HOST SPECIFICITY AND PROACTIVE SURVEILLANCE OF INFECTIOUS DISEASES

Maxwell J. Farrell

Doctor of Philosophy

Department of Biology

McGill University

Montreal, Quebec

2018-11-20

A thesis submitted to McGill University in partial fulfillment
of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

Abstract

Host specificity – the number and types of species that a parasite infects – can influence disease dynamics and the probability of emergence in novel hosts. The majority of human and domesticated animal diseases can infect more than one host, but often the range of susceptible species is unknown. Developing theory and tools for studying multi-host parasites can help reduce the burden of neglected diseases by identifying undocumented reservoir species, tracking pathogens in biodiversity hotspots, predicting the outcome of novel host-parasite associations, and increasing our understanding of how human impacts alter host-parasite systems. Studying the ecology and evolution of multi-host parasites provides baseline knowledge of parasites before they emerge, potentially as threats to global health, and forms the basis of proactive surveillance. In Chapter 1, I present a general framework for proactive disease surveillance through the study of parasite sharing. The tendency for parasites to infect closely related host species is a common property of diverse host-parasite systems, and allows for the development of theory that can be applied to the surveillance of a broad range of organisms. In Chapter 2, I use the evolutionary relationships among hosts to identify gaps in a global database of host-parasite associations for mammals. With this approach I generate ranked lists of likely, yet currently undocumented host-parasite associations, which may be targets for future surveillance. One method to fill these gaps in our knowledge of host-parasite associations and discover cryptic parasite biodiversity is the sequencing of DNA present in environmental samples. In Chapter 3, I use molecular and bioinformatic approaches to explore the bacterial diversity across waterholes in the Kruger National Park, South Africa. This area, marked by high mammal diversity, frequent cross-species contact, and endemic multi-host diseases, is an ideal location to develop these methods while generating baselines for biodiversity monitoring using environmental DNA. In addition to identifying undocumented host-parasite associations, the evolutionary relationships among hosts may

be used to predict the impact parasites will have on a given host species. While much research has focused on the drivers of host mortality in parasites infecting single host species, we lack theory for predicting the mortality of multi-host parasites. In Chapter 4, I use a global database of domesticated mammal diseases to show that the evolutionary relationships among hosts can be used to predict disease-induced mortality. I find parasites infecting distantly related hosts are more likely to result in fatal infections. Many of these domesticated mammal diseases regularly spill over and contribute to population declines in wild species. As wildlife are driven extinct, theory predicts that single-host parasites are at greatest risk of coextinction following declines in their host species. In Chapter 5, I test this prediction in a comparative study of threatened and non-threatened mammals and find, counter to prediction, that threatened ungulates are associated with fewer multi-host parasites. This indicates that the response of parasites to host endangerment varies with host life history, and that human activities driving species to extinction can modify the specificity of parasite assemblages. Overall, my work highlights the importance of host specificity for understanding the ecology and evolution of parasitism, how human pressures on ecosystems can alter disease ecology, and offers novel approaches for expanding our knowledge of host-parasite associations.

Abrégé

La spécificité des parasites chez les hôtes - le nombre et les types d'espèces qu'un parasite peut infecter - peut influencer la dynamique de la maladie et la probabilité d'émergence chez de nouvelles espèces. La majorité des maladies chez les humains et les animaux domestiques peuvent infecter plus d'une espèce d'hôte, mais l'aire de répartition des espèces sensibles est souvent inconnue. En développant la théorie et les outils pour étudier les parasites multi-hôtes, nous pouvons contribuer à la réduction du fardeau des maladies négligées de plusieurs manières: l'identification des espèces de réservoir non documentées, le suivi des agents pathogènes dans les régions à haute biodiversité, la prévision des résultats des nouvelles associations hôte-parasite et une meilleure compréhension de l'impact des impacts humains sur les systèmes hôte-parasite. L'étude de l'écologie et de l'évolution des parasites multi-hôtes fournit des connaissances de base sur les parasites avant leur apparition, et constitue la base d'une surveillance proactive qui peut éviter des menaces potentielles pour la santé mondiale. Au chapitre 1, je présente un cadre général pour la surveillance proactive des maladies infectieuses en étudiant le partage des parasites entre les hôtes. Les parasites sont connus pour infecter des hôtes étroitement apparentés dans une diversité de systèmes hôte-parasite. Cela nous permet d'élaborer des théories pouvant être appliquées à la surveillance de nombreux types d'organismes. Au chapitre 2, j'utilise les relations évolutives entre les hôtes pour identifier les lacunes dans une base de données mondiale sur les associations hôte-parasite chez les mammifères. Avec cette méthode, je génère des listes classées d'associations hôte-parasite probables, mais actuellement non documentées, qui pourraient être des cibles pour une surveillance future. Une méthode pour combler les lacunes dans notre connaissance des associations hôte-parasite et découvrir la biodiversité cryptique des parasites est le séquençage de l'ADN présent dans les échantillons environnementaux. Au chapitre 3, j'utilise des approches moléculaires et bioinformatiques pour explorer la diversité bactérienne à travers les trous

d'eau du parc Kruger, en Afrique du Sud. Cette zone, marquée par la diversité élevée des mammifères, les contacts fréquents entre les différentes espèces et les maladies multi-hôtes endémiques, est un endroit idéal pour développer ces méthodes, tout en générant des données de base pour la surveillance de la biodiversité en utilisant l'ADN environnemental. En plus d'identifier les associations hôte-parasite non documentées, les relations évolutives entre les hôtes peuvent être utilisées pour prédire l'impact que les parasites auront sur certaines espèces hôdémontretes. Bien que de nombreuses recherches aient porté sur les facteurs de mortalité chez les parasites infectant des espèces hôtes uniques, nous manquons de théorie pour prédire la mortalité des parasites multi-hôtes. Au chapitre 4, j'utilise une base de données mondiale sur les maladies des mammifères domestiqués pour montrer que les relations évolutives entre les hôtes peuvent être utilisées pour prédire les infections mortelles. Je démontre que les parasites infectant des hôtes évolutifs éloignés sont plus susceptibles de provoquer des infections mortelles. Bon nombre de ces maladies peuvent entraîner un déclin des populations d'espèces sauvages. Lorsque la faune est éteinte, la théorie prédit que les parasites à un seul hôte sont les plus exposés au risque de co-extinction après le déclin de leur seule espèce hôte. Au chapitre 5, je teste cette prédiction dans une étude comparative de mammifères menacés et non menacés. Contrairement aux prévisions, je trouve que les ongulés menacés sont associés à moins de parasites multi-hôtes. Cela indique que la réponse des parasites au processus d'extinction de l'hôte varie selon les caractéristiques de l'hôte et que les activités humaines qui causent l'extinction des espèces peuvent modifier la spécificité des assemblages de parasites. Ensemble, mes travaux soulignent l'importance de la spécificité de l'hôte pour comprendre l'écologie et l'évolution du parasitisme, comment les pressions humaines sur les écosystèmes peuvent modifier les interactions hôte-parasite et offrent de nouvelles approches pour élargir nos connaissances des associations hôte-parasite.

Acknowledgements

Throughout my studies I have been blessed to be a part of an amazingly supportive community, not only in terms of my academic pursuits, but for any endeavour I've wanted to undertake. With my friends, family, collaborators, and broader academic community, I have felt encouraged to explore any subject my curiosity led me to, and for this I will forever be grateful.

Above all I need to thank my supervisor Jonathan Davies for taking a chance on what started as an overambitious and seemingly far-fetched idea. With his unwavering guidance and inspiration this idea evolved into a set of projects that I would not have dreamt to be possible. I have grown immensely as a researcher, communicator, and collaborator under his supervision. The incredible set of collaborations he has forged, and the amazing group of people he brought together in his lab is a testament to his kindness, thoughtfulness, and brilliance. It was a privilege be a member of the Davies lab and to have worked alongside Hercules Araclides, Michael Becker, Frederic Boivin, Sofia Carvajal Endara, Chelsea Chisholm, Ted Daigle, Tammy Elliott, Jean-Philippe Lessard, Malie Lessard-Therrien, Emily Meineke, Nicolas Perrault, and Stephanie Shooner. I need to give special thanks to Ria Ghai and Jan Gogarten for stimulating conversations and advice in all topics disease related, and to Ignacio Morales-Castilla and Will Pearse for their mentorship and for being role models of remarkable scientists.

I am grateful for the support of Lea Berrang-Ford, and my incredible supervisory committee members Melania Cristescu and Virginie Millien for their thoughtful comments and encouragement. I thank Anthony Ricciardi and Lauren Chapman, who offered supervision and exposure to research while I was an undergraduate. I also thank Steven Kembel, Frederic Guichard, Pedro Peres-Neto, and Elizabeth Wolkovich who welcomed and adopted me as an honorary member of their fantastic labs. At McGill, and through interaction with the QCBS, I was able to learn from and engage with an impressive group

of emerging scholars. In particular I would like to thank Sarah Balwdin, Marc-Oliver Beausoleil, Pedro Henrique Pereira Braga, Charles Cong Xu, Vincent Fugere, Dalal Hanna, Emma Hudgins, Logan James, Joanne Littlefair, Elizabeth Nyboer, and Tim Thurman for their insights and engagement throughout the writing of this thesis. I also need to thank Zofia Taranu, Amanda Winegardner, and Russell Steele for organizing the Statistics-Biology Exchange, a remarkable initiative and setting for the sharing of ideas. Though this group I met Mohamad Elmarsi who has become one of my closest collaborators, someone who I am constantly learning from, and who I hope to work with on many more projects. I would also like to thank the members of the Macroecology of Infectious Disease Research Coordination Network for allowing me to participate and collaborate with some of the greatest minds in disease ecology.

Chapter 3 would not have been possible without the logistical support of the African Centre for DNA Barcoding, Ronny Mukala Kabongo, Purvance Shikwambana, molecular work by Shannon Eagle, encouragement and support from Lea Berrang-Ford, and the incredible knowledge and protection provided by Kruger game guards Thomas, Velly, and Martin. I also thank Steven Kembel, his lab, Alexis Carteron, and Benjamin Callahan for constructive feedback on bioinformatic analyses. Special thanks are due to B.S. Bezeng and Cody Danaher for volunteering their friendship, dedication, and long days in the field. This work was performed with the support and permission of SANParks (project DAVIJ1256) with project funding supported by the Quebec Centre for Biodiversity Science, the McGill Biology Department, and an NSERC Discovery Grant awarded to Jonathan Davies.

For Chapter 4 I thank Vanessa Ezenwa, Charlie Nunn, Elizabeth Wolkovich, Ria Ghai, Jan Gogarten, and Carl Boodman for helpful feedback on the manuscript. Special thanks are due to Elizabeth Wolkovich, Margaret Kosmala, the Harvard Stanleyi group, Will Pearse, and Bob Carpenter for feedback on the model, and to Debarun Gupta & Madeleine McGreer for help with data entry.

For Chapter 5 I thank the Altizer and Ezenwa lab groups at Odum School of Ecology,

Amy Pedersen, Liam Brierley, and the Woolhouse lab group at the University of Edinburgh, Chris Solomon, and the Davies lab at McGill University for insightful comments on initial analyses.

In addition to my academic community and wonderful collaborators, I wouldn't have been able to finish this thesis without the support of my family and friends. To my parents, Carolyn and Stephen, thank you for supporting me, for fostering a passion for learning, and for feeding my natural curiosity from a young age. Thank you to my aunt Kathy, my uncle David, my cousin Matthew, and my grandmother Margo for providing me a home away from home as I became my own person in Montreal. Thank you to my uncle Michael, and my cousin Rose for your positivity, and a special thank you to my aunt Nancy for your unwavering support and your keen eye in proof-reading. Thank you to Liam Blackwell for living with me through this entire journey and for the sanctuary we have created together. Thanks to Carl Boodman for your friendship, stimulating chats, and for inspiration in all aspects of life. Thank you to the friends and loved ones who have been there for me throughout – there are too many of you to thank, but you know who you are. I also am forever indebted to Andrea Reid for allowing me to be her undergraduate field assistant in Uganda, an experience that changed my life and provided the initial inspiration for this entire thesis. Finally, I need to thank my musical and artistic community in Montreal for allowing me to express myself, for driving me to create, and for bringing balance to my life.

Throughout my studies I was fortunate to receive funding from the Natural Science and Engineering Research Council of Canada, the Vanier Canada Graduate Scholarship, McGill University, the McGill Biology Department, the Canadian Institutes of Health Research Systems Biology Training Program, the Quebec Centre for Biodiversity Science, the Queen Elizabeth II Diamond Jubilee Scholarship.

Contribution to Original Knowledge

This thesis takes an interdisciplinary approach to the study of host-parasite interactions, framed in the context of proactive approaches for the surveillance of infectious diseases. Using global databases, phylogenetic information, hierarchical Bayesian modelling, field surveys, and next generation sequencing, I develop theory and methods to better understand the ecology and evolution of multi-host parasites.

My first chapter is a literature-based review of approaches to strengthen disease surveillance through the study of parasite sharing among species. This chapter outlines a general framework for proactive surveillance of infectious diseases, and informs the structure of my subsequent thesis chapters. My proposed approach involves identifying gaps in current knowledge of host-parasite associations, using host specificity to better understand disease dynamics, and documenting the potential for disease transmission through tracking of animal contact.

My second chapter applies a novel method for predicting host-parasite interaction using evolutionary relationships among hosts. Host phylogeny is a strong predictor of parasite community similarity, and using a novel method developed in collaboration with Mohamad Elmasri (McGill PhD Statistics 2017), we leverage this information to produce lists of highly probable yet previously undocumented host-parasite interactions. I apply this method to a global database of host-parasite associations and present approaches for recursively filling in these missing links.

My third chapter explores the potential for next generation sequencing to describe bacterial diversity from environmental samples. Using the watering holes of the Kruger National Park, South Africa as a system with frequent cross-species contact, I characterize bacterial diversity present across the watering holes, and describe community variation across space, time, and environmental factors.

My fourth chapter asks the question "why are some diseases deadly?". The factors

determining host impact resulting from infection, termed virulence, has long been a topic of study in the evolution of infectious diseases. However, we currently lack theory for predicting the virulence of diseases that infect multiple host species. Using global case-fatality data for multiple diseases of domestic mammals I show that evolutionary relationships among infected and susceptible hosts is a strong predictor of disease-induced mortality.

Finally, my fifth chapter adopts a comparative approach to the study of host-parasite coextinction. Traditional coextinction theory predicts that parasites infecting single host species are the most vulnerable to extinction following severe declines in the abundance of their sole host. By comparing the relative proportions of single-host and multi-host parasites across threatened and non-threatened hosts, I show that among ungulates the decline to extinction is associated with a loss of multi-host parasites.

Thesis Format

This thesis is written in a manuscript-based format, and consists of five manuscripts for which I am the lead author. Between each chapter, short linking statements are included to review the findings of the previous chapter and connect them to topics explored in the following chapter. Two of the included manuscripts have been published in peer-reviewed journals, one has been accepted pending revisions, and the remaining two will be submitted for publication. Throughout I use the Chicago citation style.

Chapter 1: Farrell, M.J., Berrang-Ford, L., Davies, T.J. (2013), The study of parasite sharing for surveillance of zoonotic disease. *Environmental Research Letters* 8 (1), 015036.

Chapter 2: Farrell, M.J., Elmasri, M., Stephens, D., Davies, T.J. Link Prediction in Global Host-Parasite Networks. *Prepared for submission to PLoS Neglected Tropical Diseases*.

Chapter 3: Farrell, M.J., Govender, D., Hajibabaei, M., van der Bank, M., Davies, T.J. Bacterial diversity in the waterholes of the Kruger National Park: an eDNA metabarcoding approach. *Accepted to Genome pending revisions*.

Chapter 4: Farrell, M.J., Davies, T.J., Disease mortality in domesticated animals is predicted by host evolutionary relationships. *Prepared for submission to Science*.

Chapter 5: Farrell, M.J., Stephens, P.R., Berrang-Ford, L., Gittleman, J.L., Davies, T.J. (2015), The path to host extinction can lead to loss of generalist parasites. *Journal of Animal Ecology* 84(4), 978–984.

Contribution of Authors

I (MJF) am the first author for all chapters and the appendix in this thesis, and the sole author on all non-manuscript portions of this thesis, including the General Introduction, Linking Statements, and General Discussion & Conclusion. During the generation of these manuscripts I received assistance from a number of co-authors who contributed their insights, technical expertise, and offered logistic support:

Chapter 1: MJF wrote the manuscript with input from LBF and TJD.

Chapter 2: MJF prepared the data and wrote the manuscript with input from TJD. ME and MJF performed the analyses with input from TJD. MJF and TJD designed the study.

Chapter 3: MJF wrote the manuscript with input from TJD. MJF conducted the field work and bioinformatic and statistical analyses. DG and MvdB provided logistical support and expertise for sample collection and processing. MH designed the molecular workflow and performed the DNA extractions, amplifications, and sequencing. MJF and TJD designed the study with input from MH, DG, and MvdB.

Chapter 4: MJF wrote the manuscript with input from TJD. MJF prepared the data and performed the analyses. MJF and TJD designed the study.

Chapter 5: MJF wrote the manuscript with input from TJD, LBF, PRS, and JG. MJF prepared the data and performed the analyses. MJF and TJD designed the study.

List of Figures

List of Tables

General Introduction

Infectious diseases are responsible for severe health burdens in humans, domesticated animals, and wildlife around the globe (Hotez et al., 2014; Grace et al., 2012; Smith et al., 2009; Daszak et al., 2000). The majority of these diseases have the ability to infect more than one species (Taylor et al., 2001; Cleaveland et al., 2001), but despite the severe burdens imposed by multi-host parasites, we often lack fundamental knowledge about their basic ecologies and life histories. Throughout this thesis I use *parasite* and *pathogen*, and *infectious disease* interchangeably to describe any disease-causing organism ranging from viruses, bacteria, and protozoa (commonly termed *micro-parasites* or *pathogens*) to helminths and arthropods (commonly termed *macro-parasites*) (Anderson and May, 1979; May and Anderson, 1979; Lafferty and Kuris, 2002; Stephens et al., 2016). Although I use these terms interchangeably, typically *parasite* describes an ecological relationship, *pathogen* refers to a disease causing organism, and *infectious disease* is the disorder caused by infection.

A broad scale approach to the study of multi-host parasites may help reduce the impacts these organisms have on wildlife conservation, human health via direct infection, and loss of livelihoods resulting from disease in domesticated animals. Identifying overarching patterns in the ecology and evolution of multi-host diseases can contribute theory and tools for strengthening the surveillance of infectious diseases around the world. By understanding the forces that shape the interactions among hosts and parasites, we can move towards a proactive approach to surveillance in which we study infectious organisms before they emerge as threats to people, domesticated animals, or wildlife.

The transition from a nomadic to a sedentary lifestyle and the domestication of animals beginning roughly 10,000 years ago resulted in a major shift in human infections (Armelagos et al., 1996). Diseases responsible for some of the greatest contemporary human health burdens, such as tuberculosis, measles, pertussis, and falciparal malaria,

are thought to have originated from wildlife and emerged as human diseases as population densities dramatically increased, and our reliance on domestic species created opportunities for sustained transmission of animal diseases (Pearce-Duvet, 2006; Wolfe et al., 2007). Although the origins of many major human diseases can be traced back to the onset of agriculture, the rapid human development and increased global connectedness that has occurred over the past century has facilitated the emergence of a new set of human diseases (Woolhouse and Gowtage-Sequeria, 2005; Woolhouse and Gaunt, 2007; Jones et al., 2008). These diseases are caused primarily by pathogens that are transmitted to humans directly from wildlife, or are shared among humans, wildlife, and domestic species (Cleaveland et al., 2001; Jones et al., 2008). Domesticated animals can facilitate the transmission of parasites that originate in wildlife (Daszak et al., 2000) and the progressive transformation of natural habitats has led to the emergence of diseases that are either newly recognized, newly evolved, or have undergone recent expansions into a new areas, host species or vectors (Schlundt et al., 2004; Jones et al., 2013). While much research into the drivers of disease emergence are framed in the context of zoonoses (human diseases of animal origin), human activities have also led to the emergence of devastating diseases in wild species (Dobson and Foufopoulos, 2001; Daszak et al., 2001), many of which are attributed to the transmission of diseases from domesticated animals (Pedersen et al., 2007; Smith et al., 2009). As the demand for livestock products increases around the world, we are likely to see the continued emergence of multi-host diseases with considerable global impact (Perry et al., 2011; Jones et al., 2013). This trend is also likely to facilitate the evolution of antibiotic resistant parasites (Jones et al., 2013), another major concern for disease surveillance (Morens et al., 2004), though I do not focus on it here.

At the forefront of these interfaces are often subsistence livestock keepers, many of which live on less than $2 USD per day and collectively suffer the greatest burdens of zoonotic diseases (International Livestock Research Institute, 2012). These populations lack the basic infrastructure necessary for disease diagnosis, reporting, and control, and

as a consequence suffer heavy burdens from emerging diseases as well as long standing neglected zoonoses (Perry and Grace, 2009; Maudlin et al., 2009; Molyneux et al., 2011). Poor livestock keepers often live in areas of high biodiversity (Fisher and Christopher, 2007; International Livestock Research Institute, 2012), highlighting the importance of programs that support the surveillance of endemic and emerging multi-host diseases that also infect wildlife (Halliday et al., 2012). Effective monitoring of these diseases requires interdisciplinary research that brings together the expertise of ecologists, evolutionary biologists, veterinarians, epidemiologists, statisticians, molecular biologists, social scientists, and public health officials (Kruse et al., 2004; Daszak et al., 2013; Wood et al., 2012; Lebov et al., 2017). By integrating knowledge from across these fields we can generate the theory and infrastructure necessary to mitigate the impacts of multi-host infectious diseases. Typically, investigation into the ecology of emerging infectious diseases and their susceptible hosts is undertaken only after they have been found to infect humans (Wolfe et al., 2007; Brownstein et al., 2008; Chan et al., 2010). To more effectively respond to infectious disease threats, the global community needs to adopt a paradigm of proactive surveillance in which we study infectious organisms before they shift to infect humans, domesticated animals, or endangered wildlife. The threat of emerging diseases at human-livestock-wildlife interfaces has led to a substantial body of research, but these studies are heavily biased geographically and towards a handful of zoonotic diseases (Wiethoelter et al., 2015), leaving a significant knowledge gap, especially with regard to diseases that impact livestock and wildlife health.

Advances in addressing this gap range from identifying molecular changes in parasites that facilitate the transmission of diseases to novel hosts (Longdon et al., 2014) to describing patterns in the global biogeography of infectious diseases (Dunn et al., 2010). Determining the specific ecology, transmission routes, and susceptible hosts of a given parasite is essential for designing effective control programs (Viana et al., 2014), and may also be informed by studying the distributions and impacts of parasites at the broadest scales of

organization (Stephens et al., 2016). Essential knowledge for the proactive surveillance of infectious diseases includes approaches for describing the diversity of host-parasite interactions and determining the drivers and impacts of parasite sharing across hosts. The number and types of hosts a pathogen infects, broadly termed host specificity, can influence the dynamics of transmission, outbreak, and the likelihood of emergence in novel hosts (Daszak et al., 2000; Woolhouse, 2002; Dobson, 2004; Woolhouse and Gowtage-Sequeria, 2005; Parrish et al., 2008; Allison et al., 2012). The ability to infect multiple species can also facilitate the persistence of parasites as their wild hosts decline towards extinction (Woolhouse et al., 2001; Deredec and Courchamp, 2003; Koh et al., 2004; Pedersen et al., 2007). Therefore, an important first step in the proactive surveillance of infectious diseases is to gain a clearer picture of the susceptible hosts of contemporary multi-host parasites.

In addition to identifying susceptible hosts, it is important to describe the diversity of potential pathogens in areas of high biodiversity. The encroachment of humans and domesticated animals into natural systems facilitates host shifts of wildlife diseases into domestic and human hosts (Daszak et al., 2013; Jones et al., 2013; Faust et al., 2018). In many regions considered to have a high risk of harbouring future emerging diseases, even baseline surveys of biological diversity are lacking (Hopkins and Nunn, 2007; Pedersen and Davies, 2009; Allen et al., 2017). While effective diagnostic tests have been developed for many important diseases, the surveillance of parasitic biodiversity may be expanded rapidly through the adoption of recent advances in genetic sequencing. The sequencing of DNA present in environmental samples is revolutionizing modern biodiversity surveys (Lodge et al., 2012; Taberlet et al., 2012; Bohmann et al., 2014; Deiner et al., 2017). Applying this approach to systems of high cross-species contact will be useful for building baseline surveys of cryptic biodiversity, tracking the distributions of multiple parasites simultaneously, and discovering previously unknown organisms that may have potential to emerge as infectious disease threats in the future.

Diseases that infect both domestic and wild species cause significant economic losses

(Dehove et al., 2012), and can contribute to severe declines in host populations (Heard et al., 2013). Yet some diseases rarely harm their hosts, while others nearly always fatal (Bisson et al., 2015). The ability to infect multiple hosts may contribute to the evolution of virulence and increased mortality due to infection (Brown et al., 2002; Osnas and Dobson, 2012; Alizon, 2013). At a fundamental level, parasites must do some harm to their hosts in order to successfully reproduce and transmit to infect new individuals (Alizon et al., 2009; Cressler et al., 2016). However, for multi-host parasites, conflicting trade-offs may select for greater severity of disease in some host species (Woolhouse et al., 2001; Gandon, 2004; Antonovics et al., 2013). Despite decades of research on mechanisms determining the outcome of infection in single-host single-parasite systems, we still lack a robust framework for predicting the impact of multi-host parasites in different host species (Leggett et al., 2013).

Highly virulent diseases have been implicated in the declines of endangered wildlife (Smith et al., 2009; Heard et al., 2013); however, infectious diseases also play key roles in healthy ecosystems, including the regulation of host populations (Wood et al., 2007), and promoting host genetic diversity (Altizer et al., 2003). Parasites comprise a major component of biodiversity (Dobson et al., 2008), but are often neglected as conservation targets (Gómez and Nichols, 2013). As humans push wildlife towards extinction, this not only creates opportunities for the transmission of wildlife diseases to humans and domestic animals, but may also cause the extinction or extirpation of parasites that infect threatened hosts. The extinction of parasites in natural ecosystems may promote diseases shifting to infect novel species, and their loss has been suggested as a cause of recent increases in the number of emerging infectious diseases (Dunn et al., 2009). Theory predicts that single-host parasites are more susceptible to coextinction than multi-host parasites (Anderson and May, 1979; Koh et al., 2004; Lafferty, 2012), but this is also likely to depend on the particular life histories of the hosts declining towards extinction (Colwell et al., 2012). Testing and generating novel theories of host-parasite coextinction is critical not only for

prioritizing surveillance efforts in the face of shifting species ranges and anthropogenic alteration of natural ecosystems, but also for monitoring ecosystem health and stability.

This thesis presents an interdisciplinary approach to inform the proactive surveillance of infectious diseases. In Chapter 1, I review recent advances in the study of parasite sharing across species that may be useful for the proactive surveillance of zoonotic diseases. The framework is presented in the context of emerging diseases in humans, but can be applied to any host-parasite system. The core of the approach involves identifying gaps in our current knowledge of host-parasite interactions, and developing technologies to track contact among hosts and parasites. In Chapter 2, I build on this framework by using a novel statistical model to predict undocumented links in a global database of host-parasite interactions for mammals. I then conduct targeted literature searches of the top "missing" links, and highlight interactions that should be the focus of future surveillance efforts. In Chapter 3, I explore the potential for using the sequencing of environmental DNA to track infectious organisms and build baseline surveys of bacterial diversity in areas of high cross-species contact. In Chapter 4, I address the problem of predicting disease impacts in different host species by exploring the relationship between host specificity and disease-induced mortality for domesticated animal diseases. In Chapter 5, I use a comparative approach to look for evidence of parasite coextinction in wild carnivores and ungulates, and test the theory that single-host parasites are more likely to be lost as their hosts decline to extinction. Finally, I conclude by reviewing the results from each chapter, and discuss future applications and avenues of research that expand on my findings.

CHAPTER 1
The study of parasite sharing for surveillance of zoonotic diseases

Maxwell J. Farrell[1], Lea Berrang-Ford[2] & T. J. Davies[1,3]

Author Affiliations:

[1]Department of Biology, McGill University

[2]Department of Geography, McGill University

[3]African Centre for DNA Barcoding, University of Johannesburg

This chapter is published in *Environmental Research Letters* 8 (1), 015036.

1.1 Abstract

Determining the factors that influence the transmission of parasites among hosts is important for directing surveillance of animal parasites before they successfully emerge in humans, and increasing the efficacy of programs for the control and management of zoonotic diseases. Here we present a review of recent advances in the study of parasite sharing, wildlife ecology, and epidemiology that could be extended and incorporated into proactive surveillance frameworks for multi-host infectious diseases. These methods reflect emerging interdisciplinary techniques with significant promise for the identification of future zoonotic parasites and unknown reservoirs of current zoonoses, strategies for the reduction of parasite prevalence and transmission among hosts, and decreasing the burden of infectious diseases.

1.2 Introduction

The majority of human emerging infectious diseases are of animal origin (Taylor et al., 2001), and zoonotic diseases, infectious diseases caused by parasites transmissible between humans and animals, contribute significantly to the global health burden and can impose severe economic losses (Greger, 2007; Grace et al., 2012). Many current

diseases causing significant global burden likely crossed the species barrier from animal populations to humans thousands of years ago (e.g. malaria, tuberculosis, measles) (Wolfe et al., 2007; Pearce-Duvet, 2006), while others have emerged more recently in human populations (e.g. human immunodeficiency virus (HIV), severe acute respiratory syndrome (SARS), influenza A/ H1N1). Considerable effort has focused on identifying the drivers facilitating infectious disease emergence from animal hosts, which include a diverse array of interacting social, political, environmental, biological, and ecological factors (Woods and Versalovic, 1993; Dunn et al., 2010; Daszak et al., 2013; Plowright et al., 2008; Daszak et al., 2001; Weiss and McMichael, 2004). Analyses of historical patterns of zoonotic disease emergence and identification of these drivers has formed the basis for surveillance of both novel and re-emerging zoonotic parasites (Flanagan et al., 2012), which has resulted in an increased emphasis on interdisciplinary research that bridges taxonomic divides and incorporates the overarching drivers of emergence for the effective prediction, surveillance, and management of zoonotic diseases (Borer et al., 2011; Woolhouse, 2011; Wilkinson et al., 2011). Defining the environmental and biological factors that facilitate zoonotic disease emergence is an important first step for prediction of future infectious disease risks, but it is challenging to include large scale drivers such as land use, climate change, and globalization within an explicit mechanistic framework of disease emergence (Woolhouse et al., 2012) unless the causal pathways influencing emergence can be teased apart from the networks of indirect effects associated with these drivers (Plowright et al., 2008; Eisenberg et al., 2007). In addition, the historical context and associated mechanisms behind emergence events are useful when determining appropriate actions for responding to disease outbreaks and minimizing the impacts of previously emerged diseases, but may be less useful for identifying novel infectious agents before they shift from animal reservoirs to human hosts.

Zoonotic disease surveillance is typically undertaken only after the detection of a novel illness in humans (Chan et al., 2010; Brownstein et al., 2008) and has predominantly focused on identifying human actions that promote contact with animals, which include bushmeat hunting, handling livestock, the wildlife trade, and expansion of land use practices into previously wild regions that facilitate disease emergence (Wolfe et al., 2005; Gómez and Aguirre, 2008; Tomley and Shirley, 2009; Patz et al., 2004; Chomel et al., 2007). Situations such as these, which increase the probability of human exposure to animal parasites, have been prioritized for surveillance of novel zoonotic diseases. Contemporary surveillance approaches have proven successful in documenting novel simian immunodeficiency viruses (Aghokeng et al., 2010) and have contributed to an increased understanding of transmission risk for early detection. Despite these advances, surveillance and monitoring of zoonoses remains largely reactive in that typically the emergence of a parasite into human populations must occur before research is conducted to determine its patterns of transmission, the health impacts on infected hosts, or the suite of hosts it is able to utilize. Many recent human viruses including SARS coronavirus, Ebola and Marburg viruses, Nipah virus, Hendra Virus, and simian variants of human immunodeficiency virus types 1 and 2 were not known to infect wildlife until after first being documented in humans (Parrish et al., 2008).

Wolfe et al. (2007) have highlighted the reactive nature of the current surveillance paradigm, noting the need to move from opportunistic sampling of wildlife to new systematic efforts to detect infectious pathogens prior to shifts into human populations which would allow for increased efficiency of control programs and permit accelerated responses in the face of novel epidemics. A more proactive approach to surveillance would facilitate the precursory development of vaccines or other treatments, highlight potential transmission routes and reservoir species to efficiently isolate disease spread after an initial epidemic, and aid in classification of sentinel species used to monitor outbreaks of zoonotic diseases before appearance in human populations. A shift towards proactive surveillance

and early detection necessitates baseline documentation of the variety and nature of multi-host animal parasites, including knowledge of contemporary infectious diseases of wildlife and domesticated animals, and an understanding of the ecology of these parasites and their known hosts. Identifying factors that promote parasite expansion, either geographically, in abundance, or in host range, will help prioritize the monitoring of these parasites and further development of proactive control and management programs. Here we present a review of recent advances in the study of parasite sharing, wildlife ecology, and epidemiology that provide promise for advancing surveillance frameworks for the control of zoonotic diseases. We focus on four priority areas of research and methodological development: (1) Identification of host species that are understudied and may harbour future zoonotic parasites, (2) identification of unknown reservoirs of current zoonotic parasites, (3) prediction of parasites that are likely to be transmissible to humans, and (4) monitoring the movements of potential reservoir populations to inform actions for limiting future contact with humans or other susceptible hosts that may promote emergence.

1.3 Identifying gaps in baseline knowledge

Identification of parasites that pose a risk for emergence in human populations requires knowledge of existing host-parasite associations from which we can infer future human transmission potential. This necessitates the systematic documentation of host infection by parasites. A complete knowledge of all parasites and the susceptibilities of hosts to infection is beyond our reach, but existing datasets provide useful starting points for gathering such information. These data can be used to produce a list of known animal parasites and allow profiling of important traits such as parasite type (virus, bacteria, protist, helminth, fungi, etc.), transmission mode (sexually transmitted, vector borne, water borne, etc.), genomic or proteomic markers for rapid identification or development of treatments, and the range of hosts that are known to be susceptible. Importantly, such data can be used to identify gaps in the sampling of wildlife hosts and their associated parasites.

Hopkins and Nunn (2007) illustrate one method for identifying taxonomic and geographic gaps in parasite sampling within non-human primates. Using the primate subsection of a comprehensive database of host-parasite associations for free-living mammals (Nunn and Altizer, 2005), and maps of primate geographical distributions, they highlight geographical regions where sampling of primate parasites is most lacking with respect to the diversity, taxonomy, and threat status of hosts, as well as the taxonomy of parasites. Such gap analyses are useful for revealing hosts and regions where future sampling is most likely to uncover previously undocumented parasites. The technique of gap analysis could also be conducted on a smaller scale. For example, a recent Ouranos funded project is interested in predicting the spread of Lyme disease in Canada in the face of climate change. Lyme disease, caused by the bacterium *Borrelia burgdorferi*, is transmitted via the tick vector *Ixodes scapularis* to a variety of vertebrate reservoirs and hosts, including humans. The preferred host of *I. scapularis* is the white footed mouse (*Peromyscus leucopus*), which is known to transmit the disease effectively, although studies have found that other small mammals may be important in the transmission cycle of Lyme (Bouchard et al., 2011). One aspect of the project is to identify the diversity of small mammal hosts in southern Quebec and target sampling of these species to determine the differential preference of *I. scapularis*, and prevalence of Lyme, in order to predict patterns of expansion and emergence under different climate change scenarios. Targeted sampling such as this will greatly contribute to baseline data on parasites, including associations with hosts, and hence to our knowledge of the evolutionary and ecological factors that influence the dynamics of parasite distributions, prevalence, host-shifts, and disease outbreak.

1.4 Host specificity

The range of hosts that a parasite infects, also known as host specificity, can influence the dynamics of parasite transmission, disease outbreak, and emergence in novel hosts

(Woolhouse, 2002; Daszak et al., 2000; Allison et al., 2012; Dobson, 2004). The transmissibility and virulence of a parasite can differ dramatically among hosts, and the utilization of multiple hosts may help parasites avoid extinction by not being tied to the fate of one host species (Woolhouse et al., 2001). Host specificity is traditionally defined as the absolute number of host species utilized (Mouillot et al., 2006), but alternative methods have been proposed which take into account the geography, ecology, and taxonomic or evolutionary distances among hosts (Poulin and Mouillot, 2003, 2005). Most metrics are derived from presence/absence data for host-parasite associations, but can be modified to incorporate information on differential parasite prevalence among hosts (structural specificity), quantify changes in host use across the geographic range of the parasite (β-specificity), or compounded into metrics that quantify the phylogenetic turnover of utilized host species over geographic space (Poulin et al., 2011).

Phylogenetic metrics of host specificity are particularly useful when the host traits determining parasite preferences are unmeasured and/or unknown. Phylogeny, a representation of the evolutionary relationships among species, provides a means to quantify species similarity: closely related species are more likely to share physiological, biochemical, or behavioural traits that influence the successful infection, development, and transmission of parasites, although evolutionarily more labile traits might co-vary only poorly with phylogeny. One example of a trait that influences successful parasite sharing is the presence of phylogenetically conserved cell receptors for viral pathogens, which has been proposed as a useful tool for predicting whether or not a novel virus will be able to infect humans (Woolhouse et al., 2012). Phylogeny might, therefore, act as a proxy for cell receptor similarity between potential hosts. Experimental studies which cross-infected hosts and their specific parasites found that decreasing phylogenetic distance between hosts promoted successful parasite infection and reproduction (Perlman and Jaenike, 2003; Gilbert and Webb, 2007; de Vienne et al., 2009; Longdon et al., 2011), and comparative studies of free-living primates have shown that the phylogeny and geographic distribution of hosts

are strong predictors of parasite sharing (Cooper et al., 2012; Davies and Pedersen, 2008; Pedersen et al., 2005).

Determining the factors influencing the sharing of parasites among host species in ecological communities can allow the prediction of the range of hosts that a particular parasite might be able to infect. Predicting the potential host range of a parasite is critical for prioritizing surveillance efforts in the face of shifting animal ranges and the expansion of human land use practices, which have the potential to bring previously isolated host populations into contact and create novel opportunities for cross-species transmission (Plowright et al., 2008; Parrish et al., 2008; Reluga et al., 2007; Martin et al., 2011). For example, the phylogenetic relationship between hosts could be used as an index for intrinsic susceptibility to infection based on distance from a known host, and geographic overlap could be used as a proxy for the likelihood of contact between potential hosts (Pedersen and Davies, 2009). Under these assumptions, host species that are recently diverged and have large overlaps in their geographic ranges are most likely to share similar suites of parasites. This model may be helpful in the identification of previously undocumented reservoirs for current zoonotic parasites, or prioritization of monitoring for host species that are likely to become future reservoirs after a successful host shift.

