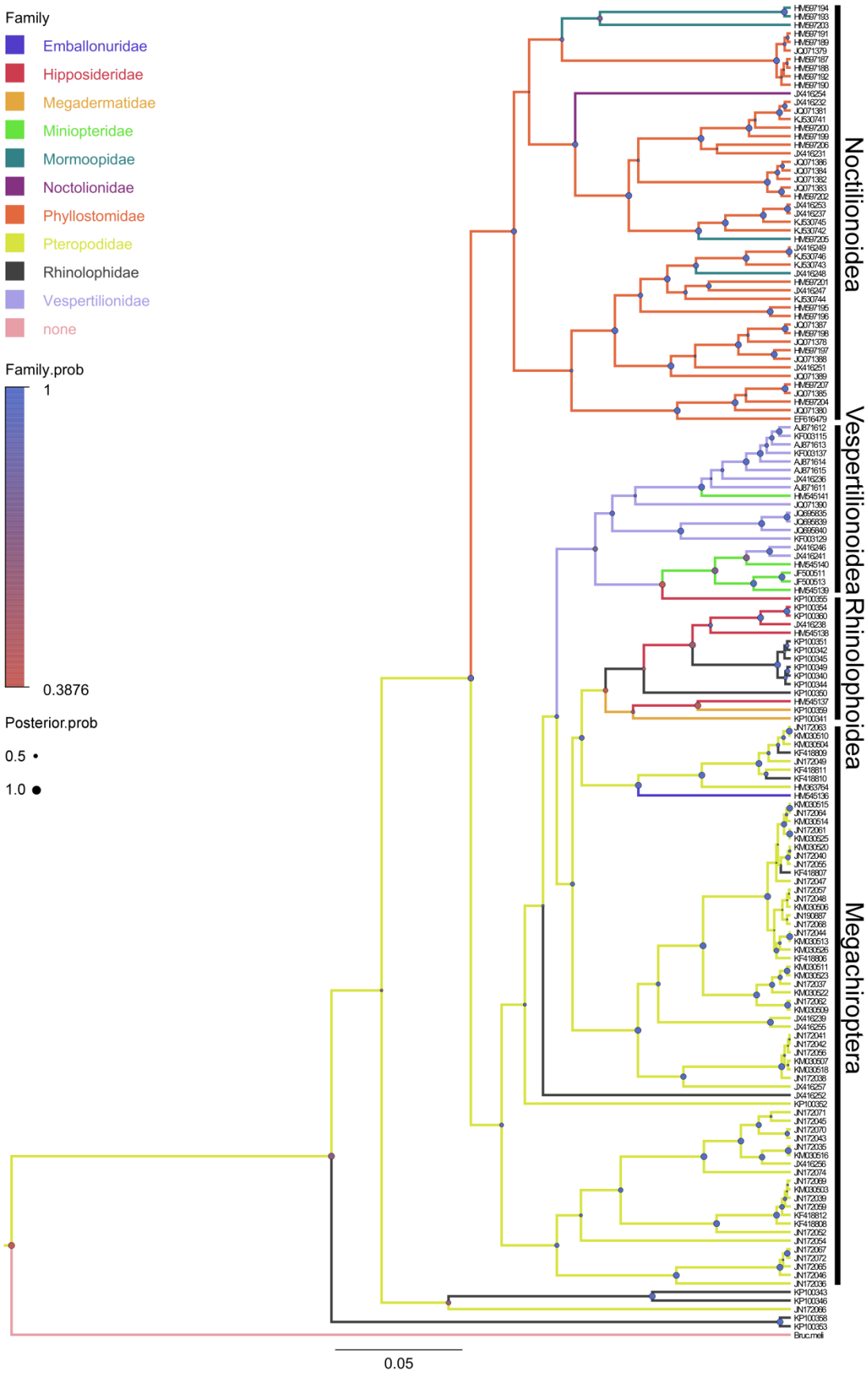
Family
Emballonuridae
Hipposideridae
Megadermatidae
Miniopteridae
Mormoopidae
Noctolionidae
Phyllostomidae
Pteropodidae
Rhinolophidae
Vespertilionidae
none
Family.prob
1
0.3876
Posterior.prob
0.5
1.0
Noctilionoidea
Vespertilionoidea
Rhinolophoidea
Megachiroptera
0.05

Figure 3.3 Bayesian phylogeny of *Bartonella* genotypes reconstructing bat host families, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of *Bartonella* citrate synthase (*gltA*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the bat host family at that node (state.prob). Clades of *Bartonella* genotypes are grouped by recognized bat superfamilies (Noctilionoidea, Vespertilionoidea, and Rhinolophoidea) and the megabats (Megachiroptera). Maximum tree likelihood (-ln) = 7941.6, ESS = 1630; maximum family tree likelihood (-ln) = 108.1, ESS = 4861. Details on tip labels for *Bartonella* genotypes and associated host species are listed in Table A2.1 and A2.2.

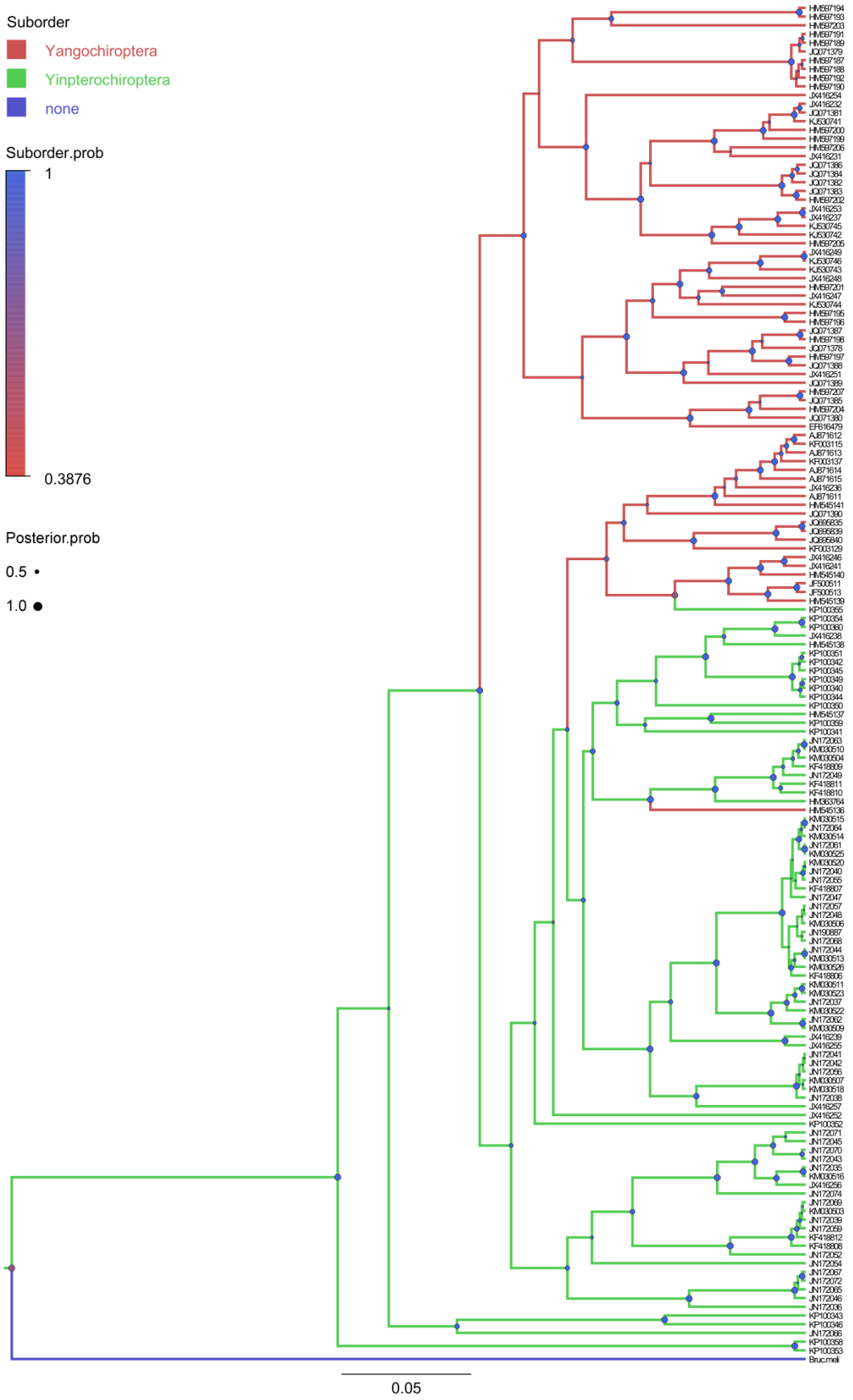
Suborder
Yangochiroptera
Yinpterochiroptera
none
Suborder.prob
1
0.3876
Posterior.prob
0.5
1.0
HM597194
HM597193
HM597203
HM597191
HM597189
JQ071379
HM597187
HM597188
HM597192
HM597190
JX416254
JX416232
JQ071381
KJ530741
HM597200
HM597199
HM597206
JX416231
JQ071386
JQ071384
JQ071382
JQ071383
HM597202
JX416253
JX416237
KJ530745
KJ530742
HM597205
JX416249
KJ530746
KJ530743
JX416248
HM597201
JX416247
KJ530744
HM597195
HM597196
JQ071387
HM597198
JQ071378
HM597197
JQ071388
JX416251
JQ071389
HM597207
JQ071385
HM597204
JQ071380
EF616479
AJ871612
KF003115
AJ871613
KF003137
AJ871614
AJ871615
JX416236
AJ871611
HM545141
JQ071390
JQ695835
JQ695839
JQ695840
KF003129
JX416246
JX416241
HM545140
JF500511
JF500513
HM545139
KP100355
KP100354
KP100360
JX416238
HM545138
KP100351
KP100342
KP100345
KP100349
KP100340
KP100344
KP100350
HM545137
KP100359
KP100341
JN172063
KM030610
KM030604
KF418809
JN172049
KF418811
KF418810
HM363764
HM545136
KM030615
JN172064
KM030614
JN172061
KM030625
KM030620
JN172040
JN172055
KF418807
JN172047
JN172057
JN172048
KM030606
JN190887
JN172068
JN172044
KM030513
KM030526
KF418806
KM030511
KM030523
JN172037
KM030622
JN172062
KM030609
JX416239
JX416255
JN172041
JN172042
JN172066
KM030607
KM030618
JN172038
JX416257
JX416252
KP100352
JN172071
JN172045
JN172070
JN172043
JN172035
KM030516
JX416256
JN172074
JN172069
KM030603
JN172039
JN172059
KF418812
KF418808
JN172052
JN172054
JN172067
JN172072
JN172065
JN172046
JN172036
KP100343
KP100346
JN172066
KP100358
KP100353
Bruc.mel
0.05

Figure 3.4 Bayesian phylogeny of *Bartonella* genotypes reconstructing bat host suborders, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of *Bartonella* citrate synthase (*gltA*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the bat host suborder at that node (state.prob). Suborders are based on current taxonomic classifications for bats (Teeling *et al.* 2002; Agnarsson *et al.* 2011). Maximum tree likelihood (-ln) = 7941.6, ESS = 1630; maximum suborder tree likelihood (-ln) = 24.1, ESS = 251. Details on tip labels for *Bartonella* genotypes and associated host species are listed in Table A2.1 and A2.2

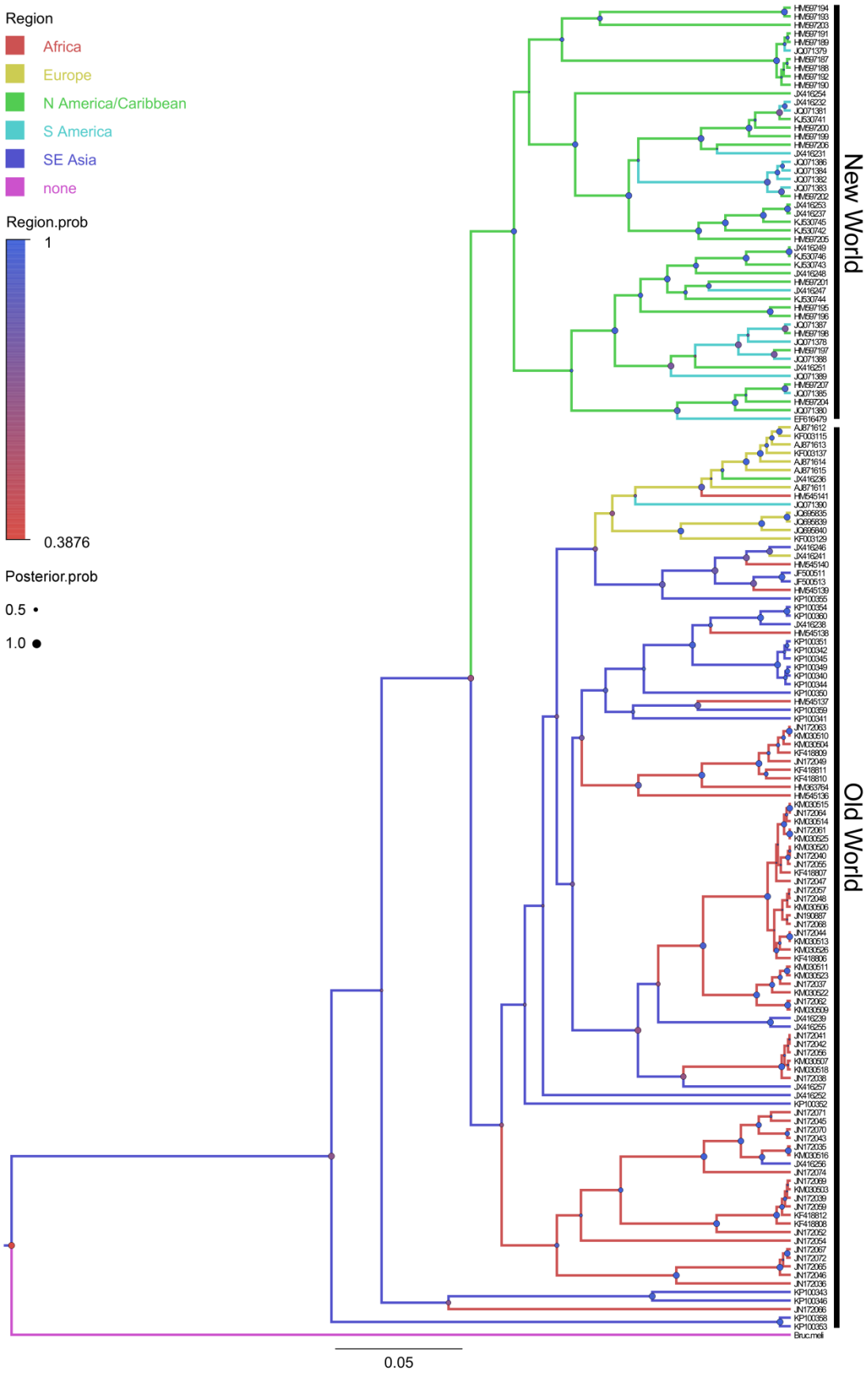
Region
Africa
Europe
N America/Caribbean
S America
SE Asia
none
Region.prob
1
0.3876
Posterior.prob
0.5
1.0
New World
Old World
0.05

Figure 3.5 Bayesian phylogeny of *Bartonella* genotypes reconstructing bat host geographic regions, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of *Bartonella* citrate synthase (*gltA*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the bat host geographic region at that node (state.prob). Clades of *Bartonella* genotypes are separated into Old World and New World groups. Note the geographic region represents where the bat host was captured, which may not reflect its total range. Maximum tree likelihood (-ln) = 7941.6, ESS = 1630; maximum region tree likelihood (-ln) = 119, ESS = 2856. Details on tip labels for *Bartonella* genotypes and associated host species are listed in Table A2.1 and A2.2

Table 3.1 Posterior state transition rate estimates from the Bayesian analysis of *Bartonella gltA* sequences, with data partitions for bat host family, suborder, and geographic region. Only transition rates with a median rate greater than one are shown, indicating that at least one state transition happened between the listed groups. Probability estimates indicate the likelihood of the median number of transition occurring since the time of the common ancestor of the 155 genotypes, as tested against a null gamma distribution. Underlined probability values are statistically significant ($\alpha < 0.05$). Clock rates reflect the mean number of state transitions occurring across all nodes of the tree.

States	Median rate	95% HPD interval	Probability
Family transitions			
Hipposideridae-Megadermatidae	1.5	(5.3E-4, 4.4)	0.21
Hipposideridae-Rhinolophidae	1.5	(2.3E-3, 4.2)	0.22
Miniopteridae-Vespertilionidae	2.5	(0.34, 5.6)	0.085
Mormoopidae-Phyllostomidae	2.6	(0.50, 5.6)	0.074
Noctolionidae-Phyllostomidae	1.0	(2.1E-4, 3.0)	0.36
Pteropodidae-Rhinolophidae	4.2	(1.3, 8.0)	<u>0.015</u>
Family clock rate	1.6	(0.93, 2.3)	
Suborder transitions			
Yangochiroptera-Yinpterochiroptera	1.4	(0.074, 3.9)	0.24
Suborder clock rate	0.31	(0.078, 0.63)	
Region transitions			
Africa-SE Asia	3.4	(1.2, 6.2)	<u>0.035</u>
N America/Caribbean-S America	4.3	(1.7, 7.7)	<u>0.013</u>
Region clock rate	2.1	(1.3, 3.0)	

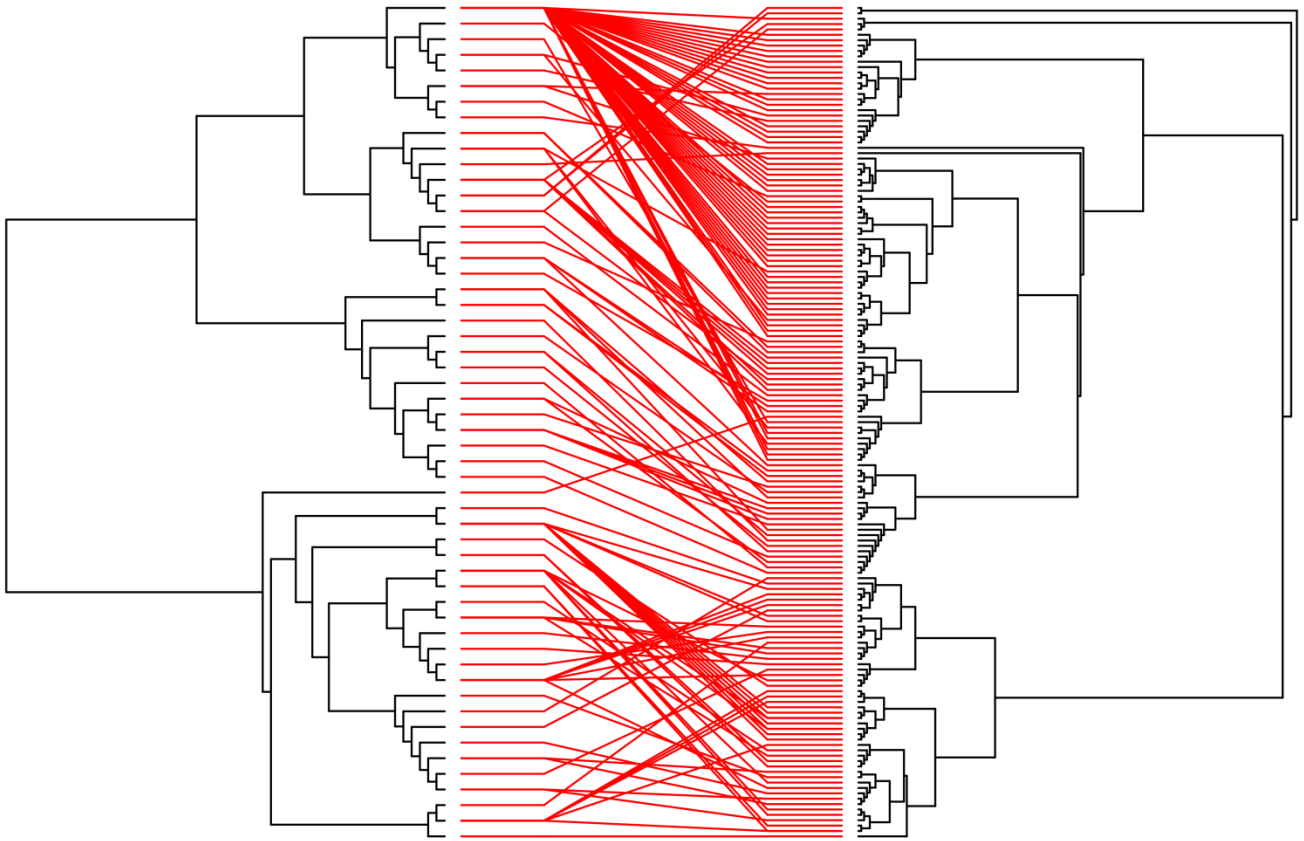

Figure 3.6 Tanglegram showing associations between bat host species (left) and *Bartonella* genotypes (right) using Bayesian phylogenies.

REFERENCES

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ (2011) A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). PLoS Currents, 3, RRN1212.

Bai Y, Hayman DT, McKee CD, Kosoy MY (2015) Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLoS Neglected Tropical Diseases, 9, e0003478.

Bai Y, Kosoy MY, Boonmar S *et al.* (2010) Enrichment culture and molecular identification of diverse *Bartonella* species in stray dogs. Veterinary Microbiology, 146, 314-319.

Bai Y, Kosoy MY, Recuenco S *et al.* (2011) *Bartonella* spp. in bats, Guatemala. Emerging Infectious Diseases, 17, 1269–1272.

Bai Y, Recuenco S, Gilbert AT *et al.* (2012) Prevalence and diversity of *Bartonella* spp. in bats in Peru. The American Journal of Tropical Medicine and Hygiene, 87, 518–523.

Balbuena JA, Míguez-Lozano R, Blasco-Costa I (2013) PACo: A novel procrustes application to cophylogenetic analysis. PLoS ONE, 8, e61048.

Berglund EC, Frank AC, Calteau A *et al.* (2009) Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. PLoS Genetics, 5, e1000546.

Berglund EC, Ellegaard K, Granberg F *et al.* (2010) Rapid diversification by recombination in *Bartonella grahamii* from wild rodents in Asia contrasts with low levels of genomic divergence in Northern Europe and America. Molecular Ecology, 19, 2241–2255.

Billeter SA, Hayman DT, Peel AJ *et al.* (2012) *Bartonella* species in bat flies (Diptera: Nycteribiidae) from western Africa. Parasitology, 139, 324–329.

Billeter SA, Levy M, Chomel BB, Breitschwerdt EB (2008) Vector transmission of *Bartonella* species with emphasis on the potential for tick transmission. Medical and Veterinary Entomology, 22, 1–15.

Bivand RS, Pebesma EJ, Gomez-Rubio V (2013) Applied spatial data analysis with R, second edition. Springer, NY.

Bivand RS, Rundel C (2014) rgeos: interface to geometry engine – open source (GEOS). R package version 0.3-8. <http://CRAN.R-project.org/package=rgeos/>.

Bradley RD, Baker RJ (2001) A test of the genetic species concept: cytochrome-b sequences and mammals. Journal of Mammalogy, 82, 960–973.

Buffet J-P, Pisanu B, Brisse S *et al.* (2013) Deciphering bartonella diversity, recombination, and host specificity in a rodent community. PLoS ONE, 8, e68956.

Concannon R, Wynn-Owen K, Simpson V, Birtles RJ (2005) Molecular characterization of haemoparasites infecting bats (Microchiroptera) in Cornwall, UK. Parasitology, 131, 489–496.

Cui J, Han N, Streicker D, Li G, Tang X, *et al.* (2007) Evolutionary relationships between bat coronaviruses and their hosts. Emerging Infectious Diseases, 13, 1526-1532.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969–1973.

Faria NR, Suchard M a, Rambaut A, Streicker DG, Lemey P (2013) Simultaneously reconstructing viral cross-species transmission history and identifying the underlying

constraints. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 368, 20120196.
Hayman DT, McDonald KD, Kosoy MY (2013) Evolutionary history of rat-borne *Bartonella*: the importance of commensal rats in the dissemination of bacterial infections globally. Ecology and Evolution, 3, 3195–3203.
Hughes GJ, Orciari L a., Rupprecht CE (2005) Evolutionary timescale of rabies virus adaptation to North American bats inferred from the substitution rate of the nucleoprotein gene. Journal of General Virology 86, 1467–1474.
IUCN (2014) The IUCN Red List of Threatened Species. Version 2014.1. <http://www.iucnredlist.org/>. Downloaded on 1 April 2015.
Kamani J, Baneth G, Mitchell M *et al.* (2014) *Bartonella* species in bats (Chiroptera) and bat flies (Nycteribiidae) from Nigeria, West Africa. Vector-Borne and Zoonotic Diseases, 14, 625–632.
Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30, 772-780.
Kocher T, Thomas W, Meyer A *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences of the United States of America, 86, 6196–6200.
Kosoy MY (2010) Ecological associations between bacteria of the genus *Bartonella* and mammals. Biology Bulletin, 37, 716–724.
Kosoy MY, Bai Y, Lynch T *et al.* (2010) *Bartonella* spp. in bats, Kenya. Emerging Infectious Diseases, 16, 1875–1881.
Kurtenbach K, Hanincová K, Tsao JI *et al.* (2006) Fundamental processes in the evolutionary ecology of Lyme borreliosis. Nature Reviews. Microbiology, 4, 660–9.
Legendre P, Desdevises Y, Bazin E (2002) A statistical test for host-parasite coevolution. Systemic Biology, 51, 217–234.
Lei BR, Olival KJ (2014) Contrasting patterns in mammal-bacteria coevolution: *Bartonella* and Leptospira in bats and rodents. PLoS Neglected Tropical Diseases, 8, e2738.
Lin J-W, Hsu Y-M, Chomel BB *et al.* (2012) Identification of novel *Bartonella* spp. in bats and evidence of Asian gray shrew as a new potential reservoir of *Bartonella*. Veterinary Microbiology, 156, 119–126.
Longdon B, Hadfield JD, Webster CL, Obbard DJ, Jiggins FM (2011) Host phylogeny determines viral persistence and replication in novel hosts. PLoS Pathogens, 7, e1002260.
Longdon B, Brockhurst MA, Russell CA, Welch JJ, Jiggins FM (2014) The evolution and genetics of virus host shifts. PLoS Pathogens, 10, e1004395.
Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Research, 27, 209–220.
Morse SF, Olival KJ, Kosoy MY *et al.* (2012) Global distribution and genetic diversity of *Bartonella* in bat flies (Hippoboscoidea, Streblidae, Nycteribiidae). Infection, Genetics and Evolution, 12, 1717–23.
Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxfor University Press, New York.
Norman A, Regnery R, Jameson P, Greene C, Krause D (1995) Differentiation of *Bartonella*-like isolates at the species level by PCR-restriction fragment length polymorphism in the citrate synthase gene. Journal of Clinical Microbiology, 33, 1797–1803.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package. R package version 2.2-1.
O'Leary MA, Bloch JI, Flynn JJ *et al.* (2013) The placental mammal ancestor and the post-K-Pg radiation of placentals. Science, 339, 662–667.
Olival KJ, Dittmar K, Bai Y *et al.* (2015) *Bartonella* spp. in a Puerto Rican bat community. Journal of Wildlife Diseases, 51, 274-278.
Page RD (1994) Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladistics, 10, 155–173.
Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289-290.
Parrish CR, Holmes EC, Morens DM *et al.* (2008) Cross-species virus transmission and the emergence of new epidemic diseases. Microbiology and Molecular Biology Reviews, 72, 457–470.
Paziewska A, Harris PD, Zwolińska L, Bajer A, Siński E (2011) Recombination within and between species of the alpha proteobacterium Bartonella infecting rodents. Microbial Ecology, 61, 134–45.
Paziewska A, Siński E, Harris PD (2012) Recombination, diversity and allele sharing of infectivity proteins between Bartonella species from rodents. Microbial Ecology, 64, 525–36.
Pebesma EJ, Bivand RS (2005) Classes and methods for spatial data in R. R News, 5.
R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
Seabloom EW, Borer ET, Gross K *et al.* (2015) The community ecology of pathogens: coinfection, coexistence and community composition. Ecology Letters, 18, 401-415.
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 of the United States of America, 110, 17415–9.
Schaer J, Reeder DM, Vodzak ME *et al.* (2015) Nycteria parasites of Afrotropical insectivorous bats. International Journal for Parasitology, 45, 375–374.
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.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729.
Teeling EC, Madsen O, Van den Bussche RA *et al.* (2002) Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. Proceedings of the National Academy of Sciences of the United States of America, 99, 1431–1436.
Veikkolainen V, Vesterinen EJ, Lilley TM, Pulliainen AT (2014) Bats as reservoir hosts of human bacterial pathogen, *Bartonella mayotimonensis*. Emerging Infectious Diseases, 20, 960–967.
Vellend M (2010) Conceptual synthesis in community ecology. The Quarterly Review of Biology, 85, 183–206.
Vos M, Didelot X (2008) A comparison of homologous recombination rates in bacteria and archaea. The ISME Journal, 3, 199–208.

Wilson D, Reeder D (2005) Mammal species of the world: a taxonomic and geographic reference. Washington D.C.: Smithsonian Institution Press.
Zhu Q, Kosoy M, Olival KJ, Dittmar K (2014) Horizontal transfers and gene losses in the phospholipid pathway of *Bartonella* reveal clues about early ecological niches. Genome Biology and Evolution, 6, 2156-2169.

CHAPTER 4

Concluding remarks and future directions

In this thesis, I have shown that for *Eidolon* spp. fruit bats, continual mixing, vagility, and communal roosting may contribute to the homogenization of *Bartonella* communities over a broad geographic range. Furthermore, I show that bartonella prevalence varies among demographic groups, with especially low infection rates in juvenile bats. Our examination of *Bartonella* genotypes from a global sampling of bat species reveals a strong pattern of cophylogeny among bats and bartonella, contributing in part to the diversity of *Bartonella* seen in bats. I also demonstrate that sympatry has an important effect in determining whether related bat species host similar *Bartonella* genotypes. Finally, the exchange of *Bartonella* genotypes among bat species appears to be constrained by phylogeny, with decreasing rates of exchange among separate bat families and suborders. Reflecting on the framework put forth by Vellend (2010), I have focused this work on the processes of migration and speciation and their contributions to bartonella diversity. Thus far, I have not explicitly considered the influence of selection and ecological drift, and we still have little knowledge about evolutionary processes occurring within host individuals and the transmission mechanisms that link hosts.

For instance, the data indicate that individual bats may be infected by diverse bartonella communities, yet we do not know how selection driven by interactions with the host immune system or among *Bartonella* species influences infection dynamics or bacterial evolution. Chan and Kosoy (2010) hypothesized that coinfection by multiple *Bartonella* species may facilitate escape from the host immune system. This could occur by density-dependent cycling of

Bartonella species in response to selection by the host immune system through the generation of diverse subtypes with novel surface antigens. Microevolutionary processes like high mutation rates or frequent homologous recombination at loci coding for surface proteins may contribute to *Bartonella* persistence in hosts (Zhang et al. 2004; Vos 2009; Tenaillon et al. 2001). Furthermore, selection imposed by competitive interactions among bartonellae may influence dynamics within host individuals. The high genetic diversity within *Bartonella* strains (below the species level) and measurable levels of recombination among strains isolated in *Eidolon helvum* by Bai *et al.* (2015) suggest that mutation and lateral gene transfer may be ongoing processes generating and maintaining bartonella diversity in bats. Additionally, ecological drift may act to change diversity of bartonella communities through stochastic losses of genotypes with low abundance or evolutionary divergence following population bottlenecks.

Lastly, we still lack fundamental knowledge about how bartonellae are transmitted among bats. Bat flies appear to be capable of hosting bartonella (Billeter *et al.* 2012; Morse *et al.* 2012) and recent studies have shown congruence between *Bartonella* genotypes found in bats and the bat flies currently parasitizing them (Judson *et al.* 2015; Brook *et al.* 2015). There is also some evidence that bat flies may vertically transmit endosymbionts to their offspring (Morse *et al.* 2013; Hosokawa *et al.* 2012), which may provide additional opportunities for bartonella diversification. Nevertheless, no study has experimentally confirmed the vector potential of bat flies. Alternative transmission pathways are plausible, through direct transmission by aggressive encounters between bats (e.g., biting and scratching) or through vertical transmission, as has been observed in rodents (Kosoy *et al.* 1998). More experimental and modeling work must be performed to measure the contribution of these pathways to bartonella dynamics in bats.

Testing the relative importance of mechanisms contributing to bartonella diversity – mutation, lateral gene transfer, ecological drift, selection, and transmission – will require long-term sampling from a single population. Future research is proposed using a time series of blood samples from a captive colony of over 100 *E. helvum* in Ghana to study the transmission and evolutionary dynamics of bartonella infections over time. Distinguishing these diversification processes from molecular data will be challenging and will necessarily involve further method development, but the results will have far-reaching implications for understanding how processes that generate and maintain parasite diversity occur in natural systems.

REFERENCES

Bai Y, Hayman DT, McKee CD, Kosoy MY (2015) Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLoS Neglected Tropical Diseases, 9, e0003478.

Billeter SA, Hayman DT, Peel AJ *et al.* (2012) *Bartonella* species in bat flies (Diptera: Nycteribiidae) from western Africa. Parasitology, 139, 324–329.

Brook CE, Bai Y, Dobson AP *et al.* (2015) *Bartonella* spp. in fruit bats and blood-feeding ectoparasites in Madagascar. PLoS Neglected Tropical Diseases, 9, e0003532.

Chan K-S, Kosoy MY (2010) Analysis of multi-strain *Bartonella* pathogens in natural host population – do they behave as species or minor genetic variants? Epidemics, 2, 165–72.

Hosokawa T, Nikoh N, Koga R *et al.* (2012) Reductive genome evolution, host-symbiont co-speciation and uterine transmission of endosymbiotic bacteria in bat flies. The ISME Journal, 6, 577–87.

Judson S, Frank H, Hadly E (2015) Bartonellae are prevalent and diverse in Costa Rican bats and bat flies. Zoonoses and Public Health.

Kosoy M, Regnery R, Kosaya O *et al.* (1998) Isolation of *Bartonella* spp. from embryos and neonates of naturally infected rodents. Journal of Wildlife Diseases, 34, 305–309.

Morse SF, Bush SE, Patterson BD *et al.* (2013) Evolution, multiple acquisition, and localization of endosymbionts in bat flies (Diptera: Hippoboscoidea: Streblidae and Nycteribiidae). Applied and Environmental Microbiology, 79, 2952–61.

Morse SF, Olival KJ, Kosoy MY *et al.* (2012) Global distribution and genetic diversity of *Bartonella* in bat flies (Hippoboscoidea, Streblidae, Nycteribiidae). Infection, Genetics and Evolution, 12, 1717–23.

Tenaillon O, Taddei F, Radman M, Matic I (2001) Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. Research in Microbiology, 152, 11–16.

Vellend M (2010) Conceptual synthesis in community ecology. The Quarterly Review of Biology, 85, 183–206.

Vos M (2009) Why do bacteria engage in homologous recombination? Trends in Microbiology, 17, 226–32.

Zhang P, Chomel BB, Schau MK *et al.* (2004) A family of variably expressed outer-membrane proteins (Vomp) mediates adhesion and autoaggregation in *Bartonella quintana*. Proceedings of the National Academy of Sciences of the United States of America, 101, 13630–5.

APPENDIX I

SUPPLEMENTARY MATERIALS FOR CHAPTER 2

Phylogeography of *Bartonella* bacteria in *Eidolon* spp. fruit bats across Africa

APPENDIX I, SECTION A

Additional figures and tables

Table A1.1 Oligonucleotide primers used for *Bartonella* species detection via PCR amplification. Sequences designated {F} are forward primers and those designated {R} are reverse primers.