The applicability of phylogenetic host specificities for predicting host switching events will likely vary depending upon the parasite type and transmission mode, as well as the strength of phylogenetic conservatism in host defence traits. Recently emerged parasites such as the coronavirus responsible for Severe Acute Respiratory Virus (SARS) and the 2009 pandemic strain of influenza A (H1N1) are examples of extremely large phylogenetic jumps between hosts which a predictive model based on phylogenetic host affinities may not have anticipated. Rapid generation times and high mutation rates, such as typical for RNA viruses, might facilitate large host jumps. However, examining the genetic and proteomic changes coincident with these distal host switching events may allow for identification of homologous viral strains in related reservoir species which may gain the

potential to shift hosts in the future. Additionally, some viruses known to have jumped large phylogenetic distances, such as rabies viruses and lentiviruses have been found to more often make small rather than large phylogenetic jumps between hosts (Streicker et al., 2010; Charleston and Robertson, 2002). For parasites that demonstrate frequent shifts between distantly related hosts, host phylogeny may be less informative, and more appropriate predictors may be identified from the geography and ecology of potential hosts. In these cases, similarity in life history traits or overlap in geographic ranges may be essential for promoting increased contact and opportunities for parasite exposure and cross-species infection.

1.5 Animal movement and contact

Monitoring the distribution of both hosts and parasites, and understanding the forces that modify ecological interactions among potential hosts will be critical for moving towards successful proactive surveillance of zoonotic diseases because host shifts are not possible unless there is opportunity for parasites to move between individuals of different host species. The movement and co-occurrence of host species is important not only for parasite transmission at global and regional scales associated with migration, species invasions, and the wildlife trade, but also impacts local disease dynamics (Fèvre et al., 2006). Many local opportunities for cross-species transmission can have far reaching effects when they involve species with long range dispersal such as humans, migratory species, or animals that are traded as commodities. By traveling large distances, these species can connect previously isolated populations and contribute to the long-range transport of parasites, as illustrated by the transmission of haematozoan parasites of migratory waterfowl (Figuerola and Green, 2000). The global transport of passengers and goods has been implicated in the spread of influenza pandemics, introduction of mosquito vectors, and increases in the range of falciparum malaria (Tatem et al., 2006). Livestock trade and complex market systems have the potential to mix infected hosts from distant sites

and often involve frequent close contact with humans involved in the raising, pasturing, transport, trade, and butchering of these animals (Fèvre et al., 2006). Additionally, the hunting of wild animals for nutritional purposes brings human hunters into direct contact with wild species harbouring zoonotic parasites, and the carcasses of hunted animals are often subject to long distance transportation via market systems and commodity chains involving multiple vendors (Wolfe et al., 2005; Aghokeng et al., 2010; Bachand et al., 2012; Kamins et al., 2011). The tendency for migrating species to travel long distances and over international boundaries often makes the tracking of movements difficult, but recent technological advances have permitted the use of satellite telemetry to track reservoir populations, such as fruit bats responsible for transmitting zoonotic Nipah and Hendra viruses (Smith et al., 2011).

The monitoring of migrating species and increasing understanding of how human activities and climate change alter species dispersal is essential to predict changes in contact patterns among susceptible hosts and the transmission of zoonotic infections. While the movement of infected individuals obviously increases a parasites geographical range, the relationship between animal movement and parasite transmission may be more complex. Altizer et al. (2011) suggest that migration might increase or decrease parasite prevalence depending on the parasites traits, such as transmission mode and host specificity, and that long distance migration is likely to decrease the prevalence of host-specific parasites while increasing the prevalence of generalist parasites able to infect both the migratory species as well as non-migratory resident species.

Examining the genetics of hosts and parasites across heterogeneous landscapes can be used to elucidate environmental drivers of parasite genetic diversity, quantify ecological processes such as gene flow and host movements that may indirectly influence parasite prevalence, and infer transmission patterns across various temporal and spatial scales (Biek and Real, 2010). Using *Escherichia coli* as a model system, Rwego et al. (2008) generated DNA fingerprints for *E. coli* isolates from humans, livestock, and gorillas around

Bwindi Impenetrable National Park, Uganda to map transmission routes. Repetitive DNA sequences are found throughout the bacterial genome and can be used to rapidly distinguish bacterial species and strains (Woods and Versalovic, 1993). Rwego et al. (2008) found that the variance in diversity of *E. coli* strains was higher within species than between, suggesting a larger number of multi-species strains than species-specific strains. In addition, they also found that habitat overlap contributed significantly to transmission: humans and livestock shared very similar strains, reflecting their close geographical proximity and frequent interactions, whereas the similarity of strains in humans and gorillas was found to be a function of the frequency of human-gorilla contact (strains of wild groups were less similar to humans than those for eco-tourism, and research gorilla groups intermediate) which may be reflective of direct exchange of microbes or indirect transmission contact through shared environment (Rwego et al., 2008). The use of genetic markers in this manner can provide pertinent information on the transmission pathways of multi-host pathogens and allow the estimation of contact rates at the scale of individuals, populations, communities.

Heterogeneity in parasite transmission within and across susceptible groups is important to consider when modeling epidemic dynamics and investigating the potential outcomes of control strategies. Traditional epidemiological models such as Susceptible-Infected-Recovered (SIR), metapopulation, or lattice-based approaches assume that all individuals are identical in their epidemiological traits that contribute to transmission, whereas network models adapted from statistical physics allow to the explicit inclusion of variation in contact patterns, infectiveness, and recovery rates among individuals (Craft and Caillaud, 2011). Craft and Caillaud (2011) review the application of network models for investigating contact structures in wildlife epidemiology, an approach that can be used in conjunction with contemporary methods for monitoring the movement of animal populations such as behavioural observations, mark and recapture surveys, video tracking, and radio or satellite telemetry. By merging the tracking of animal movements with landscape genetics,

network models can be generated for local and regional scale processes of changing land use, increasing agricultural intensity, eco-tourism, wildlife research, bushmeat hunting, and habitat fragmentation, which have been identified as modifiers of the distributions of species that promote increased contact between hosts (Wolfe et al., 2005; Goldberg et al., 2007; Alexander and Day, 2010). These models can be used to simulate control strategies targeted at particular species, or sub-groups that have been identified as "super-spreaders" – individuals contributing disproportionately to the transmission of infectious agents.

Spatially explicit models must also consider environmental variance, seasonality, and anthropogenic change as these factors can modify contact patterns, host susceptibility, and parasite prevalence (Altizer et al., 2006). Correlations between host and parasite locations and local environmental properties have been used to predict the distributions of 15 potentially interacting reservoir and vector species of Chagas disease throughout Mexico (Peterson et al., 2002). Techniques for the collection and analysis of geographic information such as satellite imagery and remote sensing have also been used to analyze environmental changes contributing to outbreaks of waterborne and vector borne zoonoses (Ford et al., 2009). Understanding the link between the environment and the biogeography of host and parasite interactions is especially critical for predicting the effects of climate change, which has the potential to alter seasonal regimes and shift both parasite and host distributions (Lafferty, 2009).

1.6 Beyond zoonotic diseases

Wild and domesticated animals have been proposed as sentinels for zoonotic diseases (Gubernot et al., 2008; Halliday et al., 2007), the monitoring of which would allow us to recognize outbreaks before they appear in human populations. The use of sentinels such as livestock and domesticated carnivores that interact frequently with both wildlife and human populations would be useful for monitoring changes in host contact patterns, or the rapid spread of previously endemic diseases. However, surveillance of wildlife and domesticated

animal parasites that are potentially harmful to humans should extend beyond the search for the next major emerging zoonotic disease, or unknown reservoirs of current infections. In many developing countries livestock remain a major livelihood resource (Herrero et al., 2009), and the health of domesticated animals has both direct and indirect impacts on those that are dependent on livestock livelihoods (Perry and Grace, 2009; Maudlin et al., 2009). The predicted rapid population growth of developing countries and concurrent increase in the demand for livestock products (Perry et al., 2011; Thornton, 2010) coupled with the documented sharing of parasites among humans, livestock, and wildlife (Cleaveland et al., 2001) highlights the need for proactive surveillance of wildlife parasites that may emerge in livestock. The overlap of biodiversity hotspots and human poverty (Fisher and Christopher, 2007), the global correlation between bird and mammal richness and the number of human pathogens (Dunn et al., 2010), and the consistent under-reporting of infectious disease burdens in the developing world (Maudlin et al., 2009) highlights the need for surveillance of multi-host parasites in these regions to identify future emerging disease threats, as well as unknown endemic diseases that may be currently afflicting these populations. Recent efforts that have employed molecular markers and rapid genetic sequencing of retroviral and bacterial pathogens of primates targeted by the bushmeat trade are particularly successful examples of proactive surveillance. These studies have taken into account host-parasite associations, host geography and ecology, as well as the environmental and social factors that increase contact and could facilitate emergence of primate parasites in human populations (Aghokeng et al., 2010; Bachand et al., 2012; Locatelli and Peeters, 2012; Wolfe et al., 2004; Betsem et al., 2011).

We suggest that the same tools described in this paper for guiding surveillance and monitoring of zoonotic parasites could be applied to any infectious disease organism, although we recognize that there is an urgent need to first increase current baseline information on host-parasite associations across a greater breadth of host taxa. Using livestock as an example, lists of parasitic and infectious diseases have been compiled (Lefèvre

et al., 2010) and could be merged with reported host-parasite associations for ungulates (Cetartiodactlya plus Perissodactlya, minus cetaceans (Nunn and Altizer, 2005)), which represent the group of terrestrial mammals most closely related to the major five livestock species: cows, goats, sheep, pigs, and horses. This information can then be used to distinguish gaps in the sampling of ungulate parasites and quantify the applicability of host phylogenetic affinities for predicting parasite sharing among ungulates. Factors influencing the transmission of infections among wild and domestic ungulates in Europe have already been identified (Martin et al., 2011), and may help direct surveillance programs towards high risk areas. Models of infectious disease spread have been produced for parasites of some wild ungulate species, such as cervids of North America, which has informed management programs for limiting the prevalence and spread of multi-host parasites such as bovine tuberculosis and brucellosis which are also known to infect livestock (Conner et al., 2008). Analyzing the environmental and anthropogenic factors that facilitate aggregations of wild ungulates (e.g. Bhola et al. (2012)) could be used to infer previously undocumented reservoir interactions, or combined with landscape genetic techniques to estimate the degree of parasite transmission among these species, humans, livestock, or other domesticated animals. Monitoring overlaps in the distributions of livestock and related wildlife species and quantifying transmission of parasites between hosts may uncover multi-host transmission dynamics which can then be integrated with environmental and ecological data, and contemporary livestock transport networks to develop a continuously updating surveillance program that would help reduce the disease burden in livestock and improve the well-being of those reliant upon them.

1.7 Conclusion

Understanding broad patterns driving host-parasite associations can aid in the prediction of novel disease emergence for humans, domesticated animals, and wildlife, and will be essential in designing effective control programs for emerging infectious diseases as

well as neglected endemic diseases. Through the amalgamation of baseline ecological data and focusing on the four priority research areas highlighted in this review: (1) identification of understudied host species, (2) identification of unknown reservoirs of current zoonotic parasites, (3) prediction of parasites that are likely to be transmissible to humans, and (4) monitoring the movements of potential reservoir populations, we can identify areas with inadequate surveillance relative to a high probability of cross-species parasite transmission. Directing surveillance in this manner will help to generate more explicit tests of the drivers of parasite sharing between species, allow for increased accuracy when predicting novel emerging disease events, identify reservoirs of contemporary infectious agents, and decrease the burden of zoonotic diseases.

Linking Statement 1

Chapter 1 presents a framework for the proactive surveillance of infectious diseases based on the study of parasite sharing among hosts. A central focus of this framework is identifying the suite of hosts that a particular parasite is likely to infect. With this information we can direct research towards parasites that may pose a risk for emergence in humans, domesticated animals, or wildlife of interest. Based on the observation that closely related species often host similar parasites, I suggest that host phylogeny may inform the likelihood of a given host-parasite combination.

In Chapter 2, I use a novel method for link prediction in ecological networks to identify missing links in a global database of host-parasite interactions for mammals. The method, described in Elmasri et al. (2017), was developed in collaboration with Mohamad Elmasri, a former doctoral student in the Department of Mathematics & Statistics at McGill. The method uses only the structure of observed interactions and the evolutionary relationships among hosts to identify links that are highly likely, yet undocumented. I apply this method to the most comprehensive host-parasite database available, amalgamating four global databases of host-parasite interactions for mammals. After applying the link prediction method of Elmasri et al. (2017), I conduct literature searches for evidence of the top missing links. This chapter demonstrates an iterative approach to proactive disease surveillance that involves synthesizing current knowledge, identifying gaps, predicting missing links, building literature based databases, and targeting undocumented interactions that may be causing unseen disease burdens, or that could become established in the future.

CHAPTER 2
Proactive disease surveillance through link prediction in global host-parasite networks

Maxwell J. Farrell[1], Mohamad Elmasri[2], T. J. Davies[3,4]

Author Affiliations:

[1]Department of Biology, McGill University

[2]Department of Mathematics & Statistics, McGill University

[3]Botany, Forest & Conservation Sciences, University of British Columbia

[4]African Centre for DNA Barcoding, University of Johannesburg

2.1 Introduction

Infectious diseases are a significant cause of death and disability in humans, domesticated animals, and wildlife. Most disease causing organisms of human and domesticated animal can infect multiple species (Cleaveland et al., 2001; Taylor et al., 2001), which has ramifications for biodiversity conservation (Smith et al., 2009) and human health via direct infection, food insecurity, and diminished livelihoods (Grace et al., 2012). The number of emerging infectious diseases has significantly increased over the last 60 years, with the majority of human diseases caused by pathogens crossing the species boundary from animals to infect humans (Jones et al., 2008). Cross-species disease transmission has been implicated in the declines of endangered wildlife (Daszak et al., 2000) and acts in synergy with other drivers of extinction (Heard et al., 2013). The human burden of multi-host diseases falls largely on the world's poor livestock keepers (International Livestock Research Institute, 2012), many of which live in biodiversity hotspots (Fisher and Christopher, 2007) where resources for disease surveillance and reporting are lacking (Halliday et al., 2012). Despite the severe burdens they impose, we do not know the full range of susceptible host species for the vast majority of infectious organisms (Hopkins and Nunn, 2007; Dallas et al., 2017). Predicting host species that may be susceptible to infection can help to

improve disease monitoring and control programs for infectious diseases around the world. There is therefore an urgent need to fill these gaps. Here, we show how it is possible to leverage current knowledge of host-parasite associations to identify missing links in global host-parasite networks, and suggest that this offers a cost-effective approach to guide the surveillance of both emerging and neglected diseases.

For a given pathogen, knowledge of susceptible hosts is critical for understanding disease spread and persistence in multi-host systems (Viana et al., 2014). Predicting the suite of potential host species for known infectious diseases may also allow for more effective disease control, rapid response in the event of disease emergence (i.e. increases in prevalence, geographic spread, or infection of novel host species), and reduce risk of disease spillover and emergence through limiting cross-species contact of potential hosts (Roeder et al., 2013; Miller et al., 2013; Büscher et al., 2018; Molyneux and Sankara, 2017). While investigation into the specific ecologies and range of natural hosts often takes place after diseases have emerged as public health threats, iterative prediction and verification of potential host species provides one method of proactive surveillance that can strengthen capacities for disease monitoring and control of multi-host pathogens before emergence (Farrell et al., 2013).

Filling gaps in our knowledge of host-parasite networks additionally provides insight into the ecological and evolutionary forces shaping host-parasite interactions and disease spread. This includes the role of network structure for transmission (Gomez et al., 2013; Pilosof et al., 2015), the nature of "forbidden links" (Morales-Castilla et al., 2015), and the drivers of parasite richness (Nunn et al., 2003; Huang et al., 2015; Ezenwa et al., 2006; Kamiya et al., 2014) and parasite sharing across hosts (Davies and Pedersen, 2008; Huang et al., 2014; Braga et al., 2015; Luis et al., 2015). Determining the roles of wildlife hosts in disease transmission can also identify factors driving the evolution of highly virulent parasites (Shwab et al., 2018).

The search for undescribed pathogens and novel host-parasite associations forms the basis of proactive disease surveillance programs (Wolfe et al., 2007; Farrell et al., 2013). Such efforts have been adopted by researchers studying primate infectious diseases (e.g. Ghai et al. (2014)), and disease exposure associated with the handling and consumption of wild-caught meat (Wolfe et al., 2005; Kamins et al., 2011; Aghokeng et al., 2010). These studies focus on interfaces of high human-wildlife contact and attempt to discover novel pathogens, often by metagenomic analyses of blood, feces, or tissue samples. These sequencing-based approaches may be scaled up to estimate the number of unknown viruses with zoonotic potential (Anthony et al., 2013), and are being expanded to identify the majority of these viruses in an effort to mitigate the impact of future pandemics (Carroll et al., 2018). While exploratory approaches aimed at identifying unknown viruses are important for describing the depth of viral diversity and building local capacities for disease surveillance, they do not directly address the issue that we do not even know the animal reservoirs for many diseases currently of public health importance (Geoghegan and Holmes, 2017).

Here, we describe an alternative approach to identifying susceptible species of known diseases that uses existing global databases of host-parasite associations to predict missing links. By allowing for more directed surveillance efforts through targeted sampling of the most likely missing links, it may be possible to both save money and improve the efficiency of detecting undocumented host-parasite interactions. Recent efforts have employed machine learning to identify potential rodent and bat reservoirs of zoonotic diseases (Han et al., 2015, 2016), and illustrate that link prediction in ecological networks can generate accurate predictions. However, they frequently rely on detailed knowledge of ecological and functional traits of interacting species (Gravel et al., 2013; Morales-Castilla et al., 2015; Bartomeus et al., 2016; Dallas et al., 2017). In some cases only a small number of traits may be necessary to accurately predict whether two species interact (Eklöf et al., 2013; Dallas et al., 2017), yet the scale and taxonomic diversity of global host-parasite

databases make their implementation challenging. Few traits are likely be available for all species (Morales-Castilla et al., 2015) and for very diverse taxa there may only be a few comparable traits. Further, as these databases grow to include more species it becomes increasingly difficult to exhaustively search the literature for trait or interaction data if not directed, especially since much of the data can be in grey literature. Trait-based methods for link prediction are therefore powerful for smaller sized or local host-parasite networks, but can scale poorly. Algorithms such as recommender systems can be more flexible, drawing strength from the size of the network, and allowing for the identification of highly probable links in a variety of large networks (see Ricci et al. (2011) for a review). Here we employ a recently developed hierarchical Bayesian approach for link prediction in bipartite ecological networks inspired by recommender systems (Elmasri et al., 2017). This method is particularly well-suited to link prediction in global ecological networks as it does not require trait data, and allows us to make accurate predictions of missing links using only the structure of the observed host-parasite associations and the evolutionary relationships among hosts.

Our implementation draws on the evolutionary theory linking species ecologies and their phylogenetic histories (Felsenstein, 1985; Harvey and Pagel, 1991). Phylogenetic trees are representations of the evolutionary relationships among species, and provide a means to quantify species similarity (Wiens et al., 2010). Hosts may be associated with a parasite through inheritance from a common ancestor, or as a result of parasites shifting to use novel host species (Page, 1993). In both cases we expect closely related species will host similar parasite assemblages (Davies and Pedersen, 2008). Over the past decade, studies of natural and experimental host-parasite systems across diverse taxa have increasingly shown cross-species parasite sharing to be constrained by host phylogeny (Perlman and Jaenike, 2003; Gilbert and Webb, 2007; Davies and Pedersen, 2008; Streicker et al., 2010; Longdon et al., 2011; Huang et al., 2014; Braga et al., 2015; Luis et al., 2015). Host phylogenetic relationships can thus capture the influence of parasite inheritance from

common ancestors, and also serve as proxies for unmeasured physiological, biochemical, or behavioural traits that influence the successful infection, development, and transmission of parasites. By using host phylogeny to predict missing links in large host-parasite networks we can make accurate predictions with limited data, and generate ranked lists of highly likely, but currently undocumented, hosts of known infectious organisms which could be targeted for disease surveillance.

2.2 Materials and Methods

Data

We aimed to generate a list of highly probable, yet previously undocumented host-parasite interactions for humans, domesticated animals, and wildlife. First, we amalgamated four recently published global host-parasite databases for mammals derived from primary literature, genetic sequence databases, and natural history collections. The Global Mammal Parasite Database 2.0 (Stephens et al., 2016) contains records of disease causing organisms (viruses, bacteria, protozoa, helminths, arthropods, and fungi) in wild ungulates (artiodactyls and perrisodactyls), carnivores, and primates drawn from over 2700 literature sources published through 2010 for ungulates and carnivores, and 2015 for primates. The static version of the Enhanced Infectious Disease Database (EID2) (Wardeh et al., 2015) contains 22,515 host-pathogen interactions from multiple kingdoms based on evidence published between 1950 - 2012 extracted from the NCBI Taxonomy database, NCBI Nucleotide database, and PubMed citation and index. The Host-Parasite Database of the Natural History Museum, London (Gibson et al., 2005) contains over a quarter of a million host-parasite records for helminth parasites extracted from 28,000 references published after 1922, and is digitally accessible via the R package *helminthR* (Dallas, 2016). Finally, Olival et al. (2017) compiled a database of 2,805 mammal-virus associations for every recognized virus found in mammals, representing 586 unique viral species and hosts from 15 mammalian orders.

To amalgamate these databases, host names were standardized to those in Wilson and Reeder (2005) and used in the Fritz et al. (2009) mammal phylogeny. Virus names were standardized to the 2015 version of the International Convention on Viral Taxonomy (Lefkowitz et al., 2018). For non-viral parasites there exists no single accepted authority for nomenclature or species concepts, however all parasite names were thoroughly checked for typographical and formatting errors. When potentially synonymous names were identified (e.g. *Cylicocyclus insigne* and *Cylicocyclus insignis*), online searches of primary literature and taxonomic databases including the Integrated Taxonomic Information System (www.itis.gov), Catalogue of Life (http://www.catalogueoflife.org), World Register of Marine Species (www.marinespecies.org), Encyclopedia of Life (www.eol.org), NCBI Taxonomy Database (https://www.ncbi.nlm.nih.gov/taxonomy), UniProt (www.uniprot.org), and the Global Names Resolver (https://resolver.globalnames.org) were conducted to verify synonymy. Synonymous names were corrected to the name with the majority of references, or to the preferred name in recently published literature or taxonomic revision when this this information was available. Host associations for parasite names that were later split into multiple species were removed from the dataset (e.g. *Bovine papillomavirus*). All hosts and parasites reported below species level were assigned to their respective species (e.g. *Alcelaphus buselaphus jacksoni* was truncated to *Alcelaphus buselaphus*), and any species reported only to genus level or higher (e.g. *Trichostrongylus sp.*) was removed.

The resulting network includes 29,112 documented associations among 1835 host and 9149 parasite species (Fig. 2–1). To our knowledge this constitutes the largest host-parasite interaction database assembled for mammals, and includes parasites from diverse groups including viruses, bacteria, protozoa, helminths, arthropods, and fungi, and wild, domestic, and human hosts. The resulting matrix is quite sparse, with $\sim$0.17% of the $\sim$16.8 million possible links having documented interactions. Parasite species are largely represented by helminths (63.9%), followed by bacteria (13.1%) and viruses (7.89%). The number of documented interactions per species (degree distribution) varies considerably and is shown

to be linear on the log scale for both hosts and parasites (Fig. 2–2). This database comprises presence-only data as we do not have consistent information about the strength or nature of each interaction, only that it has been documented at least once by any accepted method (direct observation, genetic sequencing, or serology). Therefore interactions are taken as binary (0/1 for a given host-parasite interaction). Absences should not be considered true absences and are likely to include some host-parasite associations that (1) have been observed but which are not recorded in the original database, (2) currently exist but have not yet been observed or are undocumented, and that (3) may be realized in the future given sufficient opportunity.

Statistical analyses

We apply the link prediction model of Elmasri et al. (2017) to the amalgamated dataset of 29,112 documented and $\sim$16.8 million potential host-parasite interactions. The model has three variants: "affinity only" which generates predictions based only on the number of observed interactions for each host and parasite, "phylogeny only" in which only the host evolutionary relationships are used to make predictions, and the "full model" which layers both components. The affinity model is fit by preferential attachment whereby hosts and parasites that have many interacting species in the network are assigned higher probabilities of forming novel interactions. The phylogeny model uses the similarity of host species based on evolutionary distances to assign higher probability to parasites interacting with hosts closely related to their documented host species, and lower probability of interacting with hosts that are distantly related. To account for uncertainty in the phylogeny, and allow the model to place more or less emphasis on recent versus deeper evolutionary relationships, we fit a tree scaling parameter ("Eta") based on an accelerating-decelerating model of evolution (Blomberg et al., 2003; Harmon et al., 2010). This transformation, which allows for changes in the relative evolutionary distances among hosts, was shown to have good statistical properties for link prediction in a subset of the GMPD (Elmasri

et al., 2017). We apply all three versions of the model to the full dataset. The tree scaling parameter is applied across the whole phylogeny, but since the importance of recent versus deep evolutionary relationships among hosts is likely to vary across parasite types (Park et al., 2018), we additionally run the phylogeny only model and the full model on the dataset subset by parasite taxonomy (arthropods, bacteria, fungi, helminths, protozoa, and viruses).

For each model we determined the number of iterations required for parameter convergence by visual inspection of parameter traceplots, auto-correlation plots, and effective sample size (see Elmasri et al. (2017) for detailed discussion of convergence diagnostics). For parameter estimation and evaluation of model performance we ran each model using 5-fold cross validation holding out links for which there is a minimum of two observed interactions (the model would not be able to recover interactions for parasites that infect a single host species). Model performance was evaluated using the area under the receiver operating characteristic curve (AUC) and proportion of observed interactions recovered.

As additional validation, and to determine the utility of the model, we identify the top 10 most likely links that were not documented in the database and conducted online searches of primary and grey literature for evidence of these associations. For models run on the full dataset, we also investigate the top ten links for domesticated mammals (*Bison bison*, *Bos sp.*, *Bubalus bubalis*, *Camelus sp.*, *Capra hircus*, *Canis lupus*, *Cavia porcellus*, *Equus asinus*, *Equus caballus*, *Felis catus*, *Felis silvestris*, *Lama glama*, *Mus musculus*, *Oryctolagus cuniculus*, *Ovis aries*, *Rangifer tarandus*, *Rattus norvegicus*, *Rattus rattus*, *Sus scrofa*, *Vicugna vicugna*), and wild host species separately. Together we ran the three model variants on the full dataset and identified the top ten links overall, plus the top ten links for domesticated hosts and wild hosts, resulting in 90 links. We also ran the phylogeny only and full model variants for each of the six parasite subsets, resulting in another 120 links. In total we identified 210 highly likely undocumented links to target; however, as there was some overlap in the top links among models and data subsets, this resulted in 154

unique links that we investigated for published evidence of infection.

2.3 Results and Discussion

We identify a number of links that are missing from the original databases, indicating that link prediction can be used to improve existing literature-derived databases of host-parasite interactions. These databases provide the best estimates of the host-specificity of infectious organisms, which is important for identifying potential reservoirs of neglected diseases (Viana et al., 2014), and provide essential information for macroecological studies of infectious disease (Stephens et al., 2016). Our method demonstrates that these databases can be efficiently expanded through targeted searches, and can help to identify host-parasite interactions that may emerge in the future given sufficient contact among previously isolated hosts and parasites.

Model diagnostics

All models showed high predictive accuracy in cross-fold validation: AUC values ranging from 0.92 - 0.984, where maximum AUC signifying perfect predictive accuracy is 1, and between 88.61% and 96.02% of the held-out documented interactions recovered (Table 2–1; see Fig. 2–3 for posterior interaction matrices for the full dataset). Of the top 154 undocumented links for which literature searches were conducted, we identified 59 links with evidence of infection (direct observation, genetic sequencing, or positive serology), and identified an additional 13 links with some evidence, but for which additional confirmatory data are required (e.g. antibodies but no confirmed cases for human infections, known cross-reactivity of the serological test used, an unconfirmed visual diagnosis, or the identification of a genetically similar but previously unknown parasite) (Table 2–2, see Appendix 6.1 for lists of top links and detailed results of literature searches). Of the remaining links for which we could not find conclusive evidence, we highlight 38 that should be targeted for surveillance. These include links where there is

known geographic overlap in the ranges of the host and parasite and host ecologies likely facilitate exposure. We also identify a number of links that are highly likely in the model, but may be unlikely due to the mode of disease transmission, non-overlapping host and parasite geographies, or potential competitive interactions with closely related parasites. The top links from the affinity only and full models were largely dominated by humans and domesticated hosts, while the phylogeny only models more often included endangered and relatively poorly studied host species. The full models more often included a larger diversity of parasite species, but among all models, parasites infecting large numbers and phylogenetic ranges of hosts were most often included in the top links (ex. *Rabies lyssavirus*, *Sarcoptes scabiei*, *Toxoplasma gondii*, *Trypanosoma cruzi*). This is not surprising as these parasites are commonly cited as capable of causing disease in a large number of (and sometimes all) mammals, however we were still able to identify many hosts of these parasites that were not documented in the original database.

Across all links examined, the phylogeny only models identified a greater number of documented links that were not in the original database (36/90) compared to the full models (32/90), although the full model identified a larger number of links in some subsets (Table 2–2). The slight increase in performance by the phylogeny only models may be because the influence of sampling biases among hosts and parasites is reduced. The affinity only and full models predict that hosts and parasites with many reported interactions are also likely to interact with one another. The number of interactions will vary across species because of ecological or evolutionary reasons (Kamiya et al., 2014; Poulin and Mouillot, 2003), but it may also be influenced by research effort (Nunn et al., 2003; Ezenwa et al., 2006; Huang et al., 2015; Olival et al., 2017). Thus the affinity only and full models may underestimate the contribution of ecology and evolution, however the AUC values were consistently, although only marginally higher for full models compared to phylogeny only models (Table 2–1), indicating that they performed better in predicting the structure of

links internal to the dataset.

Phylogeny scaling

Not surprisingly, the phylogeny scaling parameter (Eta) varied when the data was subset by parasite type (Figs. 2–4 & 2–5), but in all models Eta was estimated to be positive indicating accelerating evolution and less phylogenetic non-independence among hosts compared to a pure Brownian motion model (Harmon et al., 2010). Interestingly, fungi were estimated to have the smallest Eta parameter (8.42), indicating that more recent host divergences are less influential for infection. This may reflect the tendency for fungi to include opportunistic pathogens such as *Pneumocystis carinii* and *Chrysosporium parvum*. Helminths and viruses were estimated as having similar Eta transformations (17.15 and 16.37 repectively), consistent with the observation of Park et al. (2018) that mean phylogenetic specificity is similar in these two groups, though viruses are more variable and contain more extreme specialist and generalist parasites.

Missing links with published evidence

Our ability to identify links that were not included in the original database, but for which there was some published evidence of infection demonstrates the utility of the model for guiding future surveillance efforts. Through iterative cycles of prediction, evidence gathering, and re-prediction we can more effectively build global host-parasite interaction databases and identify highly likely links that should be targeted by field-based surveillance.

Some of the links we identified as missing were only documented for the first time after the input databases were assembled (e.g. *T. cruzi* in horses (Bryan et al., 2016) and *Nematodirus spathiger* in *Gazella leptoceros* (Said et al., 2018)), providing strong support for the utility of link prediction to guide future surveillance. We also identified links reported only in literature from over 30 years ago, such as *T. cruzi* in the critically

endangered cotton-top tamarin (*Saguinus oedipus*) (Marinkelle, 1982) and the vulnerable black-crowned Central American squirrel monkey (*Saimiri oerstedii*) (Sousa, 1972). Due to their potential impact on the conservation of these species, we strongly recommend that these associations be verified with modern approaches.

Parasites of endangered and captive animals

Through our guided mining of the literature we found evidence of severe infections in several endangered species such as rabies and sarcoptic mange (*Sarcoptes scabiei*) in Dhole (*Cuon alpinus*) and *Toxoplasma gondii* in critically endangered African wild dogs (*Lycaon pictus*) which was noted as causing a fatal infection in a pup (Van Heerden et al., 1995), highlighting the potential impact of our approach for conservation. The importance of infectious diseases in conservation is often hampered by our lack of knowledge about the diversity of pathogens in natural systems (Smith et al., 2009). Our approach could help identify diseases that are current drivers of extinction in endangered species. For example, the model predicts that rabies and sarcoptic mange are likely to infect the endangered Darwin's fox (*Lycalopex fulvipes*). Diseases spread via contact with domestic dogs (notably *Canine distemper virus*) is currently one of the main threats to this species (Silva-Rodríguez et al., 2016) and considering that both rabies and sarcoptic mange from domestic dogs are implicated in the declines of other wild canids (Pence and Ueckermann, 2002; Fleming et al., 2017), they may pose a risk for the conservation of Darwin's foxes. We also found evidence that these two diseases infect the near threatened bush dog (*Speothos venaticus*) (Jorge et al., 2008; DeMatteo, 2008) and sarcoptic mange is suggested to have caused a substantial loss of individuals from a group in Brazil (de Souza Lima et al., 2012).

Occasionally our models predicted interactions that are unlikely to exist due to lack of geographic overlap, such as *T. cruzi*, which is currently restricted to the Americas (Browne et al., 2017), infecting endangered African species such as black rhinoceros (*Diceros bicornis*), lowland gorilla (*Gorilla gorilla*), and chimpanzee (*Pan troglodytes*). Although

natural infections of chimpanzees by this parasite are unlikely, we did find a report of a fatal infection of a captive individual in Texas (Bommineni et al., 2009). Furthermore, we identified infections of *Giardia intestinalis* in a captive bred black rhino calf (Wagner and Edwards, 1984), sarcoptic mange in captive *Taurotragus oryx* (Bornstein et al., 2002), and *T. gondii* infections in captive *Canis aureus* (Dubey et al., 2010) and *G. gorilla* (Akue et al., 2018). While these interactions may not occur in natural settings, they demonstrate that the model is able to accurately identify biologically plausible infection risks that might be relevant for managing captive populations.

Links with mismatched host and parasite ecologies

The top predictions also included links that are unlikely due to an mismatch in host ecology. For example, *Echinococcus granulosus* is typically maintained by a domestic cycle of dogs eating raw livestock offal (Otero-Abad and Torgerson, 2013), and while wild canids such as *Lycalopex gymnocercus* is a documented host (Lucherini and Luengos Vidal, 2008), the model predicts *Lycalopex vetulus* to be susceptible. However, this interaction is unlikely as *L. vetulus* has a largely insectivorous diet (Dalponte, 2009). Another example is the prediction that domestic cattle (*Bos taurus*) is susceptible to infection by *Anisakis simplex*. *A. simplex* is a trophically transmitted nematode that uses aquatic mammals as final hosts, with marine invertebrates and fish as intermediate hosts (Buchmann and Mehrdana, 2016), implying that cattle may only be exposed to the parasite if fed a marine-based diet. Similarly, domestic sheep (*Ovis aries*) are predicted to be susceptible to the tapeworm *Echinococcus multilocularis* and although the specific distribution and epidemiology of this parasite is relatively unknown, it is maintained by a predation cycle between carnivores and rodents, and not currently known to infect ungulates (Massolo et al., 2014).

Risks for humans and domesticated animals

The top predicted links for humans and domesticated animals represent interactions that could have large impacts on public health; however, there is a large amount of effort that goes into studying infectious diseases of humans and domestic species. Out of the 26 links involving humans, we identified 5 ($\sim$19%) links with clear evidence of infection, and another 6 ($\sim$23%) links with potential evidence that require additional confirmation. It is likely that most human-parasite associations have been well documented, even if not included in the databases we aggregated, or occur rarely. For example, in our model humans were predicted to be infected by *Bartonella grahamii* and the first documented case was in an immunocompromised patient in 2013 (Oksi et al., 2013). Similarly, humans were predicted to be susceptible to *Mycoplasma haemofelis*, which was again documented in someone who was immunocompromised (dos Santos et al., 2008), indicating that while these infections may be pose little risk for a large portion of the human population, they are a serious concern for the health of immunocompromised individuals.

Two other *Mysoplasma* species, *M. conjunctivae* and *M. mycoides*, are also predicted to cause infections in humans, and while they were not documented in immunocompromised individuals, they have caused infections in people who come into close contact with domesticated animals (Lysnyansky et al., 2007; Gonçalves, 2007). Our approach also identified a number of diseases that are currently considered a risk for zoonotic transmission such as *Alaria alata*, an intestinal parasite of wild canids – a concern as other *Alaria* species have been reported to cause fatal illness in humans (Murphy et al., 2012), and *Bovine viral diarrhea virus 1*, which is not currently considered to be a human pathogen, but is highly mutable, has the ability to replicate in human cell lines, and has been isolated from humans on rare occasions (Walz et al., 2010).

While the model predicts links that may occur only in extenuating circumstances, it also highlights the potential for previously unknown zoonoses originating from domesticated animals. For example, *Neospora caninum* is responsible for severe economic losses

by causing abortions in cattle, and although antibodies against it have been reported in humans, the zoonotic potential of this parasite is not known (Dubey et al., 2007). Similarly, *Anaplasma bovis* and *Anaplasma marginale* are not currently considered zoonoses, but are globally distributed tick-borne diseases of ruminiants, and are closely related to the zoonotic *Anaplasma phagocytophilum* which can also infect a wide range of mammals, including ruminants (Rar and Golovljova, 2011).

In addition to potential zoonotic diseases, the model predicts that domestic cattle may be susceptible to known human diseases, including *St. Louis encephalitis virus* and *T. cruzi*, the aetiological agent of Chagas disease. Birds are the primary vertebrate hosts for *St. Louis encephalitis virus*, though amplification by certain mammals has been suggested (Kopp et al., 2013), and there is some serological evidence of infection in domestic mammals, including cattle (Diaz et al., 2006). For *T. cruzi*, Browne et al. (2017) compiled over 16,000 records of infection and identified that while there are 177 alternative host species, domestic dogs are thought to be largely responsible for the longterm maintenance of local parasite populations. Currently the role of cattle in the epidemiology of Chagas disease is unknown, but the majority of cows in Latin America (comprising 280 million heads) are likely to be be exposed to the parasite (Giangaspero, 2017), indicating a major risk if they are able to promote parasite transmission. While cattle have tested positive for *T. cruzi* in some serological studies, cows and other domestic animals are also commonly infected by related species, which can cause complications due to cross-reactions in serology based diagnostic tests (Gürtler and Cardinal, 2015). This highlights the importance of verifying infections through additional methods when the only evidence of an interaction is serological.

Future extensions

We have demonstrated that missing links in global databases of host-parasite associations can be predicted using information on known associations and the evolutionary relationships among host species. Applying this method to host-parasite associations for

mammal host species we are able to make robust predictions with relatively little input data, indicating that this method may be applied to disparate host-parasite systems. Here we use information on the evolutionary relationships among hosts as this has been shown to constrain parasite sharing in a diversity of parasite systems (e.g. Perlman and Jaenike (2003); Gilbert and Webb (2007); Davies and Pedersen (2008); Braga et al. (2015)). The model represents evolutionary relationships in the form of a species-by-species distance matrix. However, the flexibility of this approach means that any information on species similarities that can be represented by a distance matrix, such as trait or spatial dissimilarities, may be used to generate predictions (Elmasri et al., 2017). Similarly, the model could be extended to incorporate weighted rather than binary associations, allowing for modelling links as a function of prevalence, intensity of infection, or to explicitly incorporate the amount of evidence supporting each link. In this way sampling biases may be directly incorporated and in addition to identifying likely missing links, we could identify weakly supported interactions or sampling artefacts that may benefit from additional investigation.

Link prediction in global host-parasite networks marks the first step in an iterative process of prediction and verification whereby likely links are identified, the published literature queried, and new links are added, allowing predictions to be updated. As we move down the list we are likely to uncover links among diseases that are less well studied, but which may emerge as public health burdens in the future. These links represent key targets for disease surveillance. By disseminating these predictions to veterinarians, conservation managers, and public health officials, we can spread awareness of potential threats and enhance surveillance by incorporating these taxa within existing disease and biodiversity monitoring programs. An important next step will be to move beyond the binary notion of infection used here and attempt to classify the nature of the association between host and parasite. In this way we may be able to predict not only the presence or absence of a particular host-parasite association, but what epidemiological role the host plays in parasite transmission, the potential impact a parasite might have in a given species, and

better understand the ecologies of reservoir versus spillover hosts (Plourde et al., 2017).

2.4 Conclusion

We suggest that the model of link prediction we present here represents a cost-effective approach for supporting disease surveillance and should be employed in active attempts to document potential host-parasite interactions, and spread awareness of their potential in the wake of future movements of livestock and wildlife species. Global change in the form of shifting climates and alteration of natural habitats has the potential to bring in contact previously isolated host and parasite populations, increasing opportunities for disease spillover. Strengthening our knowledge of the potential for cross-species disease transmission is an essential step toward building effective methods to mitigate the impacts of these diseases.

Data subset	Model	AUC	% 1s recovered
Full dataset	Affinity only	0.934	91.73
	Phylogeny only	0.949	93.10
	Full model	0.966	90.91
Arthropods	Phylogeny only	0.928	94.49
	Full model	0.945	92.13
Bacteria	Phylogeny only	0.952	91.19
	Full model	0.984	95.37
Fungi	Phylogeny only	0.963	93.02
	Full model	0.982	96.02
Helminths	Phylogeny only	0.947	92.29
	Full model	0.970	94.23
Protozoa	Phylogeny only	0.965	92.60
	Full model	0.975	94.75
Viruses	Phylogeny only	0.925	89.49
	Full model	0.949	88.61

Table 2–1: Average model performances diagnostics after 5-fold cross validation: area under the receiver operating characteristic curve (AUC), and percent documented interactions (1s) correctly recovered from the held-out portion.

Dataset	Model	Unique Hosts	Unique Parasites	Links with evidence
Full	Affinity only	1	10	1
Full-Domestics	Affinity only	3	9	3
Full-Wild	Affinity only	10	1	7
Full	Phylogeny only	10	2	4
Full-Domestics	Phylogeny only	9	2	7
Full-Wild	Phylogeny only	10	2	4
Arthropod	Phylogeny only	9	2	4
Bacteria	Phylogeny only	9	4	0
Fungi	Phylogeny only	10	1	1
Helminths	Phylogeny only	8	8	2
Protozoa	Phylogeny only	10	2	8
Viruses	Phylogeny only	10	1	6
Full	Full model	2	10	2
Full-Domestics	Full model	3	9	3
Full-Wild	Full model	10	1	7
Arthropod	Full model	7	3	3
Bacteria	Full model	3	10	3
Fungi	Full model	4	8	4
Helminths	Full model	3	8	2
Protozoa	Full model	7	4	7
Viruses	Full model	10	1	1

Table 2–2: Numbers of unique hosts, unique parasites, and links for which evidence was identified in the literature for each of the top 10 lists

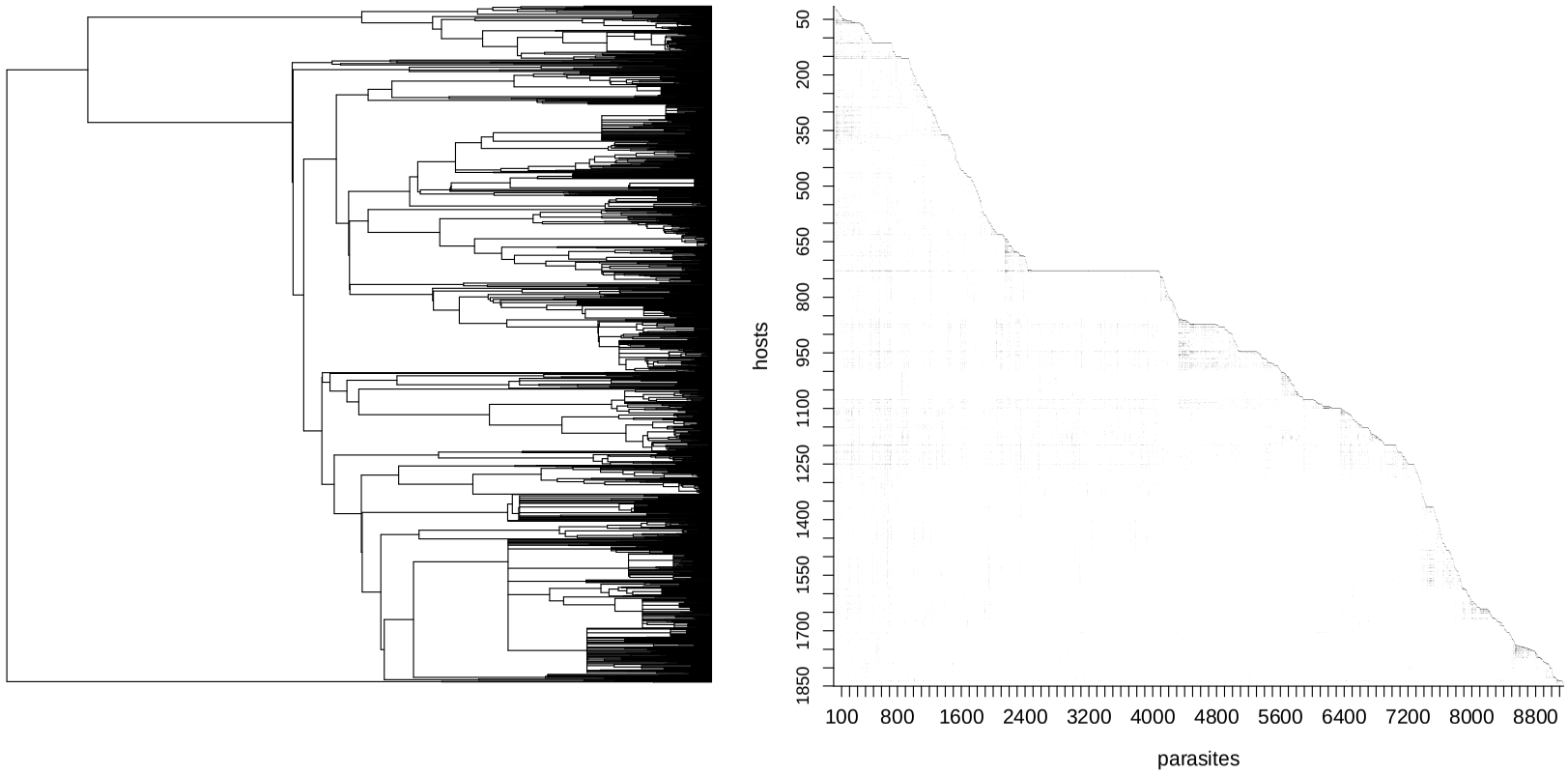


Figure 2–1: Host-parasite association matrix with phylogeny of host species

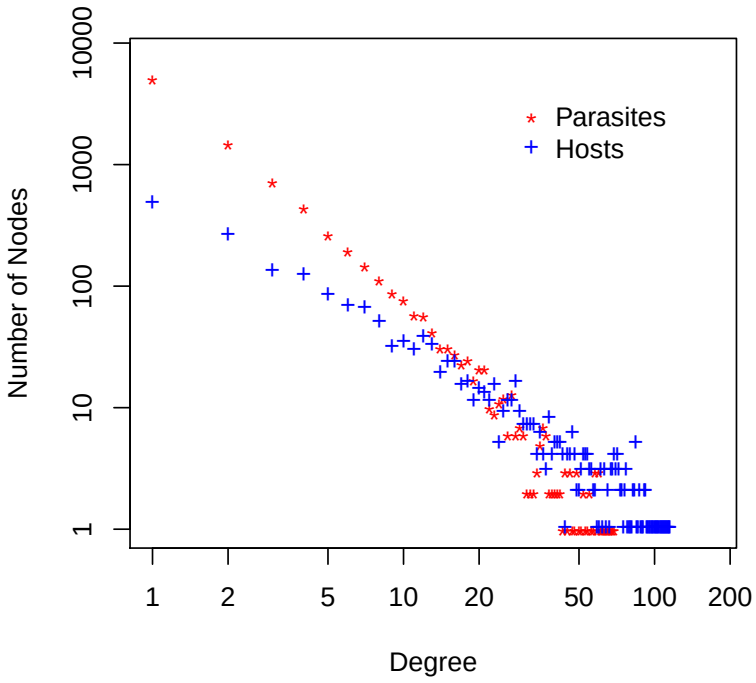


Figure 2–2: Degree distributions of the number of associations (degree) for hosts and parasites (nodes). The distributions are linear on the log scale, indicating a power-law in the number of observed associations per species across the network.

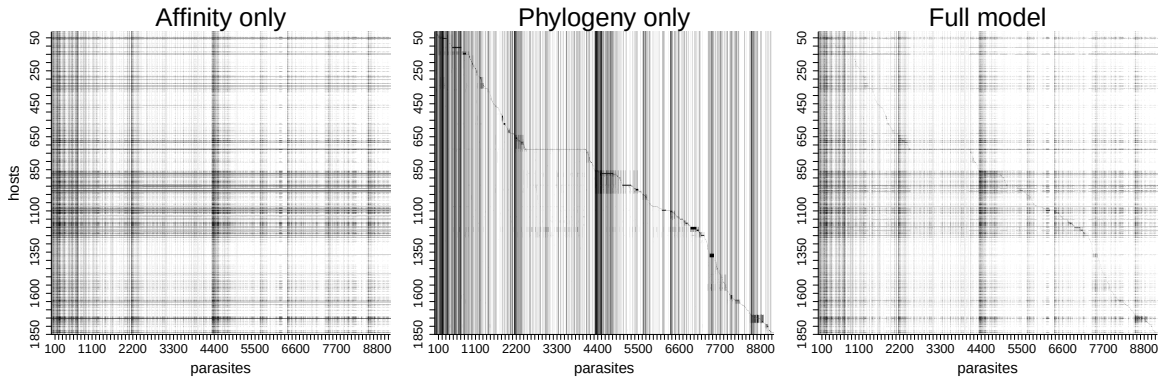


Figure 2–3: Posterior interaction matrices for the affinity only, phylogeny only, and full models run on the full dataset. Rows are ordered to match Figure 2–1.

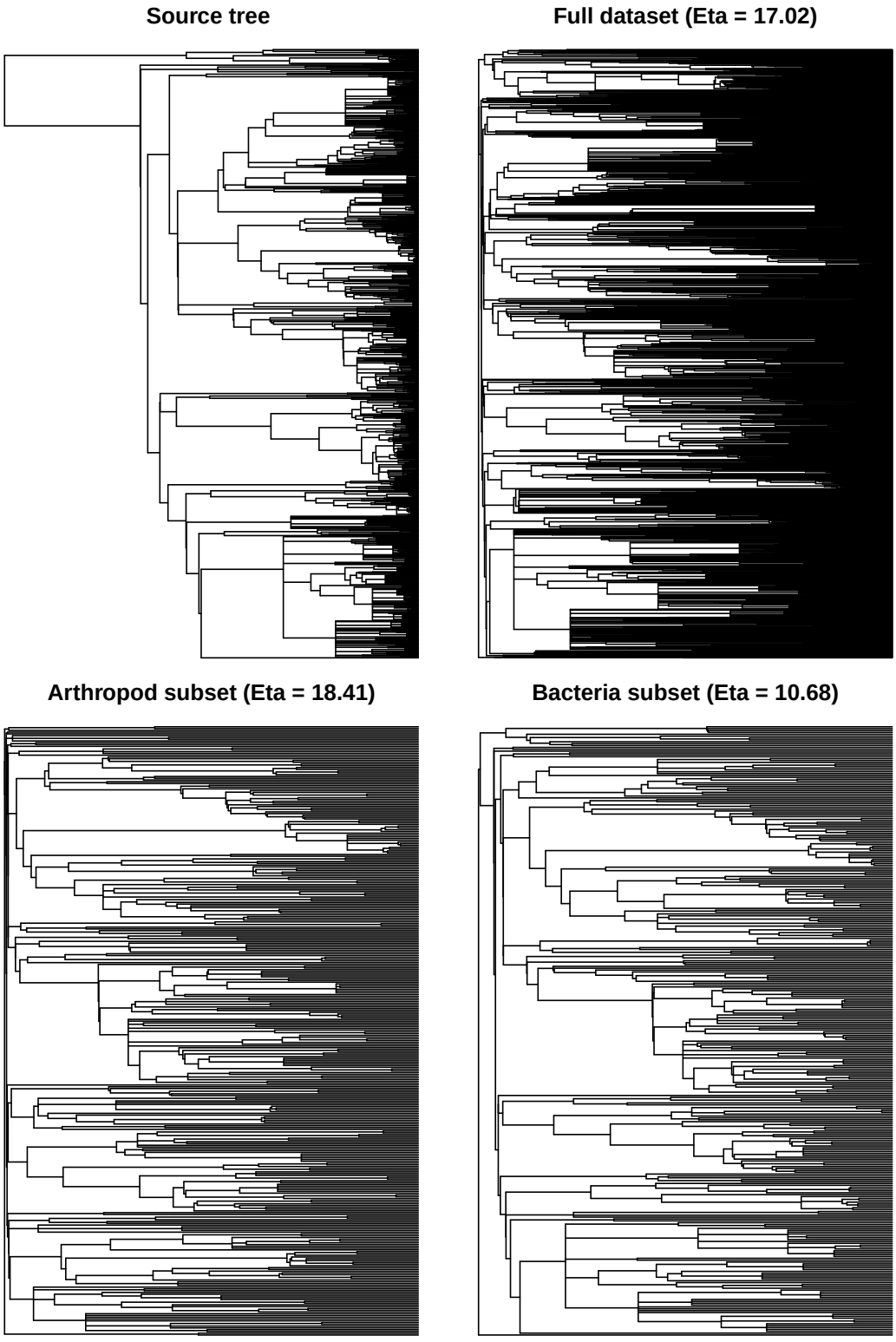


Figure 2–4: The source phylogeny pruned to include only hosts in the amalgamated dataset, and the phylogenies scaled by mean estimated Eta for the phylogeny only models applied to the full dataset, and arthropod and bacteria subsets.

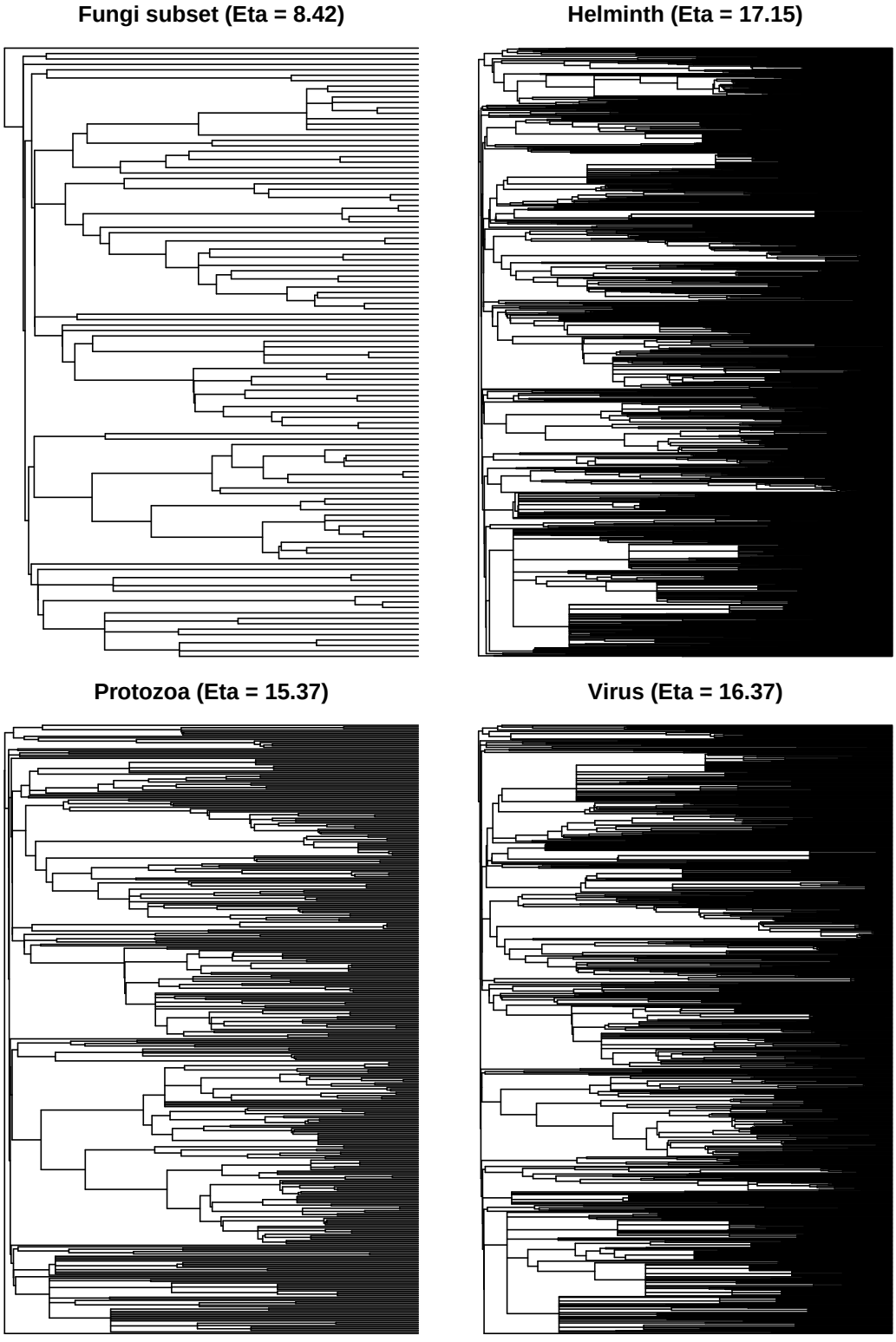


Figure 2–5: Phylogenies scaled by mean estimated Eta for the phylogeny only models applied to the fungi, helminth, protozoa, and virus subsets.

Linking Statement 2

In Chapter 2, I use a novel prediction method to generate lists of highly likely, yet undocumented host-parasite interactions that should be targeted for verification and future surveillance. I show that some of these predicted links have been recorded and are documented in primary or grey literature, but they were not included in the original host-parasite databases. For other highly likely links, there was no currently published documentation. Likely links with no support could be targeted by field-based surveys. Recent advances in genetic sequencing technologies allow for unprecedented descriptions of biological diversity, and may be used to simultaneously track host-parasite interactions and gather baseline data on the diversity of potentially pathogenic organisms in an ecosystem. Such approaches provide potential for rapid assaying of biodiversity to screen for novel host-parasite interactions, and search for evidence in support of missing links.

In Chapter 3, I conduct a survey of bacterial diversity in the Kruger National Park, South Africa, a biodiversity hotspot known for its high richness of mammals. In African savannah ecosystems such as the Kruger, waterholes are vital resources for many animals, but can also facilitate the spread devastating infectious diseases. By examining DNA present in these water sources, I describe patterns of bacterial diversity from across the park, and explore the utility of this approach for identifying infectious organisms. Developing methods for DNA-based biodiversity surveys can support disease monitoring programs by generating baseline data on bacterial diversity, a taxonomic group for which most species are not described and cannot be cultured using traditional techniques. In addition to identifying potentially pathogenic bacteria, the results of this study may provide improved understanding of the roles micro-organisms play in ecosystem stability and resilience, and offer an approach for monitoring shifting species interactions in the face of environmental change.

CHAPTER 3
Bacterial diversity in the waterholes of the Kruger National Park: an eDNA metabarcoding approach

Maxwell J. Farrell[1], D. Govender[2,3],

M. Hajibabaei[4], M. van der Bank[5], T. J. Davies[5,6]

Author Affiliations:

[1]Department of Biology, McGill University

[2]Scientific Services, Kruger National Park, SANParks

[3]Department of Paraclinical Sciences, Faculty of Veterinary Science, University of Pretoria

[4]Integrative Biology & Biodiversity Institute of Ontario, University of Guelph

[5]African Centre for DNA Barcoding, University of Johannesburg

[6]Botany, Forest & Conservation Sciences, University of British Columbia

A version of this chapter has been accepted for publication in *Genome* pending revisions.

3.1 Abstract

Bacteria are essential components of natural environments. They contribute to ecosystem functioning through roles as mutualists and pathogens for larger species, and as key components of food webs and nutrient cycles. Bacterial communities respond to environmental disturbances, and the tracking of these communities across space and time may serve as indicators of ecosystem health in areas of conservation concern. Recent advances in DNA sequencing of environmental samples allow for rapid and culture-free characterization of bacterial communities. Here we conduct the first metabarcoding survey of bacterial diversity in the waterholes of the Kruger National Park, South Africa. We show that eDNA can be amplified from waterholes and find strongly structured microbial communities, likely reflecting local abiotic conditions, animal ecology, and anthropogenic disturbance. Over timescales from days to weeks we find increased turnover in community

composition, indicating bacteria may represent host-associated taxa of large vertebrates visiting the waterholes. Through taxonomic annotation we also identify pathogenic taxa, demonstrating the utility of eDNA metabarcoding for surveillance of infectious diseases. These samples serve as a baseline survey of bacterial diversity in the Kruger, and in the future, spatially distinct microbial communities may be used as markers of ecosystem disturbance, or biotic homogenization across the park.

3.2 Introduction

Traditional programs that monitor for early signs of ecosystem degradation require baseline data on the distributions and ecology of species in an ecosystem. DNA barcoding uses differences in conserved regions of genomes to classify sequences as belonging to particular taxonomic units, regardless of whether or not they have been described formally by taxonomists (Hebert et al., 2003; Blaxter et al., 2005; Ratnasingham and Hebert, 2013). DNA barcoding is thus a particularly powerful tool for exploring microbial diversity, where there are many undescribed taxa that cannot be cultured using traditional methods (Rappé and Giovannoni, 2003). Molecular barcoding coupled with recent advances in genetic sequencing have allowed for unprecedented exploration of microbial communities and the ability to characterize organisms of interest from environmental samples with great sensitivity (Shokralla et al., 2012). In particular, sequencing of cellular and extracellular DNA that can be extracted from environmental samples, collectively known as environmental DNA (eDNA) (Taberlet et al., 2012), is an emerging approach for exploring diversity in aquatic ecosystems (Rees et al., 2014; Lodge et al., 2012).

Microbial diversity in freshwater systems responds to environmental conditions (Lozupone and Knight, 2007), and perturbations (Zeglin, 2015) including multiple anthropogenic impacts such as urbanization (Fisher et al., 2015) and pollution (Bouskill et al., 2010). In addition to acting as indicators of ecosystem health, changes in microbial diversity may be important in themselves. Bacteria are essential components of ecosystems and play

important roles in food webs, nutrient recycling, disease, and as important mutualists for larger multicellular species. Bacteria are thus integral to maintaining the natural balance of ecosystems and shifts in taxonomic composition due to environmental change may severely impact connectivity, functioning (Laforest-Lapointe et al., 2017; Delgado-Baquerizo et al., 2016) and increase exposure to pathogens (Cabral, 2010). However, in most ecosystems of conservation priority, microbial diversity is poorly described.

Surface waters are a vital resource for savannah ecosystems (Redfern et al., 2005; Owen-Smith, 1996), but frequent use by a large variety of species means they can also be a source of cross-species infection and spread of harmful pathogens (Bengis and Erasmus, 1988). These ecosystems provide an ideal context for refining eDNA metabarcoding approaches as they act as sources and sinks of microbial species for larger animals, however baseline information about microbial diversity in these systems is lacking. Here we conduct a survey of bacterial diversity among watering holes of Kruger National Park, South Africa (KNP) through spatio-temporal sampling and sequencing of the V3-V4 region of 16S genes present in water. Water can be scarce in the KNP throughout the dry season and periods of drought (Redfern et al., 2005), and the park has a long history of water provisioning that included the construction of a series of more than 300 artificial waterholes beginning in the 1930s (Smit et al., 2007). These waterholes were intended to increase game numbers by stabilizing water availability year-round and are frequently visited by a diversity of birds and mammals (Smit and Grant, 2009). However, they have proven to alter the distributions of wildlife, which in turn have negative impacts on vegetation dynamics and the park-wide ecosystem (Smit et al., 2007; Smit and Grant, 2009). As a result, a number of artificial waterpoints have been closed since 1994 as the park began reverting to a more natural cycle of water availability (Smit et al., 2007; Van Wyk, 2011). A subset of the waterholes still open are small concrete troughs which are well mixed, largely mud and silt-free, and experience limited inflow from nearby surface waters. This means that eDNA samples will

largely represent microbes in sourcewater and those dispersed by air and animals, allowing us to capture snapshots of local bacterial diversity across the park.

This study provides the first survey of bacterial diversity in the waterholes of the KNP, and is among the first studies using next-generation sequencing to describe aquatic microbial diversity in Africa (see Jordaan and Bezuidenhout (2016, 2013); Tekere et al. (2012); Mwirichia et al. (2011); Tekere et al. (2011)). Here we explore bacterial diversity across the southern half of the park and describe variation across, space, time, sample volume, and abiotic influences.

3.3 Methods

Study Site

Waterholes were sampled in June and July of 2015 in the Kruger National Park, South Africa (KNP), a large protected savannah ecosystem and a global diversity hotspot (Lahaye et al., 2008). Sampling was conducted during the dry season when natural sources of surface water are largely dry and watering hole visitation rates by medium and large vertebrates are highest. The park is divided into twenty-two ranger sections, which range in size from roughly 520 to 1,170 square kilometers. Across the southern half of the reserve below the Olifants river, ten concrete bottom artificial waterholes were selected from five of these sections (Table 3–1, Fig. 3–1). The waterholes varied in shape with some mimicking the contours of natural pans, making volume estimations difficult. However, the generic design included longer and shorter axes, with comparable dimensions across waterholes. Each waterhole is equipped with a ball-valve, which regulates water levels and re-fills the trough from nearby reservoirs when water levels drop. Water is sourced predominantly from groundwater via boreholes, but three sites use pipeline troughs filled with diverted river water.

Water Sampling and Processing

At each site, samples were taken once per week for three weeks. For one site (NWA), samples were also taken every day for five consecutive days. Sampling consisted of two 1L water samples collected in autoclaved, UV sterilized glass jars from opposite ends of the waterhole, approximately one foot from the nearest edge. These two within-waterhole samples (A/B) were taken along the waterhole's longest axis that maximized the distance and upwind-downwind gradient between them, if a strong wind was present. Water samples were placed on icepacks in a cooler and kept between 4-8°C until returning to the laboratory, where they were placed in the fridge.

Water quality parameters were taken during each sampling period using a YSI 650QS multi-parameter sonde. Temperature (°C), conductivity (mS/cm), dissolved oxygen (in mg/L and % saturation), and pH were recorded. Three measurements were taken along the same axis that the A/B water samples were drawn, and then averaged to measure water quality per sample-time.

In the lab, the outside of water sample collection bottles were washed with ELIMINase (Decon Labs) and rinsed with deionized (DI) water to limit contamination. For each A/B sample, 150 mL of water was sub-sampled and filtered through gamma-irradiated 0.2 μm Supor hydrophilic polyethersulfone membranes (Pall no. 66234). The filtration apparatus consisted of three 300 mL Advantec polysulfone 47mm filter funnels fitted to a Pall vacuum manifold with vacuum pressure maintained by a Pall filtration vacuum/pressure pump (model no. 13158). After filtration, filters were stored in sterile 15 mL Falcon tubes and placed in a freezer at -60°C. On one sampling date for six sites, additional volumes of 50 mL and 15 mL were filtered from each 1L sample to asses the impact of sample volume filtered. Twice throughout sampling, BLANK samples were generated by filtering 1L of deionized water used in the laboratory.

Prior to and between filtrations, all funnel components and tweezers used to manipulate the filters were sterilized by soaking with 10% bleach for 10 minutes, rinsing with

DI water, washing with ELIMINase, rinsing with DI water, and subsequent exposure to UV radiation for a minimum of 30 minutes. Gloves were worn at all times and changed between samples to minimize cross-sample contamination. To avoid sample freezing and bacterial growth in collection jars, all samples were processed within 12 hours of collection. Frozen 50 mL unfiltered voucher samples were kept and placed at -80 °C at the University of Johannesburg's African Centre for DNA Barcoding for long term storage.

DNA Extraction, Amplification, and Sequencing

DNA was isolated from filter papers using MO BIO PowerWater DNA Isolation Kits. Universal bacterial primer sets designed by Sundquist (2007) (V3-F: 5'ACTCCTACGGGAG-GCAGCAG 3'; V4-R: 5'GGACTACARGGTATCTAAT 3') tagged with an Illumina adapter sequence were used to amplify the V3-V4 hypervariable region of the 16S ribosomal RNA gene through polymerase chain reaction (PCR). The PCR used a standard mix of 17.8μL molecular grade water, 2.5μL 10 reaction buffer (200mM Tris HCl, 500mM KCl, pH 8.4), 1μL $MgCl_2$ (50mM), 0.5μL dNTP (10mM), 0.5μL forward primer (10mM), 0.5μL reverse primer (10mM), 0.2μL Platinum Taq DNA polymerase (Invitrogen), and 2μL DNA as template for a total volume of 25μL. PCRs underwent the following cycler conditions: initial 94°C for 5 minutes, then 30 cycles of 94°C for 40 seconds, 46°C for 1 minute, 72°C for 30 seconds, and a final temperature of 72°C for 2 minutes. Amplification success was confirmed through gel electrophoresis, using a 1.5% agarose gel. PCR products were purified using MinElute PCR purification kit (Qiagen), and quantified through flurometry using a Quant-iT PicoGreen dsDNA assay kit (Invitrogen). Samples were normalized, then multiplexed with the Nextera XT Index kit (96 indexes) (Illumina) and sequenced on an Illumina MiSeq flowcell using a V2 sequencing chemistry kit (2 x 250) making up approximately 1/8th of the run.

Sequence Processing, Taxonomy Assignment, and Phylogeny Construction

Across all samples, we generated a total of 2,164,262 Illumina reads. Primer sequences were removed using the trim.seqs function in mothur (Schloss et al., 2009). Reads were then processed in R version 3.4.3 (R Development Core Team, 2008) using the package *dada2* version 1.6.0 (Callahan et al., 2016) following a modified version of the DADA2 Bioconductor workflow (Callahan et al., 2017) and online tutorials v1.6 and workflow for big data v1.4 (benjjneb.github.io/dada2/tutorial.html). Reads were filtered by quality, removing sequences with maximum expected error (maxEE) greater than 6 for both forward and reverse reads, and reads with any base pair having Q of 6 or lower. Reads were truncated to a length of 230bp and 220bp for forward and reverse reads respectively, consistent with dropoffs in quality profiles, and reads shorter than this were removed. Since the samples were sequenced across four different runs, subsequent steps of learning error rates, dereplication, denoising and Amplicon Sequence Variant (ASV) calling (Callahan et al., 2017) using pooled samples, and merging of paired reads were performed separately for each run. Tables of ASV sequences per sample within each run were then combined and chimera detection using all pooled samples was performed (see Table 6–23 for the number of reads retained across each step). In total 1,184,831 reads were retained, representing 3533 ASVs.

Taxonomy assignment from kingdom to genus was performed using the RDP classifier and SILVA nr v128 reference database (Quast et al., 2013) formatted for DADA2 (available at benjjneb.github.io/dada2/training.html), using the assignTaxonomy function (Fig. 3–5). ASVs assigned as Archaea, Eukarya, Chloroplast, or Mitochondria were removed. Species level assignments were added by exact sequence matching using the addSpecies function. ASV sequences were aligned with the pynast algorithm via align_seqs.py in QIIME (Caporaso et al., 2010) and sequences with poor alignment automatically removed. A phylogenetic tree was constructed using the GTRCAT model in FastTree version 2.1.3

(Price et al., 2010) after filtering nucleotides with greater than 90% gap fraction and removing the 5% highest entropy positions with filter_alignment.py in QIIME (Caporaso et al., 2010). This resulted in a phylogenetic tree of 3393 ASVs which were used in subsequent community analyses.

Community Analyses

The ASV sequence table was merged with the phylogeny and sample metadata using the R pacakge *phyloseq* version 1.22.3 (McMurdie and Holmes, 2013). Negative controls (BLANK samples) used to investigate contamination during sample filtration contained 43 ASVs collectively (with 9 ASVs found in both samples). The sequence reads in each filtration blank were both dominated by the same ASV (53% and 86% respectively), however none of the 43 ASVs identified in the blanks were identified in any of the other samples. These control samples were removed prior to community analyses.

A subset of core samples was created by removing the first four daily NWA samples and samples of differential volume (S & XS samples), resulting in 54 samples of 150 mL each (Table 3–2). An ASV accumulation curve for core samples was generated using the specaccum function in the R package *vegan* version 2.4.6 (Oksanen et al., 2018) using the "exact" method, and extrapolated to total ASV richness using the Chao and Bootstrap methods in *vegan*'s specpool function. Alpha diversity was calculated for the core samples as observed ASV richness and Shannon diversity using the *phyloseq* package. Additive partitioning of Shannon diversity across core samples was investigated using the adipart function in *vegan* (Table 3–3).

Taxonomic composition was assessed by merging core samples at each site, and plotting relative abundances of reads for the most common taxa at levels of phylum, class, and order (Fig. 3–6). Temporal variation in taxonomic composition across core samples was assessed by merging A and B samples and plotting relative abundances of reads for sites

with two or more weekly samples, for the levels of phylum (Fig. 3–7), and class (Fig. S6–6). To further investigate fine-scale temporal variation in taxonomic composition (phylum, class, and order), relative abundance of reads were plotted across the daily samples at site NWA (Fig. 3–8). We also explored temporal turnover among samples with Sorensen's dissimilarity calculated using the beta.pair function from the *betapart* package (Baselga et al., 2018) (Fig. 6–9) and significant differences among daily and weekly samples was assessed using multivariate ANOVAs (anosim in *vegan*) with 999 permutations each.

Community composition across sites in the core samples was described with non-metric multidimentional scaling (NMDS) ordinations on relative ASV abundances per sample using the Bray-Curtis dissimilarity, and the abundance weighted Unifrac dissimilarity (Figs. 3–9, 6–10 & 6–11). Statistically, associations between dissimilarities and both water quality properties and common taxonomic groups were assessed using the envfit function in *vegan* for bacterial classes (Fig. 3–10) and orders (Fig. 6–12).

Phylogenetic community structure across core samples was calculated using standardized effect sizes of mean pairwise phylogenetic distances (MPD) and mean nearest taxon distances (MNTD) in the R package *picante* version 1.6.2 (Kembel et al., 2010) using the abundance weighted "richness" null model and 999 randomizations in the ses.mpd and ses.mntd functions (Fig. 6–13). For a given sample, MPD calculates the mean phylogenetic distance among each pair of taxa present, while MNTD calculates the mean phylogenetic distance from each taxa to its closest relative. These raw metrics give an estimate of how closely related community members are to each other, and are then compared to randomized communities to determine whether the observed metrics are different than what would expected if communities were assembled at random from taxa pooled across all samples.

To assess the effect of differential sample volumes, S (50 mL) and XS (15 mL) samples were subset along with their corresponding full volume samples (150 mL). Alpha diversity, calculated as observed ASV richness and Shannon diversity, were calculated as described

above (Fig. 6–14). Variation in taxonomic composition was investigated by comparing relative read abundances of bacterial phyla in A/B samples across sites and differential volumes (Fig. 6–15).

3.4 Results

Our sampling design aimed to sequence a core set of samples with all ten sites being sampled once per week for three weeks. Due to logistic constraints of sample storage and extremely low water levels from drawdown by animals, we were only able to process 27 of the planned 30 weekly samples (Table 3–2). In addition to this core sampling, we sequenced differential volumes for six samples, and an additional four daily samples from Nwaswitshaka (site NWA). A/B samples were taken at each site-time, resulting in a total of 88 sequenced samples, including the two filtration blanks. Across all 88 samples, we identified a total of 3393 ASVs. Roughly 15% of ASVs (n=524) were represented by a single read, together comprising fewer than 0.05% of all reads. The DADA2 approach infers the biological sequences in the sample prior to the introduction of amplification and sequencing errors, and can distinguish sequence variants differing by as little as one nucleotide. As such, we included all ASVs, including those represented by single reads, in subsequent analyses of biodiversity.

The ASV accumulation curve generated for the core sample set (2603 ASVs) does not appear to saturate (Fig. 3–2). Estimates of total richness using the Chao estimator predicts 6164 ASVs (+/- 262 SE) among the core samples, indicating we may be capturing less than half of the total bacterial diversity present among our sites. However, estimates of total diversity using the Bootstrap method were more conservative, with 3260 (+/- 146 SE) estimated ASVs. ASV diversity varied across sites (Figs. 3–3, Fig. 3–4), but the largest turnover (β diversity) was observed among park sections (Table 3–3). Variation among A/B samples contributed very small amounts to β diversity, indicating that at a particular time, microbial diversity within each waterhole was fairly well mixed.

In terms of taxonomic composition, 99.2% of ASVs were assigned to a known phylum, with the proportion of assignments decreasing at lower taxonomic levels (Fig. 3–5). The majority of bacterial ASVs were classified as Proteobacteria ($\sim$ 59%), followed by Bacteroidetes ($\sim$ 14%), Firmicutes ($\sim$ 9%), Actinobacteria ($\sim$ 6%), and Verrumicrobia ($\sim$ 2%). For bacterial classes, ASVs were largely classified as Betaproteobacteria ($\sim$ 34%), Alphaproteobacteria ($\sim$ 11%), Gammaproteobacteria ($\sim$ 10%), and Sphingobacteriia ($\sim$ 6%). Among core samples, relative abundances of phyla, classes, and orders varied across sites (Figs. 3–6, S6–4, S6–5). Across weeks, relative abundances of phyla varied within each site (Fig. 3–7), with some sites displaying more stability (IMB & HOY) compared to others (NYA & NGO). Patterns among bacterial classes (Fig. 6–6) largely reflected variation among phyla, though one site (HOY) displayed much more variation in relative abundances among classes, reflecting substantial turnover within Proteobacteria. Comparing weekly turnover with the five daily samples taken at Nwaswitshaka (NWA) (Fig. 3–8), taxonomic composition appeared more stable across days than weeks. Using hierarchical clustering of Sorensen's dissimilarity, we find that samples taken within a single week cluster together (Fig. 6–9). Multivariate ANOVAs on these distances revealed a significant difference in beta diversity among weekly samples (NWA 2,7,8; $Pr(> F) = 0.02$), with 49% of the variance explained by sample date, but no significant difference among the additional daily samples (NWA 3,4,5,6; $Pr(> F) = 0.54$), with 14% of the variance explained by sample date.

Community composition visualized through NMDS ordinations reflected results from the additive partitioning of diversity, with core samples clustering by site (Fig. 3–9) and section (Fig. 6–10) for both Bray-Curtis and abundance weighted UniFrac dissimilarities. Interestingly, waterholes filled by water from pipeline troughs (NGO, NYA, WIT) grouped together (Fig. 6–10), although these three sites are situated on a different geological type than sites fed by boreholes, making us unable to differentiate the effects of each factor (Fig. 6–11). Bacterial community composition was significantly structured by conductivity and

pH for both Bray-Curtis and UniFrac dissimilarities, and dissolved oxygen also had an influence on UniFrac dissimilarity (Figs. 3–10 & 6–12).The dissimilarity of high conductivity sites (particularly HOY & IMB in the Kingfisherspruit section) was associated with high abundances of Clostridia, Gammaproteobacteria, and Bacteroidia, while sites with high pH and dissolved oxygen were positively related to the abundances of Actinobacteria and Alphaproteobacteria (Fig. 3–10).