Locus	PCR round	Primer sequence	Primer name	Size (bp) of product	Reference
ftsZ		ATTAATCTGCAYCGGCCAGA {F}	Bfp1		
	1	ACVGADACACGAATAACACC {R}	Bfp2	~880	Zeaiter *et al.* 2002
	2	ATATCGCGGAATTGAAGCC {F}	ftsZ R83		Colborn *et al.* 2010;
		CGCATAGAAGTATCATCCA {R}	ftsZ L83	~670	this study
gltA		GCTATGTCTGCATTCTATCA {F}	CS443f		Birtles and Roult 1996;
	1	GATCYTCAATCATTTCTTTCCA {R}	CS1210r	~767	Gundi *et al.* 2012
	2	GGGGACCAGCTCATGGTGG {F}	BhCS781.p		
		AATGCAAAAAGAACAGTAAACA {R}	BhCS1137.n	~356	Norman *et al.* 1995
ITS	1	CTTCAGATGATGATCCCAAGCCTTCTGGCG {F}	325s		
		GAACCGACGACCCCCTGCTTGCAAAGA {R}	1100as	~300	Diniz *et al.* 2007
nuoG	1	GGCGTGATTGTTCTCGTTA {F}	nuoG1f		
		CACGACCACGGCTATCAAT {R}	nuoG1r	~360	Colborn *et al.* 2010
rpoB		CGCATTGGCTTACTTCGTATG {F}	1400F		
	1	GTAGACTGATTAGAACGCTG {R}	2300R	~1000	Renesto and Gouvernet 2001
	2	GGCAATCGTCGCGTTCGTTC {F}	1350F		
		CTACCCGATCACCAACATGC {R}	2350F	~900	This study
ssrA	1	GCTATGGTAATAAATGGACAATGAAATAA {F}	ssrA-F		
		GCTTCTGTTGCCAGGTG {R}	ssrA-R	~280	Diaz *et al.* 2012

Table A1.2 Thermocycler protocols used for *Bartonella* species detection via PCR amplification.

Locus	PCR round	Thermal program
ftsZ	1	95°C 4:00, (95°C 0:30, 55°C 0:30, 72°C 1:00)x40, 72°C 10:00, 4°C ∞
	2	95°C 4:00, (95°C 0:30, 55°C 0:30, 72°C 1:00)x40, 72°C 10:00, 4°C ∞
gltA	1	95°C 2:00, (95°C 0:30, 48°C 0:30, 72°C 2:00)x40, 72°C 7:00, 4°C ∞
	2	95°C 3:00, (95°C 0:30, 55°C 0:30, 72°C 2:00)x40, 72°C 10:00, 4°C ∞
ITS	1	95°C 3:00, (95°C 0:30, 66°C 0:30, 72°C 0:30)x55, 72°C 7:00, 4°C ∞
nuoG	1	95°C 2:00, (95°C 0:30, 55°C 1:00, 72°C 1:00)x45, 72°C 10:00, 4°C ∞
rpoB	1	95°C 2:00, (95°C 0:30, 53°C 1:00, 72°C 2:00)x45, 72°C 10:00, 4°C ∞
	2	95°C 2:00, (95°C 0:30, 55°C 1:00, 72°C 2:00)x35, 72°C 7:00, 4°C ∞
ssrA	1	95°C 2:00, (95°C 0:30, 55°C 0:30, 72°C 1:00)x45, 72°C 7:00, 4°C ∞

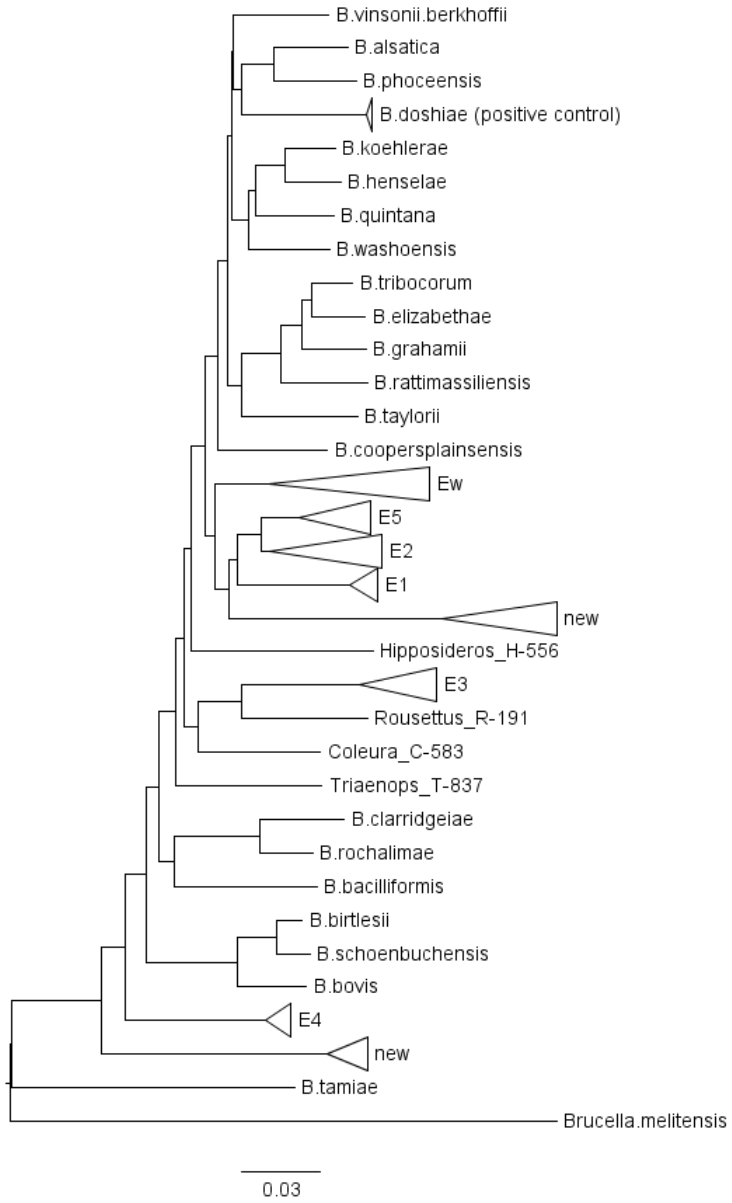

Figure A1.1 Phylogenetic tree for *ftsZ* sequences. Sequences were aligned using MAFFT (Katoh and Standley 2013). The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

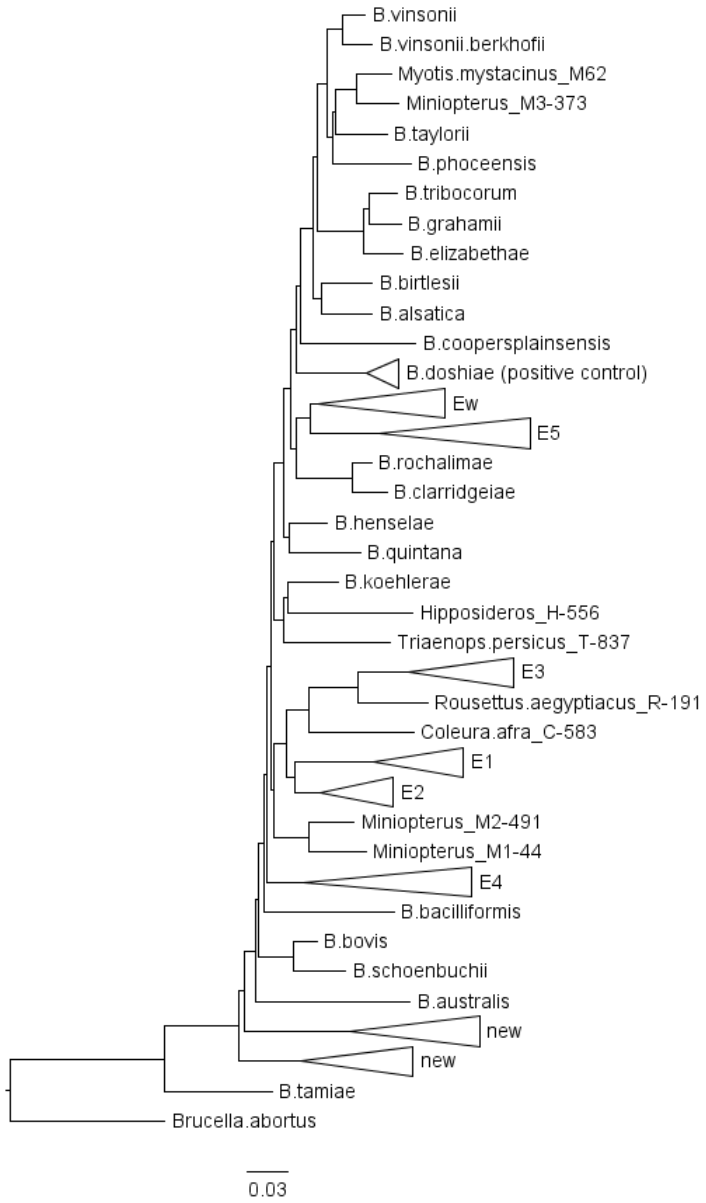

Figure A1.2 Phylogenetic tree for *gltA* sequences. Sequences were aligned using MAFFT (Katoh and Standley 2013). The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

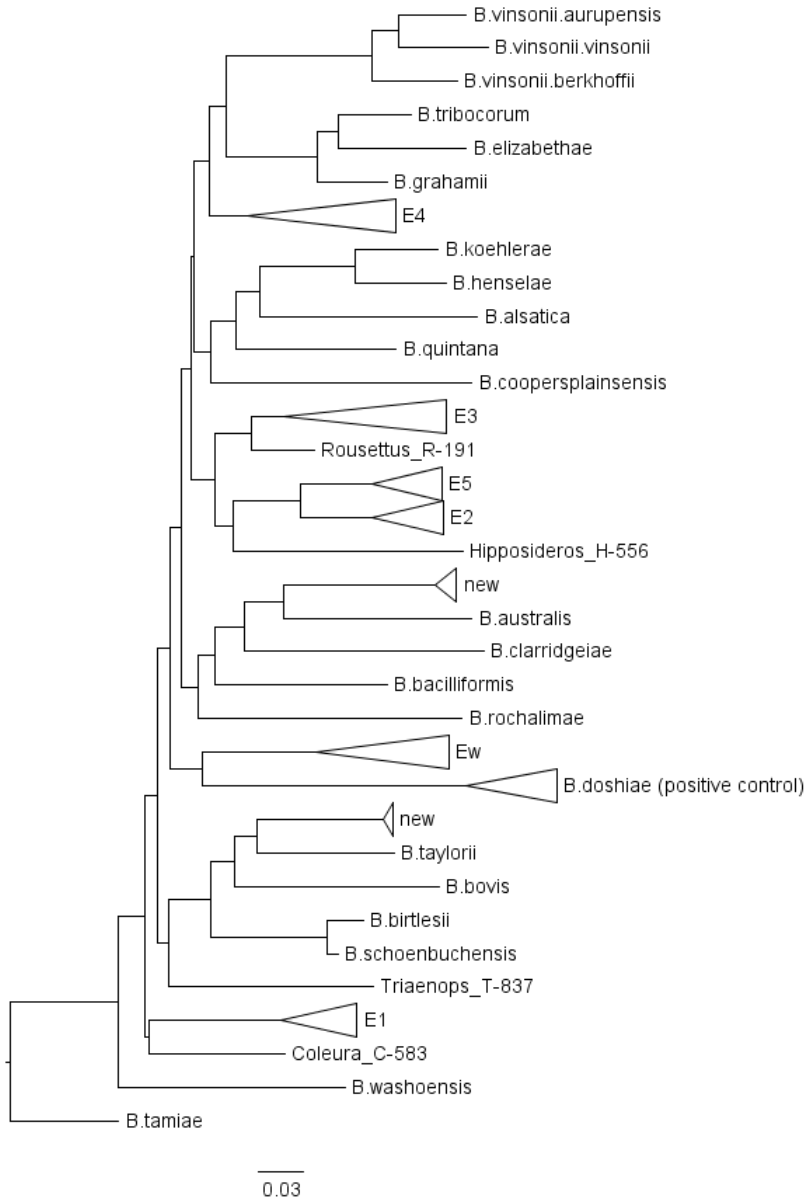

Figure A1.3 Phylogenetic tree for ITS sequences. The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

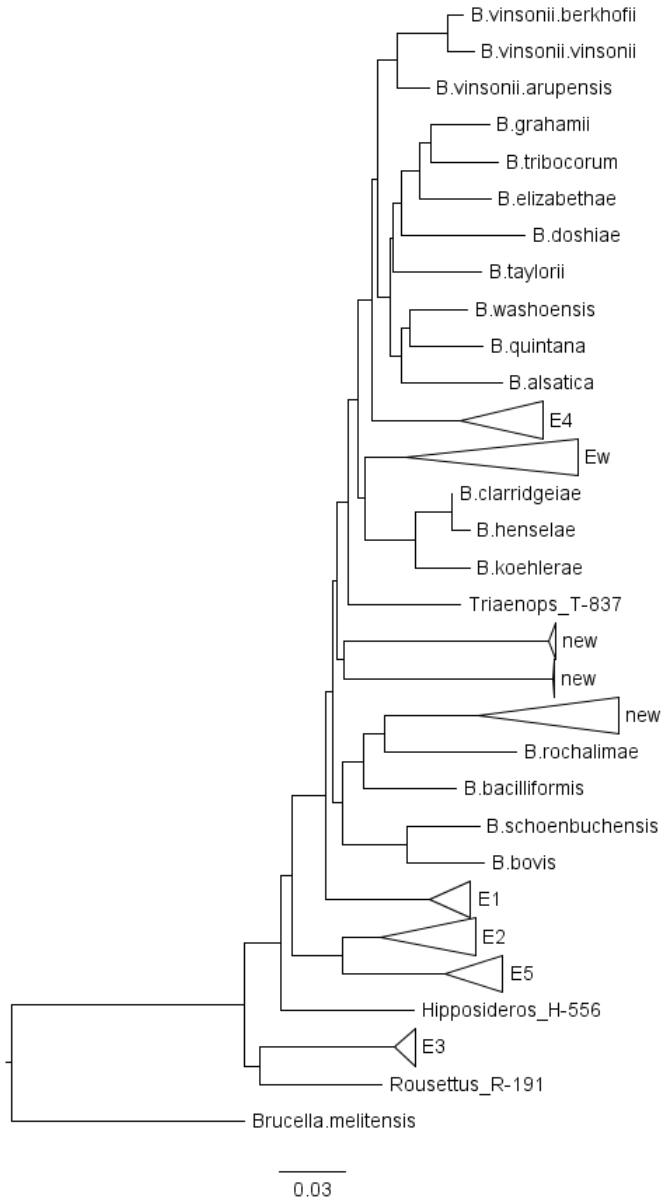

Figure A1.4 Phylogenetic tree for *nuoG* sequences. The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

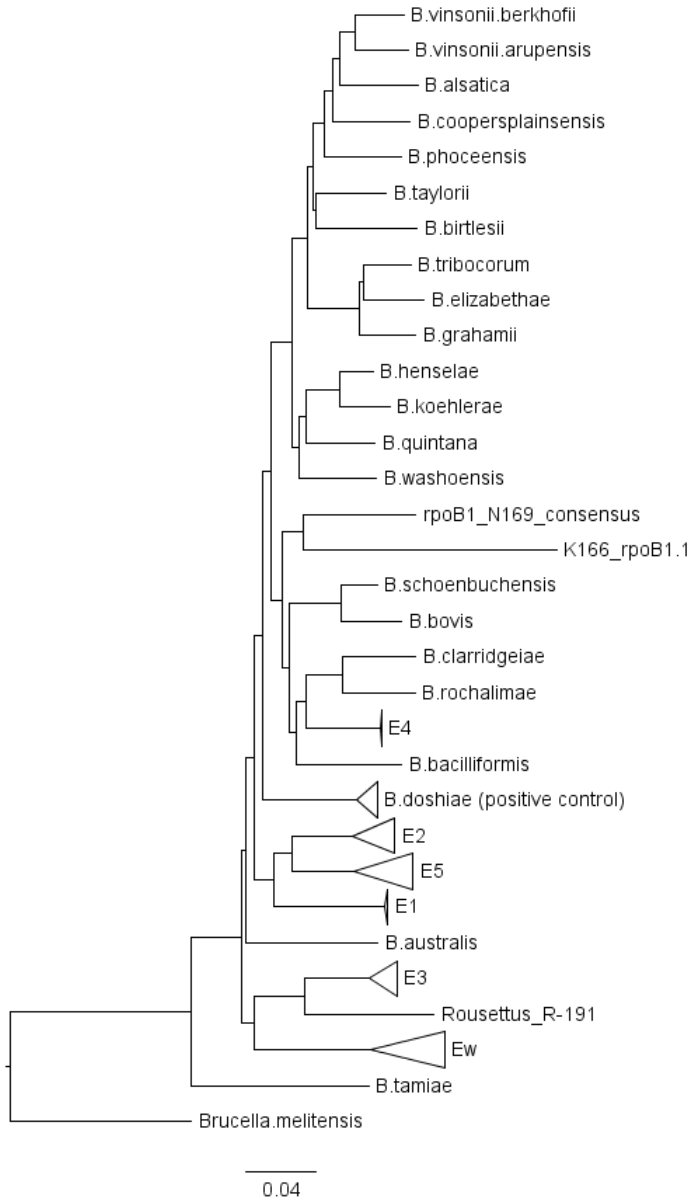

Figure A1.5 Phylogenetic tree for *rpoB* sequences. The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

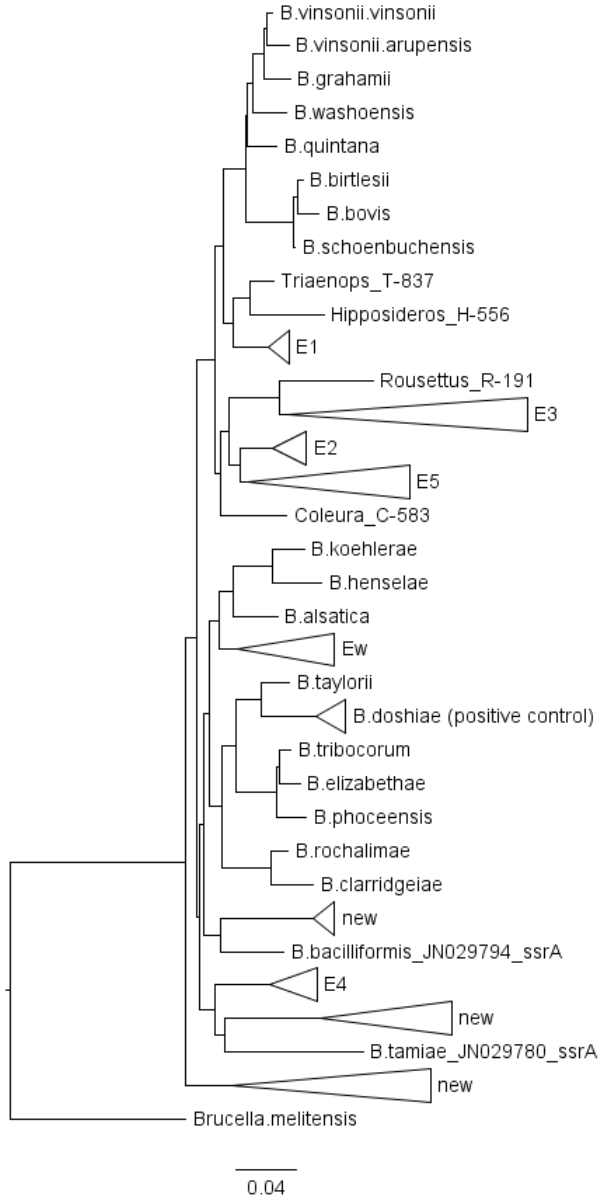

Figure A1.6 Phylogenetic tree for *ssrA* sequences. The tree was assembled using the neighbor-joining method (Saitou and Nei 1987) in MEGA6 (Tamura *et al.* 2013). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. Individual sequences are collapsed into clades represented by triangle using FigTree version 1.4.2 (University of Edinburgh).