Reflecting the NMDS structure of the abundance weighted UniFrac dissimilarities, MNTD, which is most sensitive to phylogenetic structure towards the tips of the tree (Mazel et al., 2016), indicated strong phylogenetic clustering within the majority of samples (Fig. 6–13). The strength of clustering was weaker for MPD, which is more sensitive to phylogenetic structure deeper in the tree (Mazel et al., 2016).

We did not find any clear decrease in alpha diversity with smaller sample volumes (Fig. 6–14), and one of the 15 mL samples returned the largest richness of ASVs, though the median value for 15 mL samples was lower and had a larger interquartile range than the 50 mL and 150 mL samples. The major phyla detected within samples was also relatively consistent, with most groups represented across different sample volumes, though not always in the same proportions (Fig. 6–15).

3.5 Discussion

Biological monitoring is an essential aspect of conservation for tracking contemporary changes in ecosystems as well as providing a historical baseline for making management decisions. The Kruger National Park, established in 1898, has a long history of management practices revolving around maintenance of large mammals (Venter et al., 2008). While bacterial diversity has been explored for important infectious agents in the system (Michel et al., 2007; Bengis and Erasmus, 1988; Smith et al., 2000), recent advances in next generation sequencing methods now allow for the rapid and culture-free description of bacterial diversity throughout the park.

Here we present the first description of bacterial diversity in the waterholes of the Kruger National Park. In total we identified over 3000 unique taxa (referred to as amplicon sequence variants, or ASVs), only about half of which could be assigned to a previously described genus. The relative dominance of bacterial phyla was consistent with bacterial surveys of the Vaal River in central South Africa (Jordaan and Bezuidenhout, 2016). However, bacterial diversity was strongly structured across space, with the largest turnover in diversity occurring among park sections. This is not surprising considering the distances from site to site range from 3km to 115km and represent a gradient in large animal density, rainfall, vegetation, and major subsurface geology (Chirima et al., 2012; Van Wilgen et al., 2000; Smit and Grant, 2009; Smit et al., 2013). Samples also clustered by site, displaying substantial variation in taxonomic composition across sites. This variation was associated with physico-chemical properties of the water, with conductivity and pH being important explanatory variables. In addition to water quality, variability in taxonomic composition is likely influenced by the origin of the water used to fill each waterhole, differences in the surrounding soil and vegetation types, and the particular species and populations of animals using the waterholes.

We assessed daily turnover in composition at Nwaswitshaka, which appeared to be more stable over this shorter timescale when compared to turnover across weeks. However, Nwaswitshaka was less variable across weeks than other sites, indicating that daily variation in bacterial communities could be greater in other locations. Important water quality variables (conductivity and pH) were largely consistent across weeks (Table 6–22), suggesting that the observed temporal heterogeneity may be driven by differences in external factors influencing bacterial input and removal from the system, such as variation in animal visitation throughout the sampling period. Between sampling events, water levels would sometimes drop substantially, indicating major drawdown by animals and likely removing bacteria deposited by animals visiting earlier in the week. Large mammal communities vary across the sampled regions of the park (Chirima et al., 2012), which

may contribute to observed spatial variation in bacterial communities. However, different species also differ in their dependence on water, which is reflected in their rates of visitation to water points (Redfern et al., 2005). Variation in samples taken across subsequent weeks may therefore reflect different components of local animal communities, each with their unique host-associated bacterial taxa (Ley et al., 2008). By pairing bacterial composition with animal visitation prior to sampling (either through direct observation, or presence of genetic material), it may be feasible to build statistical probabilities of associations using co-occurrences of microbial and animal signatures.

Across samples, patterns of phylogenetic clustering were consistent with observed taxonomic variation. Multiple phyla were present in all samples, consistent with an even representation of deep bacterial lineages. However, turnover at lower taxonomic levels shown by significantly low mean nearest taxon distances indicate that there are distinct subsets of closely related taxa present at each site. This structuring may reflect filtering of bacterial communities by local environmental conditions, or the deposition of microbes by particular animal populations or individuals. Many vertebrate species have expansive home ranges, but during the dry season drought-intolerant animals will restrict their movement so as to stay close to permanent water bodies (Redfern et al., 2005). Thus the maintenance of major bacterial taxa may reflect both free-living environmental bacteria, and the core microbiome of water-dependent species. By taking repeat temporal samples, it may be possible to build association networks between bacterial taxa and host species, or even their local populations, solely from environmental DNA.

Through examination of taxonomic assignments we identified taxa belonging to genera that include important pathogens (*Arcobacter, Bacillus, Burkholderia, Coxiella, Legionella, Neisseria, Pasteurella, Rickettsia*, and *Yersinia*). While many of these genera include both pathogenic as well as benign species found in environmental samples, some of these genera are comprised solely of pathogenic species. For example, the genus *Coxiella* is represented by one species, *Coxiella burnetii*, the causative agent of Q fever, which has

previously been documented as causing disease in the park (Van Heerden et al., 1995). Additionally, taxa in the order *Chlamydiales* are all obligate intracellular pathogens of eukaryotes (Ball et al., 2015), and taxa in the genus *Neisseria* colonize mucosal surfaces of animals, some of which are pathogenic in humans (Liu et al., 2015). Interestingly, we also identified sequences classified as *Streptococcus urinalis*, a recently described species linked to urinary tract infections in humans (Peltroche-Llacsahuanga et al., 2012). While hypervariable regions of the 16S gene may not be the optimal genomic regions for detecting the presence of particular pathogenic species or strains, our findings indicate that broad scale surveys of microbial diversity may be useful in determining the presence of potential pathogens across vastly divergent groups of bacteria. This can in turn guide more targeted sampling of both pathogenic and commensal bacteria across the park.

In addition to the presence of pathogens, surveys of bacterial diversity may be used to detect anthropogenic influences in the park. For example, we found that bacterial diversity was quite different for the two sites sampled in the Kingfisherspruit section (Hoyo Hoyo and Imbali). These sites were dominated by an ASV assigned to the genus *Arcobacter*, but which did not exactly match any sequence in the SILVA reference database. Three of the five described members of this genus are known to be pathogenic (Fera et al., 2004) and include *A. butzleri*, which can cause severe diarrhea (Lerner et al., 1994) and was detected in two of the weekly samples at Hoyo Hoyo. The two waterholes in Kingfisherspruit are fed in part with greywater that is passed through reed beds. Greywater is untreated household wastewater that typically has not been contaminated by toilet waste and is often used as year-round sources of water, especially in water scarce areas (Nganga et al., 2012). Compared to source water, kitchen and laundry greywater can have elevated conductivity (Nganga et al., 2012), which may explain the high conductivity of water at these sites, and present a strong selective environment driving their unique bacterial communities.

We did not have sufficient sampling of sites to explore all possible drivers of differences in bacterial communities. Nonetheless, some features differed obviously among

sites. For example, Witpens is a heavily vulture-dominated site. Bathing by vultures likely results in large influxes of nutrients such as blood, and vulture feces has been found to alter soil bacterial communities through elevated nitrogen and decreased pH (Ganz et al., 2012). Visually, water from Witpens was bright green, indicating high abundance of photosynthetic species and consistent with the large variations observed in dissolved oxygen (Table S1). However, we found no evidence of elevated abundances of *Microcystis* or other Cyanobacteria, though samples from Witpens strongly clustered together and had relatively high abundances of Rhizobiales and Rhodocylales, both of which include species known to fix nitrogen (Carvalho et al., 2010; Loy et al., 2005).

Witpens, along with Nyamarhi and Ngotso North are filled by pipeline troughs that divert river water to waterholes many kilometers away. While pipeline troughs are likely to reflect a subsample of the diversity found in river water, the acts of pipeline transport and storage themselves may have strong filtering effects on bacterial communities. Our results indicate that waterholes represent locally unique bacterial communities, thus the practice of diverting river water to waterholes kilometers away may homogenize microbial diversity across the landscape, ultimately disrupting local communities. The consequences of such shifts in community structure are difficult to assess without comparing pipeline troughs with their source waters, but the diversion of river water may have unintended impacts on microbial diversity. For example, genes conferring antimicrobial resistance have been shown to spread from river water to impala in the Kruger National Park (Mariano et al., 2009).

Here we show that eDNA can be amplified from waterholes in the Kruger National Park, and find strongly structured microbial communities, likely reflecting local abiotic conditions, animal ecology, and anthropogenic disturbance. We suggest that disruption of spatially distinct microbial communities may be used as a marker of ecosystem disturbance, or biotic homogenization across the park. We find that for artificial waterholes, bacterial diversity is surprisingly insensitive to sample volume, with even small volumes

useful for capturing major components of bacterial communities, though larger volumes are necessary to detect rare taxa. Replicating this study across different seasons, and expanding sampling to include natural waterpoints will provide improved understanding of the roles micro-organisms play in ecosystem stability and resilience, and offer an effective method for monitoring of shifting species interactions in the face of environmental change. Just as studies of the microbiome have revolutionized our understanding of human health, metagenomic analysis of environmental DNA have the potential to revolutionize our understanding of ecosystem health. Tracking of bacterial communities can provide a template for monitoring ecosystem disturbance through their response to biological contaminants, documenting the spread of invasive species or infectious organisms, and better understanding the impacts ecological disturbances have on the composition of native communities.

Figures & Tables

Section	Site	Type	Geology
Tshokwane (TSH)	Nhlanguleni (NHL)	Borehole	Granite
Skukuza (SKZ)	Nwaswitshaka (NWA)	Borehole	Granite
Skukuza (SKZ)	De LaPorte (DLP)	Borehole	Granite
Skukuza (SKZ)	Kwaggas Pan (KWA)	Borehole	Granite
Satara (SAT)	Girivana (GIR)	Borehole	Granite
SaTara (SAT)	Witpens (WIT)	Pipeline trough	Basalt
Kingfisherpruit (KFI)	Imbali (IMB)	Borehole	Granite
Kingfisherpruit (KFI)	Hoyo Hoyo (HOY)	Borehole	Granite
Houtboschrand (HOU)	Nyamarhi (NYA)	Pipeline trough	Basalt
Houtboschrand (HOU)	Ngosto North (NGO)	Pipeline trough	Basalt

Table 3–1: Sample locations with section, waterhole type, and geology.

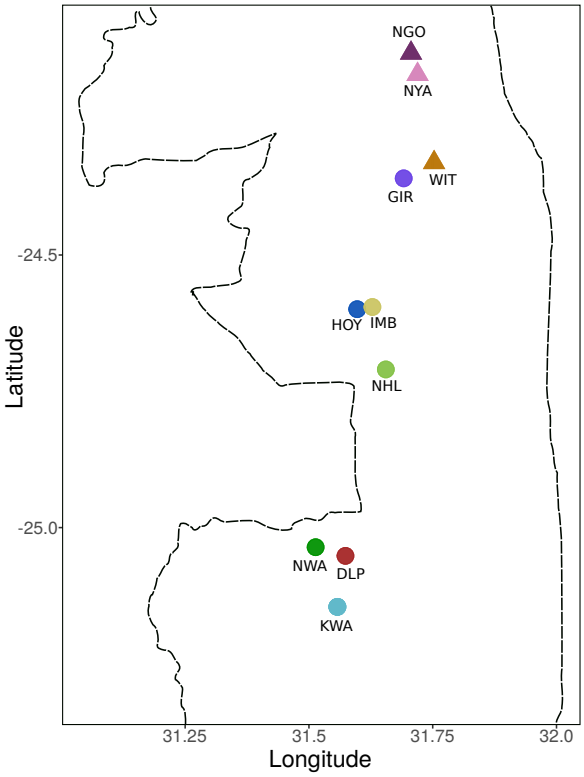


Figure 3–1: Map of site locations with park boundary indicated by dashed line. Circles represent sites filled by boreholes while triangles represent sites filled by river water via pipeline troughs.

Site	Weeks	S	XS	Daily	A/B	Total
Nhlanguleni (NHL)	3	0	0	0	Yes	6
Nwaswitshaka (NWA)	3	1	1	4	Yes	18
De LaPorte (DLP)	1	1	1	0	Yes	6
Kwaggas Pan (KWA)	2	1	1	0	Yes	8
Girivana (GIR)	3	0	0	0	Yes	6
Witpens (WIT)	3	0	0	0	Yes	6
Imbali (IMB)	3	0	0	0	Yes	6
Hoyo Hoyo (HOY)	3	1	1	0	Yes	10
Nyamarhi (NYA)	3	1	1	0	Yes	10
Ngosto North (NGO)	3	1	1	0	Yes	10
BLANK	2	0	0	0	No	2
	29	6	6	4		88

Table 3–2: Samples sequences, broken down by number of weekly samples, number of site-times for which S (50 mL) and XS (15 mL) samples were filtered, additional daily samples taken, whether A/B samples were taken, and the resulting total number of samples sequenced per site.

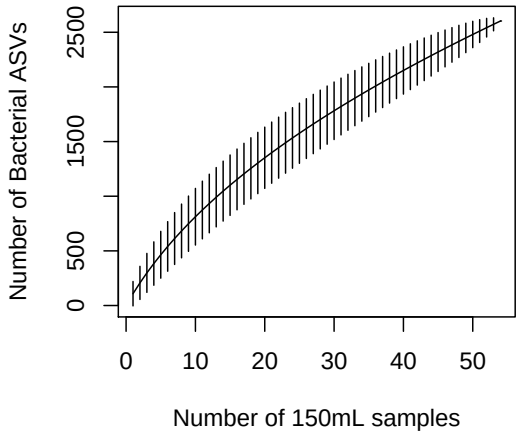


Figure 3–2: ASV accumulation curve of bacterial ASV richness using the “exact” method. Bars represent two standard deviations around mean estimates.

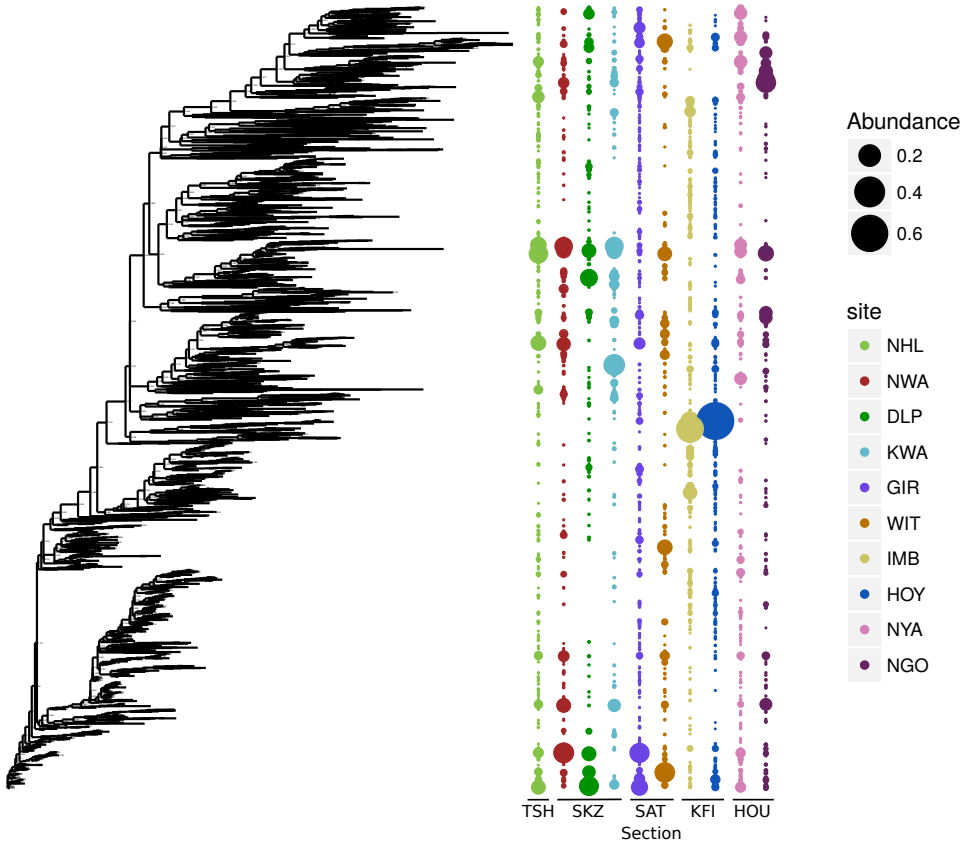


Figure 3–3: Phylogenetic tree of 16S ASV sequences in the core samples, paired with their relative abundances at each site. Sites are ordered by section.

Diversity	Level	Shannon	SES	2.5%	97.5%	Pr(sim.)
α	A/B samples	2.93	-9335.9	5.10	5.10	0.01
α	Temporal samples	2.99	-12535.9	5.13	5.13	0.01
α	Sites	3.39	-16873.6	5.15	5.15	0.01
α	Sections	3.94	-14919.9	5.16	5.16	0.01
γ	(Total)	5.17	0.0	5.17	5.17	1.00
β	A/B samples	0.05	153.3	0.027	0.027	0.01
β	Temporal samples	0.40	2688.2	0.022	0.022	0.01
β	Sites	0.55	6528.7	0.008	0.008	0.01
β	Sections	1.23	14919.9	0.008	0.008	0.01

Table 3–3: Additive partitioning of Shannon diversity into α, β, and γ diversities across sections, sites, temporal samples, and within site-time samples (A/B) as components of the total diversity observed across all core samples. Observed diversity is compared to 99 simulations using "r2dtable" null model with the *adipart* function in the R package *vegan*.

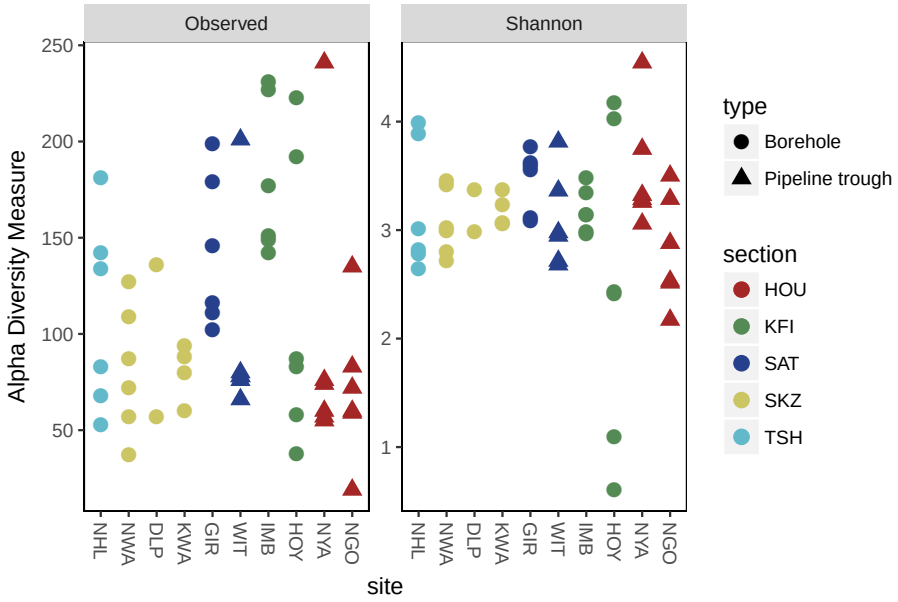


Figure 3–4: Plots of observed ASV richness and Shannon diversity across samples. Samples are grouped and coloured by park section and with shape indicating waterhole type.

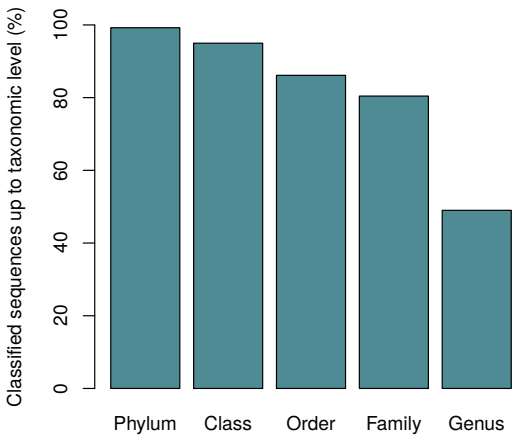


Figure 3–5: The proportion of ASVs assigned at a given taxonomic level using the SILVA database v128 and a pre-trained RDP classifier with minimum 50% bootstrap support.

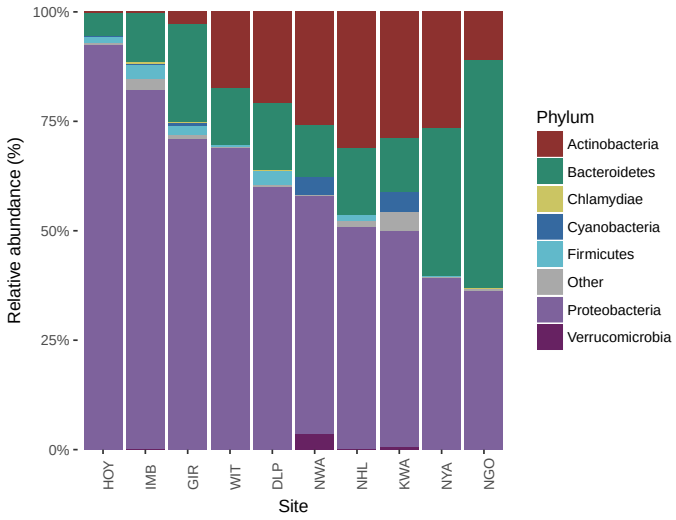


Figure 3–6: Relative abundances of bacterial phyla across sites. Sites are ordered by relative abundance of phylum Proteobacteria.

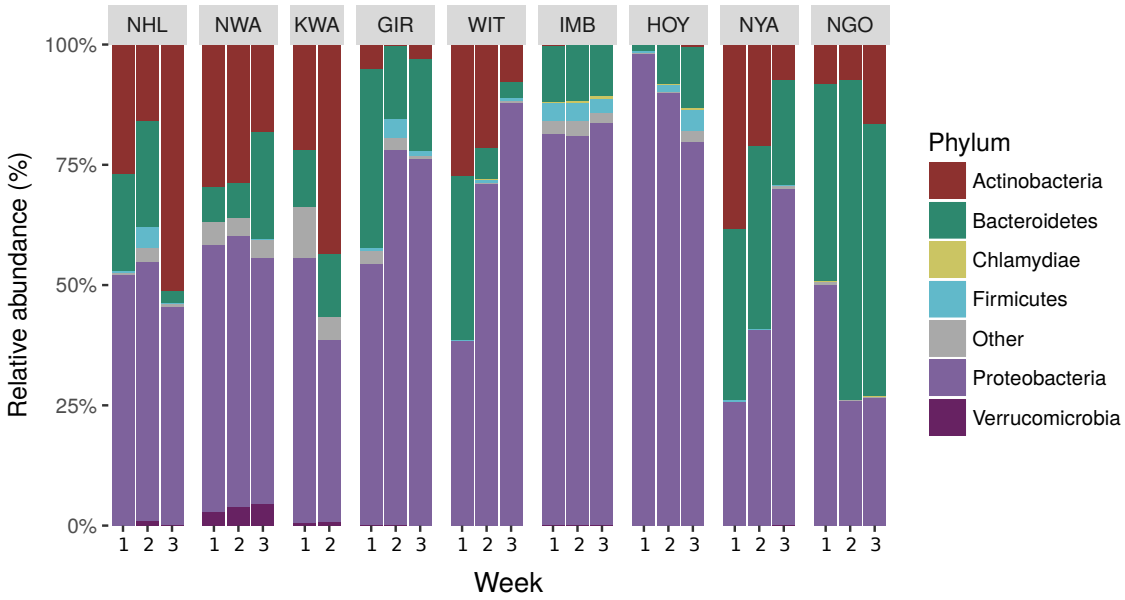


Figure 3–7: Relative abundances of bacterial phyla across weekly samples.

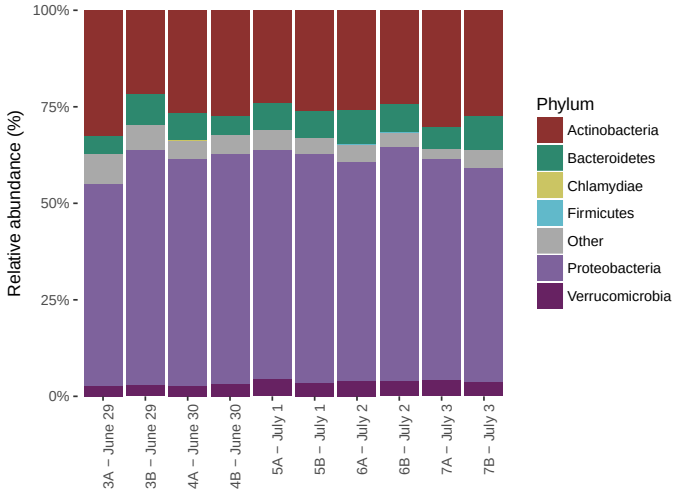


Figure 3–8: Relative abundances of bacterial phyla across five days at a single site (NWA).

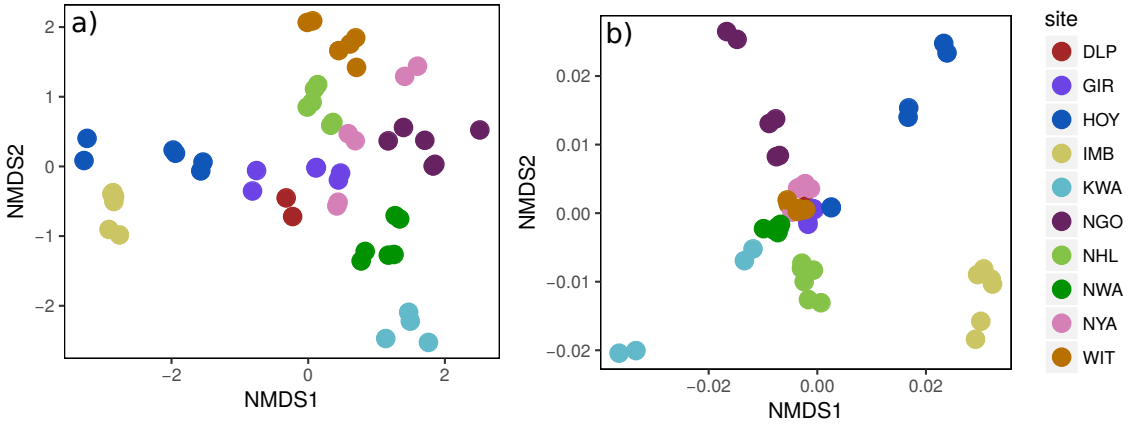
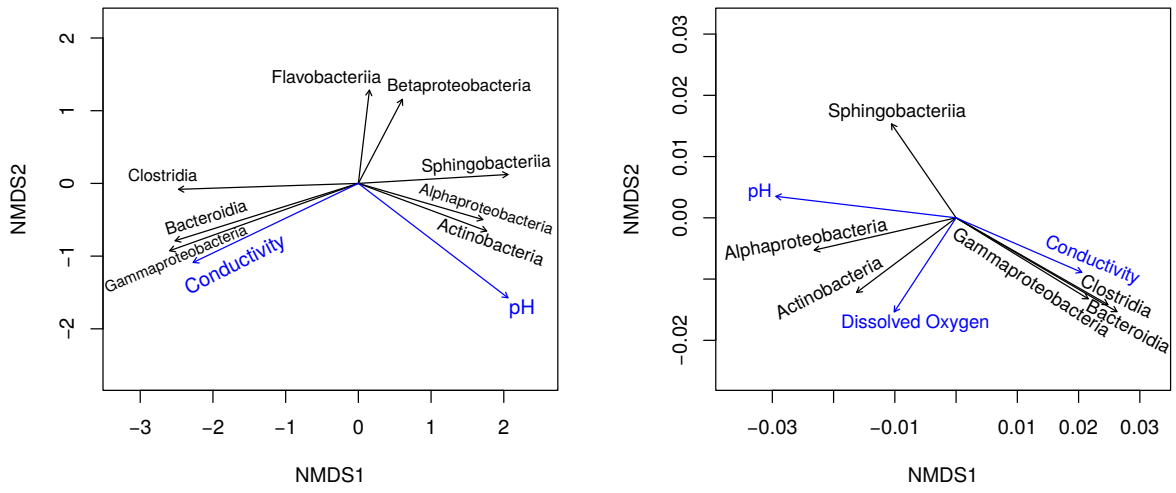


Figure 3–9: Nonmetric multidimensional scaling (NMDS) ordination of variation in bacterial community structure across 54 samples based on a) Bray-Curtis and b) abundance-weighted UniFrac distances. Colours represent site.



(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

Figure 3–10: NMDS ordinations of variation in bacterial community structure across 54 samples based on a) Bray-Curtis and b) abundance-weighted UniFrac distances. Arrows indicate the direction of significant ($p < 0.05$) correlations among variables and the NMDS axes, with arrow length indicating the strength of the correlation. Blue arrows indicate environmental variables, while black arrows indicate relative abundances of sequences from different microbial classes. The ordination axes explain 96.8% (a) and 98.1% (b) of the variance in the dissimilarities (Fig. 6–17).

Linking Statement 3

In Chapters 2 & 3, I apply statistical and molecular tools to gather baseline information that may be useful for the proactive surveillance of infectious diseases. This includes using evolutionary relationships among hosts to predict likely host-parasite interactions, and developing molecular approaches for monitoring microbes in shared water sources. As we continue to discover potentially pathogenic organisms and build our knowledge of the host ranges of poorly studied parasites, a critical next step for prioritizing research and surveillance efforts is to identify the set of potential host-parasite interactions that likely impose the greatest disease burden. Some parasites are notorious for causing severe disease in infected hosts, while others are relatively benign. In Chapter 2, I use evolutionary relationships among hosts to predict likely host-parasite associations. Phylogeny may provide a good proxy for latent traits, and also reflects the co-evolutionary histories of hosts and parasites. It seems reasonable, therefore, that phylogeny might also be useful for predicting disease outcomes.

In Chapter 4, I use evolutionary relationships among susceptible and infected hosts to predict disease-induced mortality in domesticated animals. Predicting the impact a given parasite will have on host fitness is a major challenge in disease ecology (Osnas and Dobson, 2012; Leggett et al., 2013), but developing theory for multi-host parasites has been limited by the availability of comparable data across a range of host-parasite combinations. The severe economic burdens caused by domesticated animal diseases have encouraged the establishment of organizations that collate information on the numbers of cases and deaths due to infectious diseases from nations around the world. These data provide a unique opportunity to examine parasite virulence through a comparative lens, and identifying predictors of disease mortality in domesticated animals has the potential to benefit global animal health as well as contribute to ecological and evolutionary theory of infectious diseases.

CHAPTER 4
Evolutionary relationships among hosts predicts mortality of infectious diseases

Maxwell J. Farrell[1] & T. J. Davies[2,3]

Author Affiliations:

[1]Department of Biology, McGill University

[2]Botany, Forest & Conservation Sciences, University of British Columbia

[3]African Centre for DNA Barcoding, University of Johannesburg

4.1 Abstract

Infectious diseases vary in the impacts they have on their hosts; some are nearly always fatal, whereas others rarely result in death. Identifying the drivers of mortality for diseases that infect multiple host species may help to reduce the burden of infectious diseases across the globe, and predict the severity of novel host-parasite combinations. Using a global dataset of over 4000 case-fatality rates for 65 parasites and 12 domesticated host species, we show that the average evolutionary distance from an infected host to other mammal host species is a strong predictor of disease-induced mortality. We find that as parasites infect species outside of their typical phylogenetic host range, they are more likely to result in lethal infections, the odds of host death doubling with each additional 10 million years of evolutionary distance. The study of domesticated animal diseases reveals patterns in the evolution of highly lethal parasites that may be difficult to observe in the wild, and indicates that the severity of emerging infectious diseases may be predicted from the evolutionary relationships among hosts.

4.2 Introduction

Infectious diseases that cross species barriers are responsible for severe human health burdens (Hotez et al., 2014), and act as direct and synergistic drivers of species extinctions (Heard et al., 2013). Many of these diseases are also shared with domesticated animals

and additionally impact human well-being via loss of food security, labour and livelihoods, costs of prevention and control programs, and increased human infection (Dehove et al., 2012). However, the severity of disease can vary dramatically among parasites. Canine rabies alone results in approximately 59,000 human deaths and 8.6 billion USD in economic losses annually (Hampson et al., 2015). By contrast, other diseases rarely result in death. For example, bovine brucellosis largely impacts cattle by causing abortion, infertility and reduced growth, but disease induced mortality in adult cows is uncommon (McDermott et al., 2013).

For parasites restricted to a single host species, the reduction in host fitness caused by infection, termed virulence, should evolve to an optimal level determined by a trade-off with transmission (Cressler et al., 2016; Alizon et al., 2009). For multi-host parasites, conflicting trade-offs may select for high or low virulence depending on the evolutionary histories and ecological backgrounds of the parasite and each susceptible host species (Woolhouse et al., 2001). Despite a large body of work on mechanisms driving virulence in single-host single-parasite systems, we still lack a framework for predicting virulence of multi-host parasites (Leggett et al., 2013). This gap in our understanding of parasite virulence is a major concern given recent increases in the emergence of diseases that transmit among humans and animals (Jones et al., 2008).

In the absence of constraints, the expansion of a parasite's host range should provide a larger pool of susceptible hosts and increased opportunities for transmission and persistence, in turn allowing for higher levels of virulence to evolve (Barrett et al., 2009). However, as parasites adapt to infect novel hosts they may encounter trade-offs such that the ability to utilize resources of their original host is reduced (Ebert, 1998; Longdon et al., 2014), ultimately resulting in limited replication and decreased virulence among more generalist parasites (Antonovics et al., 2013). This trade-off is supported by comparative studies of plant RNA viruses and avian malaria parasites in which specialist parasites tended to be more virulent than generalists (Garamszegi, 2006; Agudelo-Romero and Elena, 2008).

Yet some generalist parasites remain highly virulent, potentially due to increased transmission opportunities, or to a decoupling of virulence with transmission. This decoupling may occur through parasites escaping costs of high virulence in hosts contributing little to transmission, or through co-evolutionary mismatch and maladaptation of parasites and novel hosts, resulting in sub-optimal virulence in some host species (Leggett et al., 2013). Thus, while virulence was historically considered a property of the parasite, it may be better understood as an outcome of the interaction between parasite and host (Poulin and Combes, 1999).

For the vast majority of host-parasite interactions, the full suite of traits underlying virulence are either poorly estimated or unknown. Our knowledge of host evolutionary relationships is often much better, and phylogeny can act as a proxy for latent traits and evolutionary histories that have shaped host-parasite associations (Davies and Pedersen, 2008). Closely related hosts suffer similar impacts for some parasites of *Drosophila* (Longdon et al., 2015; Perlman and Jaenike, 2003), consistent with the prediction that parasite virulence should co-vary with host phylogeny, but the influence of host evolutionary relationships across multiple host-parasite combinations is less well understood.

Fitness costs that parasites experience when adapting to novel hosts are expected to increase with evolutionary distance from the original host (Antonovics et al., 2013), leading to the prediction of lowered virulence following greater phylogenetic jumps. This pattern, referred to as "non-host resistance", may act in tandem with resistance evolved by hosts in response to infection. However, the strength of evolved host resistance is expected to decrease with evolutionary distance from a parasite's original host, since distantly related hosts may have experienced little selective pressure to evolve defenses (Antonovics et al., 2013). The relative strengths of these opposing relationships with evolutionary distance will influence the expressed level of virulence for a given host-parasite interaction.

To explore the link between host specificity and virulence we use a measure that takes into account both the diversity of susceptible hosts and their relative evolutionary

relationships, which we term "host evolutionary isolation" (Fig 4–1). For a given parasite, host evolutionary isolation measures the mean phylogenetic distance from all documented host species to the infected host species, and may capture the extent of adaptation of both the parasite and infected host. This metric is analagous to species-level measures of mean phylogenetic relatedness, which have been used to describe patterns of species invasion (Strauss et al., 2006), and a related metric was found to be a strong predictor of disease pressure in plant communities (Parker et al., 2015). Here we evaluate whether this relationship can inform patterns of infection-induced mortality for domesticated mammals.

The majority of domesticated mammal diseases can infect multiple species (Cleaveland et al., 2001) and in Africa, thirty-five priority livestock diseases together result in 9 billion USD in losses per year, primarily due to animal deaths (OIE, 2015). World Trade Organization member countries have undertaken systematic reporting of economically important animal diseases over several years, with numbers of cases and deaths aggregated by the World Organization for Animal Health (OIE) (OIE, 2016), providing a remarkable dataset on disease-induced mortality rates for single and multi-host parasites. These data present a unique opportunity to examine parasite virulence through a comparative lens, and identify predictors of virulence for diseases that have far reaching consequences for human and ecosystem health.

Using OIE data on infection-induced mortality rates in domesticated mammals, we construct a Bayesian hierarchical model to examine the relationship between host evolutionary isolation and virulence. The database comprises 4157 reports (after removing those for which culling was recorded) for 202 unique combinations of 65 parasites and 12 hosts, reported by 155 countries across 7 years. Among host-parasite combinations average mortality varied substantially (Fig 4–2A). While virulence can take on many forms, data on host mortality is most widely reported, and we use it here to quantify disease impact on host fitness. For each parasite, we identified the set of known domestic and wild mammal host species from two recently published global host-parasite databases (see section 4.3). This

returned 788 unique host-parasite interactions, from which we quantified host evolutionary isolation, and host species richness calculated as the total number of documented mammal host species for each parasite (Fig 4–2B).

To separate the importance of host evolutionary isolation and host species richness from other factors that might also influence host mortality, we include additional co-predictors and hierarchical terms in our model. These include at the parasite level traits for major modes of transmission (via arthropod vector, via reproduction, and production of an environmental resting stage), plus hierarchical effects of parasite type (virus, bacteria, helminth, etc.) to account for parasite traits not measured directly. We additionally include hierarchical effects for host, host taxonomic order, country, and year of reporting. Environmental conditions, which include socio-economic factors such as the ability of local peoples to maintain animal health, effects of ambient temperature on parasite growth rate, or co-infection with other parasites may also influence host mortality. To control for these country-level effects we include per capita Gross Domestic Product (GDP) and latitude per country in addition to modelling variation among countries. The virulence-transmission trade-off suggests that outbreaks resulting in large numbers of infected individuals are unlikely to be associated with high mortality, as premature host death restricts transmission rate, ultimately resulting in lower case numbers for more lethal diseases (Alizon et al., 2009). We therefore also include the number of cases per report as an offset variable. We estimate the effect sizes of these predictors on host mortality with a Bayesian hierarchical binomial-logit model.

Methods

4.3 Data

Case-fatality reports

Reports of number of cases and deaths due to infection were taken from published OIE year end reports for the years 2005-2011 (OIE, 2005, 2006, 2007, 2008, 2009, 2010,

2011). Reported by individual countries, these include information per disease-host combination on the number of cases (infected individuals), deaths due to infection, individuals destroyed, and individuals slaughtered. We included only reports of diseases in mammal hosts. We excluded any observations in which host individuals were reported as destroyed or slaughtered as this would interfere with estimates of deaths due to infection. We also excluded the few instances where the reported number of deaths due to infection exceeded the number of reported cases.

Host and parasite Latin binomials

Reported host codes used by the OIE were assigned a Latin binomial based on a combination of geographic location, OIE reports, and classifications defined by Clutton-Brock (1999) (Table 4–1). Reports that included OIE host codes "cer" (cervidae) and "o/c" (sheep or goats) could not be attributed to a single host species and were excluded.

OIE host code	Location	Binomial
bov	Global	*Bos taurus*
buf	Sub-saharan Africa	*Syncerus caffer*
buf	North America	*Bison bison*
buf	Asia, Latin America, Caribbean, Europe, N. Africa	*Bubalus bubalis*
can	Global	*Canis lupus*
cap	Global	*Capra hircus*
cml	Global	*Camelus dromedarius*
equ	Global	*Equus caballus*
fel	Global	*Felis silvestris*
lep	Global	*Oryctolagus cuniculus*
ovi	Global	*Ovis aries*
sui	Global	*Sus scrofa*

Table 4–1: Conversion table for OIE host codes to Latin binomials.