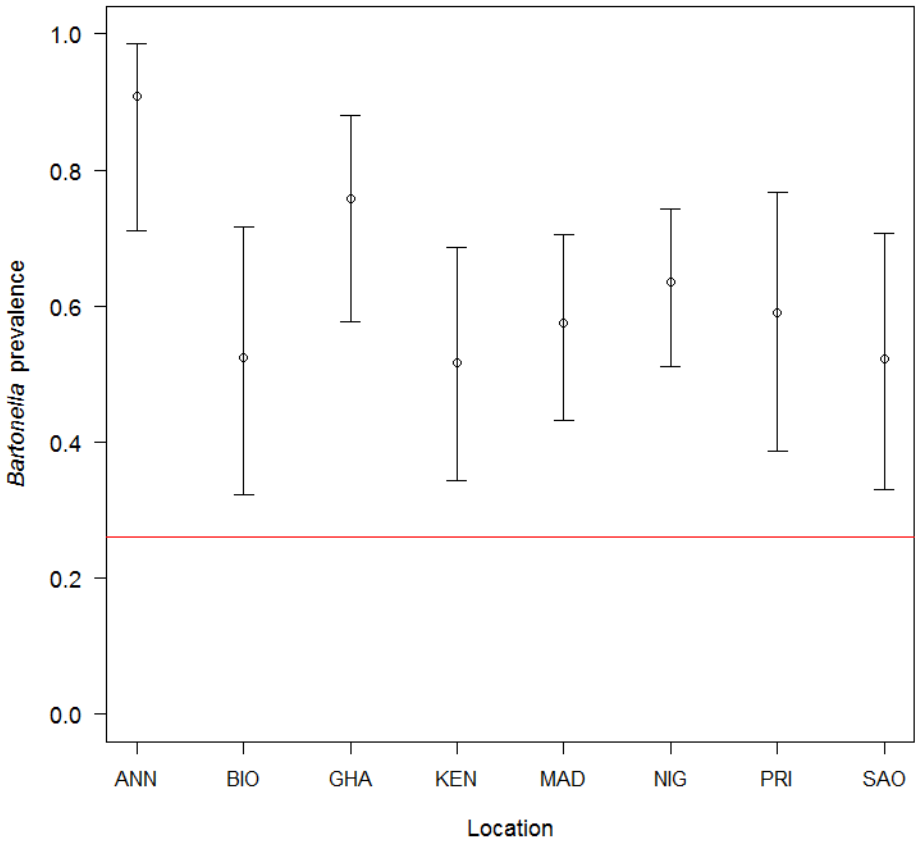

Figure A1.7 Comparison of *Bartonella* spp. prevalence in *Eidolon* spp. fruit bats across sampled locations. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé with sample sizes 22, 21, 29, 29, 47, 63, 22, and 23, respectively. Bats were considered positively infected if multiple runs of one locus yielded *Bartonella* sequences and at least one other locus yielded a *Bartonella* sequence. Point estimates represent total bartonella abundance for all bats in that location. Binomial confidence intervals are estimated using the "add 2 successes and 2 failures" method from Agresti and Coull (1998). The red line indicates the 26.1% prevalence seen in *E. helvum* by Kosoy *et al.* (2010).

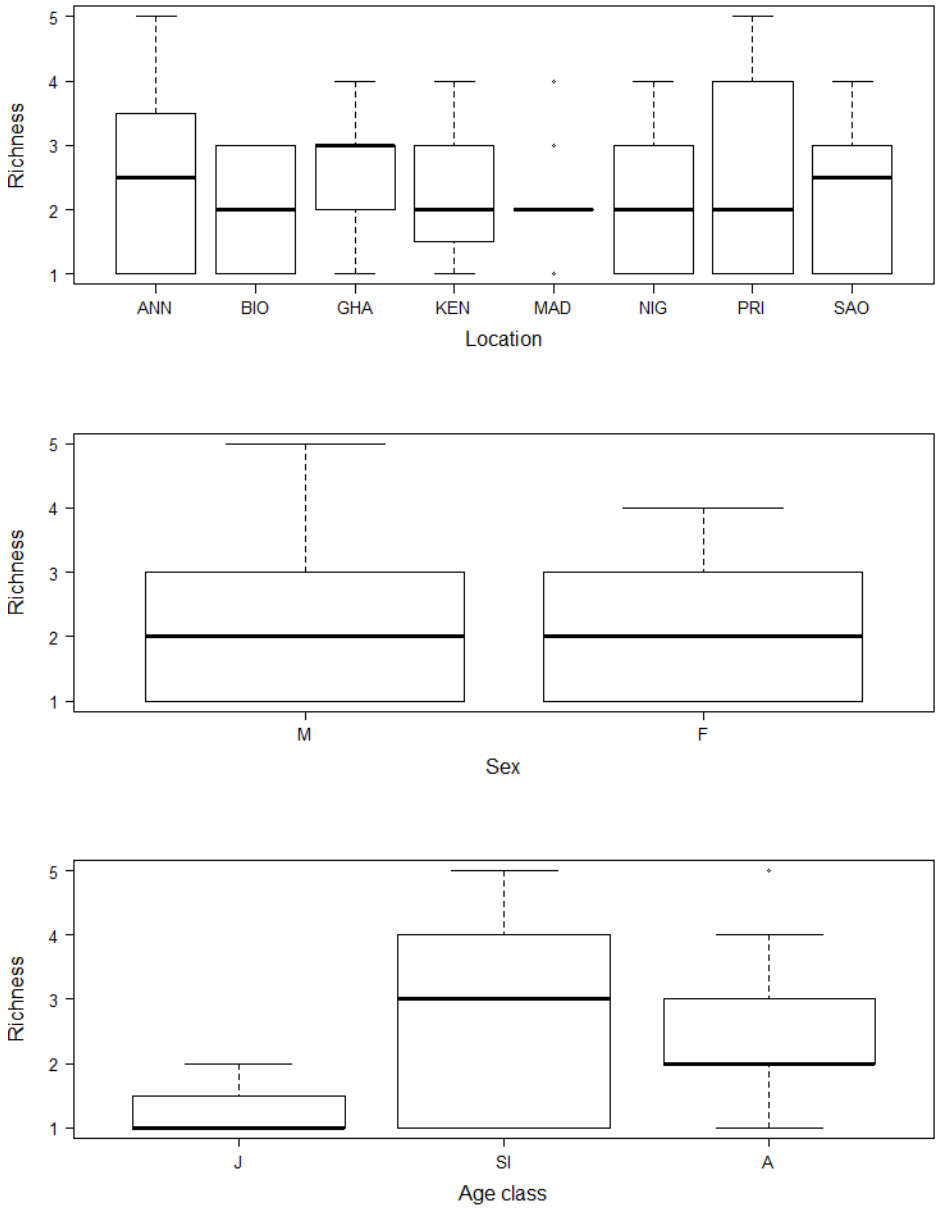

Figure A1.8 Plots of species richness within individuals across location, sex, and age class. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé; sexes are male and female; age classes are juvenile, sexually immature, and adult.

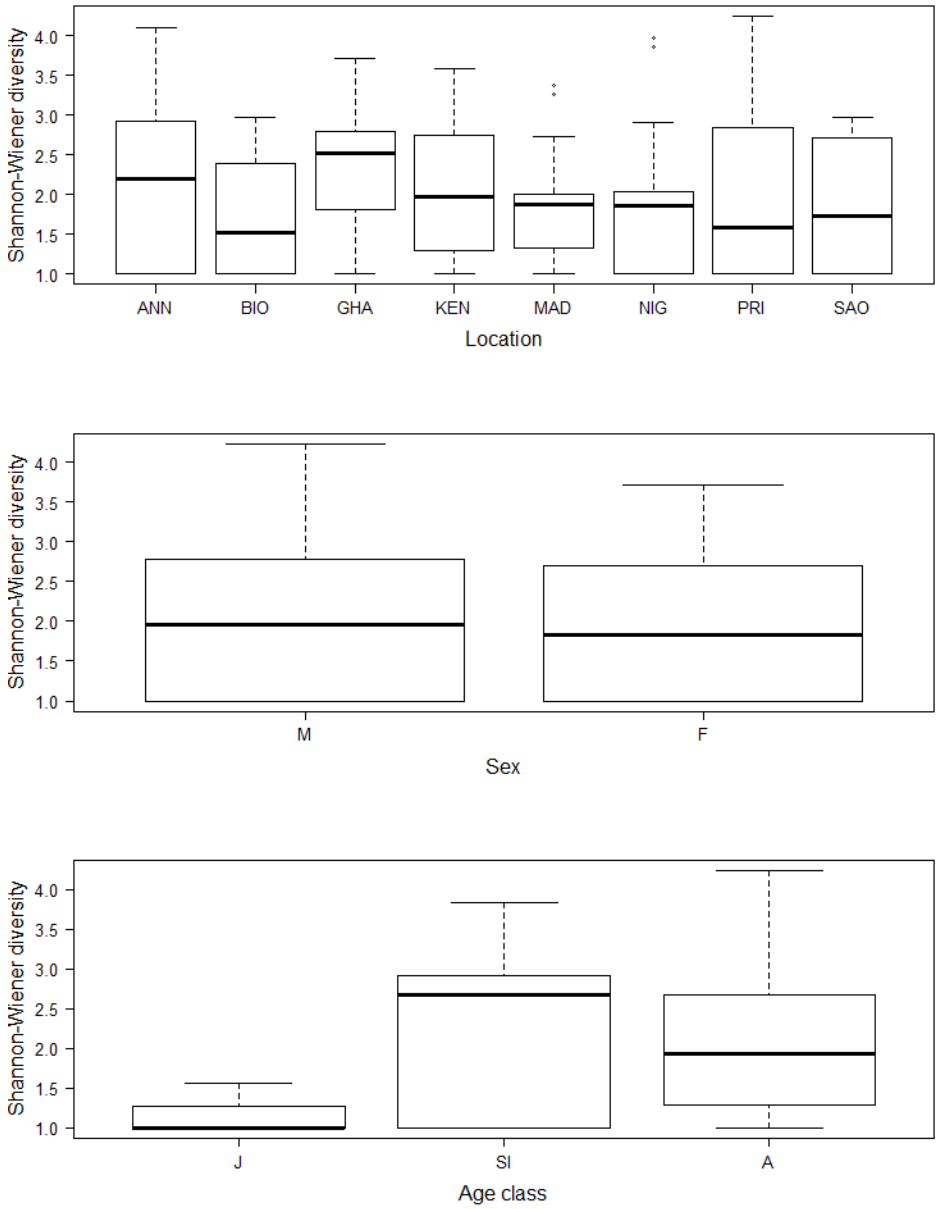

Figure A1.9 Plots of species evenness within individuals across location, sex, and age class, as measured by the Shannon-Wiener diversity index. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé; sexes are male and female; age classes are juvenile, sexually immature, and adult.

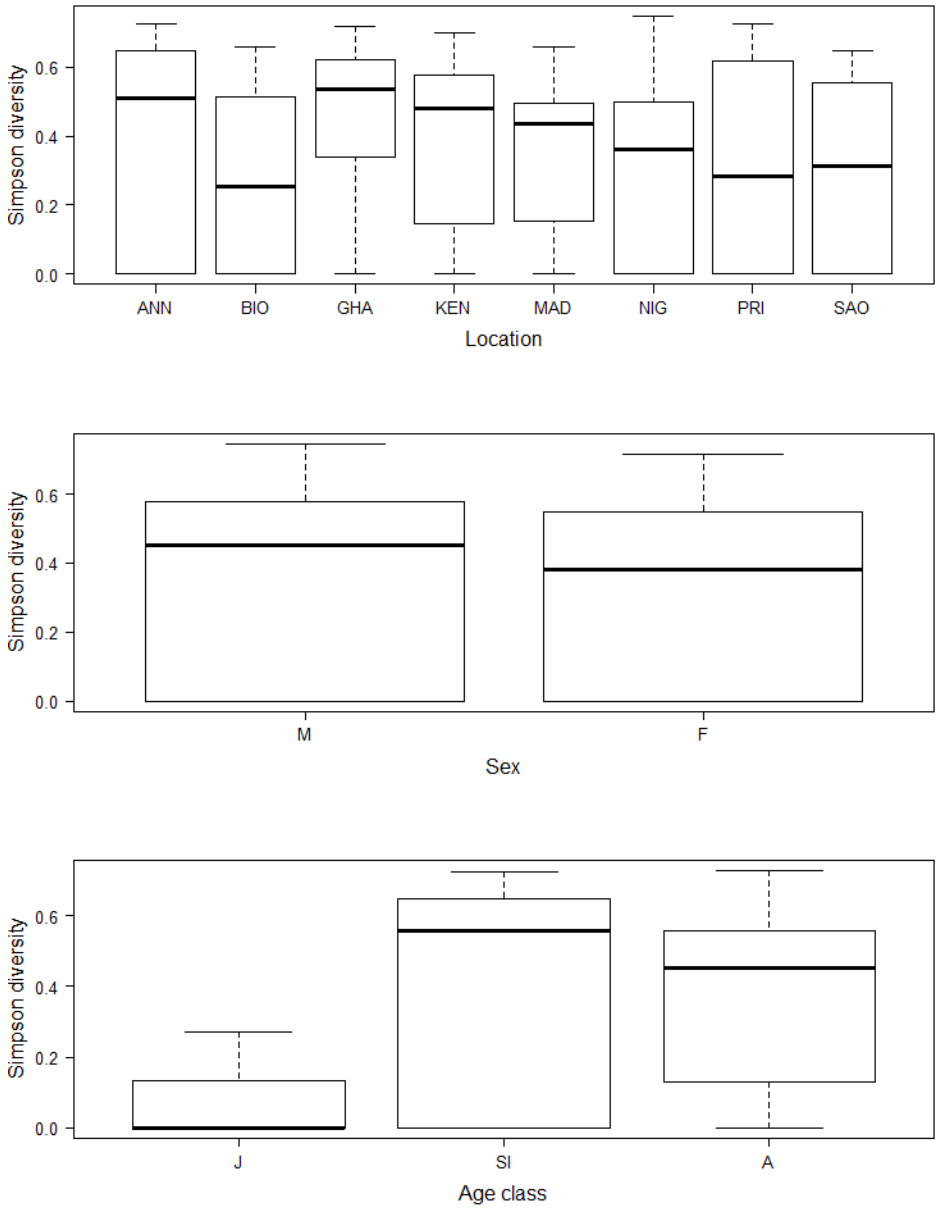

Figure A1.10 Plots of species evenness within individuals across location, sex, and age class, as measured by the Simpson diversity index. From left to right, the locations are Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé; sexes are male and female; age classes are juvenile, sexually immature, and adult.

Table A1.3 Validation of Bayesian inference on the multinomial phylogeography model for *Bartonella* species abundance distributions and covariates location, age class, and sex (A) and for the covariates location and sex with all locations included (B). Models were run using different values of binomial variance $\sigma^2{}_Z$ for τ, β, γ, δ, and ζ parameters**.**

Prior precision	**Prior variance**	**(A) Significant parameters**	**(B) Significant parameters**
0.0099	101	None	None
0.00729	137	None	None
0.00357	280	None	None
0.001	1000	None	None
0.0005	2000	None	β[KEN,E1] > 0
0.0001	10000	β[SAO,E1] > 0	β[KEN,E1] > 0, β[SAO,E1] > 0

Table A1.4 Model selection for *Bartonella* prevalence using location, age class, and sex. This test includes 6 out of 8 locations (excluding Kenya and Nigeria because age classes were not reported).

Model	AICc	ΔAICc	P(χ2)	df	R^2	AUC	ΔAUC
Age + Sex	183.7	0	<0.0001, 0.044	2, 1	0.201744	0.737521	-0.07307
Location + Age + Sex	184.2	0.43	0.11, <0.0001, 0.079	5, 2, 1	0.251495	0.787274	-0.02332
Location + Age	185.1	1.4	0.076, <0.0001	5, 2	0.23671	0.779392	-0.0312
Age	185.9	2.1	<0.0001	2	0.180688	0.698686	-0.11191
Age + Sex + Age*Sex	187.8	4.1	<0.0001, 0.11, 0.91	2, 1, 2	0.202665	0.737521	-0.07307
Location + Age + Sex + Age*Sex	188.2	4.5	0.10, 0.0001, 0.10, 0.78	5, 2, 1, 2	0.25386	0.786782	-0.02381
Location + Age + Location*Age	189.5	5.7	1, 0.99, 1	5, 2, 7	0.290758	0.792036	-0.01856
Location + Age + Sex + Location*Age	189.5	5.8	1, 0.99, 0.13, 1	5, 2, 1, 7	0.301108	0.806486	-0.0041
Location + Age + Sex + Location*Sex	192.6	8.8	0.26, <0.0001, 0.95, 0.93	5, 2, 1, 5	0.266356	0.79023	-0.02036
Location	211.0	27.3	0.067	5	0.080709	0.646798	-0.16379
Location + Sex	211.8	28.1	0.083, 0.24	5, 1	0.088547	0.659606	-0.15099
Sex	214.2	30.5	0.15	1	0.012736	0.558867	-0.25172
(Intercept)	214.2	30.5	0.0003	1	2.44E-15	0.5	-0.31059
Location + Age + Sex + Location*Age*Sex	215.3	31.6	1, 0.96, 0.89, 1	5, 2, 1, 18	0.316806	0.810591	0
Location + Sex + Location*Sex	218.9	35.2	0.62, 0.95, 0.86	5, 1, 5	0.112091	0.678489	-0.1321

Table A1.5 Model selection for *Bartonella* prevalence in females using location, age class, and pregnancy status. This test includes 6 out of 8 locations (excluding Kenya and Nigeria because age classes were not reported).

Model	AICc	ΔAICc	P(χ^2)	df	R^2	AUC	ΔAUC
Age	107.3	0	0.013	2	0.162142	0.668067	-0.11793
Age + Pregnant	110.5	3.1	0.018, 0.53	2, 2	0.174806	0.701401	-0.08459
Location + Age	111.4	4.1	0.63, 0.34	5, 2	0.23192	0.739776	-0.04622
Age + Pregnant + Age*Pregnant	112.2	4.9	0.027, 0.99, 0.96	2, 2, 1	0.180489	0.701401	-0.08459
Location + Age + Pregnant	116.4	9.1	0.79, 0.35, 0.96	5, 2, 2	0.232546	0.74902	-0.03697
Location + Age + Location*Age	118.3	10.9	1, 0.98, 1	5, 2, 4	0.263016	0.754342	-0.03165
(Intercept)	118.3	11.0	0.086	1	3.33E-16	0.5	-0.28599
Location + Age + Pregnant + Age*Pregnant	118.7	11.4	0.81, 0.99, 0.98, 0.96	5, 2, 2, 1	0.235849	0.741176	-0.04482
Pregnant	118.9	11.6	0.180	2	0.040894	0.605602	-0.18039
Location + Age + Pregnant + Location*Pregnant	120.5	13.2	1, 0.49, 1, 1	5, 2, 2, 3	0.267941	0.77423	-0.01176
Location + Age + Pregnant + Location*Age	123.5	16.2	1, 0.98, 0.83, 1	5, 2, 2, 4	0.266222	0.759664	-0.02633
Location + Pregnant + Location*Pregnant	123.5	16.2	0.98, 0.99, 0.97	5, 2, 3	0.191373	0.721289	-0.06471
Location	123.8	16.5	0.430	5	0.061589	0.637815	-0.14818
Location + Pregnant	126.8	19.5	0.67, 0.41	5, 2	0.081219	0.658543	-0.12745
Location + Age + Pregnant + Location*Age*Pregnant	127.8	20.5	1, 0.99, 1, 1	5, 2, 2, 7	0.306337	0.785994	0

Table A1.6 Model selection for *Bartonella* prevalence using location and sex. This test includes all locations (Annobón, Bioko, Ghana, Kenya, Madagascar, Nigeria, Príncipe, and São Tomé).

Model	AICc	ΔAICc	P(χ^2)	df	R^2	AUC	ΔAUC
Location + Sex	334.8	0	0.11, 0.081	7, 1	0.071843	0.641306	-0.01359
Location	335.7	0.93	0.11	7	0.060486	0.624817	-0.03008
Sex	335.9	1.1	0.08	1	0.011916	0.556396	-0.0985
(Intercept)	336.9	2.1	<.0001	1	4.44E-16	0.5	-0.1549
Location + Sex + Location*Sex	345.6	10.9	0.62, 0.97, 0.92	7, 1, 7	0.088975	0.654897	0

APPENDIX I, SECTION B

Computational details

Computational details for the multiple primer model

Metropolis-Hastings (Metropolis *et al.* 1953; Hastings 1970) and Gibbs (Geman and Geman 1984) algorithms were used to sample from the joint posterior of the parameters in the model. ψ_{ijk} parameters needed to be samples in Metropolis-Hastings steps while all other parameters could be samples in Gibbs steps. An adaptive proposal scheme for random walk Metropolis was used for Metropolis-Hastings sampling of ψ parameters. This proposal scheme ensures acceptance rates of proposed parameters values are close to the optimal acceptance rate of 0.234 (Gelman *et al.* 1996). All ψ_{ijk} are updated in a block for each ij; all other parameters are updated in series.

The 50,000 MCMC samples were generated in R version 3.0.3 (R Core Team 2015) and thinned every five iterations. Convergence diagnostics were checked using the CODA library (Plummer *et al.* 2006) in R 3.0.3. A burn-in time of 10,000 iterations was chosen for MCMC chains. Three separate chains were run using unique random number seeds. Not all chains met convergence criteria (e.g., chain autocorrelation, Geweke, Gelman-Rubin, and Heidelberg and Welch diagnostics), however they did all converge according to visual inspections. Abundance estimates for each of the three chains were similar for each sample, and were thus averaged as a measure of the relative abundances of each *Bartonella* species in each sample.

Computational details for the phylogeography model

Gibbs sampling of posterior parameters was implemented in OpenBUGS (Lunn *et al.* 2009) and the R2WinBUGS package (Sturz *et al.* 2005) in R version 3.0.3 (R Core Team 2015). 1,000,000 MCMC samples were generated in R version 3.0.3 (R Core Team 2015) with no thinning, discarding the first 10% as burn-in. Convergence diagnostics were checked using the CODA library (Plummer *et al.* 2006) in R 3.0.3. Three separate were run with different initial values of τ, β, γ, δ, and ζ parameters. All chains converged visually and met all convergence criteria. Corner point comparisons were made to species *E5* first, and then checked by making additional comparisons to species *E3*. Other variance values above and below ${\sigma^2}_Z$ were used to run models to assess the effect of the prior on parameter estimates.

REFERENCES

Agresti A, Coull BA (1998) Approximate is better than 'exact' for interval estimation of binomial proportions. The American Statistician, 52, 119–126.

Birtles RJ, Raoult D (1996) Comparison of partial citrate synthase gene (*gltA*) sequences for phylogenetic analysis of *Bartonella* species. International Journal of Systematic and Evolutionary Microbiology, 46, 891–897.

Colborn JM, Kosoy MY, Motin VL, *et al.* (2010) Improved detection of *Bartonella* DNA in mammalian hosts and arthropod vectors by real-time PCR using the NADH dehydrogenase gamma subunit (*nuoG*). Journal of Clinical Microbiology, 48, 4630-4633.

Diaz MH, Bai Y, Malania L, Winchell JM, Kosoy MY. 2012. Development of a novel genus-specific real-time PCR assay for detection and differentiation of *Bartonella* species and genotypes. Journal of Clinical Microbiology, 50, 1645-1649.

Diniz PP, Maggi RG, Schwartz DS, *et al.* (2007) Canine bartonellosis: serological and molecular prevalence in Brazil and evidence of coinfection with *Bartonella henselae* and *Bartonella vinsonii* subsp. *berkhoffii*. Veterinary Research, 38, 697-710.

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, 721-741.

Gelman A, Roberts G, Gilks W (1996) Efficient Metropolis jumping rules. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, editors, Bayesian Statistics, Volume 5. Oxford, Oxford University Press, 599-608.

Gundi VAKB, Billeter SA, Rood MP, Kosoy MY (2012) *Bartonella* spp. in rats and zoonoses, Los Angeles, California, USA. Emerging Infectious Diseases, 18, 631-633.

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57, 97-109.

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN, editor, Mammalian Protein Metabolism. New York: Academic Press, 21-132.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30, 772-780.

Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The BUGS project: Evolution, critique and future directions (with discussion). Statistics in Medicine, 28, 3049-3082.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. Journal of Chemical Physics, 21, 1087-1092.

Norman AF, Regnery R, Jameson P, *et al.* (1995) Differentiation of *Bartonella*-like isolates at the species level by PCR-restriction fragment length polymorphism in the citrate synthase gene. Journal of Clinical Microbiology, 33, 1797–1803.

Plummer M, Best N, Cowles K, Vines K (2006) CODA: Convergence diagnosis and output analysis for MCMC. R News, 6, 7-11.

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Renesto P, Gouvernet J (2001) Use of *rpoB* gene analysis for detection and identification of Bartonella species. Journal of Clinical Microbiology, 39, 430–437.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4, 406-425.
Sturtz S, Ligges U, Gelman A (2005). R2WinBUGS: A Package for Running WinBUGS from R. Journal of Statistical Software, 12, 1-16.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729.
Zeaiter Z, Liang Z, Raoult D (2002) Genetic classification and differentiation of *Bartonella* species based on comparison of partial *ftsZ* gene sequences. Journal of Clinical Microbiology, 40, 3641–3647.

APPENDIX II

SUPPLEMENTARY MATERIALS FOR CHAPTER 3

Phylogenetic and geographic constraints on *Bartonella* transmission among bat species

APPENDIX II, SECTION A

Additional figures and tables

Table A2.1 *Bartonella* citrate synthase (*gltA*) genotypes isolated from bats included in the analysis dataset with GenBank accession numbers. Host bat genus and species were extracted from GenBank metadata or from published articles (Concannon *et al.* 2005; Kosoy *et al.* 2010; Bai *et al.* 2011; Bai *et al.* 2012; Lin *et al.* 2012; Veikkolainen *et al.* 2014; Olival *et al.* 2015; Kamani *et al.* 2014; Bai *et al.* 2015).

Genotype	Host species	Location	*gltA* accession number
M406	*Myotis daubentonii*	UK	AJ871613
M62	*Myotis mystacinus*	UK	AJ871612
M207	*Pipistrellus* sp.	UK	AJ871614
M451	*Nyctalus noctula*	UK	AJ871615
M409	*Pipistrellus sp.*	UK	AJ871611
Cul-9	*Tonatia silvicola*	Peru	EF616479
R-191	*Rousettus aegyptiacus*	Kenya	HM363764
T-837	*Triaenops persicus*	Kenya	HM545138
H-556	*Hipposideros* sp.	Kenya	HM545137
C-583	*Coleura afra*	Kenya	HM545136
M1-44	*Miniopterus* sp.	Kenya	HM545139
M2-491	*Miniopterus* sp.	Kenya	HM545140
M3-373	*Miniopterus* sp.	Kenya	HM545141
B29042	*Desmodus rotundus*	Guatemala	HM597187
B29043	*Desmodus rotundus*	Guatemala	HM597188
B29044	*Desmodus rotundus*	Guatemala	HM597189
B29107	*Desmodus rotundus*	Guatemala	HM597190
B29108	*Desmodus rotundus, Carollia perspicillata*	Guatemala	HM597191
B29114	*Desmodus rotundus, Carollia perspicillata*	Guatemala	HM597192
B29102	*Pteronotus davyi*	Guatemala	HM597193
B29109	*Pteronotus davyi*	Guatemala	HM597194
B29119	*Desmodus rotundus*	Guatemala	HM597195
B29122	*Desmodus rotundus*	Guatemala	HM597196
B29116	*Phyllostomus discolor*	Guatemala	HM597198
B29126	*Carollia perspicillata*	Guatemala	HM597199
B29230	*Phyllostomus discolor*	Guatemala	HM597200
B29115	*Phyllostomus discolor*	Guatemala	HM597201
B29110	*Glossophaga soricina, Pteronotus davyi*	Guatemala	HM597202
B29105	*Pteronotus davyi*	Guatemala	HM597203
B29112	*Phyllostomus discolor*	Guatemala	HM597204
B29134	*Pteronotus davyi*	Guatemala	HM597205
B29137	*Sturnira lilium*	Guatemala	HM597206
B29172	*Micronycteris microtis*	Guatemala	HM597207
B29111	*Artibeus toltecus*	Guatemala	HM597197
B32945	*Desmodus rotundus*	Peru	JQ071379
B32947	*Phyllostomus discolor*	Peru	JQ071387
B32954	*Artibeus planirostris*	Peru	JQ071382
B32946	*Glossophaga soricina*	Peru	JQ071383
B32943	*Carollia perspicillata*	Peru	JQ071384
B32960	*Carollia perspicillata*	Peru	JQ071386
B32955	*Carollia perspicillata*	Peru	JQ071385
B32854	*Phyllostomus hastatus*	Peru	JQ071388
B32855	*Desmodus rotundus*	Peru	JQ071378
B32856	*Vampyressa bidens*	Peru	JQ071389

Genotype	Host species	Location	*gltA* accession number
B32942	*Myotis* sp.	Peru	JQ071390
B32851	*Artibeus obscurus*	Peru	JQ071380
B32953	*Artibeus planirostris*	Peru	JQ071381
No. 5	*Miniopterus schreibersii*	Taiwan	JF500511
No. 7	*Miniopterus schreibersii*	Taiwan	JF500513
AS050	*Myotis*	Poland	JQ695835
2574/1	*Myotis daubentonii*	Finland	KF003129
1157/3	*Eptesicus nilssoni*	Finland	KF003115
Mr37079	*Monophyllus redmani*	Puerto Rico	KJ530746
Mr37078	*Monophyllus redmani*	Puerto Rico	KJ530745
Mr37077	*Monophyllus redmani*	Puerto Rico	KJ530744
Mr37075	*Monophyllus redmani*	Puerto Rico	KJ530743
Bc37076	*Brachyphylla cavernarum*	Puerto Rico	KJ530742
Aj37081	*Artibeus jamaicensis*	Puerto Rico	KJ530741
B23976	*Eidolon helvum*	Kenya	KM030507
B40005	*Eidolon helvum*	Cameroon	KM030518
B23979	*Eidolon helvum*	Kenya	KM030509
B24225	*Eidolon helvum*	Kenya	KM030511
B40396	*Eidolon helvum*	Tanzania	KM030522
B40400	*Eidolon helvum*	Tanzania	KM030523
B23812	*Eidolon helvum*	Kenya	KM030504
B24163	*Eidolon helvum*	Kenya	KM030510
B39301	*Eidolon helvum*	Ghana	KM030516
B23975	*Eidolon helvum*	Kenya	KM030506
B39286	*Eidolon helvum*	Ghana	KM030514
B39296	*Eidolon helvum*	Ghana	KM030515
B40908	*Eidolon helvum*	Uganda	KM030526
B39249	*Eidolon helvum*	Ghana	KM030513
B40014	*Eidolon helvum*	Tanzania	KM030520
B40406	*Eidolon helvum*	Tanzania	KM030525
B23797	*Eidolon helvum*	Kenya	KM030503
Mi-BA38	*Micropteropus* sp.	Nigeria	KF418812
Eh-GB64	*Eidolon helvum*	Nigeria	KF418811
Rh-GB31	*Rhinolophus* sp.	Nigeria	KF418810
Ep-BA63	*Epomophorus* sp.	Nigeria	KF418808
Rh-GB59	*Rhinolophus* sp.	Nigeria	KF418809
Rh-GB1	*Rhinolophus* sp.	Nigeria	KF418807
Ep-GB65	*Epomophorus* sp.	Nigeria	KF418806
B110	*Hipposideros larvatus*	Vietnam	KP100360
B109	*Megaderma lyra*	Vietnam	KP100359
B102	*Rhinolophus chaseni*	Vietnam	KP100358
B087	*Hipposideros larvatus*	Vietnam	KP100355
B081	*Hipposideros larvatus*	Vietnam	KP100354
B079	*Rhinolophus chaseni*	Vietnam	KP100353
B072	*Megaerops niphanae*	Vietnam	KP100352
B068	*Rhinolophus acuminatus*	Vietnam	KP100351
B064	*Rhinolophus acuminatus*	Vietnam	KP100350
B063	*Rhinolophus acuminatus*	Vietnam	KP100349
B052	*Rhinolophus acuminatus*	Vietnam	KP100346

Genotype	Host species	Location	*gltA* accession number
B050	*Rhinolophus acuminatus*	Vietnam	KP100345
B049	*Rhinolophus sinicus*	Vietnam	KP100344
B047	*Rhinolophus sinicus*	Vietnam	KP100343
B006	*Rhinolophus acuminatus*	Vietnam	KP100342
B005	*Megaderma spasma*	Vietnam	KP100341
B003	*Rhinolophus acuminatus*	Vietnam	KP100340
2308	*Brucella melitensis*	outgroup	AM040264

Table A2.2 *Bartonella* citrate synthase (*gltA*) genotypes isolated from ectoparasites included in the analysis dataset with GenBank accession numbers. Ectoparasite and host bat genus and species were extracted from GenBank metadata or from published articles (Morse *et al.* 2012; Billeter *et al.* 2012; Veikkolainen *et al.* 2014).