Reported disease names were assigned a parasite Latin binomial based on OIE publications (disease summaries from the OIE Terrestrial Manual (OIE, 2012) and OIE technical disease cards (available via www.oie.int/animal-health-in-the-world/technical-disease-cards). For diseases caused by a particular subspecies or strain, this subtype was kept in

cases where susceptible host species was available for these subtypes (Equine Influenza being largely caused by strain H3N8, and Paratuberculosis caused by *Mycobacterium avium paratuberculosis*). Diseases attributed to multiple species were removed (Atrophic rhinitis of swine, Equine piroplasmosis, Equine rhinopneumonitis, Horse mange, Leishmaniosis, Leptospirosis, Sheep and goat pox, Theileriosis, Trichinellosis, and Trypanosomosis), unless the likely causative species could be identified based on geography and/or reported host species (Bovine babesiosis in Europe caused by *Babesia divergens*, Malignant catarrhal fever in sheep worldwide largely caused by *Macavirus ovine herpesvirus 2*, and Malignant catarrhal fever in African cattle caused by *Macavirus alcelaphine herpesvirus 1*). Diseases caused by prions (Scrapie, Bovine Spongiform Encephalopathy) were excluded.

Host specificity

The suite of mammalian host species infected by each parasite was taken from the Global Mammal Parasite Database 2.0 (Stephens et al., 2017) and a static version of the Enhanced Infectious Disease Database (EID2) database (Wardeh et al., 2015). Host species for *Influenza A H3N8* and *Mycobacterium avium paratuberculosis* are not included in the static version of the EID2 database, so were instead taken from EID2 online (eid2.liverpool.ac.uk, accessed June 14th 2017). We also included the host species reported as infected by each parasite in the OIE report data used in the analysis. Host Latin binomials were standardized to the 2005 Wilson Reeder taxonomy (Wilson and Reeder, 2005) using Wilson & Reeder online (www.departments.bucknell.edu/biology/resources/msw3) and the Wilson & Reeder 1993-2005 binomial synonym table included in PanTHERIA (Jones et al., 2009). Hosts reported to subspecies were collapsed to the parent binomial, and hosts not reported to species level were removed. *Homo sapiens* were excluded. Host species richness was then calculated as the number of unique host Latin binomials associated with each parasite. For each combination of host and parasite reported in the OIE data, mean phylogenetic distances from all known hosts to the infected OIE host was calculated using

the Fritz et al. mammal supertree (Fritz et al., 2009) and the R package ape version 3.4 (Paradis et al., 2004).

Parasite traits

Transmission mode is often listed as a key factor linked to virulence (Ewald, 1983; Alizon et al., 2009; Rigaud et al., 2010; Cressler et al., 2016). Here we include whether a parasite is transmitted by an arthropod vector, is transmitted as a function of reproduction (either vertically transmitted, sexually transmitted, or passed from mother to offspring via ingestion of milk or colostrum), and whether it has a resting stage capable of persisting for long periods of time in the environment (typically months to years). Binary parasite traits coding primary modes of transmission and the use of avian species as reservoir hosts were taken from OIE publications (disease summaries from the OIE Terrestrial Manual (OIE, 2012) and OIE technical disease cards), and from Lefèvre et al. (2010). Parasite-level effects were modelled as a function of these covariates plus hierarchical effects of parasite type (virus, bacteria, helminth, etc.), to account for phylogenetic non-independence and to capture additional parasite traits not measured directly.

Country-level covariates

Host mortality is also likely influenced by local environmental conditions. In our data, these may include socio-economic factors such as the ability of local peoples to maintain animal health, effects of ambient temperature on parasite growth rate, or co-infection with other parasites. While the scale of reporting does not allow us to investigate these factors directly, we include two country-level predictors: 1) per capita Gross Domestic Product (GDP) to model economic abilities to reduce host mortality, and 2) latitude as a proxy for temperature and biodiversity gradients that may reflect environmental conditions determining the strength of species interactions (Schemske et al., 2009), in addition to modelling country-level variation. To include country-level covariates from the World

Bank World Development Indicators API (https://data.worldbank.org/data-catalog/world-development-indicators), we standardized country names to those used in the WDI R package version 2.4 (Arel-Bundock, 2013). For each country we extracted mid-country latitude and per capita in current US dollars (WDI code "NY.GDP.PCAP.CD") using the WDI package. Countries that did not have reported GPD per capita from the WDI were supplemented with information from the United Nations Data Retrieval System (data.un.org) so that there was at least one estimate of per capita GDP for the period of 2005-2011. Mean gross domestic product per capita per country was then calculated across all years. We excluded records from countries with no iso3 code or for which no latitude was reported.

4.4 Methods

Model

Using a hierarchical Bayesian binomial-logit model, we model deaths ($deaths_i$) as following a binomial distribution determined by sample size per observation ($cases_i$) and a probability parameter p_i. The higher-level structure of the model is as follows:

$$deaths_i \sim Bin(cases_i, p_i) \tag{4.1}$$

Where p_i is modeled with β_0 as the grand mean plus the effects of mean phylogenetic distance from all known hosts to the species infected ($EvoIso_i$), the number of cases per observation ($cases_i$), and partially-pooled hierarchical effects for parasites (μ_{para}), hosts (μ_{host}), countries ($\mu_{country}$), and years (μ_{year}):

$$logit(p_i) = \beta_0 + \beta_1 * EvoIso_i + \beta_2 * log(cases_i) + \mu_{para} + \mu_{host} + \mu_{country} + \mu_{year} \tag{4.2}$$

Parasite level effects, μ_{para}, are defined by a normal distribution as follows:

$$\mu_{para} \sim \mathcal{N}(\beta_3 * SR_{para} + \beta_4 * aviRes_{para} + \beta_5 * vect_{para} + \beta_6 * repro_{para} + \beta_7 * envRest_{para} + \mu_{type}, \sigma_P^2) \tag{4.3}$$

Where the difference from the grand mean (β_0) for each parasite ($para$) is determined by host species richness (SR_{para}), transmission modes ($aviRes_{para}$, $repro_{para}$, $envRes_{para}$), and a hierarchical effect of the parasite type (μ_{type}), and variance parameter (σ^2_P).

Parasite taxonomic type (i.e. virus, bacteria, helminth, etc.), μ_{type}, is modelled following a normal distribution with mean of zero and variance parameter (σ^2_T) as follows:

$$\mu_{type} \sim \mathcal{N}(0, \sigma^2_T) \tag{4.4}$$

Host level effects, μ_{host}, are modelled following a normal distribution with mean determined by a hierarchical effect of the host taxonomic order (μ_{order}) and variance parameter (σ^2_H) as follows:

$$\mu_{host} \sim \mathcal{N}(\mu_{order}, \sigma^2_H) \tag{4.5}$$

Host taxonomic order level effects, μ_{order}, are modelled following a normal distribution with mean of zero and variance parameter (σ^2_O) as follows:

$$\mu_{order} \sim \mathcal{N}(0, \sigma^2_O) \tag{4.6}$$

Country level effects, $\mu_{country}$, are modelled following a normal distribution with mean determined by gross domestic product per capita (GDP_c) and latitude ($latitude_c$), and variance parameter (σ^2_C) as follows:

$$\mu_{country} \sim \mathcal{N}(\beta_8 * GDP_c + \beta_9 * latitude_c, \sigma^2_C) \tag{4.7}$$

Year level effects, μ_{year}, are modelled following a normal distribution with mean of zero and variance parameter (σ^2_Y) as follows:

$$\mu_{year} \sim \mathcal{N}(0, \sigma^2_Y) \tag{4.8}$$

Priors & Data transformations

Following the recommendations of Gelman et al. (2008), continuous predictors were normalized to mean of zero and standard deviation of 0.5. Estimated parameters were modelled using weakly informative priors as recommended by Ghosh et al. (2015) and the Stan development team (https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations):

$$\beta_{0-9} \sim Student\ t(4, 0, 1) \tag{4.9}$$

$$\sigma^2_{P,H,O,C,Y} \sim Half\ Student\ t(4, 0, 1) \tag{4.10}$$

Sampling and Convergence Diagnostics

Models were fit in Stan (Stan Development Team, 2017c; Carpenter et al., 2017) via R 3.2.3 (R Core Team, 2015) with rstan version 2.14.2 (Stan Development Team, 2017a) using 4 chains with 30,000 iterations per chain. The first 15,000 iterations per chain were used for warm-up and discarded. The remaining posterior was thinned to retain every 10th iteration, resulting in a total of 6,000 posterior draws. Convergence was diagnosed by observation of Rhat values equal to 1 (see Table 6–24) and explored with shinystan version 2.4.0 (Stan Development Team, 2017b). Posterior predictive checks were performed to ensure model validity and fit to the data. The main model was also fit with simulated data to ensure the model performs as expected and is able to recover simulated parameters.

4.5 Results

We find that mortality is highest when the infected host is evolutionarily distant from other documented hosts (Fig 4–3A, Fig 4–4, Table 6–24), with an increase of 10 million years of evolutionary isolation resulting in a two-fold increase in the odds of host death (odds ratio 50% credible interval: 1.99 - 2.15). A disease infecting only Primate hosts that shifts to infect an Artiodactyl is thus expected to have $\sim$ 4.8 times higher odds of host death than a parasite shifting from hosts in the order Carnivora. This effect becomes

stronger when excluding single-host parasites from the analysis (Appendix 6.3). Consistent with the virulence-transmission trade-off, our results also indicate that instances of high mortality are not usually associated with large numbers of infected individuals (Fig 4–3B, Fig 4–4). We find some support for a positive relationship between mortality and host species richness (50% credible interval does not overlap zero), opposite to the relationship that would be predicted if there was a trade-off between parasite generalism and virulence (parasites with larger host ranges causing lower mortality on average). There is large variability in the strength of this relationship, as is the case for all parasite-level predictors; however, host species richness is a better predictor than host taxonomic diversity (Appendix 6.3).

4.6 Discussion

Our study demonstrates that as parasites infect domesticated species outside of their typical evolutionary host range, they have a higher probability of resulting in lethal infections. This result is surprising considering that for some parasites, such as poxviruses, it has been suggested that host switches involving more distantly related species tend to result in benign infections, whereas shifts onto closely related species lead to severe disease (Haller et al., 2014). Our results, to the contrary, indicate that on average host switches involving more distantly related species have a greater potential to result in host death. Certain poxviruses, and some of the diseases included in this study have host ranges that extend beyond mammals, and expanding our framework to include non-mammal hosts may provide further insights into trade-offs for parasites exhibiting extreme phylogenetic generalism. For example, non-host resistance may dominate in parasites that undergo cross-kingdom jumps (van Baarlen et al., 2007).

There are several reasons why parasites might cause high mortality in evolutionarily isolated hosts. Mortality can result from a combination of direct damage caused by parasites, and damage caused by the host's immune response to infection, which may

impose different selective pressures on the evolution of virulence (Graham et al., 2005). For example, if parasites impose fitness costs on a host, but the host contributes little to transmission, or constitutes an epidemiological dead end, there may be little to no selection for parasites to reduce virulence (Antonovics et al., 2013). This can occur when the majority of transmission is facilitated by a reservoir host, such as has been suggested for foot and mouth disease in southern Africa which uses asymptomatic African buffalo as a reservoir, but causes severe outbreaks after spillover in domestic cattle (Michel and Bengis, 2012). This is also the case for many arboviruses, which commonly use birds as reservoir hosts, but fail to transmit after spillover into mammal hosts such as humans and horses, where they are highly virulent (Weaver and Barrett, 2004). Identifying reservoir species can be a difficult challenge, and for many parasites included here the reservoirs are unknown. However, we were able identify diseases in our data that use avian species as reservoirs, though found no strong evidence that these parasites cause higher mortality (Appendix 6.3, Table 6–27).

Virulence may also be high for host-parasite combinations if host mortality is decoupled from transmission. This decoupling may occur if parasites infect tissues unrelated to transmission, such as bacterial meningitis infection of the central nervous system (Longdon et al., 2015), if host mortality is due to hyperactive immune responses (Graham et al., 2005), or if parasites produce a long lasting environmental resting stage (Cressler et al., 2016). We did not have information on infected tissues or specific host immune responses to model these effects directly; although we were able to identify parasites capable of producing an environmental resting stage. However, this and other transmission modes included in our model had no clear relationship with host mortality, with the 50% credible intervals overlapping zero (Fig 4–4). Nonetheless, parasite identity had an important effect (Table 6–24, Fig 6–19, Fig 6–18), suggesting other parasite traits not considered here modify virulence.

Our model also reveals an important effect of country (Fig 4–3, Fig 6–19). Variation in host mortality among countries may indicate the potential to identify animal management practices that reduce infection-induced mortality in one nation, and introduce them in other nations. Nations with the largest positive country effects, indicating mortality rates higher than predicted from the rest of the model, may have difficulties building capacities for early detection and prevention of infectious disease outbreaks. For example, Sri Lanka, which ranks first, has struggled to develop adequate legislative frameworks and infrastructure for addressing veterinary public health issues, largely due to rural development driving animal health priorities rather than export-oriented animal production (Dissanayake et al., 2012). Kyrgyzstan, ranked second, has seen a severe deterioration of its veterinary and sanitation systems since independence in 1991, and also suffers from severe under-reporting of zoonotic diseases despite the majority of its citizens having livelihoods dependent on livestock farming (Counotte et al., 2016). In contrast, nations with mortality rates lower than predicted from the rest of the model (e.g. Macedonia, China, and Iran), while still suffering considerable burdens from infectious diseases, have made great strides towards improved surveillance systems, and implemented successful large-scale control and eradication programs (Stojmanovski et al., 2014; Hotez et al., 2012; Wang et al., 2008).

We have shown host evolutionary isolation to be a strong predictor of host mortality in domesticated mammals, and our results re-enforce the notion that virulence is a product of both parasite and host properties. The subset of diseases for which we have multiple estimates of case-fatality rates are also those diseases that have global health impacts. While these might not represent the full spectrum of parasite virulence, we suggest that these diseases provide a window into the evolution of virulence that is otherwise hard to observe. In natural systems, spillover of highly virulent diseases often display stuttering chains of transmission before parasites burn themselves out (Longdon et al., 2014). It is likely, therefore, that these instances of highly virulent disease in wild hosts may often go undocumented (Leggett et al., 2013). High host densities allow parasites to maintain high

transmission rates despite causing high mortality (Mennerat et al., 2010), and in human influenced systems, artificially high host densities of domesticated animals may facilitate the maintenance of highly deadly diseases, allowing us to better observe their behaviour. In addition, intensively farmed populations often have shorter lifespans and are more likely to experience co-infection, both of which can also promote parasite virulence (Mennerat et al., 2010). Farmed populations often have low levels of genetic variation which may favour the spread of disease, and on longer timescales selection for traits that promote production in livestock may lead to trade-offs with reduced resistance to parasites (Mennerat et al., 2010).

The diseases with exceptionally high mortality studied here are likely achieving such high virulence through multiple pathways including spillover from wildlife reservoirs into dead end hosts, inappropriate host immune responses, a decoupling of transmission from virulence, and maladaptation due to recent or frequent host shifts. While it is difficult to differentiate among these alternative mechanisms, we suggest that the evolutionary distances among infected and susceptible hosts can, to some extent, capture these multiple dimensions.

Predicting the likely outcomes of novel host-parasite interactions presents a major challenge in disease ecology. Our analysis showing how evolutionary relationships among hosts links to the virulence of multi-host pathogens provides an important first step towards addressing this challenge, and suggests that diseases that shift between distantly related hosts are likely to be more deadly. This is also likely to be the case for emerging animal diseases shifting to infect humans. As ecosystems are increasingly transformed by human actions we may witness the formation of communities never before seen in evolutionary history, and with them comes the opportunity for diseases to interact with novel hosts. For example, Nipah virus, a lethal zoonosis transmitted from bats to humans via domestic pigs, emerged as a result of agricultural intensification and repeated spillover events from wild bats, followed by transmission among pig farms (Daszak et al., 2013). Our results indicate that proactive approaches to fill gaps in our knowledge of the wildlife hosts (Farrell et al.,

2013) may help predict mortality of emerging diseases in novel hosts. We also suggest there may be opportunities to reduce disease mortality by identifying factors that contribute to the large variation in mortality across countries, and by close monitoring of diseases at the domestic-wildlife interface.

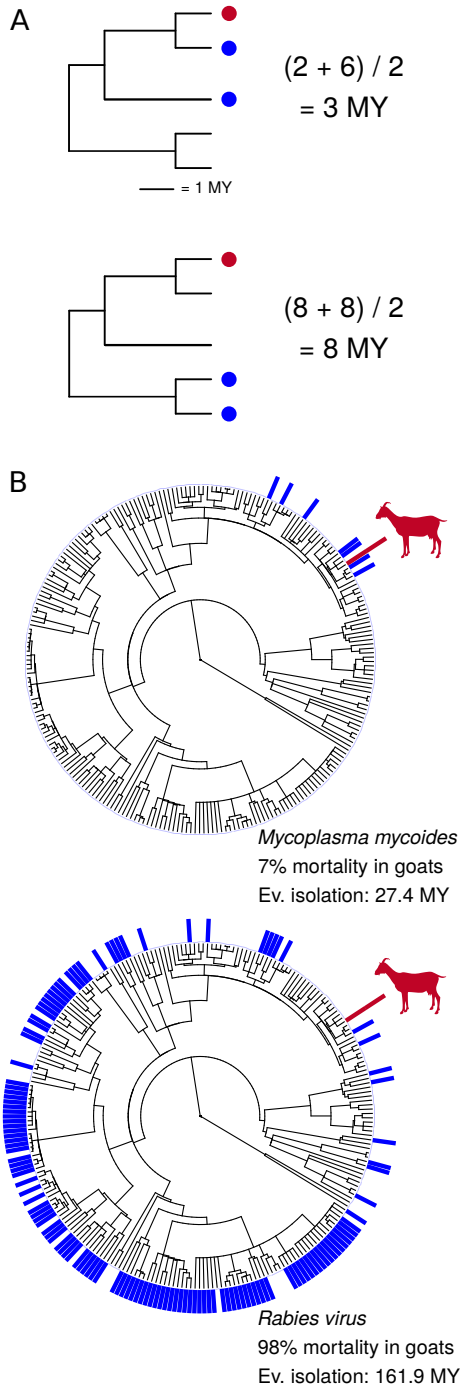


Figure 4–1: a) Example of how host evolutionary isolation is calculated. Red circles indicate the infected host, blue circles indicate documented hosts. Host evolutionary isolation is calculated as the mean phylogenetic distance from the infected host to all documented host species. b) Examples with *Mycoplasma mycoides* and *Rabies virus*. Documented hosts are indicated by blue bars on the host phylogeny, with host evolutionary isolation and average mortality calculated for goats (*Capra hircus*, shown in red).

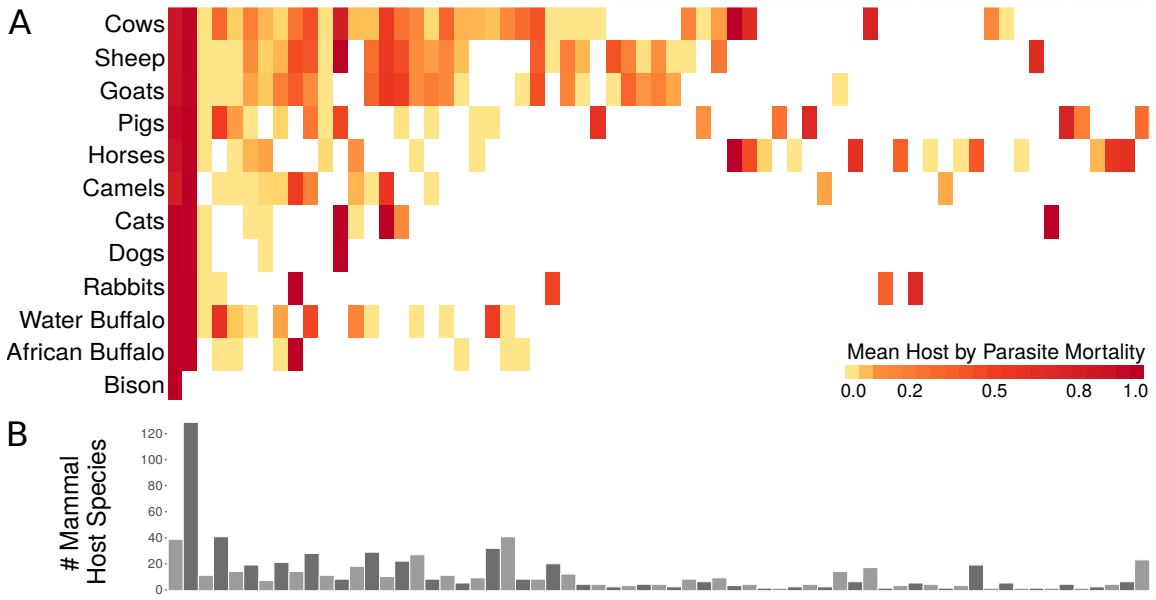


Figure 4–2: a) Heatmap of mean host by parasite mortality derived from the OIE World Animal Health yearly reports from 2005-2011 (full heatmap with disease common names included in Appendix Fig 6–20). b) Barplot of the number of documented mammal host species per parasite derived from the Global Mammal Parasite 2.0 and EID2 databases. Order of parasites matches column order in b).

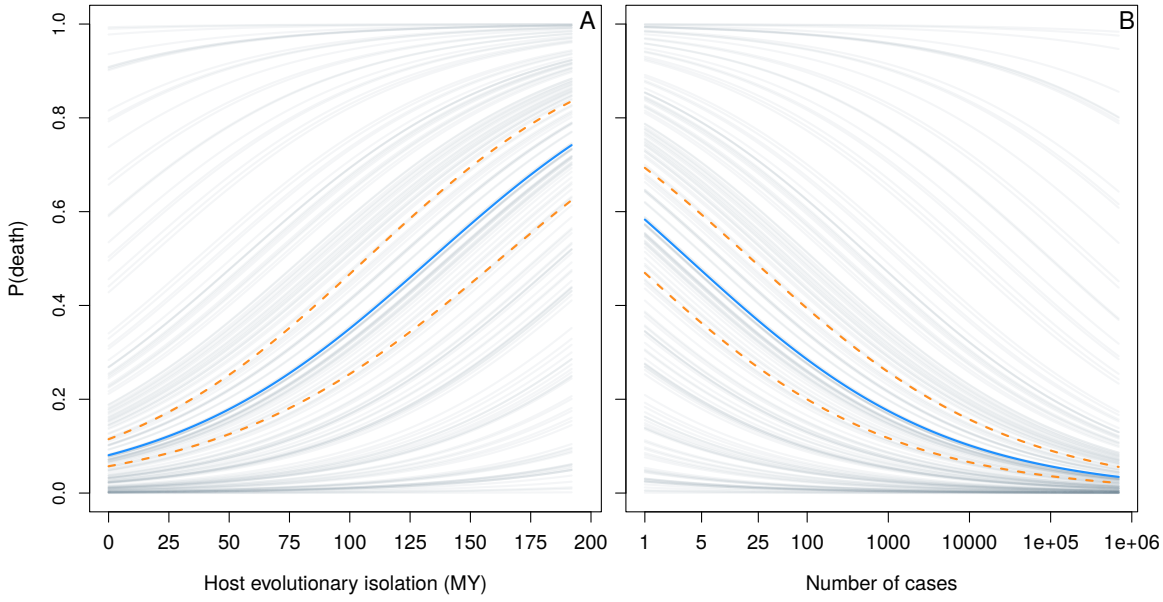


Figure 4–3: Posterior predictions of the probability of death as a function of a) host evolutionary isolation (in millions of years), and b) the number of cases. Solid blue lines represent the mean logistic curve, dashed yellow lines represent the upper and lower bounds of the 50% credible interval. Grey lines depict equivalent mean curves offset by the posterior mean effects for each country.

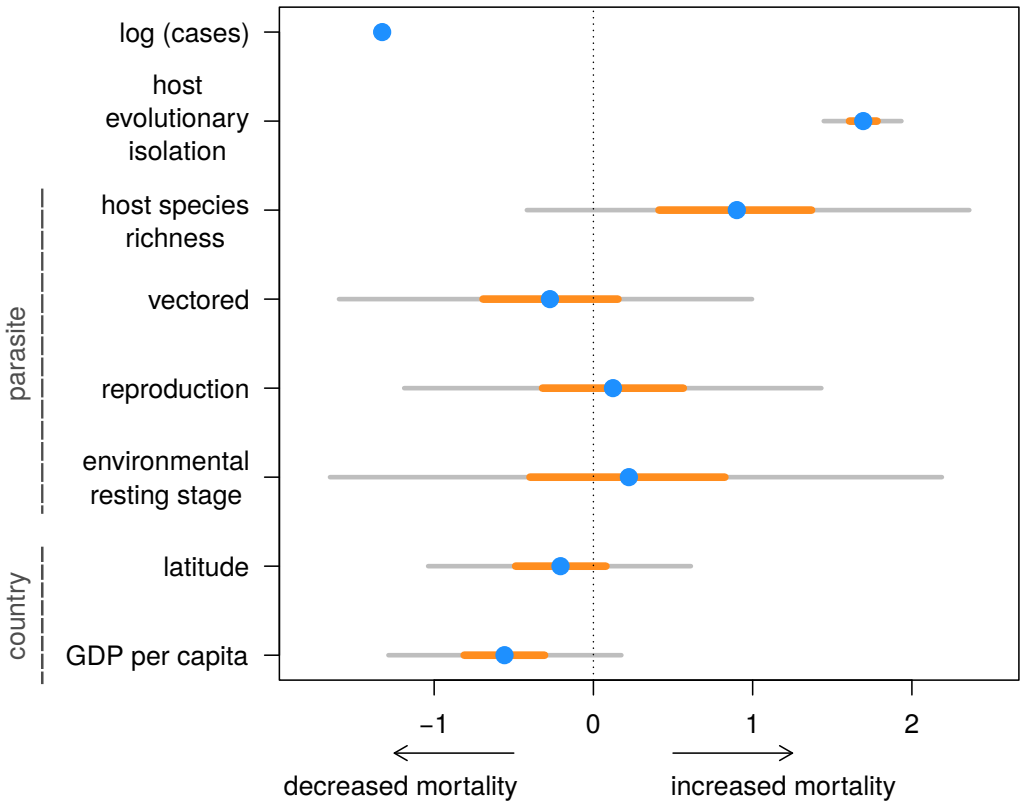


Figure 4–4: Estimated regression coefficients for continuous predictors. Blue circles represent posterior means, yellow horizontal lines represent 50% credible intervals, grey horizontal lines represent 95% credible intervals. Predictors at the parasite and country level are indicated.

Linking Statement 4

In Chapter 4, I show that the evolutionary relationships among host species can be used to predict the mortality of infectious diseases in domestic animals. I suggest that the process of domestication facilitates the evolution and maintenance of deadly diseases, many of which are known to cause severe declines in wild animals (Smith et al., 2009) and act in synergy with other human-mediated drivers of extinction such as habitat loss, hunting, and persecution (Heard et al., 2013). As human activities encroach into previously wild landscapes, this creates opportunities for contact among previously isolated hosts and facilitates the spread of diseases (Faust et al., 2018). However, as the fate of parasites is intimately linked to that of their hosts, processes that drive hosts to extinction may either cause the coextinction of parasites (Woolhouse et al., 2001; Dunn et al., 2009; Colwell et al., 2012), or alter host specificity. Because host specificity is linked to disease impact, host extinction could impose selection on parasites that lead to increases or decreases in disease burden.

In Chapter 5, I explore how host declines reshapes the parasite assemblages of endangered species. Coextinction theory predicts that parasites specializing on a single host species are more likely to go extinct following significant declines of their hosts (Koh et al., 2004). Parasites comprise a major component of biological diversity (Dobson et al., 2008), and contribute to ecosystem health (Hudson et al., 2006). The loss of parasite diversity is therefore a conservation concern, but they are often neglected when setting conservation targets (Gómez and Nichols, 2013). The extinction of co-evolved specialist parasites may facilitate the emergence of novel diseases due to loss of competition among parasites (Lloyd-Smith, 2013; Dunn et al., 2009). The loss of any single host will also alter selective pressures with respect to extant hosts, which could have important implications for both disease transmission pathways and evolution of virulence. Developing a robust theory of

coextinction is therefore essential for tracking shifting patterns of host-parasite interactions and building surveillance systems for diseases in wildlife. Here, I conduct a comparative study to investigate whether single-host parasites are being lost from threatened hosts more often than multi-host parasites.

CHAPTER 5
The path to host extinction can lead to loss of generalist parasites

Maxwell J. Farrell[1], Patrick R. Stephens[2], Lea Berrang-Ford[3],

John L. Gittleman[2], T. Jonathan Davies[1,4]

Author Affiliations:

[1]Department of Biology, McGill University

[2]Odum School of Ecology, University of Georgia

[3]Geography Department, McGill University

[4]African Centre for DNA Barcoding, University of Johannesburg

This chapter is published in the *Journal of Animal Ecology* 84(4) 978–984

5.1 Abstract

Host extinction can alter disease transmission dynamics, influence parasite extinction, and ultimately change the nature of host-parasite systems. While theory predicts that single-host parasites are among the parasite species most susceptible to extinction following declines in their hosts, documented parasite extinctions are rare. Using a comparative approach, we investigate how the richness of single-host and multi-host parasites is influenced by extinction risk among ungulate and carnivore hosts. Host-parasite associations for free-living carnivores (order Carnivora) and terrestrial ungulates (orders Perrisodactlya + Cetartiodactlya minus cetaceans) were merged with host trait data and IUCN Red List status to explore the distribution of single-host and multi-host parasites among threatened and non-threatened hosts. We find that threatened ungulates harbour a higher proportion of single-host parasites compared to non-threatened ungulates, which is explained by decreases in the richness of multi-host parasites. However, among carnivores threat status is not a significant predictor of the proportion of single-host parasites, or the richness of single-host or multi-host parasites. The loss of multi-host parasites from

threatened ungulates may be explained by decreased cross-species contact as hosts decline and habitats become fragmented. Among carnivores threat status may not be important in predicting patterns of parasite specificity because host decline results in equal losses of both single-host parasites and multi-host parasites through reduction in average population density and frequency of cross-species contact. Our results contrast with current models of parasite coextinction and highlight the need for updated theories that are applicable across host groups and account for both inter and intraspecific contact.

5.2 Introduction

Parasites are often viewed as negative aspects of ecosystems, as causes of disease and indicators of unhealthy environments. Infectious diseases of wildlife have been implicated in the declines of several free-living hosts (Smith et al., 2009; Heard et al., 2013), and as agents of many important diseases of humans and domesticated animals (Cleaveland et al., 2001). However, parasites are integral components of natural ecosystems with unique roles in food webs (Dunne et al., 2013), regulating community composition (Wood et al., 2007), and maintaining host genetic diversity (Altizer et al., 2003). Furthermore, it has been speculated that loss of parasites may predispose hosts to infection by generalist parasites or emerging diseases (Dunn et al., 2009; Lloyd-Smith, 2013), and thus have severe downstream effects. Recent efforts have therefore rallied for the inclusion of parasitic biodiversity in conservation efforts (Harris and Dunn, 2010; Mihalca et al., 2011; Moir et al., 2012; Pérez et al., 2013; Gómez and Nichols, 2013). Parasites are susceptible to many of the same threats affecting free-living species, but their added dependence on hosts suggests that parasites will be among the organisms most susceptible to extinction in the ongoing biodiversity crisis (Dunn et al., 2009; Colwell et al., 2012).

Parasite extinction risk depends on the number of hosts a parasite infects and the extinction rates of those hosts (Koh et al., 2004; Lafferty, 2012). Recent models of parasite coextinction indicate that generalist parasites (those able to utilize multiple host species)

are less susceptible to coextinction as they have alternative ways to complete their life-cycles (Dunn et al., 2009; Lafferty, 2012). However, local extirpations of infected host populations or preferred host species could result in parasite extinction well before the loss of all available host species as parasites are often unevenly distributed among host individuals, populations, and species (Shaw and Dobson, 1995; Poulin, 2007; Moir et al., 2010; Välimäki et al., 2011). As the transmission of parasites depends in part upon sufficient contact among infected and susceptible hosts (McCallum et al., 2001), extinction processes may decrease the density or abundance of hosts below critical thresholds for parasites to transmit effectively (Deredec and Courchamp, 2003; de Castro and Bolker, 2005).

Although theory suggests that coextinction of hosts and parasites is a common phenomenon, empirical support is scarce and only a few instances of coextinction have been documented (Dunn et al., 2009; Moir et al., 2010). An analysis of parasite richness among wild primates revealed that threatened host species were associated with fewer parasite species compared to non-threatened hosts (Altizer et al., 2007). This result supports coextinction theory; however, the proportion of specialist parasites was not significantly different, suggesting that host declines lead to the loss of both specialist and generalist parasites. A similar result was found among acacias and the phytophagous insects that feed on them (Powell, 2011). A recent study of fish parasites found that specialist parasites tend to use hosts with low vulnerability to extinction (Strona et al., 2013), which may be additional support for the loss of specialist parasites from threatened hosts. Conversely, a study of viral richness in bats found the opposite trend with threatened bat species harbouring increased richness of viruses (Turmelle and Olival, 2009). These conflicting results highlight the need for an improved understanding of coextinction and host-pathogen dynamics in declining host populations.

Here we contrast patterns of single-host versus multi-host parasites among threatened and non-threatened wild ungulates and carnivores. Although host range can be considered as a continuous variable, multi-host parasites may also infect hosts with different risks of

extinction, which makes it difficult to predict coextinction risk based solely on the number of hosts per parasite. However, single-host parasites must go extinct before or immediately upon the loss of their sole host. We use phylogenetic comparative methods to evaluate whether single-host parasites are less common among threatened hosts compared to non-threatened hosts, as we expect host species with small or declining populations to be less likely to support viable populations of single-host parasites. We focus on wild ungulates and carnivores because of the availability of data and their shared evolutionary histories with domesticated mammals.

5.3 Materials and methods

Parasite Records

Records of parasitic organisms reported from free-living carnivores and terrestrial ungulates (Perrisodactlya + Cetartiodactlya minus cetaceans) were obtained from the Global Mammal Parasite Database (Nunn and Altizer, 2005) updated to include papers published up to 2010 and accessed October 15th 2013. This database documents host-parasite associations and encompasses both micro and macro-parasites including viruses, bacteria, protozoa, fungi, helminths, and arthropods. We restricted reports to wild populations of hosts sampled within their native ranges. Host Latin binomials were standardized to Wilson and Reeder (2005). Parasites reported to genus level or higher were excluded from the analyses. Only reports with prevalence greater than zero were included. Parasites were classified as single-host if associated with only one host species in the dataset and multi-host if associated with more than one host species.

Host Traits

To quantify host extinction risk we employed the categorical threat status reported by the 2014 IUCN Red List of Threatened Species (International Union for the Conservation of Nature, 2014). We converted IUCN status to a binary variable with species listed as Least

Concern (LC) or Near Threatened (NT) considered Non-Threatened, and species listed as Vulnerable (VU), Endangered (EN), or Critically Endangered (CR) considered Threatened. Hosts with the status of Data Deficient (DD), Extinct in the Wild (EW) or Extinct (EX) were excluded from our analysis. To limit temporal mismatches in the reporting of parasites and changes in IUCN status, two species reported to have genuine changes in status between the 2010 and 2014 editions of the IUCN Red List were assigned their 2010 status. The IUCN also provides information on species population trends, which was converted into a binary variable with species listed as either decreasing, or not decreasing (i.e. increasing or stable) to capture decreases in host abundances where rate of decline may be below thresholds necessary to be considered threatened (Mace et al., 2008). To control for additional host traits that have been shown to correlate with both extinction risk and parasite richness in mammals (Nunn et al., 2003; Ezenwa et al., 2006; Lindenfors et al., 2007; Cardillo et al., 2008) we included data on average adult female body mass (g), geographic range area (km2), and average population density (number of individuals/km2) per host species from the PanTHERIA database of mammalian life history traits (Jones et al., 2009). Additional population density estimates for ten species in the order Carnivora were obtained from grey literature.

Statistical Analyses

Generalize Estimating Equations (GEEs) (Paradis and Claude, 2002) were used to explore relationships between host traits and univariate responses of a) proportion single-host parasites, b) richness of single-host parasites, and c) richness of multi-host parasites. GEEs offer a flexible and unified method allowing the specification of non-normal error structures for binomial and count data as well as correcting for phylogenetic non-independence. All analysis were conducted in R v.3.1.0 (R Core Team, 2014).

To control for uneven sampling of parasites among hosts we included as a covariate the number of citations per host (Altizer et al., 2007) as reported in the ISI Web of Knowledge

using the Latin binomial of each species in either the title or topic fields. Citation counts as well as estimates of body mass, geographic range, and population density were log transformed prior to analyses. As ungulates and carnivores differ significantly with respect to their life histories, ecology, and predictors of extinction risk, we conducted models separately for each group, though the number of hosts per parasite was taken from the merged host-parasite lists of both groups. Before model fitting we checked the predictors for collinearity by estimating the variance inflation factors (VIFs). In all cases VIFs were less than three, indicating no significant effect of collinearity (Belsley et al., 1980).

Full models including all predictor variables were first fit with the geeglm function in the package *geepack* (Højsgaard et al., 2006). For proportion of single-host parasites, we used counts of single-host and multi-host parasites as the response and specified a binomial error structure with a logit link function. To determine whether or not changes in the proportion of single-host parasites were being driven by changes in the richness of single-host or multi-host parasites, we fit separate models with richness of single-host and multi-host parasites, respectively, as response variables assuming a Poisson error structure.

For each of the full models (6 in total), we estimated Pagels lambda as a measure of phylogenetic signal in the residuals using the fitContinuous function in the package *geiger* (Harmon et al., 2008) and most current inclusive estimates of phylogeny. We used a species-level mammal super tree for ungulates (Fritz et al., 2009), and a more recent species-level tree for the order Carnivora (Nyakatura and Bininda-Emonds, 2012). To correct for phylogenetic non-independence in model residuals due to shared evolutionary histories among hosts, the respective trees were transformed using the estimated lambda value from model residuals (Revell, 2010) with the function transform.phylo in the package *geiger* (Harmon et al., 2008). We then refit the models using the compar.gee function in the package *ape* (Paradis et al., 2004) specifying the phylogenetic covariance structure from the transformed phylogeny.

Sensitivity Analyses

As a full suite of predictors was not available for every host species, sample sizes were maximized by removing predictors with $p > 0.1$ and the above procedure correcting for phylogenetic non-independence in the residuals repeated. The reduced models often allowed for increased sample sizes and thus comparison across a larger number of host species. Threat status and citation count were forced into all models. We additionally fit bivariate models which included only threat status and citation count as predictors to explore the effect of threat status in the absence of ecological covariates. To determine whether patterns differed by parasite type, parasites were categorized as either micro-parasites (bacteria, viruses, fungi, protozoa), or macro-parasites (helminths, arthropods), and models re-run separately for each parasite type. This distinction represents a functional difference whereby micro-parasites tend to have shorter lifecycles and directly reproduce in their definitive hosts (Anderson and May, 1991), which may cause the two groups to respond differently to host extinction.

5.4 Results

A total of 13,724 reports of host-parasite associations were used, representing 4098 unique confirmed host-parasite associations among 235 hosts and 1384 parasites. This resulted in 729 single-host and 656 multi-host parasites. 95 ungulate and 140 carnivore host species were included, of which complete covariate data was available for 68 and 64 respectively. In both groups, mean multi-host parasite richness per host appears lower among threatened hosts when examining data uncorrected for sampling effort or ecological covariates (Fig. 5–1).

In statistical tests threatened ungulates harbour a significantly higher proportion of single-host parasites compared to non-threatened ungulates (Fig. 5–2a, Table 6–28). This result is supported by the finding that threatened ungulates harbour a significantly lower richness of multi-host parasites compared to non-threatened ungulates (Fig. 5–2c), whereas

single-host parasite richness was not significantly associated with threat status (Fig. 5–2b). Multi-host parasite richness was significantly negatively associated with threat status and a decreasing population trend, and significantly positively associated with range area, body mass, and citation count (Fig. 5–2c). The significance and direction of predictors did not vary between full (n=68; Fig. 5–2, Table 6–28) and reduced models (n=68-95; Tables 6–29 & 6–30), except for Threat Status in the bivariate model of multi-host parasite richness, which became marginally non-significant ($p = 0.073$). In contrast, single-host parasite richness did not correlate significantly with any of our ecological predictors. The significance and direction of predictors did not vary between full models when classifying parasites as either micro or macro-parasites (Table 6–31).

Among carnivores we found no significant effect of host threat status on the proportion of single-host parasites, the richness of single-host, or the richness of multi-host parasites (Fig. 5–2d-f, Tables 6–29 & 6–30). However, both the proportion of single-host parasites and richness of single-host parasites were positively associated with population density (Fig. 5–2d-e). The richness of multi-host parasites was not significantly predicted by any of our ecological variables (Fig. 5–2f). As for ungulates, the significance and direction of predictors did not vary between full (n=64; Table 6–28) and reduced models (n=64-140, Tables 6–29 & 6–30). While the significance and direction of predictors was similar for models predicting the proportion of single-host parasites and the proportion of single-host macro-parasites, the proportion of single-host micro-parasites among carnivores was not significantly predicted by any of our ecological predictors (Table 6–31).

5.5 Discussion

We found host extinction risk to be a significant predictor of the proportion of single-host parasites among ungulates, but not carnivores (Fig. 5–2). Although current theory predicts that threatened hosts should have decreased proportions of single-host parasites (Dunn et al., 2009; Lafferty, 2012), we found no trend within carnivores and the opposite

trend within ungulates. For ungulates, we suggest this result can be explained by the disproportionate loss of multi-host over single-host parasites. For carnivore hosts, threat status was not a significant predictor of the proportion of single-host parasites or the richness of either single-host or multi-host parasites. Our results suggest a need for improved theory on the process of coextinction, and highlight the necessity of challenging models with empirical data.

Recent models of coextinction operate under the assumption that parasites will be lost from a system when all of their potential hosts go extinct (Colwell et al., 2012). However, these models focus only on the outcome after complete host extinction, whereas a species will frequently experience significant contractions in abundance and range size well before it finally becomes extinct. Here we used IUCN Red List status and data on host population trend to document this path towards extinction. Our study provides added evidence that models of parasite coextinction with host decline may differ from models based solely on host extinction.

Why does threat status result in a disproportionate loss of multi-host parasites in ungulates but not carnivores? It is possible that this contrast reflects differences in the major threatening processes or life-history traits between the two groups. If ungulates and carnivores are listed under different Red List criteria, this might explain differences in loss of parasites with extinction risk. Indeed, a greater proportion of threatened ungulates included in this study are listed because of an observed reduction in population size (Criterion A1, A2, or A4): 14 out of 20 ungulates opposed to 13 out of 29 carnivores. While this difference is not statistically significant (Pearson's $\chi^2 = 2.10$, $p = 0.147$), it may indicate that threatened ungulates have more often experienced significant reductions in abundance and geographic range. Host geographic range is a key predictor of parasite species richness (Kamiya et al., 2014). Broad ranging species are also more likely to overlap with other host species, increasing opportunities for infection with multi-host parasites (Gregory, 1990). Overlap in host geographic range is a significant predictor of parasite community similarity

in carnivores (Huang et al., 2014) and primates (Davies and Pedersen, 2008), and similarly among sympatric African bovids habitat overlap is positively correlated with increased prevalence of multi-host gastrointestinal parasites (Ezenwa, 2003). Geographic range is often used as a proxy for the amount of interspecific contacts experienced by a given host, while population density is used to represent the amount of intraspecific contact. However, as extinction processes can reduce both species abundances and range (Price and Gittleman, 2007; Hayward, 2009), it is not immediately obvious how population density will respond, especially when host individuals are unevenly distributed throughout a species range.

Among ungulates group living is common and is hypothesized to provide benefits via reduction in predation pressure (Averbeck et al., 2012). In some cases, human hunting and habitat degradation have even resulted in increased group sizes, which is hypothesized to allow greater vigilance and predator avoidance (Averbeck et al., 2012). It is possible, therefore, that extinction drivers might decrease total abundance of species by reducing the total number of groups, while maintaining the number of individuals per group. If this is the case, the number of intraspecific contacts among individuals may stay relatively constant as species decline. We suggest that the path to extinction in ungulates may follow this trajectory. In threatened ungulate species local densities may remain high due to pressure to maintain minimum group sizes, and thus allow intraspecific contact rates sufficient to support single-host parasites, while reduction in total range size will result in a loss of multi-host parasites. In contrast with ungulates, only 10-15% of carnivores are known to live in social groups (Gittleman, 1989). The natural rarity of carnivores may lead to species being placed on the IUCN Red List solely because of living at critically small population sizes. Additionally, species which have undergone a historical decline but currently have range and population sizes above critical thresholds will not be considered threatened (Mace et al., 2008), which may be the case for many large carnivores (Ripple et al., 2014). Consequently, threat status in carnivores may not be a reliable proxy of recent population decline. It is understandable then that direct measures of population density are

a better predictor of parasite assemblages in carnivore hosts, with those species living at higher densities able to support a higher proportion and richness of single-host parasites.

While these explanations fit the known ecologies of these two host groups, comparative analyses are inherently susceptible to issues of data quality. It is possible that insufficient sample size or mismatched data may lead us to miss important patterns, but it is unclear how these issues would bias our results so as to cause the observed differences between carnivores and ungulates. We have shown that the response of parasite assemblages to host decline depends upon the interaction between intrinsic attributes of hosts and extrinsic drivers of extinction. As changes in both geographic range and population density during host decline impact parasite transmission, gathering additional baseline data on host population densities is therefore essential for predicting coextinction events.

Considering the important roles parasites play in ecosystems and the burdens they cause for wildlife, domesticated animals, and humans, it is vital that we better understand how anthropogenic changes to natural ecosystems alter host-parasite dynamics. The loss of multi-host parasites may have detrimental outcomes to ecosystems, including facilitating disease emergence (Johnson, 2013). Determining how anthropogenic and biological factors interact to alter host-parasite systems can aid in the prediction of disease emergence through future host shifts, or increased prevalence of endemic diseases; an important consideration for the proactive surveillance of emerging pathogens (Farrell et al., 2013). Generating a broader understanding of host-parasite coextinction dynamics will be critical for prioritizing surveillance efforts in the face of shifting species ranges and expansion of human land use practices. Anthropogenic activities not only directly contribute to species loss, but have the potential to bring previously isolated host populations into contact and create novel opportunities for cross-species transmission and exacerbation of existing threats. Our results indicate that there is an urgent need to develop new theories of parasite transmission and loss in declining hosts, but more critically this theory needs to be tested against empirical data which we currently lack.

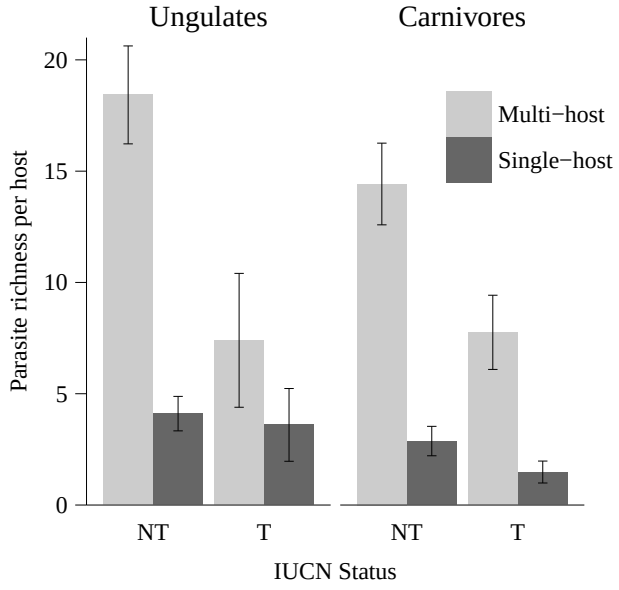


Figure 5–1: Average parasite richness per host for multi-host and single-host parasites, sorted by host group and threat status (NT = Non-threatened; T = threatened). Error bars represent standard errors.

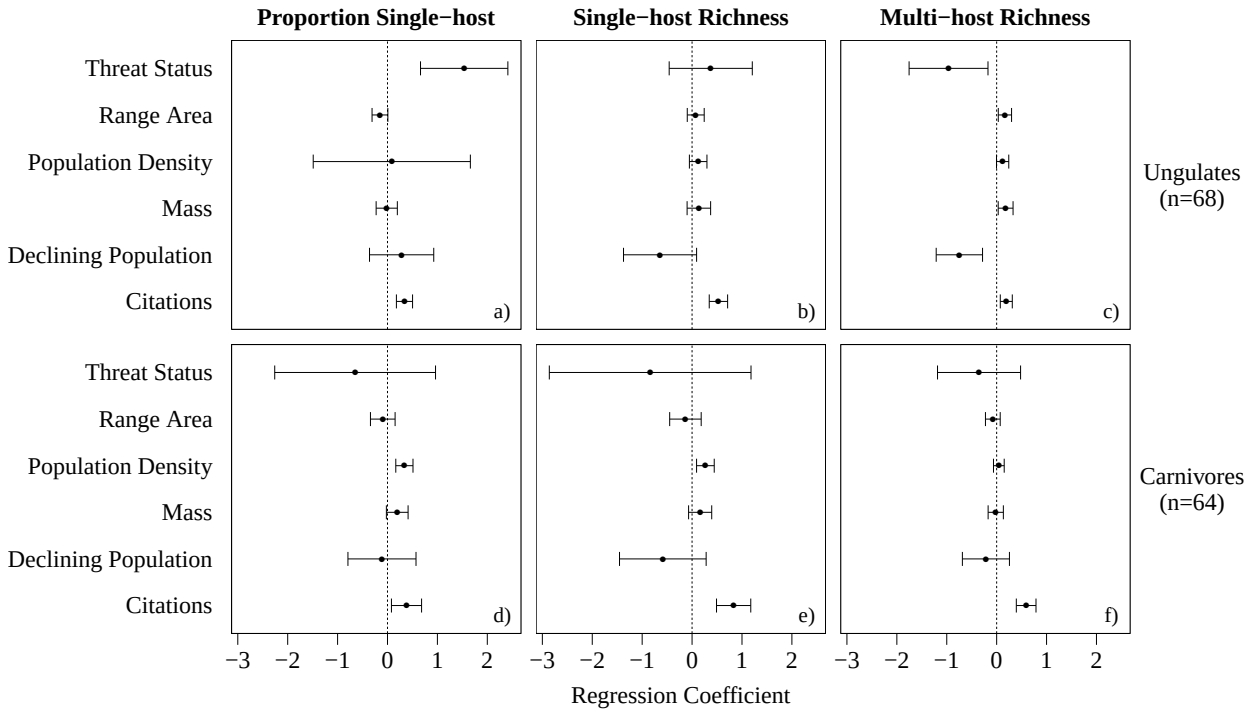


Figure 5–2: Estimated regression coefficients from Generalized Estimating Equations for ungulate (a – c) and carnivore (d – f) hosts. Columns represent response variables. Error bars represent 95% confidence intervals.

General Discussion & Conclusion

Contemporary drivers of ecological change alter disease landscapes, facilitate the spread of parasites to novel hosts and ultimately shape the ecology, evolution, and impacts of infectious diseases around the globe. In this thesis, I presented theory and tools that can help support the surveillance of infectious diseases that infect multiple species. This involved analysis of multiple data sources including global species interaction databases, phylogenetic trees, animal health reports, life history traits, and genetic sequence data, and required the integration of approaches from ecology, evolutionary biology, computational statistics, bioinformatics, and molecular biology. Together I used these approaches to identify susceptible hosts of mammal parasites, describe patterns of bacterial diversity in a mammal diversity hotspot, predict the disease-induced mortality of domesticated animal diseases, and explore how host extinction impacts parasite diversity. Each of these approaches supports proactive disease surveillance by offering ways to study neglected diseases and those that may emerge as threats to public health or conservation in the future.

In Chapter 2, I showed that by predicting undocumented host-parasite interactions we can effectively build on existing global infectious disease databases and prioritize future disease surveillance efforts. Initial efforts to search for evidence of predicted interactions may be done through targeted literature surveys, but will eventually require field based surveillance. Proactive disease surveillance may be most effective when combined with baseline biodiversity surveys in high-risk areas of disease emergence. In Chapter 3, I showed that through the sequencing of DNA present in environmental samples, we have the potential to simultaneously track the distributions of diverse sets of taxa, including currently recognized parasites, and those that may infect novel hosts in the future.

As we continue to build inventories of diseases that can infect both wild and domesticated animals, we may wish to direct research towards those parasites more likely to cause severe disease burdens. In Chapter 4, I found that domesticated animals are more likely

to die from diseases that also infect distantly related hosts. While the exact mechanisms underlying this pattern are unknown, these results indicate that evolutionary relationships among hosts can act as a proxy for disease outcomes, and thus help prioritize research and surveillance efforts towards parasites with greater potential to cause lethal or debilitating infections. The majority of the diseases examined in Chapter 4 have the ability to infect both domestic and wild species, and in some cases are responsible for severe declines in wild populations. The encroachment of humans and domesticated animals into wild habitats facilitates the spread of these diseases and may act in synergy to push wild species towards extinction. This in turn can result in the extinction of parasites that depend on them. In Chapter 5, I found that endangered ungulates are associated with fewer multi-host parasites compared to non-endangered ungulates. While the small population sizes typical of many endangered species may provide some shelter from contact with diseases harboured by domesticated animals, the extinction of co-evolved parasites may create open niches and increase the impact of diseases that are able to reach these rare species. Host extinction may also result in a narrowing of host specificity, or expansion through adaptation to alternate hosts, each having knock-on impacts on the evolution of virulence. To better understand these risks, we must expand our knowledge and understanding of diseases that infect wild species.

The results and methods presented in this thesis narrow some of the knowledge gaps described in the general introduction, and provide approaches that may be put into practice to support current disease surveillance efforts. Through an iterative process of link prediction, literature compilation, and field based surveys (as outlined in Chapters 2 & 3), we can more efficiently build the databases that are necessary to develop broad-scale theories of the ecology and evolution of infectious diseases and for guiding proactive surveillance (such as Chapters 4 & 5). The merger of pure and applied research is also vital for the development of healthy landscape management strategies, fostering sustainable livelihoods, and implementation of cost-effective programs for disease control (Fish et al.,

2011; Lebov et al., 2017). Due to the interdisciplinary nature of this approach and the need to integrate data across multiple spatial, temporal, and taxonomic scales, the task will require collaboration from multiple domains. Through the dissolution of disciplinary boundaries we have the ability to generate a holistic understanding of the mechanisms underlying disease emergence, and ultimately prevent excessive burdens of disease around the globe.

Using the most comprehensive host-parasite interaction databases available for mammals, I showed that there are still missing interactions that have some published documentation in the literature (Chapter 2). The link prediction model I implement (Elmasri et al., 2017) demonstrates how our knowledge of host-parasite interactions could be efficiently expanded by targeted literature searches. However, I identified a number of interactions that were highly likely, but for which I could not find any published evidence of infection. Some are due to geographic or ecological mismatch, indicating that contemporary infections are unlikely. Extensions of the model could include geographic or trait dissimilarities among hosts – or parasites – in addition to phylogenetic relatedness. In addition to these predicted but ecologically unlikely links, I also identified links among hosts and parasites with overlapping ranges and ecologies that would facilitate infection. For these links, it may be possible to generate a "most wanted list" which could be disseminated to veterinary and public health offices around the world, potentially in conjunction with existing platforms such as the Program for Monitoring Emerging Diseases (ProMED-Mail – an internet-based reporting system for infectious disease outbreaks) (Morse et al., 2012), or more recently developed approaches for digital-based emerging disease surveillance (Olson et al., 2015). Considering that evidence of many of the top links in Chapter 2 were found in older articles or more obscure journals, there likely exists a wealth of information collected by conservation practitioners and national veterinary laboratories that is well known locally, but does not make it into peer reviewed academic literature. In addition, there may be opportunities to target the investigation of particular diseases in coordination with local

organizations or ongoing efforts such as the USAID PREDICT program, which aims to build global capacity for infectious disease surveillance across species at the animal-human interface (Morse et al., 2012). Obtained samples could be sent to centralized facilities for expert taxonomic or molecular identification (Fonjungo et al., 2017), and integrated into a global real-time genomics-based surveillance system (Gardy and Loman, 2018).

One approach to develop this program would be to initiate DNA-based monitoring in collaboration with disease ecologists and veterinarians in conservation areas. The field work conducted in the Kruger National Park as part of my third chapter is one example of a highly collaborative effort that brought together genomics experts, veterinarians, and ecologists to develop a biodiversity monitoring program. This marked the first exploration of bacterial diversity in watering holes throughout the park, and is one of the few metagenomic studies of freshwater systems in Africa. We showed that it is possible to sequence DNA present in watering holes, however the majority of sequences could not be identified to the species level. This is common for microbial DNA in environmental samples (Solden et al., 2016), but also illustrates the potential for environmental DNA to be used for both biodiversity censuses and disease surveillance when target organisms might not be known. These sequences could be merged with efforts to create a real-time genomics-based global disease surveillance system (Gardy and Loman, 2018). A useful extension of this framework would be to sequence additional genetic markers to identify vertebrates and known disease vectors from the same samples. Through repeated spatial and temporal sampling it may be able to use the method of Elmasri et al. (2017) or other hierarchical network models (e.g. Ovaskainen et al. (2017)) to build more complete species association networks and infer potential host-parasite interactions.

Filling gaps in our current knowledge of host specificity provides critical primary data for understanding and modelling multi-host multi-parasite disease dynamics (Rigaud et al., 2010; Buhnerkempe et al., 2015). In some cases this information may be directly useful for disease prevention, such as the identification of novel reservoir hosts for neglected diseases,

but these data also play a fundamental role in generating and testing hypotheses about the ecology and evolution of infectious diseases (Stephens et al., 2016). Chapters 4 and 5 used published data to test ecological theories and identify broad scale patterns across multiple host-parasite systems. These studies contribute novel insights to the existing literature on parasite virulence and coextinction, however they also lead to new questions. Some questions may be answered with existing data, while others may require the building of novel databases. For example, for the majority of host-parasite interactions we have some evidence of exposure or infection that results in apparent clinical symptoms, but often we lack more specific information about the nature of the infection (Bisson et al., 2015). In many cases we do not know if a host is necessary for successful transmission, or is merely an accidental or spillover host that plays little role in the epidemiology of the disease (Buhnerkempe et al., 2015). Similarly, we often do not have data on the specific site of infection, induced pathologies, or how host physiology and behaviour may be modified by infection (Stephens et al., 2016). Undoubtedly these not only play a role in the cross-species transmission of diseases, but may also constrain or facilitate the evolution of parasite host specificity. Finally, we are often missing phylogenies and even basic trait data for the vast majority of parasitic organisms (Stephens et al., 2017). By gathering these data we may be able to adopt more trait-based approaches to the study of host-parasite interactions, and explore co-phylogenetic patterns to shed light on the evolution of complex disease systems (Sweet et al., 2018).

Finally, I emphasize the importance of research that builds ecological theory in an applied context. While recent disease emergence events have provided new insights into the ecology of infectious diseases, it is important to remember that full burden of infectious diseases is unknown (Halliday et al., 2012). This is largely due to a lack of infrastructure for diagnosing and reporting of diseases in the world's lowest resource areas (Perry and Grace, 2009; Chan et al., 2010; Halliday et al., 2017). In these areas, human population densities and demand for livestock are expanding at unprecedented rates (Perry et al.,

2011). This enormous increase in livestock production will inevitably increase contact rates among livestock and wildlife as agricultural lands expand into previously undeveloped territories (Perry et al., 2011; Jones et al., 2013). Increasing contact among previously isolated populations can contribute to disease emergence through parasites switching to novel hosts (Charleston and Robertson, 2002; Parrish et al., 2008; Faust et al., 2018). A holistic view of health that incorporates humans, animals, and ecosystems has come to the forefront of public health initiatives over the last twenty years (Lebov et al., 2017). One overarching theme that is integral to the study of emerging infections disease is the need for interdisciplinary research. I have been lucky to develop projects with fantastic collaborators, and by moving between disciplines I have been able to learn a little about a great diversity of topics. Interdisciplinary work is central to innovation and necessary for tackling complex problems. None of the research presented in this thesis would have been possible without the support of experts willing to learn new vocabularies and explain their work to those outside of their disciplines.

CHAPTER 6
Appendix

6.1 Chapter 2 Supplementary Data and Results

Top predicted links and literature search results

Affinity only – Full Dataset

Host	Parasite
Homo sapiens	Taenia mustelae
Homo sapiens	Bluetongue virus
Homo sapiens	Bovine viral diarrhea virus 1
Homo sapiens	Neospora caninum
Homo sapiens	Mastophorus muris
Homo sapiens	Plagiorchis vespertilionis
Homo sapiens	Alaria alata
Homo sapiens	Carnivore protoparvovirus 1
Homo sapiens	Taenia pisiformis
Homo sapiens	Physocephalus sexalatus

Table 6–1: Top 10 undocumented links with highest probablity of interaction (Affinity only model - full data set)

- A recent molecular phylogeny of the *Taenia* genus supported the creation of a new genus and renaming of *Taenia mustelae* to *Versteria mustelae* (Nakao et al., 2013). Since then there has been a report of fatal infection of a previously unknown *Versteria* species in a captive orangutan *Pongo pygmaeus* cloesly related to species found in wild mustelids, suggesting the need for increased vigilance of *Versteria* infections in humans (Lee et al., 2016).
- While bluetongue virus is known to infect a wide range of ruminants, it is not currently considered to infect humans (Spickler, 2015).
- Bovine viral diarrhea viruses are not considered to be human pathogens, but there is some concern about zoonotic potential as they are highly mutable, have the ability to replicate in human cell lines, and have been isolated from humans on rare occasions (Walz et al., 2010).
- Although antibodies to *Neospora caninum* have been reported in humans, the parasite has not been identified in human tissues and the zoonotic potential is not known (Dubey et al., 2007).

- While natural infections of *Plagiorchis* species in humans are rare, the first case of human infection by the bat parasite *Plagiorchis vespertilionis* was reported in 2007 (Guk et al., 2007). The source of infection is uncertain, it has been suggested that freshwater fish and snails may be undocumented intermediate hosts and infection was due to ingestion of raw freshwater fish.
- *Mastophorus muris* is a rodent-specific nematode that requires arthropods as intermediate hosts and while this makes it unlikely to infect humans, it was recently documented in an urban population of rats in the UK (Mcgarry et al., 2015), indicating the potential for human exposure.
- *Alaria alata*, an intestinal parasite of wild canids, has not been identified in humans, but is considered a potential zoonotic risk as other *Alaria* species have been reported to cause fatal illness in humans (Murphy et al., 2012).
- Carnivore protoparvovirus can infect a number of hosts in the order *Carnivora* (Balboni et al., 2018), though there seems to be no evidence of human infection.
- Due to the characteristics of the biological cycle of *Tenia pisiformis* and the observation that it is innocuous in humans, this parasite has been used as a model for the study of other important zoonosis relevant to human health including *T. solium* (Betancourt-Alonso et al., 2011).
- The definitive hosts of *Physocephalus sexalatus* are commonly wild and domestic pigs, but it is sometimes found in other mammals and some reptiles (McAllister et al., 2004). However the parasite uses beetles as an intermediate host, which makes human infection unlikely.

Affinity only – Full Dataset – Domestic Hosts

Host	Parasite
Bos taurus	Trypanosoma cruzi
Bos taurus	Hymenolepis diminuta
Rattus norvegicus	Rabies lyssavirus
Bos taurus	Mesocestoides lineatus
Bos taurus	Capillaria hepatica
Bos taurus	St louis encephalitis virus
Bos taurus	Canine distemper virus
Bos taurus	Anisakis simplex
Bos taurus	Taenia mustelae
Ovis aries	Trypanosoma cruzi

Table 6–2: Top 10 undocumented links with highest probablity of interaction for domesticated species (Affinity only model - full data set)

- Currently the role of cattle in the epidemiology of Chagas disease (caused by *Trypanosoma cruzi*) is unknown, though the majority of cattle in Latin America may be exposed (280 million heads; 1/4 of the world population) (Giangaspero, 2017). Browne et al. (2017) report that 177 species have been documented as susceptible to infection by *T. cruzi*, with domestic hosts in some cases being responsible for the maintenance of local parasite populations over long periods of time. While cattle have tested positive in serological studies, cows and other domestic species are also infected by *Trypanosoma* and *Phytomonas* species which can cause cross-reactions in diagnostic tests (Gürtler and Cardinal, 2015).
- The natural hosts of *Hymenolepis diminuta* are rats and cattle have not been found to be susceptible to infection, however *H. diminuta* eggs have been found in the feces of dairy cattle, likely the result of ingesting forage contaminated with rodent feces (Huang et al., 2014).
- Rabies has an extremely large host range and surprisingly rats are rarely reported as suffering from rabies, though there have been a few reported cases rabid *Rattus norvegicus* in the United States (Fitzpatrick et al., 2014).

- *Mesocestoides lineatus* has a three-stage lifecycle with two intermediate hosts and a large range of carnivorous mammals as definitive hosts (Cho et al., 2013). Human infections of the tapeworm *Mesocestoides lineatus* are rare but can occur through the consumption of chickens, snails, snakes, or frogs (Ito and Budke, 2014) and therefore it is unlikely that cows will consume to the intermediate life stages of this parasite.
- *Capillaria hepatica* (syn. *Calodium hepaticum*), is a globally distributed zoonotic parasite which uses rodents as main hosts, but is known to cause infection in over 180 mammalian species, including cattle (Fuehrer, 2014).
- Birds are the primary vertebrate hosts for St. Louis encephalitis virus, though amplification by certain mammals has been suggested (Kopp et al., 2013). There is some serological evidence of infection in domestic mammals, including cattle (Diaz et al., 2006). The common vector *Culex nigripalpus* feeds primarily on birds, but shows a seasonal shift from avian hosts in the spring to mammalian hosts in the summer, indicating it may be able to act as a bridging vector among different host species (Kopp et al., 2013).
- Canine distemper virus infects a wide range of hosts within the order Carnivora, but has also been found to cause fatal infection in some non-human primates and peccaries (Beineke et al., 2015).
- *Anisakis simplex* uses cetaceans as final hosts, with marine invertebrates and fish as intermediate hosts (Buchmann and Mehrdana, 2016). Whales are infected by ingesting that cattle may be susceptible, though they would need sufficient exposure to marine based feed.
- Ovis aries is reported to be infected by *Trypanosoma cruzi* (Browne et al., 2017).

Affinity only – Full Dataset – Wild Hosts

Host	Parasite
Cervus elaphus	Rabies lyssavirus
Ondatra zibethicus	Rabies lyssavirus
Capreolus capreolus	Rabies lyssavirus
Sorex araneus	Rabies lyssavirus
Apodemus sylvaticus	Rabies lyssavirus
Odocoileus virginianus	Rabies lyssavirus
Pan troglodytes	Rabies lyssavirus
Syncerus caffer	Rabies lyssavirus
Dama dama	Rabies lyssavirus
Microtus arvalis	Rabies lyssavirus

Table 6–3: Top 10 undocumented links with highest probablity of interaction for wild host species (Affinity only model - full data set)

- For the affinity only model run on the full dataset, all of the top ten interactions involve rabies, which is unsurprising as rabies is commonly said to be capable of infecting all mammal species (Mollentze et al., 2014) and is the parasite with the largest number of documented hosts in the database (176). Of these ten most likely hosts, we found evidence of rabies infection in eight (*Cervus elaphus* and *Odocoileus virginianus* (Birhane et al., 2017), *Ondatra zibethicus* (World Health Organization, 1983), *Capreolus capreolus* (Sempere et al., 1996), *Pan troglodytes* (Gautret et al., 2014), *Apodemus sylvaticus* (Steck and Wandeler, 1980), *Dama dama* (Zhu et al., 2015), and successful experimental infection of *Microtus arvalis* (Schindler, 1957)).

Phylogeny only – Full Dataset

Host	Parasite
Vulpes ferrilata	Rabies lyssavirus
Vulpes rueppellii	Rabies lyssavirus
Vulpes macrotis	Rabies lyssavirus
Lycalopex culpaeus	Rabies lyssavirus
Lycalopex fulvipes	Rabies lyssavirus
Lycalopex griseus	Rabies lyssavirus
Lycalopex gymnocercus	Rabies lyssavirus
Cuon alpinus	Rabies lyssavirus
Speothos venaticus	Rabies lyssavirus
Holochilus chacarius	Schistosoma mansoni

Table 6–4: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - full data set)

- Domestic dogs are considered a major predation threat to the Tibetan fox (*Vulpes ferrilata*) (Wang et al., 2007), and rabies is confirmed to circulate in wild and domestic animals in Tibet (Tao et al., 2015).
- We did not find any record of rabies infecting *Vulpes rueppellii*, but this should be investigated as this disease is known to cause severe declines in wild canids (Fleming et al., 2017).
- Rabies has been documented to infect the endangered San Joaquin kit fox (*Vulpes macrotis mutica*) and is suggested to have caused a catastrophic decline of the species in the 1990s (White et al., 2000).
- Domestic dogs alter the ecology of Andean foxes (*Lycalopex culpaeus*), have been observed hunting them, and are at risk of disease transmission from dogs (Zapata-Ríos and Branch, 2016).
- Diseases from domestic dogs (largely canine distemper) is considered a major threat to the endangered *Lycalopex fulvipes* (Silva-Rodríguez et al., 2016), indicating that rabies may also pose a risk.

- There is one report of serological evidence of rabies in *Lycalopex griseus* in Chile in 1989 (Juan, 1989) (reported as *Pseudalopex griseus*, a formerly accepted name (Wilson and Reeder, 2005)).
- We did not find evidence of rabies infection in *Lycalopex gymnocercus*, however its distribution overlaps with species known to be important in the transmission of rabies in Brazil (Carnieli et al., 2008).
- The endangered Dhole (*Cuon alpinus*) is known to suffer from rabies and was a source of fatal human infections during an outbreak in the 1940s (Durbin et al., 2005).
- Rabies has been reported as potentially infecting *Speothos venaticus* (DeMatteo, 2008) and there is a report of an individual with positive serology (Jorge et al., 2010).
- We could not find evidence of *Schistosoma mansoni* infection in *Holochilus chacarius*. Congener *Holochilus braziliensis* was experimentally shown to be a viable host, although infection resulted in host death (Borda and Rea, 2006).

Phylogeny only – Full Dataset – Domestic Hosts

Host	Parasite
Bison bison	Rabies lyssavirus
Bos grunniens	Rabies lyssavirus
Bos frontalis	Rabies lyssavirus
Bos javanicus	Rabies lyssavirus
Vicugna vicugna	Rabies lyssavirus
Rattus rattus	Rabies lyssavirus
Rattus norvegicus	Rabies lyssavirus
Cavia porcellus	Rabies lyssavirus
Oryctolagus cuniculus	Rabies lyssavirus
Bos grunniens	Toxoplasma gondii

Table 6–5: Top 10 undocumented links with highest probablity of interaction for domesticated species (Phylogeny only model - full data set)

- Rabies in *Bison bison* is considered rare, but there are multiple cases reported (Stoltenow et al., 2000).
- Rabies has been reported to infect yak (*Bos grunniens*) in Nepal (Joshi, 1982).
- We could not find evidence of rabies infection in *Bos frontalis*, but as this is a semi-wild and endangered species (Mei et al., 2016) and other *Bos* species are susceptible, the disease may pose a conservation risk. This may also be the case for the endangered *Bos javanicus*.
- We did not find a specific report of rabies infection in *Vicugna vicugna*, although all South American camelids are noted to be susceptible and display clinical signs of infection (Fowler, 1996).
- There are documented cases of rabies infecting *Rattus rattus* (ex. Carey and Mclean (1983)), though it appears to be rare.
- Rabies in *Rattus norvegicus* was predicted by the affinity only model with the full dataset for domestic hosts.

- Pet guinea pigs *Cavia porcellus* have been infected with rabies after being bitten by a raccoon (Eidson et al., 2005).
- There are reported cases of rabid *Oryctolagus cuniculus* in the United States (Fitzpatrick et al., 2014).
- *Toxoplasma gondii* is known to infect *Bos grunniens* and cause severe economic losses (Li et al., 2014).

Phylogeny only – Full Dataset – Wild Hosts

Host	Parasite
Vulpes ferrilata	Rabies lyssavirus
Vulpes rueppellii	Rabies lyssavirus
Vulpes macrotis	Rabies lyssavirus
Lycalopex culpaeus	Rabies lyssavirus
Lycalopex fulvipes	Rabies lyssavirus
Lycalopex griseus	Rabies lyssavirus
Lycalopex gymnocercus	Rabies lyssavirus
Cuon alpinus	Rabies lyssavirus
Speothos venaticus	Rabies lyssavirus
Holochilus chacarius	Schistosoma mansoni

Table 6–6: Top 10 undocumented links with highest probablity of interaction for wild host species (Phylogeny only model - full data set)

- These predictions are the same as for the phylogeny only model on the full dataset (all wild host species in the top 10 links).

Full model – Full Dataset

Host	Parasite
Homo sapiens	Taenia mustelae
Homo sapiens	Bluetongue virus
Homo sapiens	Bovine viral diarrhea virus 1
Homo sapiens	Neospora caninum
Homo sapiens	Mastophorus muris
Homo sapiens	Simian immunodeficiency virus
Homo sapiens	Plagiorchis vespertilionis
Homo sapiens	Alaria alata
Bos taurus	Trypanosoma cruzi
Homo sapiens	Carnivore protoparvovirus 1

Table 6–7: Top 10 undocumented links with highest probablity of interaction (Full model - full data set)

- All of the top links were previously discussed except for human infection with simian immunodeficiency virus (SIV). Notably SIV strains of wild primates have infected humans and are responsible for the AIDS pandemic (from HIV-1) (Van Heuverswyn et al., 2006). While HIV and SIV are now considered different species, SIV is clearly a risk for human populations.

Full model – Full Dataset – Domestic Hosts

Host	Parasite
Bos taurus	Trypanosoma cruzi
Rattus norvegicus	Rabies lyssavirus
Bos taurus	Hymenolepis diminuta
Bos taurus	Capillaria hepatica
Bos taurus	Anisakis simplex
Bos taurus	Mesocestoides lineatus
Bos taurus	St louis encephalitis virus
Bos taurus	Canine distemper virus
Bos taurus	Taenia mustelae
Ovis aries	Trypanosoma cruzi

Table 6–8: Top 10 undocumented links with highest probablity of interaction for domesticated species (Full model - full data set)

- These links have already been predicted in models discussed above.

Full model – Full Dataset – Wild Hosts

Host	Parasite
Cervus elaphus	Rabies lyssavirus
Capreolus capreolus	Rabies lyssavirus
Sorex araneus	Rabies lyssavirus
Ondatra zibethicus	Rabies lyssavirus
Apodemus sylvaticus	Rabies lyssavirus
Odocoileus virginianus	Rabies lyssavirus
Dama dama	Rabies lyssavirus
Apodemus agrarius	Rabies lyssavirus
Syncerus caffer	Rabies lyssavirus
Microtus arvalis	Rabies lyssavirus

Table 6–9: Top 10 undocumented links with highest probablity of interaction for wild host species (Full model - full data set)

- All links were predicted by models discussed above except for rabies infection in *Apodemus agrarius*, which was documented in China in 2009 (Wang et al., 2014).

Arthropods – Phylogeny only

Host	Parasite
Canis adustus	Sarcoptes scabiei
Lycalopex fulvipes	Sarcoptes scabiei
Lycalopex vetulus	Sarcoptes scabiei
Cerdocyon thous	Sarcoptes scabiei
Chrysocyon brachyurus	Sarcoptes scabiei
Cuon alpinus	Sarcoptes scabiei
Speothos venaticus	Sarcoptes scabiei
Vulpes lagopus	Pulex irritans
Ovis ammon	Sarcoptes scabiei
Vulpes lagopus	Sarcoptes scabiei

Table 6–10: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Arthropod subset)

- Bornstein et al. (2002) compiled a list of documented host species of sarcoptic mange (caused by *Sarcoptes scabiei*) which includes *Cerdocyon thous* and *Vulpes lagopus* (reported as *Alopex lagopus*, a common synonym).
- There does not appear to be a published record of sarcoptic mange in *Canis adustus*, however in areas with sympatric jackal species *C. adustus* usually display ecological segregation through preferring denser vegetation (Loveridge and Macdonald, 2003). This may indicate that while *C. adustus* may be susceptible to sarcoptic mange, differences in the ecologies of this species relative to other canids may limit transmission making overt infections difficult to document.
- *Lycalopex fulvipes* is endangered (Silva-Rodríguez et al., 2016), meaning that its small population sizes and restricted geographic range may reduce exposure to *S. scabiei*, however as sarcoptic mange is implicated in the declines of other wild canids, it should be targeted in disease monitoring programs for this species.
- While *Lycalopex vetulus* is not endangered, it displays some adaptability to anthropogenic disturbance (Dalponte and Courtenay, 2008), which may expose it to sarcoptic mange through

contact with domestic dogs. In addition, *Lycalopex vetulus* is sympatric with the crab eating fox (*Cerdocyon thous*) – a documented host of *S. scabiei* (Bornstein et al., 2002). The IUCN reports a gap in conservation actions for *L. vetulus* regarding the role of disease in population regulation, and their status as reservoirs of scabies, canine distemper, leishmaniasis, and rabies (Dalponte and Courtenay, 2008).

- *Chrysocyon branchyurus* has one report of clinical signs suggestive of sarcoptic mange-like infestation (Luque et al., 2014).
- The endangered Dhole (*Cuon alpinus*) has been documented as suffering from mange as early as 1937 (Durbin et al., 2005) and appear to be especially susceptible to disease outbreaks due to their large group sizes and amicable behaviour within packs.
- *S. scabei* was identified in *Speothos venaticus* (Jorge et al., 2008), and identified as potentially contributing to the loss of individuals from a group in Mato Grosso, Brazil (de Souza Lima et al., 2012).
- While the Arctic fox (*Vulpes lagopus*) is considered the most important terrestrial game species in the Arctic (Angerbjörn and Tannerfeldt, 2014), I cannot find documented infection by the "human flea" (*Pulex irritans*). *P. irritans* is thought to be unable to persist in Arctic envrionments due to the lack of artificial warmth necessary for breeding (Buckland and Sadler, 1989), though this may change in the future with continued Arctic warming.
- I cannot find a record of mange in *Ovis ammon*, though outbreaks of sarcoptic mange have been documented in ibex and blue sheep in the Taxkorgan Reserve, China, in which *O. ammon* are also present, although this population has received little study (Schaller and Kang, 2008).