Genotype	Ectoparasite	Host species	Location	*gltA* accession number
NB-1.2	*Siphonaptera* sp.	*Myotis brandtii*	Finland	KF003137
AS025	*Spinturnix myoti*	*Myotis*	Poland	JQ695839
AS067	*Spinturnix myoti*	*Myotis*	Poland	JQ695840
Cg 462	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172074
Cg 401	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172072
Cg 454	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172071
Cg 414	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172070
Cg 433	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172069
Cg 443	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172068
Cg 465-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172067
Cg 713-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172066
Cg 405	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172065
Cg 424	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172064
Cg 417-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172063
Cg 426-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172062
Cg 436-3	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172061
Cg 418	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172060
Cg 423-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Bioko	JN172059
Cg 364	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172049
Cg 358-3	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172051
Cg 366-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172052
Cg 315-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172054
Cg 303-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172055
Cg 303-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172056
Cg 366	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Equatorial Guinea: Annobón	JN172057
Cg K1-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172035
Cg K5-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172036

Genotype	Ectoparasite	Host species	Location	*gltA* accession number
Cg K8-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172037
Cg Q22-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172038
Cg Q98-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172039
Cg Q100-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172040
Cg Q130	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172041
Cg G35-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172042
Cg G31-1	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172043
Cg G38-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172044
Cg GG236	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172045
Cg GG48	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172046
Cg GG243-2	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172047
Cg GG243-3	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN172048
E-124	*Cyclopodia greeffi greeffi*	*Eidolon helvum*	Ghana	JN190887
E7	*Cyclopodia horsfieldii*	*Pteropus hypomelanus*	Malaysia	JX416257
E5	*Cyclopodia horsfieldii*	*Pteropus hypomelanus*	Malaysia	JX416256
P2874	*Cyclopodia simulans*	*Ptenochirus jagori*	Philippines	JX416255
27_3_4	*Paradyschiria lineata*	*Noctilio leporinus*	Panama	JX416254
DR0583	*Trichobius adamsi*	*Macrotus waterhousii*	Dominican Republic	JX416253
05_01_07	*Phthiridium* sp. *scissa* group	*Rhinolophus pearsoni*	Laos	JX416252
23_03_04	*Strebla diaemi*	*Diaemus youngi*	Panama	JX416251
DR05259	*Trichobius adamsi*	*Phyllonycteris poeyi*	Dominican Republic	JX416249
CWD974	*Trichobius johnsonae*	*Pteronotus personatus*	Mexico	JX416248
RCO934	*Pseudostrebla ribeiroi*	*Lophostoma silvicolum*	Peru	JX416247
Mala11	*Basilia (Tripselia) coronata*	*Tylonycteris* sp.	Malaysia	JX416246
ZAG03	*Basilia nattereri*	*Myotis nattererei*	Slovenia	JX416241
JAE1033	*Leptocyclopodia* sp. *nov.*	*Harpionycteris whiteheadi*	Philippines	JX416239
Mala15	*Phthiridium (Stylidia) fraterna*	*Hipposideros* sp.	Malaysia	JX416238
DR05241	*Trichobius adamsi*	*Macrotus waterhousii*	Dominican Republic	JX416237
ZAG01	*Trichobius corynorhinus*	*Corynorhinus townsendii*	USA	JX416236
FG13	*Paratrichobius longicrus* complex	*Artibeus lituratus*	French Guyana	JX416232

Genotype	Ectoparasite	Host species	Location	*gltA* accession number
FG10	*Paratrichobius longicrus* complex	*Artibeus lituratus*	French Guyana	JX416231
2308	*Brucella melitensis*	-	outgroup	AM040264

Table A2.3 Cytochrome b (*cytb*) sequences for bat species included in the analysis dataset with GenBank accession numbers. An asterisk (*) indicates that the species is a representative for sequences attributed only to the bat genus. A dagger (†) indicates that the species a replacement for an original host species with no suitable *cytb* sequence in GenBank. Host bat family and suborder were recorded based on the IUCN Red List of Threatened Species (IUCN 2014) and Agnarsson *et al.* (2011). Web of Science citations were recorded based on a search of the binomial species name. Study sample sizes for each species were quantified from original articles, using the original host species wherever a species-level replacement was made.

Host species	Abbreviated name	Family	Suborder	*cytb* accession number	Web of Science citations	Study sample size	Number of links
Myotis daubentonii	Myot.daub	Vespertilionidae	Yangochiroptera	AB106589	234	6	2
Myotis mystacinus	Myot.myst	Vespertilionidae	Yangochiroptera	AB106605	78	2	1
Nyctalus noctula	Nyct.noct	Vespertilionidae	Yangochiroptera	JX570902	186	1	1
Myotis myotis	Myot.myot	Vespertilionidae	Yangochiroptera	AM261883	2751	unpublished	3
Eptesicus nilssoni	Epte.nils	Vespertilionidae	Yangochiroptera	GQ272582	48	1	1
*Pipistrellus pipistrellus**	Pipi.pipi	Vespertilionidae	Yangochiroptera	KF874521	949	36	2
*Myotis nigricans**	Myot.nigr	Vespertilionidae	Yangochiroptera	KP134584	50	7	1
Myotis brandtii	Myot.bran	Vespertilionidae	Yangochiroptera	AF376844	58	1	1
*Tylonycteris pachypus**	Tylo.pach	Vespertilionidae	Yangochiroptera	EF517315	15	1	1
Myotis nattereri	Myot.natt	Vespertilionidae	Yangochiroptera	JF412413	140	2	1
Corynorhinus townsendii	Cory.town	Vespertilionidae	Yangochiroptera	KC747680	53	1	1
Rhinolophus pearsonii	Rhin.pear	Rhinolophidae	Yinpterochiroptera	JX502551	2	1	1
*Rhinolophus landeri**	Rhin.land	Rhinolophidae	Yinpterochiroptera	EU436668	5	18	3
Rousettus aegyptiacus	Rous.aegy	Pteropodidae	Yinpterochiroptera	JF728760	249	105	1
Eidolon helvum	Eido.helv	Pteropodidae	Yinpterochiroptera	JN398200	94	383	53
*Micropteropus pusillus**	Mcrp.pusi	Pteropodidae	Yinpterochiroptera	JF728734	3	11	1
*Epomophorus gambianus**	Epom.gamb	Pteropodidae	Yinpterochiroptera	JF728757	8	53	2

Host species	Abbreviated name	Family	Suborder	*cytb* accession number	Web of Science citations	Study sample size	Number of links
Pteropus hypomelanus	Pter.hypo	Pteropodidae	Yinpterochiroptera	AB062472	43	2	2
Ptenochirus jagori	Pten.jago	Pteropodidae	Yinpterochiroptera	AB046325	7	1	1
Harpyionycteris whiteheadi	Harp.whit	Pteropodidae	Yinpterochiroptera	DQ445708	1	1	1
Tonatia saurophila†	Tona.saur	Phyllostomidae	Yangochiroptera	FJ155488	7	1	1
Desmodus rotundus	Desm.rotu	Phyllostomidae	Yangochiroptera	FJ155477	349	49	10
Carollia perspicillata	Caro.pers	Phyllostomidae	Yangochiroptera	KF019723	332	43	6
Glossophaga soricina	Glos.sori	Phyllostomidae	Yangochiroptera	FJ392516	203	154	2
Sturnira lilium	Stur.lili	Phyllostomidae	Yangochiroptera	KC753849	106	13	1
Micronycteris microtis	Mcny.micr	Phyllostomidae	Yangochiroptera	AY380756	17	3	1
Artibeus toltecus	Arti.tolt	Phyllostomidae	Yangochiroptera	FJ376728	3	1	1
Artibeus planirostris	Arti.plan	Phyllostomidae	Yangochiroptera	KP134540	35	16	2
Phyllostomus hastatus†	Pyst.hast	Phyllostomidae	Yangochiroptera	FJ155479	91	13	6
Vampyressa bidens	Vamp.bide	Phyllostomidae	Yangochiroptera	FJ154181	4	3	1
Artibeus obscurus	Arti.obsc	Phyllostomidae	Yangochiroptera	KP134536	18	10	1
Monophyllus redmani	Mono.redm	Phyllostomidae	Yangochiroptera	AF382888	17	20	4
Brachyphylla cavernarum	Brac.cave	Phyllostomidae	Yangochiroptera	AY572383	14	2	1
Artibeus jamaicensis	Arti.jama	Phyllostomidae	Yangochiroptera	GQ861667	230	17	1
Macrotus waterhousii	Macr.wate	Phyllostomidae	Yangochiroptera	AY380745	21	3	2
Diaemus youngi	Diae.youn	Phyllostomidae	Yangochiroptera	FJ155475	19	1	1
Phyllonycteris poeyi	Pyny.poey	Phyllostomidae	Yangochiroptera	GU937240	7	1	1

Host species	Abbreviated name	Family	Suborder	*cytb* accession number	Web of Science citations	Study sample size	Number of links
Lophostoma silvicolum	Loph.silv	Phyllostomidae	Yangochiroptera	JF923862	15	1	1
Artibeus lituratus	Arti.litu	Phyllostomidae	Yangochiroptera	KP134571	135	1	2
Noctilio leporinus	Noct.lepo	Noctilionidae	Yangochiroptera	JX257161	62	1	1
Pteronotus davyi	Ptnt.davy	Mormoopidae	Yangochiroptera	AF338672	17	10	5
Pteronotus personatus	Ptnt.pers	Mormoopidae	Yangochiroptera	KC011599	13	1	1
*Miniopterus natalensis**	Mnpt.nata	Miniopteridae	Yangochiroptera	AJ841977	26	87	3
Miniopterus schreibersii	Mnpt.schr	Miniopteridae	Yangochiroptera	EF530348	204	14	2
Triaenops persicus	Tria.pers	Hipposideridae	Yinpterochiroptera	EU798758	7	8	1
*Hipposideros ruber**	Hipp.rube	Hipposideridae	Yinpterochiroptera	FJ347996	10	4	1
*Hipposideros diadema**	Hipp.diad	Hipposideridae	Yinpterochiroptera	DQ219421	13	1	1
Hipposideros larvatus	Hipp.larv	Hipposideridae	Yinpterochiroptera	EU434949	18	unpublished	3
Megaderma lyra†	Mgdm.lyra	Megadermatidae	Yinpterochiroptera	DQ888678	121	unpublished	2
Rhinolophus affinis†	Rhin.affi	Rhinolophidae	Yinpterochiroptera	DQ987605	19	unpublished	2
Megaerops ecaudatus†	Mgps.ecau	Pteropodidae	Yinpterochiroptera	GQ410214	3	unpublished	1
Rhinolophus macrotis†	Rhin.macr	Rhinolophidae	Yinpterochiroptera	JX465355	13	unpublished	7
Rhinolophus sinicus	Rhin.sini	Rhinolophidae	Yinpterochiroptera	HM134917	17	unpublished	2
Coleura afra	Cole.afra	Emballonuridae	Yangochiroptera	JQ710752	6	9	1
Ornithorhynchus anatinus	Orni.anat	outgroup	-	HQ379928	-	-	-

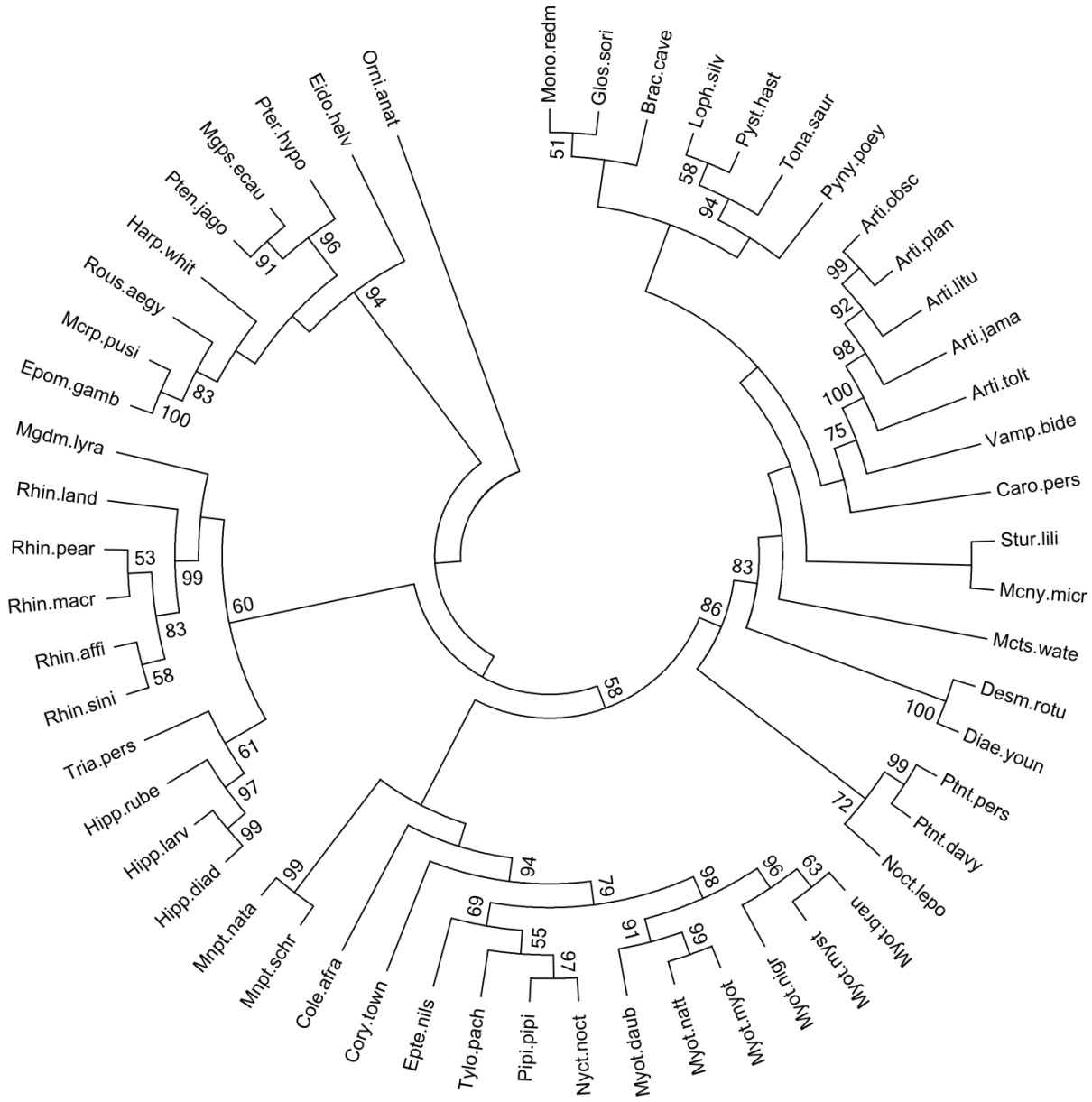

Figure A2.1 Maximum likelihood (-ln likelihood = 23862.8) phylogenetic tree of bat species using 1140 base pair sequences of the mitochondrial *cytb* gene aligned using MAFFT (Katoh and Standley 2013). The tree was assembled in MEGA6 (Tamura *et al.* 2013) using the GTR+G substitution model with five gamma categories (Nei and Kumar 2000). Node support values were estimated from 1000 bootstrap replicates; only support values >50% are shown. Abbreviated species names are listed in Supplementary Table 3.

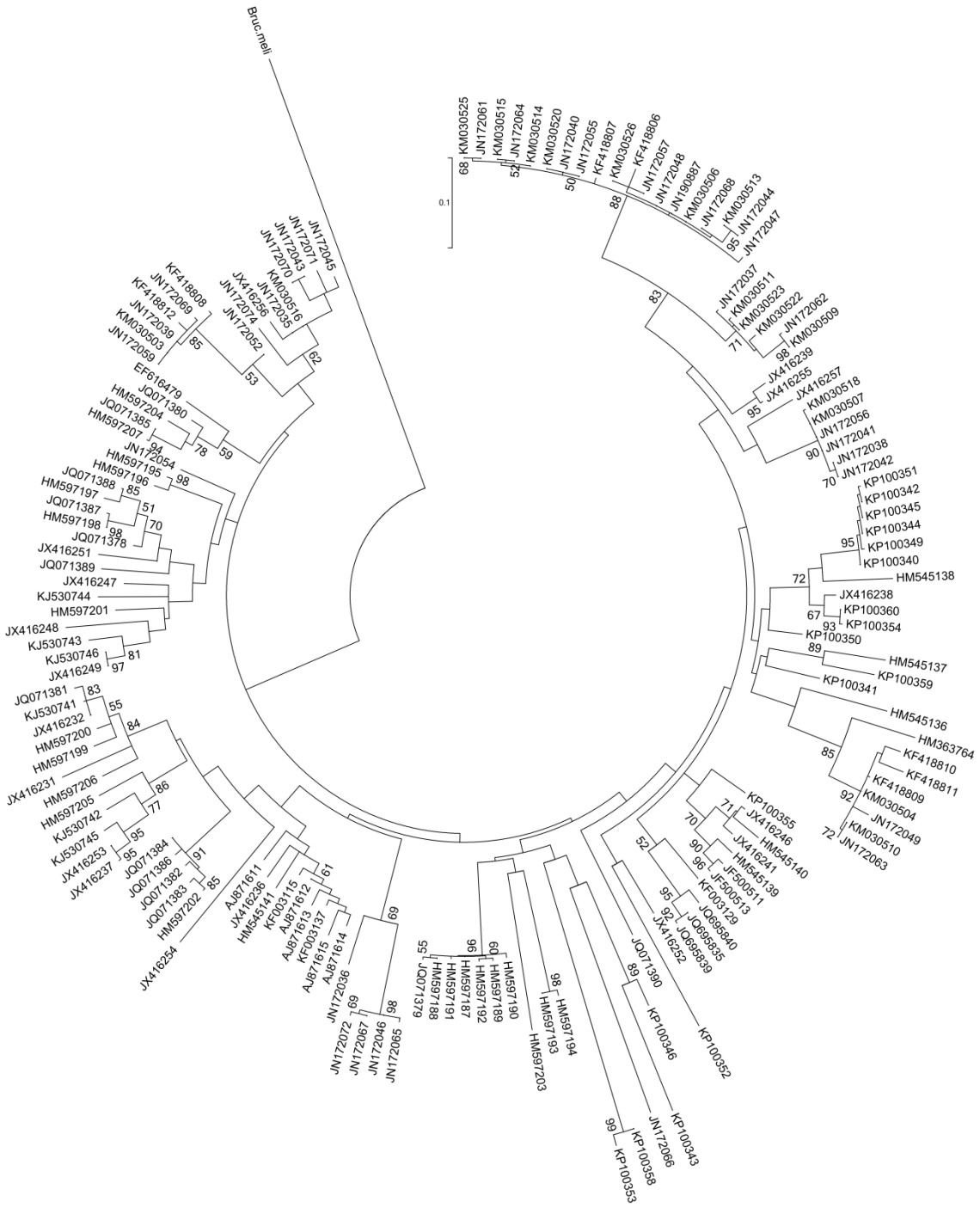

Figure A2.2 Maximum likelihood (-ln likelihood = 7744.3) phylogenetic tree of *Bartonella* genotypes using 334 base pair sequences of the genomic *gltA* gene aligned using MAFFT (Katoh and Standley 2013). The tree was assembled in MEGA6 (Tamura *et al.* 2013) using the GTR+G substitution model with five gamma categories (Nei and Kumar 2000). Node support values were estimated from 1000 bootstrap replicates; only support values >50% are shown.