Arthropods – Full model

Host	Parasite
Cervus elaphus	Rhipicephalus evertsi
Taurotragus oryx	Sarcoptes scabiei
Vulpes vulpes	Rhipicephalus evertsi
Sus scrofa	Rhipicephalus evertsi
Cervus nippon	Sarcoptes scabiei
Odocoileus virginianus	Sarcoptes scabiei
Vulpes vulpes	Rhipicephalus appendiculatus
Sus scrofa	Rhipicephalus appendiculatus
Cervus elaphus	Rhipicephalus appendiculatus
Cerdocyon thous	Sarcoptes scabiei

Table 6–11: Top 10 undocumented links with highest probablity of interaction (Full model - Arthropod subset)

- *Rhipicephalus evertsi* and *R. appendiculatus* are common ticks in East and Southern Africa (Jongejan and Uilenberg, 1994) and are unlikely to interact with non-African hosts such as *Cervus elaphus* and *Vulpes vulpes* (though there are some populations of *Vulpes vulpes* in North Africa). Future iterations of the link prediction model by Elmasri et al. (2017) may benefit from the inclusion of information on geographic range overlap among host species, however this may reduce the ability to identify future host-parasite associations that may occur given range expansions or species translocations.
- *R. appendiculatus* was found to be the most prevalent tick species on domestic pigs in the Busia District of Kenya (Kagira et al., 2013), indicating that increased monitoring may also identify *R. evertsi* on domestic pigs.
- *Sarcoptes scabiei* has been documented in captive *Taurotragus oryx* in Israel (Bornstein et al., 2002).
- *Sarcoptes scabiei* has been documented to infect *Cervus nippon* in Japan (Chen et al., 2012).

- While multiple cervids have been reported with sarcoptic mange (Bornstein et al., 2002), we cannot find any report of infection in white-tailed deer *Odocoileus virginianus*, although there are numerous reports of infection with mange caused by *Demodex sp.*, including the host specific *Demodex odocoilei* (Nemeth et al., 2014), potentially indicating competition among *Sarcoptes* and *Demodex* species.
- *Cerdocyon thous* found to have sarcoptic mange (Bornstein et al., 2002), which was also in the top ten predictions made by the phylogeny only model for the arthropod subset.

Bacteria – Phylogeny only model

Host	Parasite
Canis aureus	Leptospira interrogans
Canis mesomelas	Leptospira interrogans
Canis mesomelas	Anaplasma phagocytophilum
Lycaon pictus	Leptospira interrogans
Saguinus geoffroyi	Escherichia coli
Equus burchellii	Escherichia coli
Equus zebra	Escherichia coli
Vulpes lagopus	Leptospira interrogans
Vulpes velox	Leptospira interrogans
Equus zebra	Leptospira interrogans

Table 6–12: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Bacteria subset)

- *Leptospira sp.* are commonly regarded as infecting a wide range of mammals (Siembieda et al., 2011). van der Hoeden (1955) identified *Leptospira interrogans* serovar *canicola* in the urine of jackals in Israel though did not identify the particular species. However, *Canis aureus* is the only jackal species present in the country (Dayan et al., 1992) providing some support for this host-parasite association, though the findings of van der Hoeden (1955) should be verified.
- We did not find any documentation of leptospirosis in *Canis mesomelas* or *Lycaon pictus* but considering it is found in multiple species in Africa including domestic dogs (Allan et al., 2015), wild canids are likely to be exposed.
- Penzhorn et al. (2018) sequenced DNA from blood samples of *Canis mesomelas* in South Africa and identified 16S rDNA sequences very similar to *Anaplasma phagocytophilum*. This study also identified other *Anaplasma* species indicating the potential for 16S rDNA sequencing to gather evidence of predicted host-parasite and discover previously unknown pathogens.

- *E. coli* is ubiquitous commensal microbe of vertebrates (Tenaillon et al., 2010) and pathogenicitiy is linked to particular strains, indicating that our approach may be expanded by identifying the host ranges of particular subspecies or virulent strains of common commensal bacteria.
- We did not find any evidence of *Leptospira* infections in *Vulpes lagopus*. We did find one report of positive serology for *Leptospira interrogans* in *Vulpes velox macrotis* (Standley and McCue, Standley and McCue), though there is debate as to whether this subspecies is actually its own species *Vulpes macrotis* (Wilson and Reeder, 2005).
- We did not find evidence of *Leptospira interrogans* in *Equus Zebra* although one study reported a low prevalence of *Leptospira* antibodies in *Equus burchellii* in Zimbabwe (Anderson and Rowe, 1998).

Bacteria – Full model

Host	Parasite
Homo sapiens	Bartonella grahamii
Homo sapiens	Anaplasma bovis
Homo sapiens	Anaplasma marginale
Homo sapiens	Mycoplasma haemofelis
Homo sapiens	Mycoplasma mycoides
Canis lupus	Yersinia enterocolitica
Homo sapiens	Chlamydophila pecorum
Ovis aries	Mycobacterium bovis
Homo sapiens	Lawsonia intracellularis
Homo sapiens	Mycoplasma conjunctivae

Table 6–13: Top 10 undocumented links with highest probablity of interaction (Full model - Bacteria subset)

- *Bartonella grahamii* is a pathogen of rodents worldwide, but was first identified as causing an infection in an immunocompromised human in 2013 (Oksi et al., 2013).
- *Anaplasma bovis*, causal agent of bovine anaplasmosis, is not currently considered zoonotic (Rar and Golovljova, 2011), but *Anaplasma phagocytophilum* the causative agent of human anaplasmosis placed is as sister taxa to *A. bovis* in a recent phylogeny (Yang et al., 2017). Similarly, *A. marginale*, the causative agent of anaplasmosis in cattle, is also not considered zoonotic, but it reaches high prevalence in cattle and humans are likely exposed to the tick vector (Rar and Golovljova, 2011).
- There has been one documented case of infection in an immunocompromised human with a *Mycoplasma haemofelis*-like bacteria (dos Santos et al., 2008) and the authors note that disease-causing latent mycoplasma infections in immunocompromised and non-immunocompromised patients are an emerging issue.

- While there is some debate whether sheep are relatively immune or highly susceptible to infection by *Mycobacterium bovis*, spillover infections can occur when animals are exposed to contaminated pasture (Cousins, 2001).
- *Mycoplasma mycoides* is not typically thought to infect humans, but there is one report of disease and positive serology in a farm worker exposed to multiple calves infected with *M. mycoides subsp. mycoides LC* (Gonçalves, 2007).
- Pathogenic strains of *Yersinia enterocolitica*, the main cause of yersiniosis in Europe and one of the five main bacterial gastrointestinal diseases of humans, were found in the feces of dogs from several European countries (Stamm et al., 2013).
- The zoonotic potential of *Chlamydophila pecorum* is not known, although it is associated with abortions in small ruminants and related *Chlamydophila psittaci* is a known zoonotic disease from birds (Barati et al., 2017).
- *Mycobacterium bovis* is known to cause infection in goats as a result of contact with infected cattle, although the importance of the disease varies across countries and production systems (Cousins, 2001).
- *Lawsonia intracellularis* was recently recognised as the cause of an emerging intestinal disease in horses (Equine proliferative enteropathy), but is currently not considered to be zoonotic (Pusterla and Gebhart, 2009).
- *Mycoplasma conjunctivae* causes a highly contagious ocular infection of sheep, goats, and wild Caprinae, and is possibly zoonotic as it has been associated with eye inflammation in young children (Lysnyansky et al., 2007).

Fungi – Phylogeny only

Host	Parasite
Macaca sylvanus	Pneumocystis carinii
Felis silvestris	Pneumocystis carinii
Ateles paniscus	Pneumocystis carinii
Mustela lutreola	Pneumocystis carinii
Pan troglodytes	Pneumocystis carinii
Gorilla beringei	Pneumocystis carinii
Mustela erminea	Pneumocystis carinii
Ovis canadensis	Pneumocystis carinii
Nyctereutes procyonoides	Pneumocystis carinii
Rattus rattus	Pneumocystis carinii

Table 6–14: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Fungi subset)

- *Pneumocystis carinii* belongs to a genus that normally reside in the pulmonary parenchyma of a wide range of mammals (Danesi et al., 2016). It is capable of causing life threatening pneumonia in immunocompromised hosts is documented as causing natural infections in *Rattus rattus* (Palmer et al., 2000) as well as species in the genera *Macaca* and *Mustela* (Laakkonen, 1998).

Fungi – Full model

Host	Parasite
Homo sapiens	Geomyces destructans
Homo sapiens	Chrysosporium parvum
Homo sapiens	Neocallimastix frontalis
Bos taurus	Pneumocystis carinii
Homo sapiens	Pilobolus kleinii
Sus scrofa	Pneumocystis carinii
Phascolarctos cinereus	Pneumocystis carinii
Homo sapiens	Chaetomidium arxii
Homo sapiens	Trichophyton terrestre
Homo sapiens	Entomophthora coronata

Table 6–15: Top 10 undocumented links with highest probablity of interaction (Full model - Fungi subset)

- *Geomyces destructans* is the cause of white nose syndrome in multiple bat species (Warnecke et al., 2012), but is not documented to be zoonotic.
- *Chrysosporium parvum* and related species are soil fungi that cause pulmonary infections in rodents, fossorial mammals, their predators, and occasionally humans, though the taxonomy of these pathogens is muddied in the literature (Anstead et al., 2012).
- *Neocallimastix frontalis* appears to be a commensal fungi of bovid rumens (Gleason and Marano, 2011) and we cannot find documentation of zoonotic infection.
- *Pneumocystis carinii* has been documented to infect cattle and pigs (Settnes and Henriksen, 1989).
- *Pilobolus kleinii* play a role in the decomposition of herbivore dung and although they are non-pathogenic to herbivores, they can facilitate the spread of attached parasitic lungworms because of their projectile dispersal system (Aluoch et al., 2017).
- We cannot find any documentation of *Chaetomidium arxii* infection in humans, although this genus is well known for its opportunistic animal and human pathogens (Ma et al., 2018).

- *Trichophpyton terrestre* is part of a large species complex with some variants documented to cause human infection (Campbell et al., 2006).
- *Entomophthora coronata* is primarily a parasite of insects, but can cause sinus infections in humans and was first reported in 1965 (Su et al., 1997).

Helminths – Phylogeny only

Host	Parasite
Holochilus chacarius	Schistosoma mansoni
Onychogalea unguifera	Echinococcus granulosus
Canis adustus	Echinococcus granulosus
Gazella leptoceros	Nematodirus spathiger
Kobus vardonii	Cotylophoron cotylophorum
Onychogalea unguifera	Rugopharynx australis
Lycalopex vetulus	Echinococcus granulosus
Canis mesomelas	Trichinella spiralis
Kobus vardonii	Paramphistomum cervi
Gazella leptoceros	Trichostrongylus vitrinus

Table 6–16: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Helminth subset)

- *Schistosoma mansoni* infection in *Holochilus chacarius* was predicted by the phylogeny only model in the full dataset (discussed above).
- *Echinococcus granulosus* has not been reported to infect northern nail-tail wallabies (*Onychogalea unguifera*), other wallaby species including endangered bridled nail-tailed wallaby (*Onychogaela fraenata*) are involved in the transmission the parasite in Australia (Jenkins and Macpherson, 2003). *Onychogaela unguifera* may be also be involved in *Echinococcosus* transmission, but its parasites may not be as well studied compared to the bridled nail-tail wallaby due to its stable conservation status.
- Similarly, *Rugopharynx australis* is known to infect multiple wallaby species, however the diversity of *Rugopharynx* and their susceptible hosts is still being discovered (Chilton et al., 2016), suggesting that *Onychogalea ungifera* may be a promising target for future study.
- While there does not appear to be evidence of *Echinococcosus granulosus* infection in *Canis adustus*, other *Canis* species in Africa are known hosts (Otero-Abad and Torgerson, 2013), indicating that this should be a target for future surveillance.

- A recent molecular survey of gastrointestinal parasites of wild ruminants in Tunisia identified *Nematodirus spathiger* in engandered *Gazella leptoceros* that were genetically identical from those found in other domestic and wild ruminants (Said et al., 2018). This is an example of a successful exploratory study aimed at describing the diversity of parasites in threatened species.
- We did not find much information on the parasites of the near threatened *Kobus vardonii*, however it is known to inhabit floodplains and grasslands near permanent water in south-central Africa (IUCN SSC Antelope Specialist Group, 2016) where is likely to be exposed to *Cotylophoron cotylophoron*, a "rumen fluke" which emerge from snails intermediate hosts and encyst on vegetation, later being ingested by ruminant definitive hosts in East Africa (Laidemitt et al., 2017). Similarly, the related stomach fluke *Paramphistomum cervi* has been found in Kenya (Dinnik, 1951), though additional surveillance may identify a distribution overlapping with that of *Kobus vardonii* in Zambia or Tanzania.
- *Echinococcus granulosus* is usually maintained by a domestic cycle of dogs eating raw livestock offal (Otero-Abad and Torgerson, 2013), and while its vertebrate-eating congener *Lycalopex gymnocercus* has been documented to host the parasite (Lucherini and Luengos Vidal, 2008), *Lycalopex vetulus* is unlikely to become infected with *E. granulosus* as it has a largely insectivorous diet (Dalponte, 2009).
- *Canis mesomelas* has been reported with infection of *Trichinella spiralis* in the Kruger National Park, South Africa (Young and Kruger, 1967).
- *Gazella leptoceros* is also predicted to be susceptible to *Trichostrongylus vitrinus*. Although Said et al. (2018) did not identify this parasite in their study, *T. vitrinus* has been documented in lambs in Tunisia (Akkari et al., 2012), indicating potential range overlap with *G. leptoceros*.

Helminths – Full model

Host	Parasite
Homo sapiens	Taenia mustelae
Bos taurus	Hymenolepis diminuta
Homo sapiens	Mastophorus muris
Ovis aries	Hymenolepis diminuta
Bos taurus	Anisakis simplex
Bos taurus	Mesocestoides lineatus
Bos taurus	Capillaria hepatica
Homo sapiens	Plagiorchis vespertilionis
Ovis aries	Echinococcus multilocularis
Bos taurus	Taenia mustelae

Table 6–17: Top 10 undocumented links with highest probablity of interaction (Full model - Helminth subset)

- We did not find any documented infections of sheep with *Hymenolepis diminuta*.
- Although the distribution, ecology, and epidemiology of *Echinococcus multilocularis* in North America is still largely unknown, it does not appear to infect any ungulates as it is maintained in a carnivore-rodent prey cycle (Massolo et al., 2014).
- The other top links were predicted in other models and discussed above.

Protozoa – Phylogeny only

Host	Parasite
Canis aureus	Toxoplasma gondii
Saguinus oedipus	Trypanosoma cruzi
Cuon alpinus	Toxoplasma gondii
Lycaon pictus	Toxoplasma gondii
Saimiri oerstedii	Trypanosoma cruzi
Mazama gouazoubira	Toxoplasma gondii
Nyctereutes procyonoides	Toxoplasma gondii
Saguinus niger	Trypanosoma cruzi
Capricornis swinhoei	Toxoplasma gondii
Panthera tigris	Toxoplasma gondii

Table 6–18: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Protozoa subset)

- *Canis aureus* with antibodies against *T. gondii* have been identified in captive animals in the United Arab Emirates (Dubey et al., 2010).
- Two recent reviews of parasites in non-human primates find no documented infection of the critically endangered Cotton-top tamarin (*Saguinus oedipus*) by *Trypanosoma cruzi*, although multiple *Saguinus sp.* have been documented with infections (Strait et al., 2012; Solórzano-García and Pérez-Ponce de León, 2018). However, a 1982 study of Colombian monkeys and marmosets identified *S. oedipus* as a host for *T. cruzi* for the first time (Marinkelle, 1982). This highlights the potential conservation importance of this parasite for *S. oedipus* and the need for periodic disease surveys of critically endangered species.
- Similarly, *Saimiri oerstedii* is not listed by these reviews as a host of *T. cruzi*, although a 1972 study identifies *S. oerstedii* as a reservoir for the parasite in Panama (Sousa, 1972). While this report should be follwed up with contemporary diagnostic methods, this reiterates the difficulty of exhaustively searching the literature for interaction data and the utility of link prediction methods to for directing these efforts.

- *Toxoplasmoa gondii* infection in *Cuon alpinus* has rarely been investigated, except for one captive individual which tested negative in serological testing (Zhang et al., 2000).
- High prevalence of antibodies against *Toxoplasma gondii* was found in wild dogs (*Lycaon pictus*) in the Kruger National Park, South Africa, and was documented as causing a fatal infection in one pup (Van Heerden et al., 1995), indicating that this parasite has the potential to influence the population dynamics of this endangered canid.
- *T. gondii* has been identified in *Mazama gouanzoubira* from French Guiana (Mercier et al., 2011) and *Nyctereutes procyonoides* (Zhou et al., 2017) in China via genetic sequencing.
- *Trypanosoma cruzi* is known to infect *Saguinus niger* (Solórzano-García and Pérez-Ponce de León, 2018).
- We did not find any reports of *T. gondii* infection in Taiwan serow *Capricornis swinhoei*, although direct evidence of infection has been found in Japanese serow (Sakae and Ishida, 2012) and *T. gondii* has been found to infect multiple animals in Taiwan (Chen et al., 2015).
- The Siberian tiger *Panthera tigris altaica* acts as a definitive host for *T. gondii* and observed to naturally shed oocysts (Elmore et al., 2010).
- *Saguinus niger* is documented to be infected by *Trypanosoma cruzi* (Solórzano-García and Pérez-Ponce de León, 2018).

Protozoa – Full model

Host	Parasite
Bos taurus	Trypanosoma cruzi
Ovis aries	Trypanosoma cruzi
Homo sapiens	Neospora caninum
Pan troglodytes	Toxoplasma gondii
Diceros bicornis	Trypanosoma cruzi
Equus caballus	Trypanosoma cruzi
Gorilla gorilla	Toxoplasma gondii
Diceros bicornis	Giardia intestinalis
Pan troglodytes	Trypanosoma cruzi
Gorilla gorilla	Trypanosoma cruzi

Table 6–19: Top 10 undocumented links with highest probablity of interaction (Full model - Protozoa subset)

- *Toxoplasma gondii* has been documented to infect chimpanzees (*Pan troglodytes*), and interestingly appears to mirror the behaviour induced in rodents and humans to infection, with infected chimpanzees attracted to the urine of leopards, their only natural predator (Poirotte et al., 2016).
- Recent finding of a *Gorilla gorilla* individual seropositive for *T. gondii* at a primate center in Gabon (Akue et al., 2018).
- *T. cruzi* was recently identified in *Equus caballus*, marking the first evidence of infection in equids (Bryan et al., 2016).
- As *T. cruzi* is currently restricted to the Americas (Browne et al., 2017), it is unlikely to infect black rhinos (*Diceros bicornis*) or gorillas (*Gorilla gorilla*) in natural conditions, unless exported through human movement.
- *Giardia* has been identified in a captive bred *Diceros bicornis* calf (Wagner and Edwards, 1984) in San Diego, indicating the potential for grey literature from zoo and captive breeding facilities to inform potential host-parasite interactions.

- Although *T. cruzi* naturally occurs in the Americas, and thus natural infection of chimpanzees is unlikely, a fatal infection was documented in a captive individual in Texas (Bommineni et al., 2009).
- The remaining links were predicted by models discussed above.

Viruses – Phylogeny only

Host	Parasite
Vulpes macrotis	Rabies lyssavirus
Lycalopex culpaeus	Rabies lyssavirus
Lycalopex griseus	Rabies lyssavirus
Lycalopex gymnocercus	Rabies lyssavirus
Myotis nattereri	Rabies lyssavirus
Myotis blythii	Rabies lyssavirus
Myotis myotis	Rabies lyssavirus
Myotis macrodactylus	Rabies lyssavirus
Myotis mystacinus	Rabies lyssavirus
Myotis dasycneme	Rabies lyssavirus

Table 6–20: Top 10 undocumented links with highest probablity of interaction (Phylogeny only model - Virus subset)

- The first four links were predicted by previous models and discussed above.
- Rabies has been isolated from a single *Myotis nattereri* individual in France (Picard-Meyer et al., 2014).
- In 2017, an inidivual *Myotis blythii* from Croatia tested positive for antibodies against rabies (Šimić et al., 2017).
- Rabies has been detected in *Myotis myotis* in a few European countries (Schatz et al., 2013).
- We did not find evidence of rabies infection in *Myotis macrodactylus* or *Myotis mystacinus*.
- Identification of rabies positive *Myotis dasycneme* in the Netherlands (Nieuwenhuijs et al., 1992).

Viruses – Full model

Host	Parasite
Mus musculus	Crimean congo hemorrhagic fever nairovirus
Rattus tiomanicus	Crimean congo hemorrhagic fever nairovirus
Gerbilliscus kempi	Crimean congo hemorrhagic fever nairovirus
Lepus californicus	Crimean congo hemorrhagic fever nairovirus
Oryx beisa	Crimean congo hemorrhagic fever nairovirus
Oryx leucoryx	Crimean congo hemorrhagic fever nairovirus
Hippotragus equinus	Crimean congo hemorrhagic fever nairovirus
Alcelaphus lichtensteinii	Crimean congo hemorrhagic fever nairovirus
Connochaetes gnou	Crimean congo hemorrhagic fever nairovirus
Connochaetes taurinus	Crimean congo hemorrhagic fever nairovirus

Table 6–21: Top 10 undocumented links with highest probablity of interaction (Full model - Virus subset)

- A recent review of 50 years of seroepidemiological studies of Crimean-Congo hemorrhagic fever virus in domestic and wild species found that a large number of bird and mammal species may be infected (Spengler et al., 2016). This study found positive serology in *Alcelaphus lichtenshetinii* (reported as *Alcelaphus buselaphus*, a synonym (Wilson and Reeder, 2005)). They also report positive serology for species in the genera *Rattus*, *Lepus*, *Oryx*, and *Hippotragus*, although the particular species identified by our model were not identified with positive serology.

6.2 Chapter 3 Supplementary Data and Results

Table 6–22: Water quality measurements

Sample code	Site	Date	Temp (°C)	mS/cm	DO (%)	DO (mg/L)	pH
DLP_8	DLP	July 10	15.27	3.11	83.37	39.67	9.16
GIR_1	GIR	June 24	18.58	1.95	50.83	42.00	9.27
GIR_2	GIR	July 1	21.85	1.80	74.47	41.00	9.24
GIR_3	GIR	July 8	20.72	1.90	88.47	39.00	9.35
HOY_2	HOY	June 22	17.59	3.18	14.53	43.43	8.14
HOY_3	HOY	June 29	17.84	3.01	42.53	40.00	8.25
HOY_4	HOY	July 6	16.83	2.96	39.27	35.90	8.39
IMB_2	IMB	June 22	15.17	2.46	74.80	46.77	8.19
IMB_3	IMB	June 29	16.07	2.43	45.23	40.67	8.16
IMB_4	IMB	July 6	15.56	2.46	35.30	35.57	8.13
KWA_5	KWA	June 19	19.20	1.80	154.90	55.30	9.92
KWA_6	KWA	June 26	16.50	1.75	111.93	45.10	9.51
NGO_2	NGO	June 24	14.97	0.48	111.23	44.10	9.44
NGO_3	NGO	July 1	17.26	0.48	108.43	40.00	9.45
NGO_4	NGO	July 8	18.22	0.51	94.50	40.33	9.21
NHL_2	NHL	June 22	17.80	1.99	118.23	50.20	8.45
NHL_3	NHL	June 29	25.52	1.97	134.97	49.87	8.08
NHL_4	NHL	July 6	22.38	2.03	125.73	40.93	8.23
NWA_2	NWA	June 26	16.14	0.90	106.67	44.77	9.77
NWA_3	NWA	June 29	24.14	0.82	199.27	53.60	9.90
NWA_4	NWA	June 30	18.91	0.90	124.63	44.43	9.66
NWA_5	NWA	July 1	23.40	0.93	180.93	49.20	9.75
NWA_6	NWA	July 2	18.65	0.91	114.90	41.00	9.66
NWA_7	NWA	July 3	17.90	0.91	104.30	40.00	9.73
NWA_8	NWA	July 10	18.06	0.87	68.53	37.27	9.24

NYA_2	NYA	June 24	14.92	0.52	54.90	39.67	8.64
NYA_3	NYA	July 1	17.51	0.54	70.27	38.00	8.64
NYA_4	NYA	July 8	18.30	0.55	76.93	37.63	8.82
WIT_2	WIT	June 24	15.31	0.58	173.60	48.87	9.40
WIT_3	WIT	July 1	18.94	0.58	139.73	43.73	9.11
WIT_4	WIT	July 8	18.89	0.69	63.37	36.23	8.54

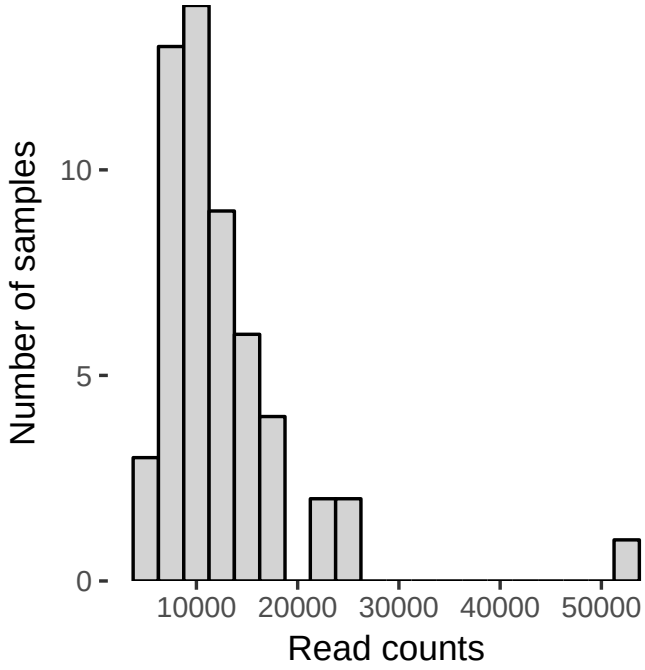


Figure 6–1: Histogram of retained sequencing depth across samples.

Table 6–23: Read counts tracked through the DADA2 pipeline

Sample code	Input	Filtered	Denoised	Merged	No Chimeras	Final Count	ASVs
BLANK_2	33813	24642	24642	22763	22468	22468	38
BLANK_3	20304	17219	17219	16166	16130	16130	14
DLP_8A	27355	26994	26994	22875	15843	15767	136
DLP_8A_S	31978	30863	30863	25222	19664	19461	179
DLP_8A_XS	38337	36972	36972	31551	27325	27087	274
DLP_8B	48282	40544	40544	34017	22401	22401	57

DLP_8B_S	28748	27713	27713	22625	17294	17119	176
DLP_8B_XS	27466	26385	26385	22277	19193	19138	168
GIR_1A	17079	16649	16649	12463	8752	8658	146
GIR_1B	15040	14314	14314	10702	7928	7865	102
GIR_2A	19825	18985	18985	14444	10003	9974	179
GIR_2B	25000	23992	23992	18315	11595	11558	199
GIR_3A	19792	19024	19024	15107	10745	10479	111
GIR_3B	23759	22888	22888	18333	12489	12184	116
HOY_2A	33258	32209	32209	15900	13831	13831	83
HOY_2B	67589	61567	61567	55496	52321	52321	58
HOY_3A	13966	13303	13303	6637	5364	5364	38
HOY_3A_S	31543	29962	29962	19154	15199	15152	128
HOY_3A_XS	34251	32392	32392	20304	16688	16599	130
HOY_3B	46882	40568	40568	31787	25515	25470	87
HOY_3B_S	14641	13727	13727	7053	5936	5908	24
HOY_3B_XS	39784	37517	37517	23198	18706	18531	168
HOY_4A	21589	19517	19517	13064	10283	10278	192
HOY_4B	26018	23698	23698	15860	11959	11959	223
IMB_2A	25930	22856	22856	15788	13076	13076	149
IMB_2B	32851	29178	29178	21347	17355	17349	227
IMB_3A	28191	25483	25483	16789	12998	12998	231
IMB_3B	21749	18798	18798	11600	9444	9444	177
IMB_4A	33674	30905	30905	19937	15659	15659	151
IMB_4B	32822	30115	30115	18815	14584	14578	142
KWA_5A	46221	39502	39502	29959	25293	24740	88
KWA_5B	44409	38312	38312	28420	22721	22216	80
KWA_6A	15998	15708	15708	12820	10654	9660	94
KWA_6A_S	21283	20439	20439	17043	15071	13822	98
KWA_6A_XS	31733	30158	30158	25821	24649	20123	162

KWA_6B	29050	24618	24618	17083	15058	12912	60
KWA_6B_S	18791	18050	18050	15145	13585	11870	100
KWA_6B_XS	44397	42687	42687	24822	24362	19676	63
NGO_2A	23500	22834	22834	18953	15565	10494	135
NGO_2B	34501	29490	29490	23297	20611	16238	72
NGO_3A	23725	23459	23459	19390	13459	8635	83
NGO_3A_S	20459	19686	19686	15621	12597	10548	67
NGO_3A_XS	25913	25002	25002	14069	13273	7472	28
NGO_3B	19885	16860	16860	11958	9884	6077	19
NGO_3B_S	23204	22345	22345	17909	14328	10529	73
NGO_3B_XS	23313	22525	22525	13626	12918	9092	24
NGO_4A	22054	21360	21360	16498	12455	10794	59
NGO_4B	22608	21868	21868	17077	13071	11159	60
NHL_2A	30627	29976	29976	26084	17626	17156	134
NHL_2B	15758	15015	15015	11797	8362	8185	53
NHL_3A	20174	18424	18424	13628	10730	9355	181
NHL_3B	17707	15991	15991	11497	9145	8746	142
NHL_4A	16675	15894	15894	12917	8961	8700	68
NHL_4B	20089	19130	19130	15677	10697	10307	83
NWA_2A	26097	22271	22271	16953	12983	12405	57
NWA_2A_S	19149	18326	18326	15107	12871	11840	116
NWA_2A_XS	20071	19273	19273	10180	9890	6330	20
NWA_2B	20898	17634	17634	13251	10543	10330	37
NWA_2B_S	15648	14878	14878	12009	9819	9052	104
NWA_2B_XS	27808	26838	26838	13564	12591	10713	38
NWA_3A	10226	9735	9735	7264	6029	4895	67
NWA_3B	28430	26973	26973	21273	15416	13726	119
NWA_4A	17654	17123	17123	14118	11196	10140	124
NWA_4B	15907	15184	15184	12240	9456	8760	93

NWA_5A	14896	14275	14275	11127	9207	8777	91
NWA_5B	17054	16316	16316	13064	10570	10189	89
NWA_6A	19653	18822	18822	14894	12020	11380	98
NWA_6B	15184	14591	14591	11804	9839	9355	85
NWA_7A	13767	13177	13177	10932	9017	8445	72
NWA_7B	16831	16102	16102	12835	9883	9231	87
NWA_8A	15266	14481	14481	10673	8879	8531	109
NWA_8B	17882	16958	16958	12503	10249	10038	127
NYA_2A	33395	28245	28245	21413	18148	18143	60
NYA_2B	36205	30519	30519	23387	18678	18631	57
NYA_3A	29675	25050	25050	16505	13118	13019	55
NYA_3B	30193	25942	25942	17723	14146	13985	76
NYA_4A	17561	16928	16928	12689	8760	7377	241
NYA_4A_S	13136	12431	12431	9507	7404	5624	157
NYA_4A_XS	13283	12762	12762	5737	4120	3220	24
NYA_4B	24397	20885	20885	14109	9512	7813	74
NYA_4B_S	15610	14774	14774	11127	8646	6365	169
NYA_4B_XS	16154	15502	15502	8077	6182	4653	36
WIT_2A	14307	13757	13757	9617	7282	7100	78
WIT_2B	13757	13277	13277	8501	6283	6283	80
WIT_3A	25714	25291	25291	16183	10271	9958	201
WIT_3B	14067	13479	13479	7899	5799	5600	76
WIT_4A	18581	17947	17947	13911	8764	8701	66
WIT_4B	22216	21441	21441	16481	11437	11333	78

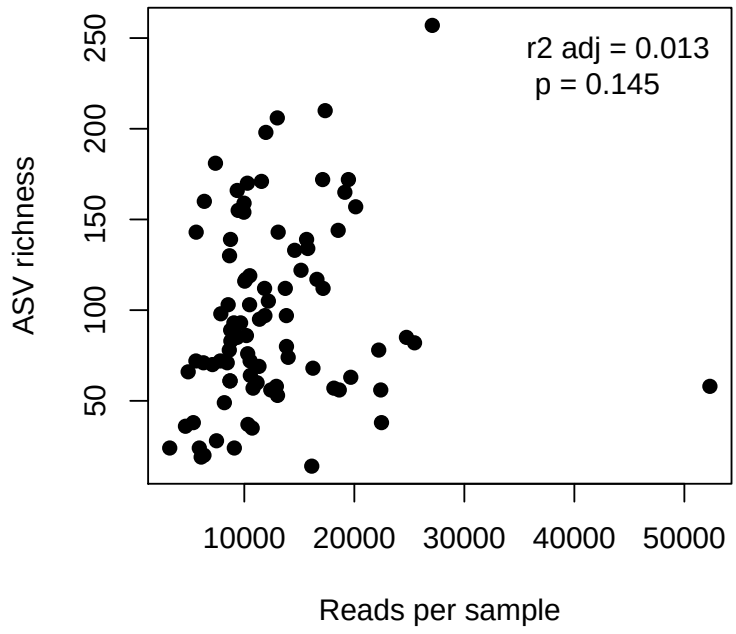


Figure 6–2: Scatterplot of reads and ASV richness per sample.

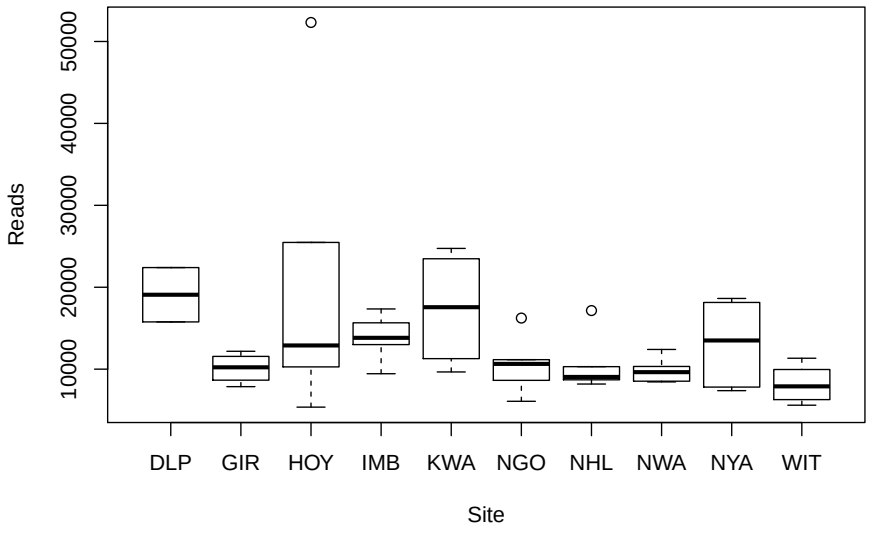


Figure 6–3: Box and whisker plots of reads among the core samples, grouped by site.

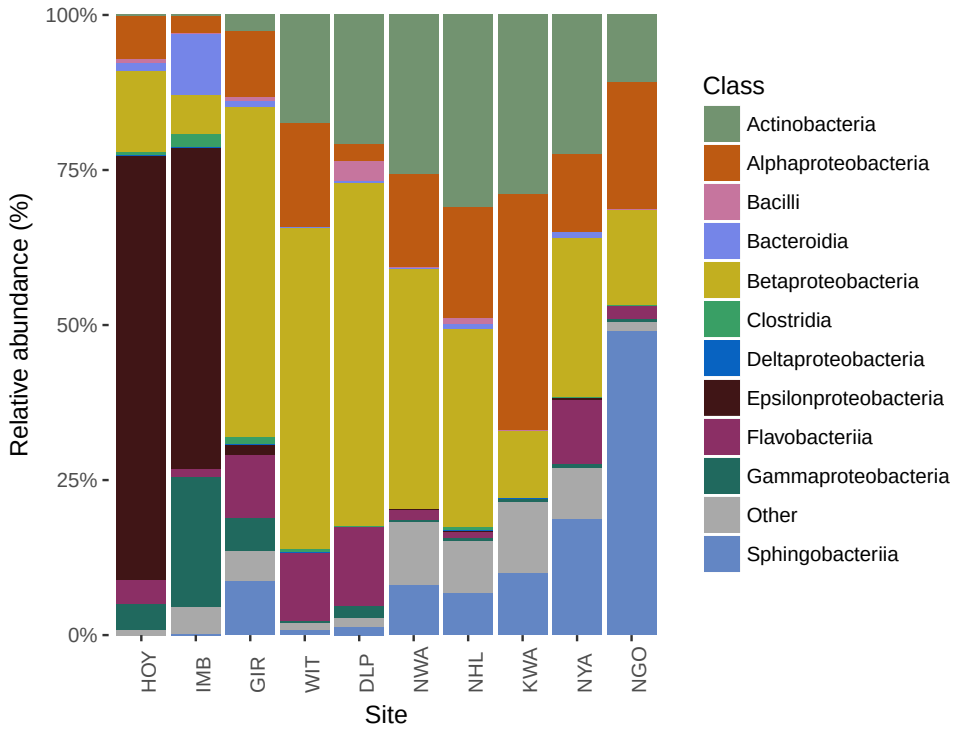


Figure 6–4: Relative abundances of bacterial classes across sites. Sites are ordered by relative abundance of phylum Proteobacteria.

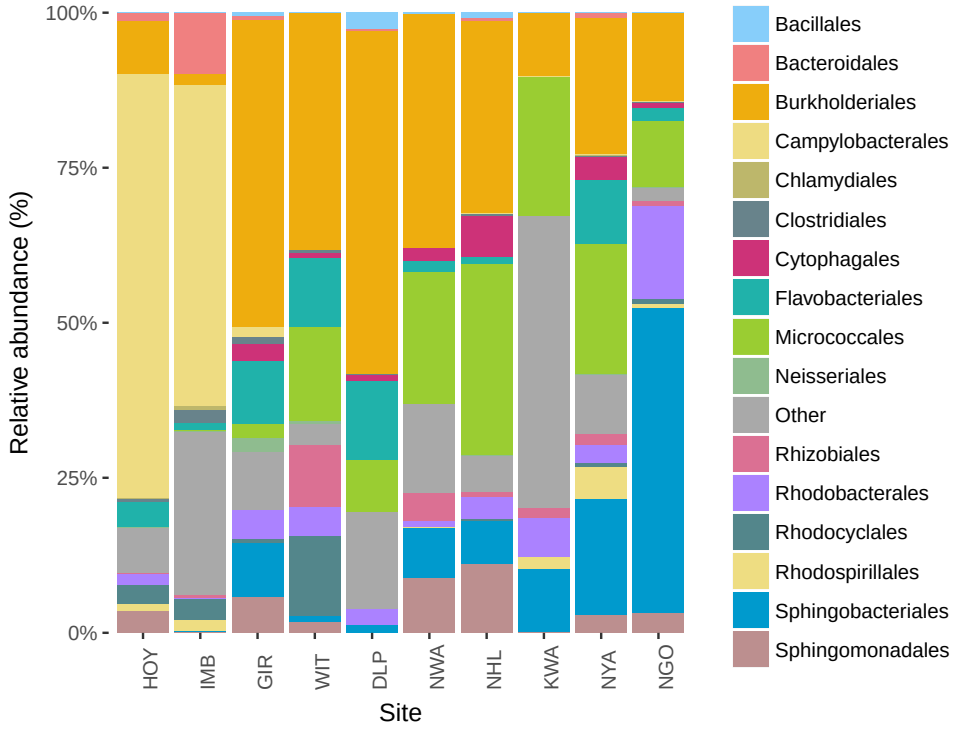


Figure 6–5: Relative abundances of bacterial orders across sites. Sites are ordered by relative abundance of phylum Proteobacteria.