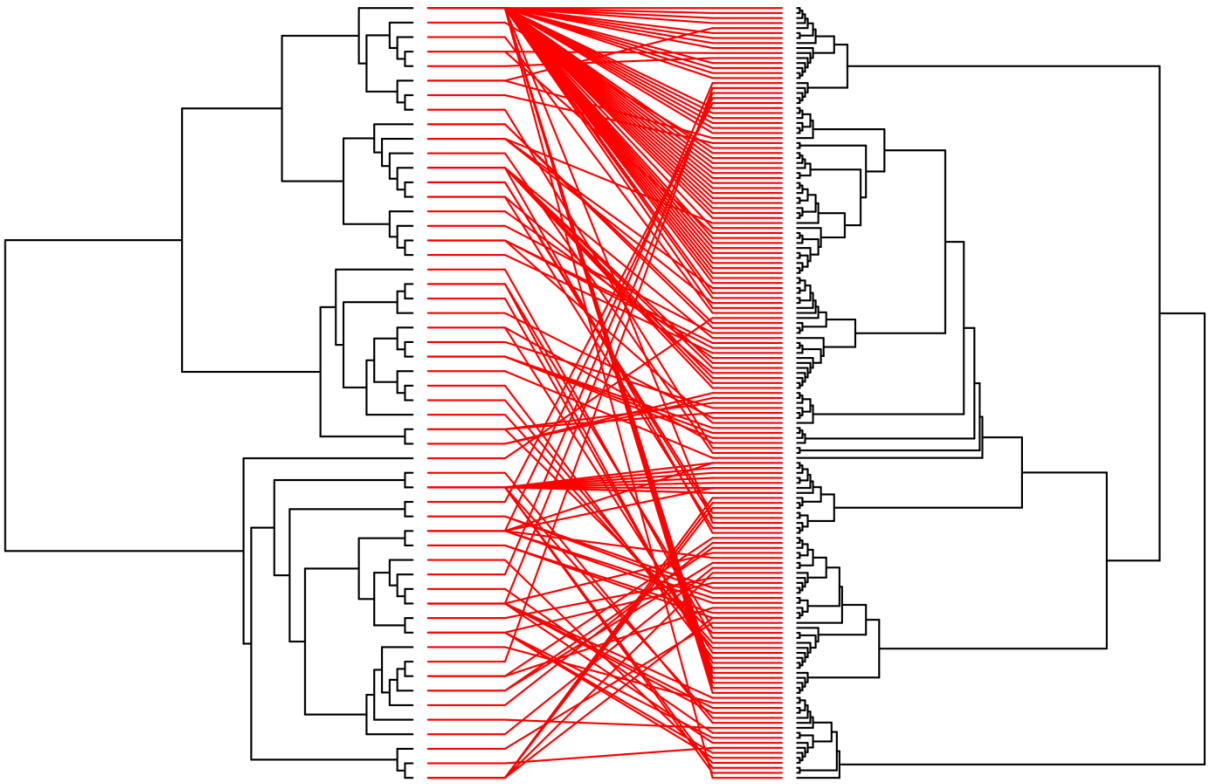

Figure A2.3 Tanglegram showing associations between bat host species (left) and *Bartonella* genotypes (right) using maximum likelihood phylogenies.

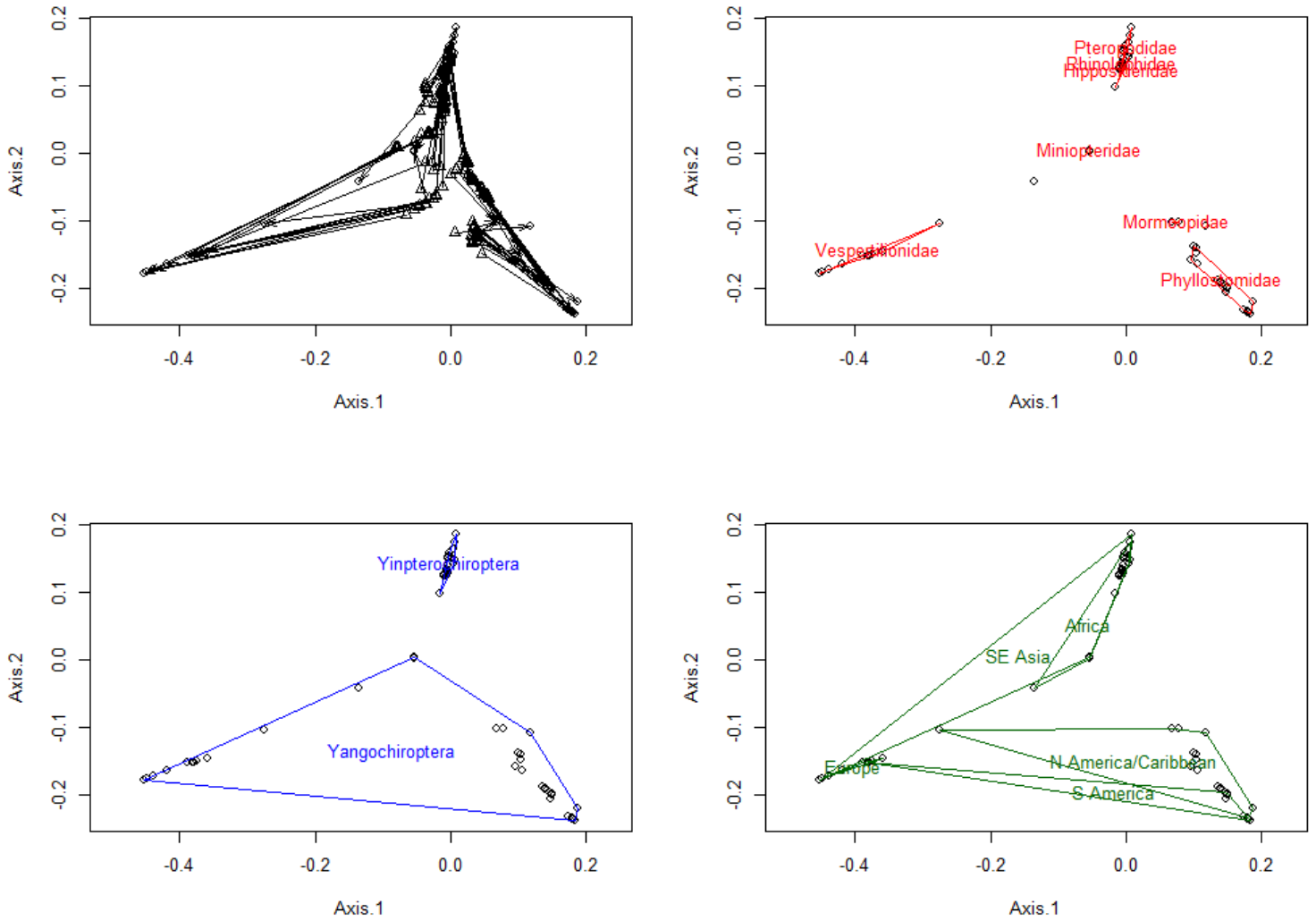

Figure A2.4 Procrustes ordination (Balbuena *et al.* 2013) using distances calculated from maximum likelihood trees of bat and *Bartonella* sequences. *Bartonella* genotypes are represented as triangles and bat species are represented as circles. Hulls around bat species are drawn using bat families (red), suborders (blue), and geographic regions (green).

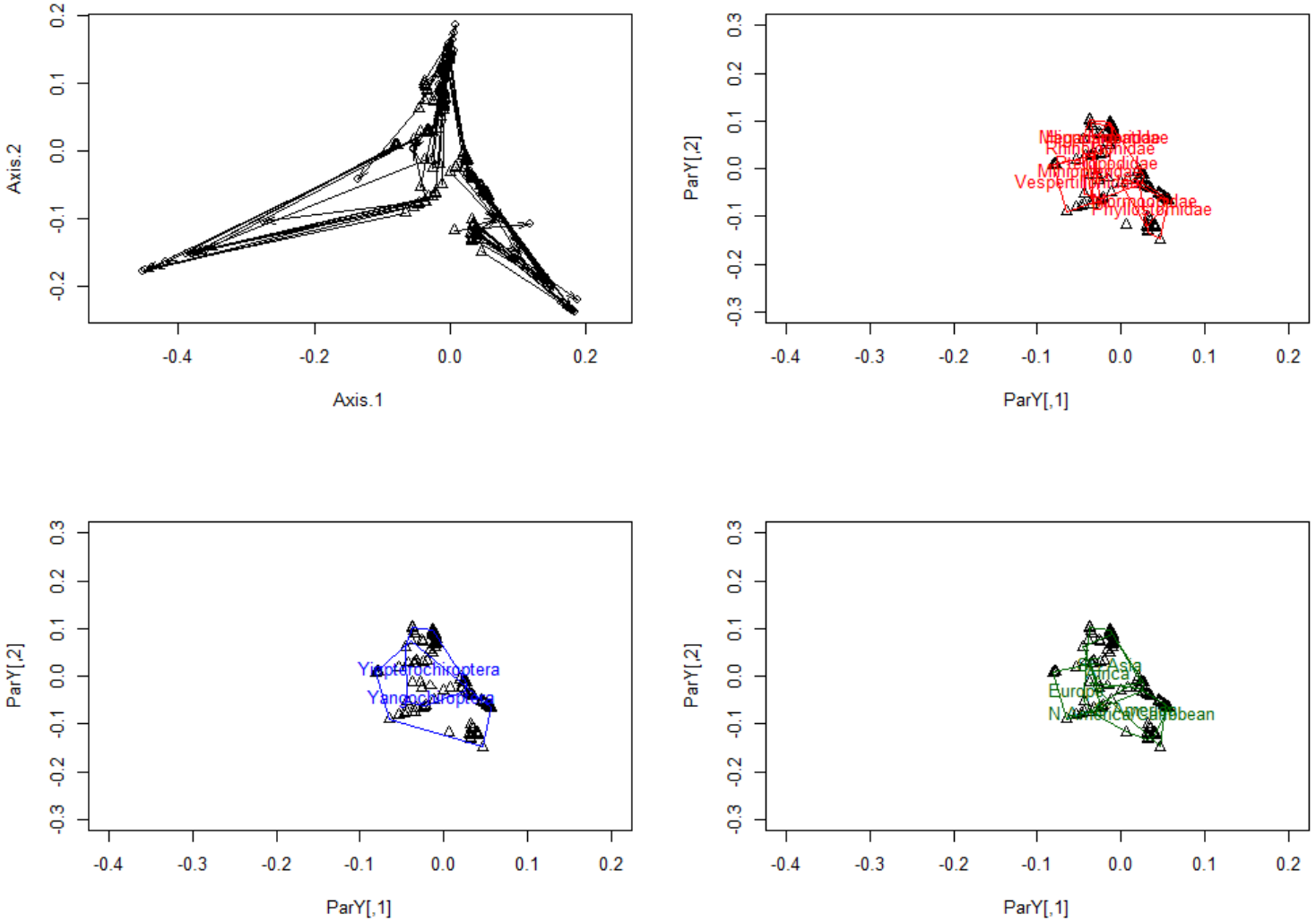

Figure A2.5 Procrustes ordination (Balbuena *et al.* 2013) using distances calculated from maximum likelihood trees of bat and *Bartonella* sequences. *Bartonella* genotypes are represented as triangles and bat species are represented as circles. Hulls around *Bartonella* genotypes are drawn using host bat families (red), suborders (blue), and geographic regions (green).

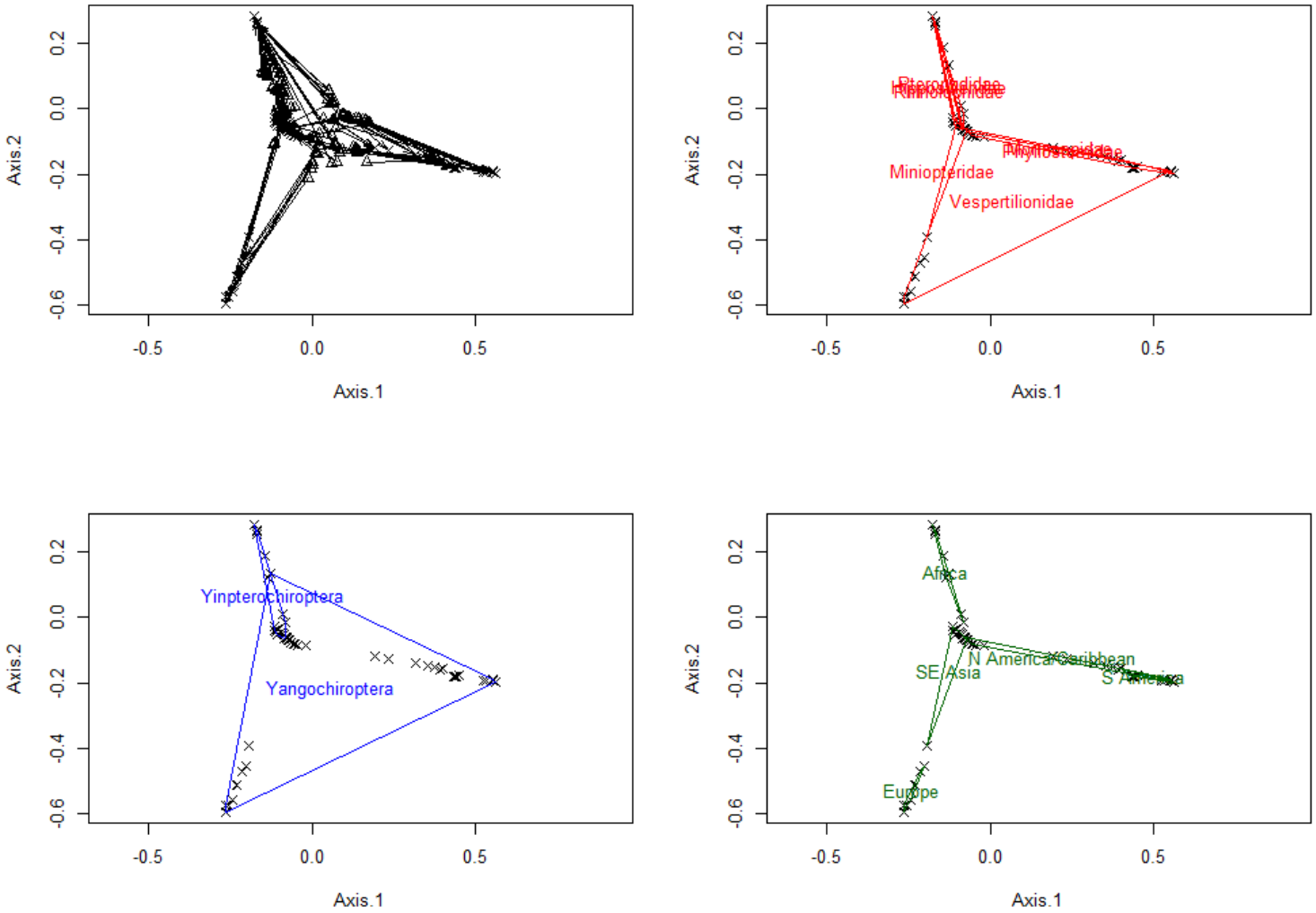

Figure A2.6 Procrustes ordination (Balbuena *et al.* 2013) using distances calculated from maximum likelihood trees of *Bartonella* sequences and host sympatry. *Bartonella* genotypes are represented as triangles and bat species range overlaps are represented as x symbols. Hulls around *Bartonella* genotypes are drawn using host bat families (red), suborders (blue), and geographic regions (green).

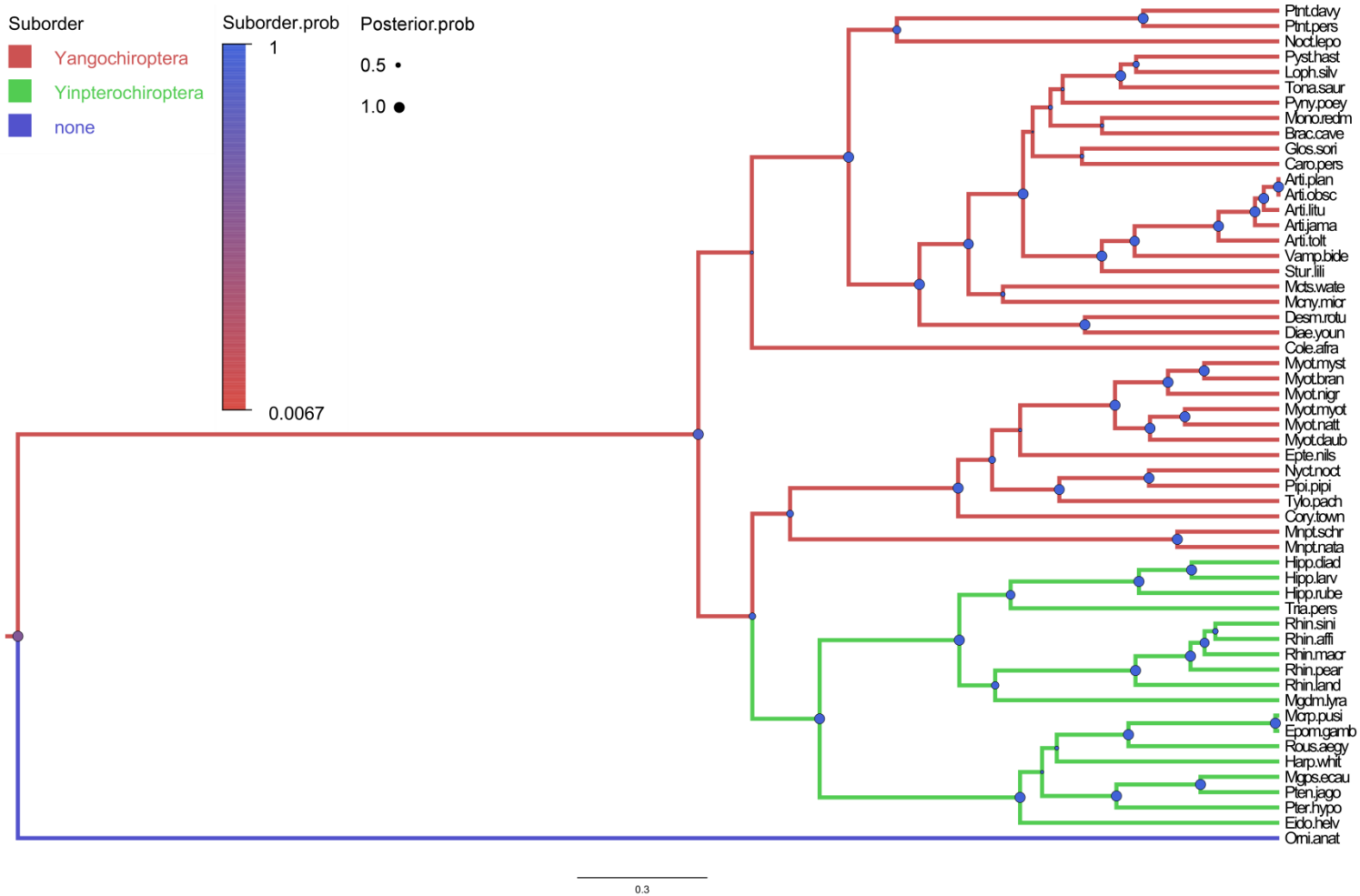

Figure A2.7 Bayesian phylogeny of bat host species reconstructing bat suborders, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of bat cytochrome *b* (*cytb*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the bat suborder at that node (state.prob). Suborders are based on current taxonomic classifications for bats (Teeling *et al.* 2002; Agnarsson *et al.* 2011). Mean tree likelihood (-ln) = 23077.7, ESS = 7073; mean suborder tree likelihood (-ln) = 11, ESS = 8357. Details on tip labels for bat species are listed in Table A2.3.

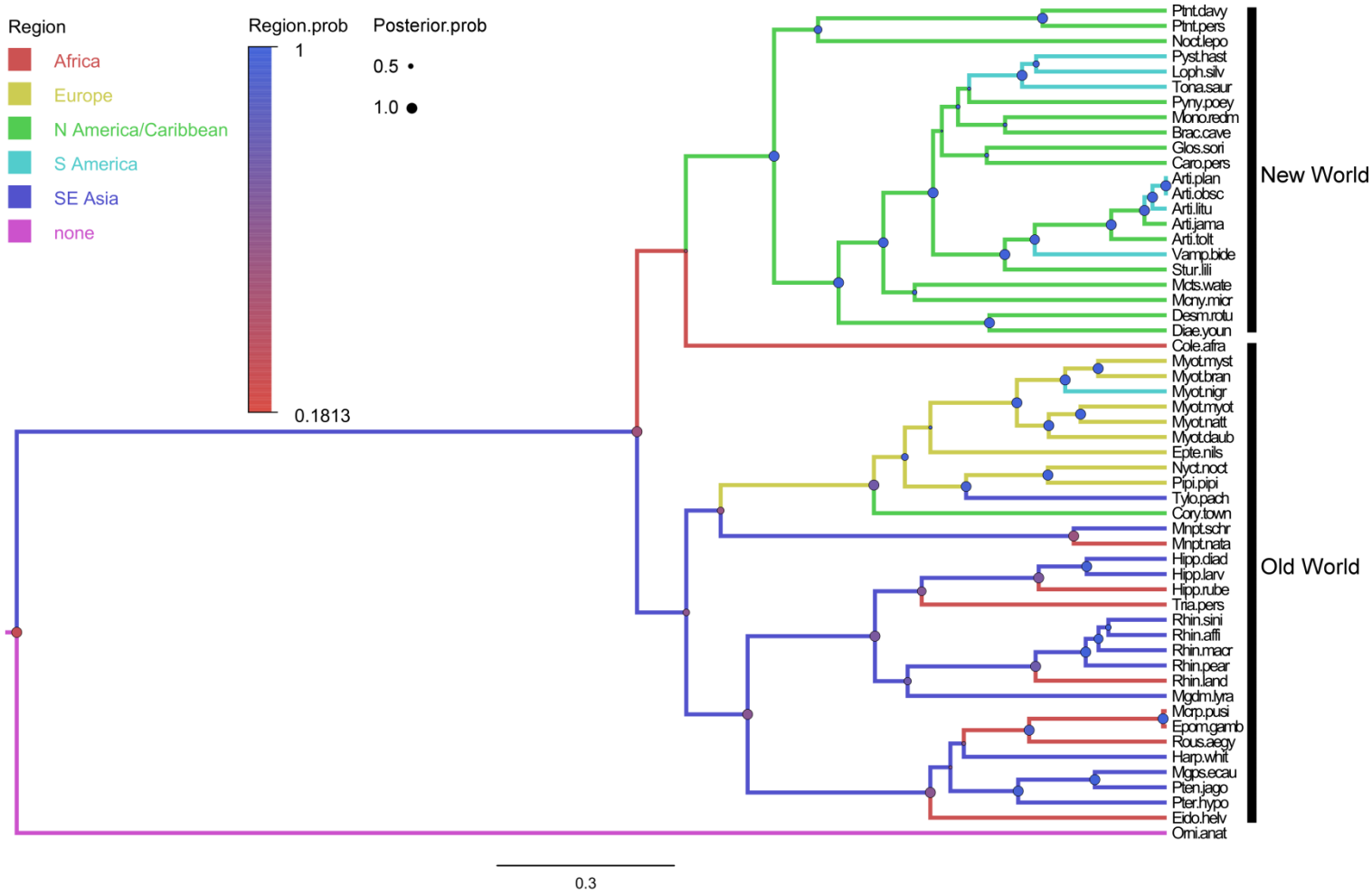
Region
Africa
Europe
N America/Caribbean
S America
SE Asia
none
Region.prob
1
0.1813
Posterior.prob
0.5
1.0
Ptnt.davy
Ptnt.pers
Noct.lepo
Pyst.hast
Loph.silv
Tona.saur
Pyny.poey
Mono.redm
Brac.cave
Glos.sori
Caro.pers
Arti.plan
Arti.obsc
Arti.litu
Arti.jama
Arti.tolt
Vamp.bide
Stur.lili
Mcts.wate
Mcny.micr
Desm.rotu
Diae.youn
Cole.afra
Myot.myst
Myot.bran
Myot.nigr
Myot.myot
Myot.natt
Myot.daub
Epte.nils
Nyct.noct
Pipi.pipi
Tylo.pach
Cory.town
Mnpt.schr
Mnpt.nata
Hipp.diad
Hipp.larv
Hipp.rube
Tria.pers
Rhin.sini
Rhin.affi
Rhin.macr
Rhin.pear
Rhin.land
Mgdm.lyra
Mcrp.pusi
Epom.gamb
Rous.aegy
Harp.whit
Mgps.ecau
Pten.jago
Pter.hypo
Eido.helv
Omi.anat
New World
Old World
0.3

Figure A2.8 Bayesian phylogeny of bat host species reconstructing bat geographic regions, shown by colored branches. The tree was assembled from a MAFFT (Katoh and Standley 2013) alignment of bat cytochrome *b* (*cytb*) sequences using the GTR+G (Nei and Kumar 2000) substitution model in BEAST 1.8.2 (Drummond *et al.* 2012) and visualized using FigTree version 1.4.2 (University of Edinburgh). Posterior probabilities for nodes are shown as circles (●) scaled by size from 0 to 1 (posterior.prob) and colored by the support for the geographic region at that node (state.prob). Clades of bat species are separated into Old World and New World groups. Note the geographic region represents where the bat host was captured, which may not reflect its total range. Mean tree likelihood (-ln) = 23077.7, ESS = 7073; mean region tree likelihood (-ln) = 62.2, ESS = 9001. Details on tip labels for bat species are listed in Table A2.3.

REFERENCES

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ (2011) A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). PLoS Currents, 3, RRN1212.

Bai Y, Hayman DT, McKee CD, Kosoy MY (2015) Classification of *Bartonella* strains associated with straw-colored fruit bats (*Eidolon helvum*) across Africa using a multi-locus sequence typing platform. PLoS Neglected Tropical Diseases, 9, e0003478.

Bai Y, Kosoy MY, Recuenco S et al. (2011) *Bartonella* spp. in bats, Guatemala. Emerging Infectious Diseases, 17, 1269–1272.

Bai Y, Recuenco S, Gilbert AT et al. (2012) Prevalence and diversity of *Bartonella* spp. in bats in Peru. The American Journal of Tropical Medicine and Hygiene, 87, 518–523.

Balbuena JA, Míguez-Lozano R, Blasco-Costa I (2013) PACo: A novel procrustes application to cophylogenetic analysis. PLoS ONE, 8, e61048.

Billeter SA, Hayman DT, Peel AJ *et al.* (2012) *Bartonella* species in bat flies (Diptera: Nycteribiidae) from western Africa. Parasitology, 139, 324–329.

Concannon R, Wynn-Owen K, Simpson V, Birtles RJ (2005) Molecular characterization of haemoparasites infecting bats (Microchiroptera) in Cornwall, UK. Parasitology, 131, 489–496.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969–1973.

Kamani J, Baneth G, Mitchell M et al. (2014) *Bartonella* species in bats (Chiroptera) and bat flies (Nycteribiidae) from Nigeria, West Africa. Vector-Borne and Zoonotic Diseases, 14, 625–632.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30, 772-780.

Kosoy MY, Bai Y, Lynch T et al. (2010) *Bartonella* spp. in bats, Kenya. Emerging Infectious Diseases, 16, 1875–1881.

IUCN (2014) The IUCN Red List of Threatened Species. Version 2014.1. <http://www.iucnredlist.org/>. Downloaded on 1 April 2015.

Lin EY, Tsigrelis C, Baddour LM *et al.* (2010) Candidatus *Bartonella mayotimonensis* and endocarditis. Emerging Infectious Diseases, 16, 500-503.

Lin J-W, Hsu Y-M, Chomel BB et al. (2012) Identification of novel *Bartonella* spp. in bats and evidence of Asian gray shrew as a new potential reservoir of *Bartonella*. Veterinary Microbiology, 156, 119–126.

Morse SF, Olival KJ, Kosoy MY *et al.* (2012) Global distribution and genetic diversity of *Bartonella* in bat flies (Hippoboscoidea, Streblidae, Nycteribiidae). Infection, Genetics and Evolution, 12, 1717–23.

Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxfor University Press, New York.

Olival KJ, Dittmar K, Bai Y et al. (2015) *Bartonella* spp. in a Puerto Rican bat community. Journal of Wildlife Diseases, 51, 274-278.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30, 2725-2729.

Teeling EC, Madsen O, Van den Bussche RA *et al.* (2002) Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. Proceedings of the National Academy of Sciences of the United States of America, 99, 1431–1436.

Veikkolainen V, Vesterinen EJ, Lilley TM, Pulliainen AT (2014) Bats as reservoir hosts of human bacterial pathogen, *Bartonella mayotimonensis*. Emerging Infectious Diseases, 20, 960–967.

APPENDIX II, SECTION B

Species replacement details for cophylogeny analysis

Some host species in the *Bartonella* dataset were only identified to the genus level, but these genera were important to include because they represented unique families and their exclusion would have reduced the power of the analysis. Thus, a representative species was chosen based on a) geographic range overlap with the study capture location and b) the availability of *cytb* sequences on GenBank with similar length to others in the dataset (~1000 base pairs). With these criteria, *Pipistrellus pipistrellus* was chosen to represent *Pipistrellus* sp. from the UK, *Myotis nigricans* for *Myotis* sp. from Peru, *Tylonycteris pachypus* for *Tylonycteris* sp. from Malaysia, *Rhinolophus landeri* for *Rhinolophus* sp. from Nigeria, *Micropteropus pusillus* for *Micropteropus* sp. from Nigeria (misidentified as "*Micropterus* sp." in Kamani *et al.* 2014), *Epomophorus gambianus* for *Epomophorus* sp. from Nigeria, *Miniopterus natalensis* for *Miniopterus* sp. from Kenya, *Hipposideros ruber* for *Hipposideros* from Kenya, and *Hipposideros diadema* for *Hipposideros* sp. from Malaysia. These replacements are marked with an asterisk (*) in Table A2.3. For other bat species, no *cytb* sequences could be found or they were too short (much less than 1000 base pairs). Thus, a suitable replacement was found based on a) geographic range overlap with the study capture location and b) close relatedness to the species caught in the study. With these criteria, I replace *Tonatia silvicola* with *Tonatia saurophila*, *Phyllostomus discolor* with *Phyllostomus hastatus*, *Megaderma spasma* with *Megaderma lyra*, *Rhinolophus affinis* for *Rhinolophus borneensis chaseni*, *Megaerops ecaudatus* for *Megaerops niphanae*, and *Rhinolophus macrotis* for *Rhinolophus acuminatus*. These replacements are marked with a dagger (†) in Table A2.3. Lei and Olival (2014) made similar replacements for species in their analysis, although without the stipulation that the geographic range of the substitute species should overlap with the capture location. The inclusion of this

criterion is important in the present study because of the dual focus on bat phylogeny and sympatry.

There were several options substitute species of hosts that had a) geographic ranges overlapping with the capture location and b) *cytb* sequences near 1000 base pairs long:

Pipistrellus sp. (UK) – *P. pipistrellus*, *P. nathusii*, *P. pygmaeus*

Myotis sp. (Peru) – *M. nigricans, M. albescans, M. keaysi, M. riparius, M. simus*

Rhinolophus sp. (Nigeria) – *R. landeri, R. fumigatus*

Miniopterus sp. (Kenya) – *M. natalensis, M. fraterculus*

Hipposideros sp. (Kenya) – *H. ruber, H. caffer, H. camerunen, H. cyclops, H. gigas*

Hipposideros sp. (Malaysia) – *H. diadema, H. armiger, H. ater, H. bicolor, H. cervinus, H. cineraceus, H. larvatus, H. pomona, H. ridleyi*

Tonatia silvicola (Peru) – *T. saurophila*, *Lophostoma silvicolum*, *L. brasiliense*, *L. carrikeri*

For *Micropteropus* sp. and *Epomophorus* sp. from Nigeria, only *Micropteropus pusillus* and *Epomophorus gambianus* had ranges that overlapped with the capture location. *Tylonycteris pachypus* was the only species with a suitable *cytb* sequence for *Tylonycteris* sp. *Phyllostomus discolor* is present in both Guatemala and Peru and only *Phyllostomus hastatus* is also present in both locations. *Megaderma lyra* and *Megaerops ecaudatus* were the only other member of their genera in Vietnam to replace *Megaderma spasma* and *Megaerops niphanae*, respectively. For *Rhinolophus borneensis* subsp. *chaseni* and *R. acuminatus*, the search for replacement species could not be narrowed by the capture location within Vietnam the article associated with these sequences has not been published. However, there were short *cytb* sequences available on GenBank, so these sequences were searched using the BLAST program (NCBI, Bethesda, MD).

The most closely related species also found in Vietnam were chosen as replacements, *R. affinis* and *R. macrotis*, respectively.

It would be infeasible to test how each choice of replacement affects the results of the cophylogenetic analysis, so I perform a sensitivity analysis by using a second set of suitable replacement species in the global fit tests and correlation between bat phylogeny and sympatry. For the sensitivity analysis, I choose *Pipistrellus pygmaeus* (AJ504442) to represent *Pipistrellus* sp. from the UK, *Myotis riparius* (JX130570) for *Myotis* sp. from Peru, *Rhinolophus fumigatus* (FJ457614) for *Rhinolophus* sp. from Nigeria, *Miniopterus fraterculus* (AJ841975) for *Miniopterus* sp. from Kenya, *Hipposideros caffer* (FJ347980) for *Hipposideros* sp. from Kenya, *Hipposideros pomona* (DQ054810) for *Hipposideros* sp. from Malaysia, and *Lophostoma carrikeri* (JF923844) for *Tonatia silvicola* from Peru.

Using the maximum likelihood trees of bat species and *Bartonella* genotypes, global fit tests find a strongly significant cophylogenetic signal (ParaFitGlobal = 16.6, P = 1E-5; m^2 global value = 11.6, P = 1E-4). The correlation between bat phylogeny and sympatry remains strong with the substitute species (Pearson correlation coefficient = 0.34, P = 1E-5). Global fit tests were repeated using bat sympatry and the maximum likelihood tree of *Bartonella* genotypes, finding a signficant trend (ParaFitGlobal = 72.3, P = 1E-5; m^2 global value = 45.2, P = 1E-4). Our sensitivity analysis indicates that the choice of substitute species does not significantly affect the results of the cophylogeny tests.

REFERENCES

Kamani J, Baneth G, Mitchell M *et al.* (2014) *Bartonella* species in bats (Chiroptera) and bat flies (Nycteribiidae) from Nigeria, West Africa. Vector-Borne and Zoonotic Diseases, 14, 625–632.

Lei BR, Olival KJ (2014) Contrasting patterns in mammal-bacteria coevolution: *Bartonella* and *Leptospira* in bats and rodents. PLoS Neglected Tropical Diseases, 8, e2738.