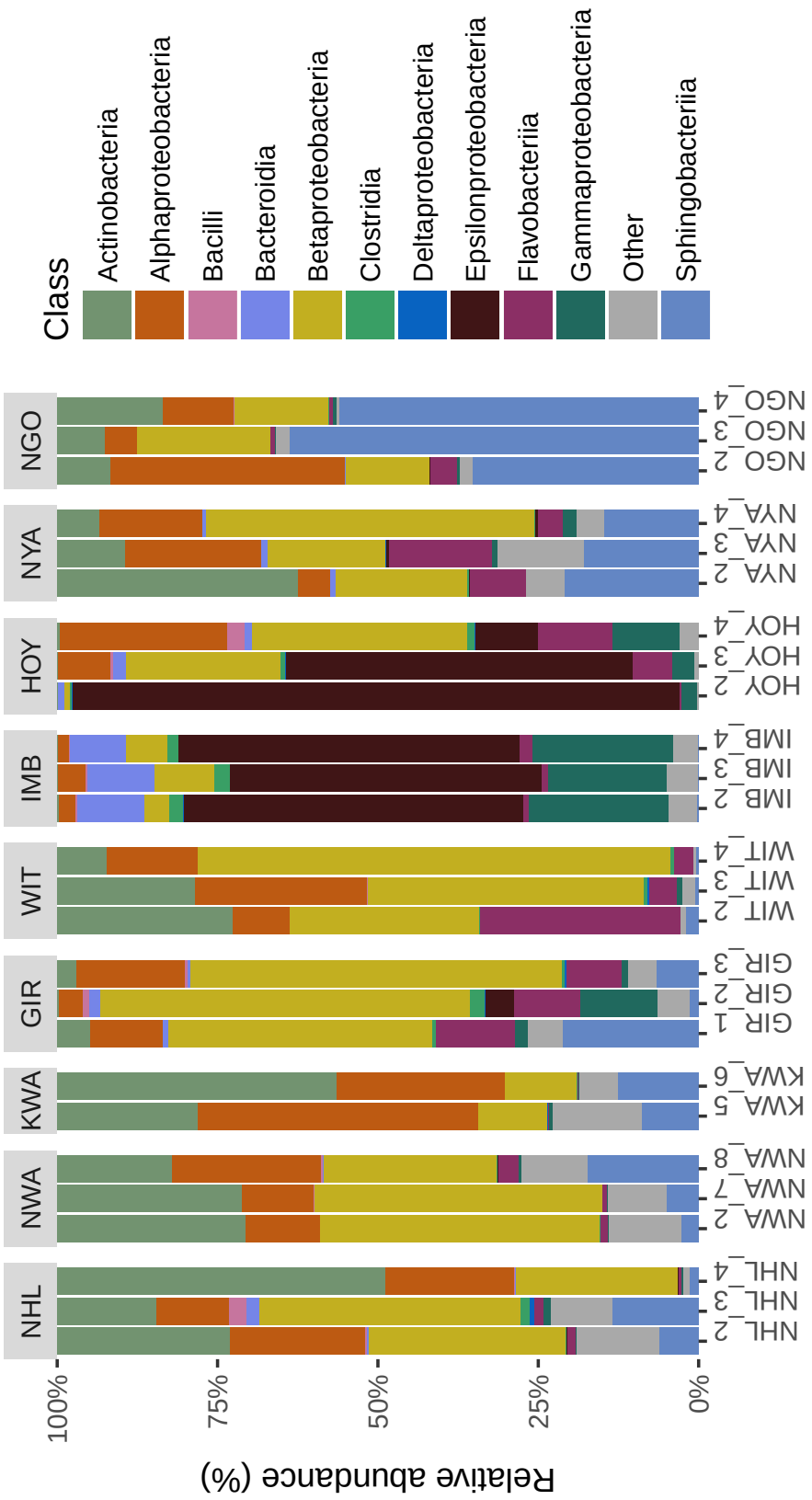


Figure 6–6: Relative abundances of bacterial classes across weekly samples.

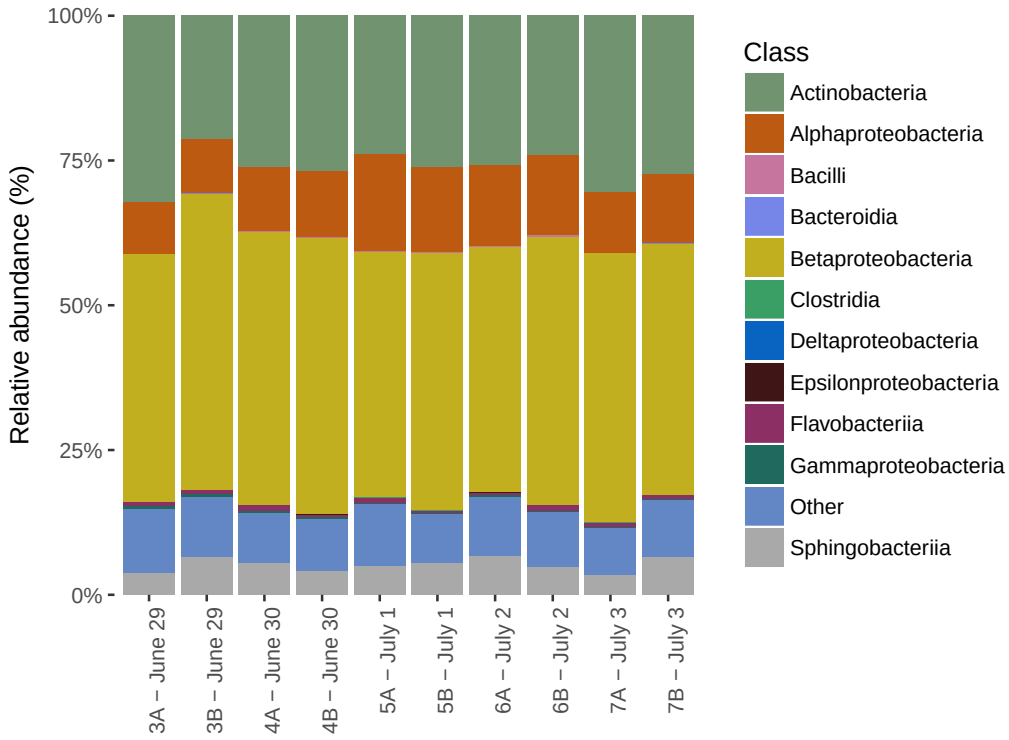


Figure 6–7: Relative abundances of bacterial classes across five days at a single site (NWA).

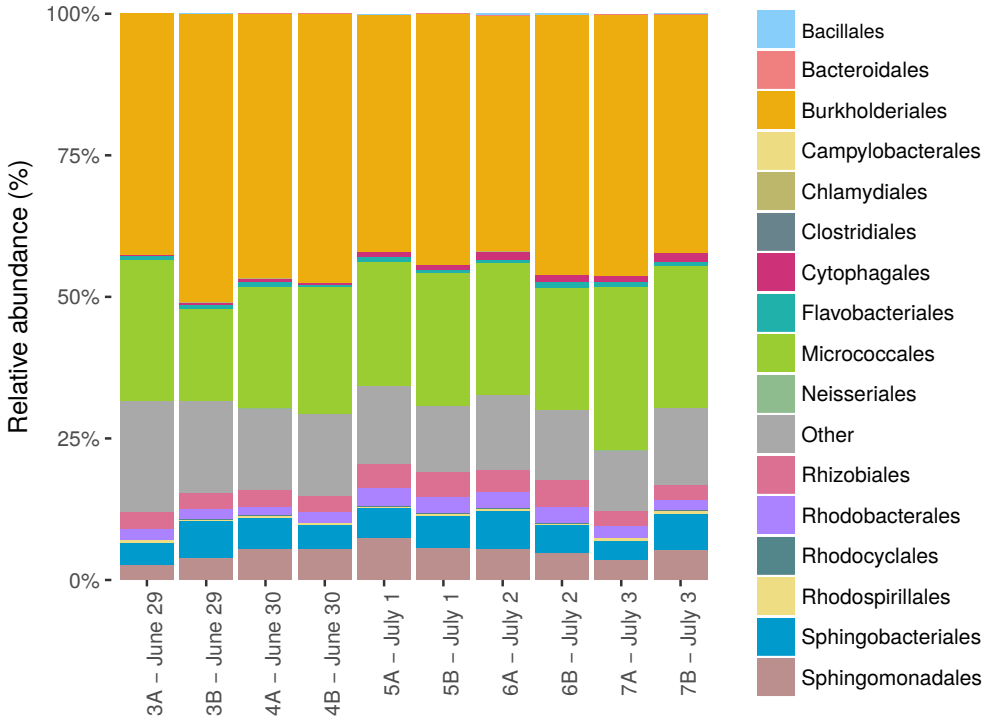


Figure 6–8: Relative abundances of bacterial orders across five days at a single site (NWA).

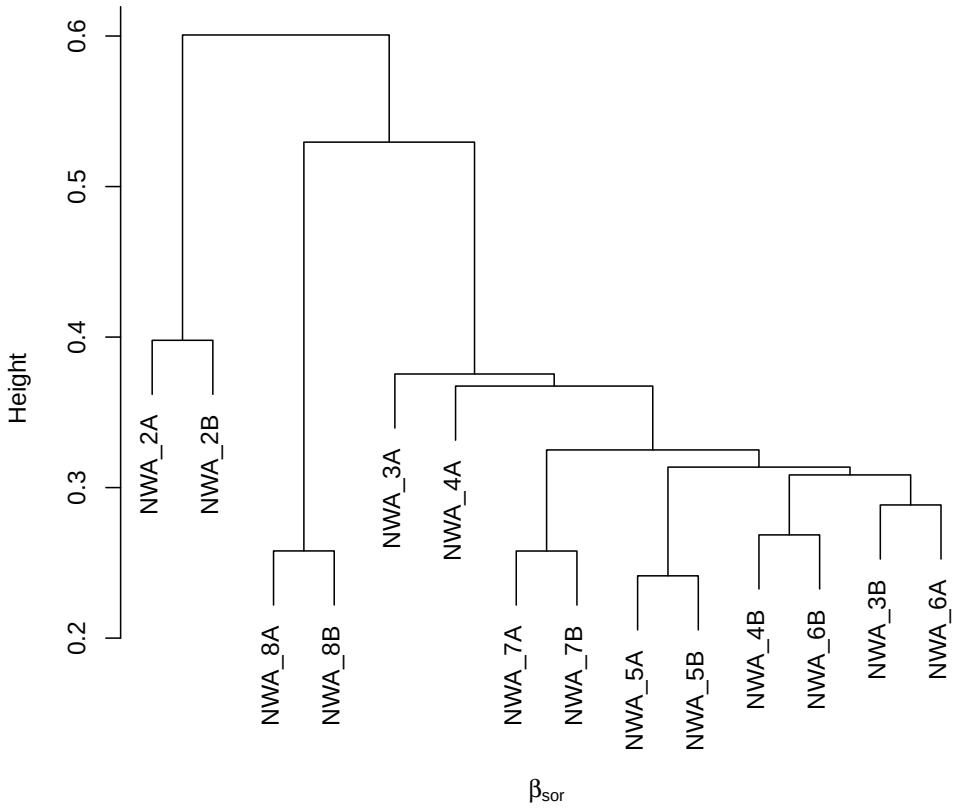
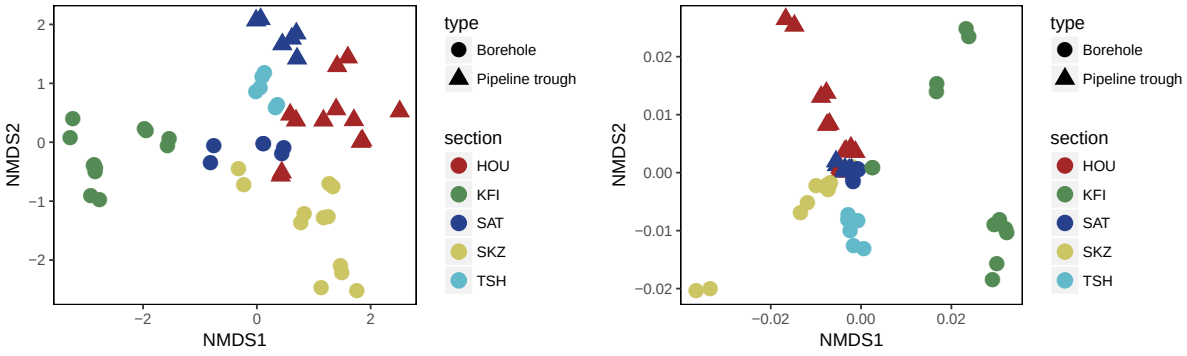


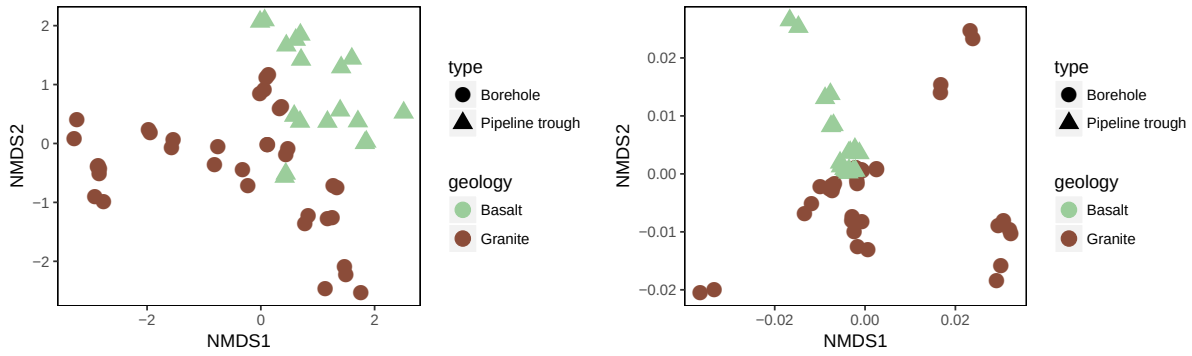
Figure 6–9: Hierarchical clustering of 150mL NWA samples based on pairwise Sorensen's beta diversity.



(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

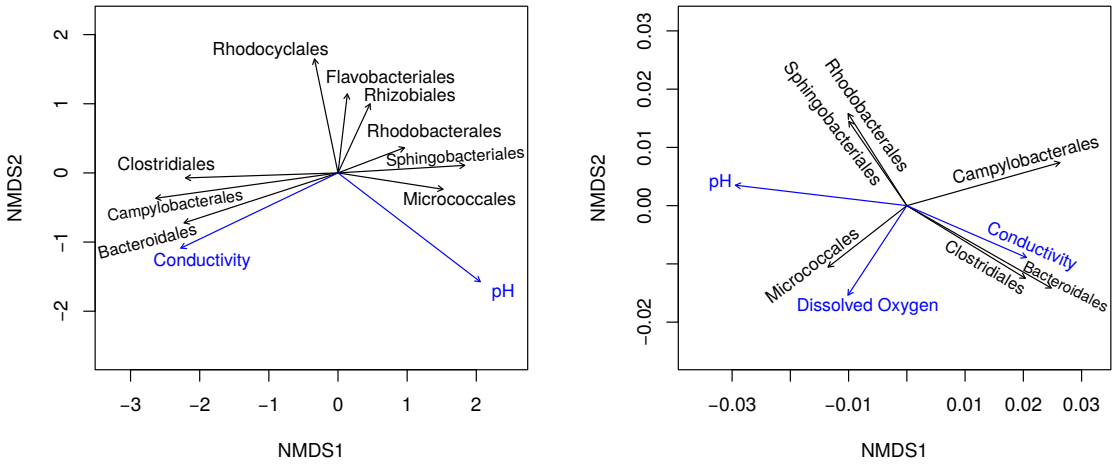
Figure 6–10: NMDS plots of a) Bray-Curtis and b) abundance-weighted UniFrac distances. Colours represent section, shapes represent waterhole type.



(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

Figure 6–11: NMDS plots of a) Bray-Curtis and b) abundance-weighted UniFrac distances. Colours represent subsurface geology, shapes represent waterhole type.



(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

Figure 6–12: NMDS ordinations of variation in bacterial community structure across 54 samples based on a) Bray-Curtis and b) abundance-weighted UniFrac distances. Arrows indicate the direction of significant ($p < 0.05$) correlations among variables and the NMDS axes, with arrow length indicating the strength of the correlation. Blue arrows indicate environmental variables, while black arrows indicate relative abundances of sequences from different microbial orders. The ordination axes explain 96.9% (a) and 98.1% (b) of the variance in the dissimilarities (Fig. 6–17).

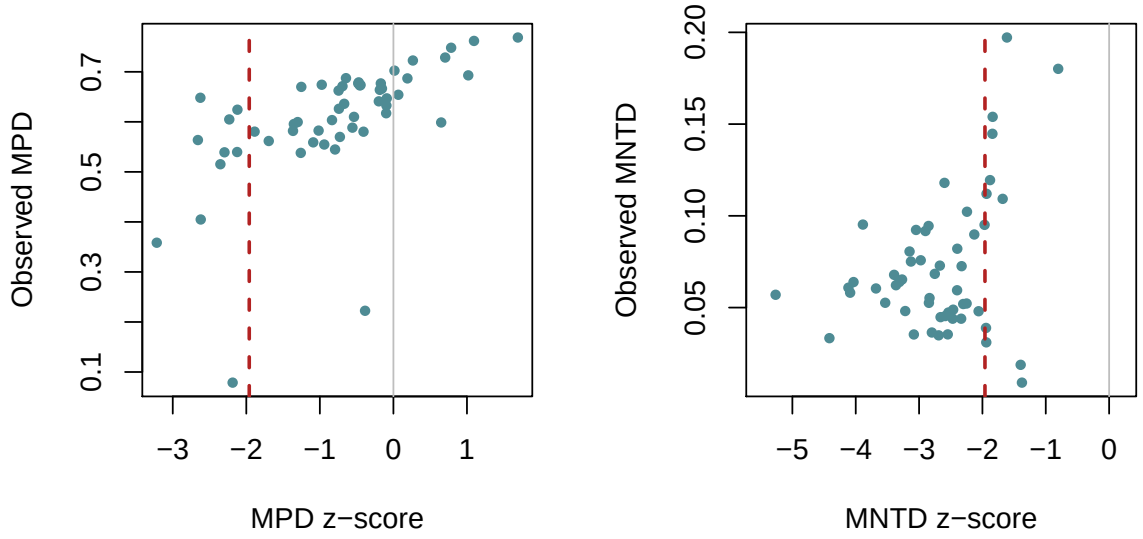


Figure 6–13: Phylogenetic community structure of bacterial ASVs across samples based on a) mean phylogenetic pairwise distance (MPD) and b) mean nearest taxon distance (MNTD). Each plot depicts the relationship between observed values (y-axis), and the standardized effect size (z-score) for each community following 999 permutations using the "richness" null model. Gray lines depict z-scores of 0. Red dashed lines represent critical values for a two-tailed z-test (α=0.05), with points lesser than these values indicating significant phylogenetic clustering.

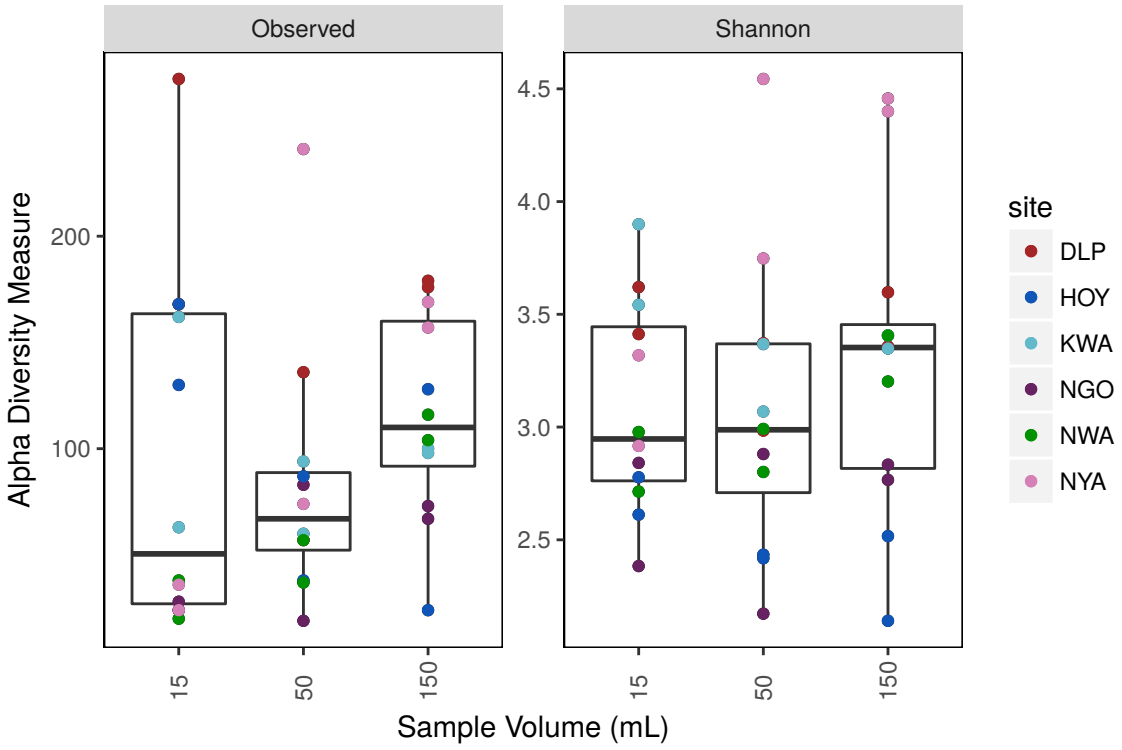


Figure 6–14: Alpha diversity as measured by observed number of ASVs and Shannon diversity for samples in which 150mL, 50mL and 15mL volumes were filtered. Colours represent sites.

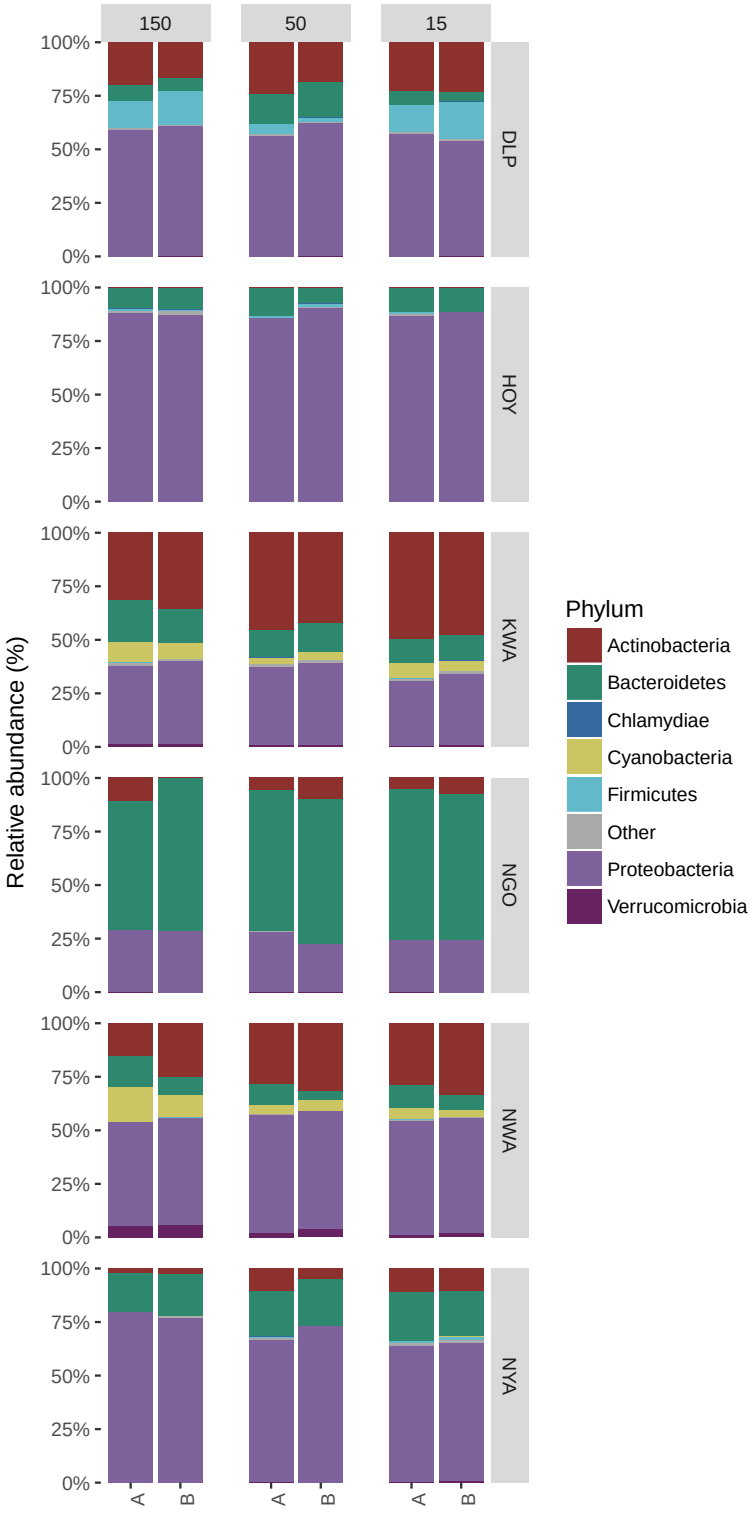


Figure 6–15: Relative abundances of bacterial phyla across different sample volumes.

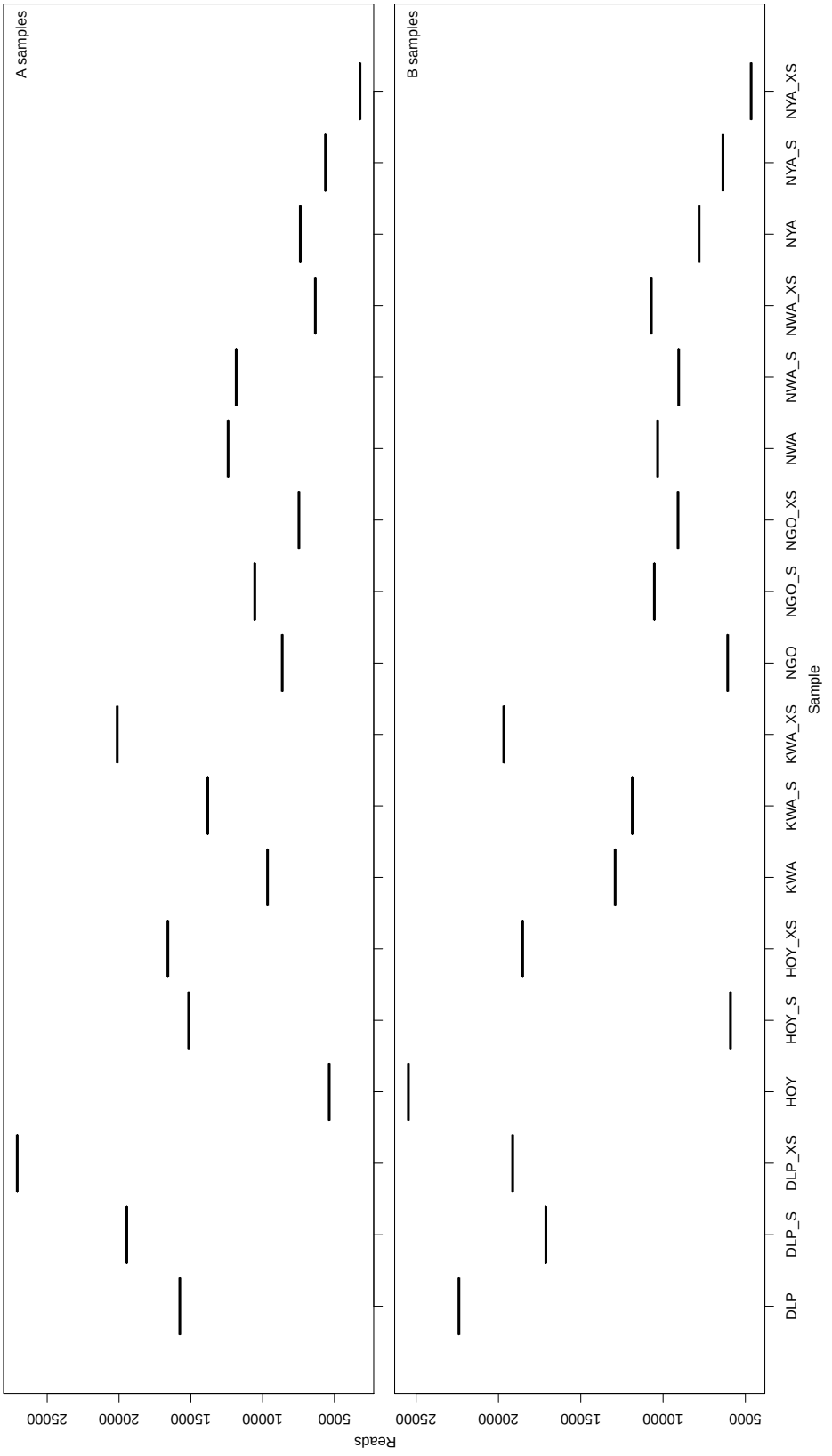
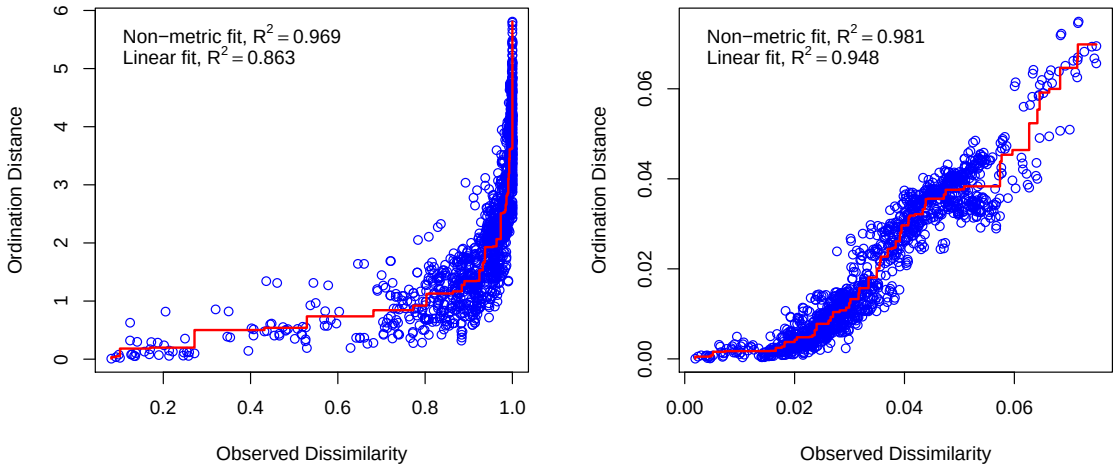


Figure 6–16: Reads for samples in which different volumes were filtered.



(a) Bray-Curtis NMDS

(b) Weighted UniFrac NMDS

Figure 6–17: Stressplots for NMDS plots using a) Bray-Curtis and b) abundance-weighted UniFrac distances

6.3 Chapter 4 Supplementary Data and Results

Main model

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.57	0.74	-2.06	-1.03	-0.57	-0.08	0.87	1338	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	6000	1.00
	Evolutionary Isolation	1.69	0.12	1.45	1.61	1.70	1.78	1.93	6000	1.00
Parasite	Host Species Richness	0.90	0.71	-0.42	0.41	0.88	1.37	2.36	3335	1.00
	Vectored	-0.27	0.65	-1.60	-0.69	-0.26	0.15	1.00	3631	1.00
	Reproduction	0.12	0.66	-1.19	-0.32	0.13	0.56	1.43	4317	1.00
	Environmental Resting Stage	0.22	0.96	-1.65	-0.40	0.20	0.83	2.19	6000	1.00
Country	Latitude	-0.21	0.42	-1.04	-0.49	-0.21	0.08	0.61	3796	1.00
	GDP per capita	-0.56	0.37	-1.29	-0.81	-0.56	-0.31	0.18	6000	1.00

Table 6–24: Summary of model output for continuous predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

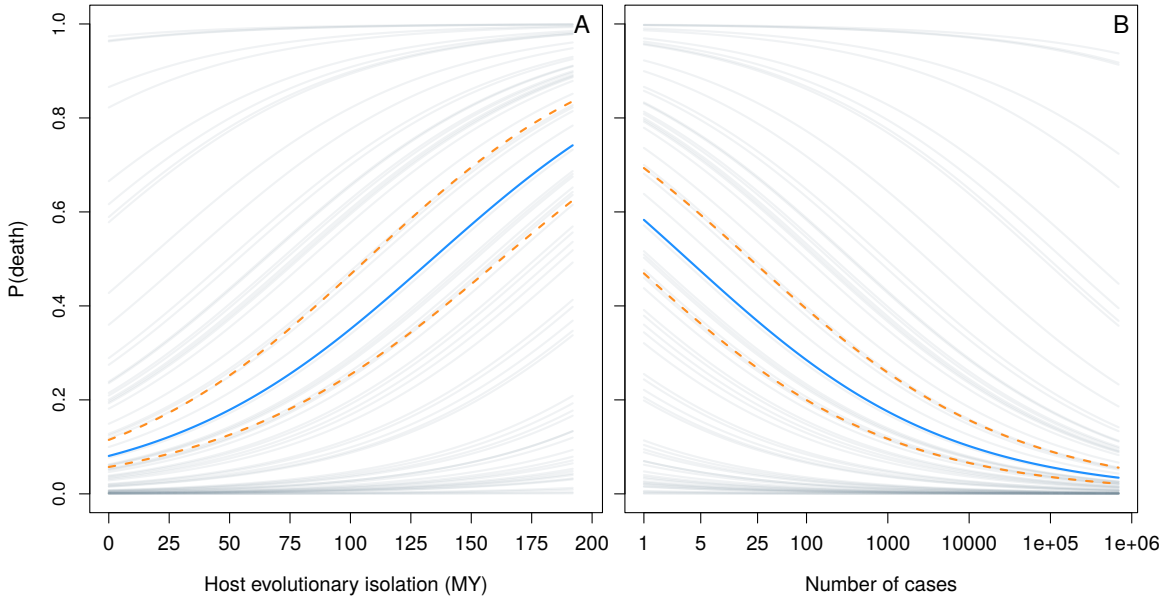


Figure 6–18: Posterior predictions of the probability of death as a function of a) host evolutionary isolation (in millions of years), and b) the number of cases. Solid blue lines represent the mean logistic curve, dashed yellow lines represent the upper and lower bounds of the 50% credible interval. Grey lines depict equivalent mean curves offset by the posterior mean effects for each parasite.

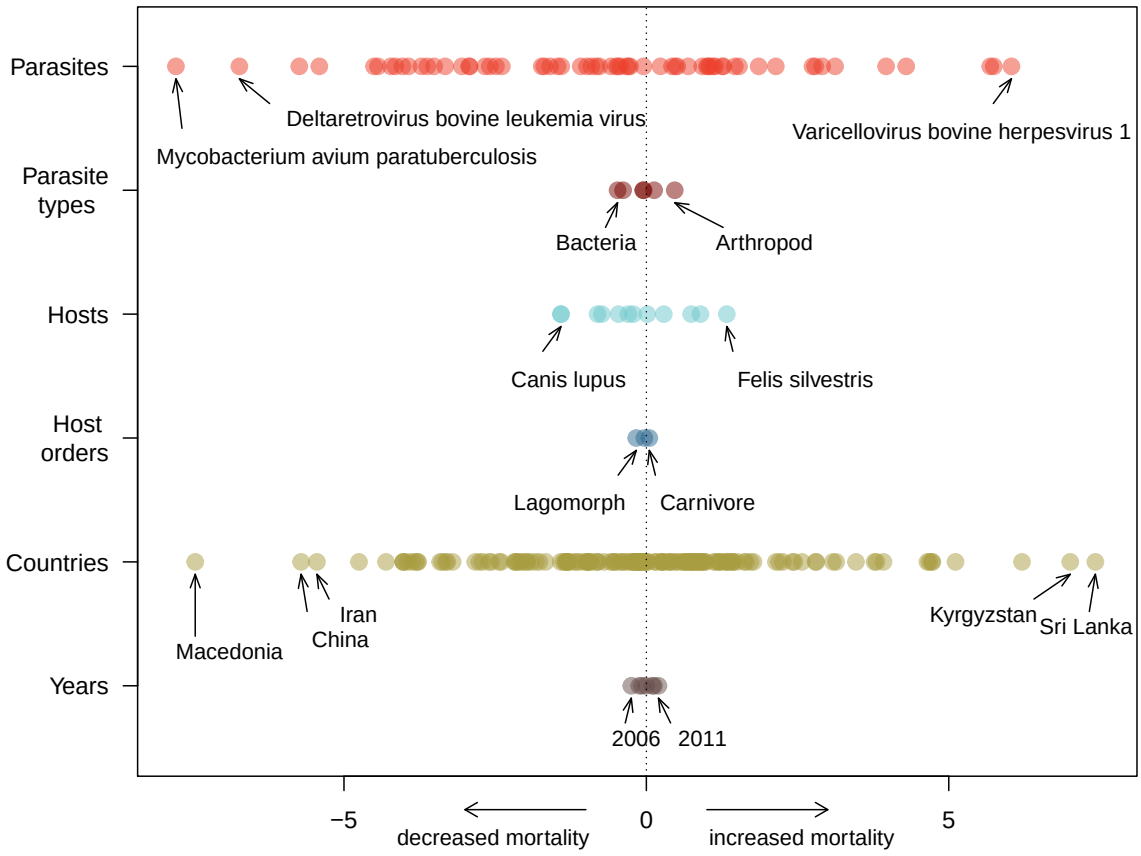


Figure 6–19: Mean estimated effects for individual hierarchical terms (parasites, parasite types, hosts, host orders, countries, and years). Plotted estimates have been set to 50% transparency to visualize overlapping points, and extreme estimates in each group have been identified.

Sensitivity Analyses and Alternative Models

Excluding single-host parasites

As selective pressures driving virulence evolution are likely to differ among single and multi-host parasites, the main model fit again after removing single-host parasites from the data.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.83	0.80	-2.50	-1.32	-0.81	-0.30	0.70	1075	1.00
	log (Cases)	-1.30	0.01	-1.33	-1.31	-1.30	-1.29	-1.28	6000	1.00
	Evolutionary Isolation	1.74	0.11	1.52	1.66	1.74	1.81	1.96	5847	1.00
Parasite	Host Species Richness	0.55	0.75	-0.86	0.06	0.53	1.04	2.09	3668	1.00
	Vectored	-1.05	0.76	-2.61	-1.53	-1.02	-0.54	0.36	2760	1.00
	Reproduction	0.47	0.71	-0.87	-0.02	0.44	0.92	1.92	3564	1.00
	Environmental Resting Stage	0.26	0.96	-1.60	-0.34	0.24	0.83	2.26	6000	1.00
Country	Latitude	-0.25	0.42	-1.08	-0.52	-0.24	0.04	0.57	4573	1.00
	GDP per capita	-0.58	0.37	-1.32	-0.83	-0.58	-0.33	0.13	5092	1.00

Table 6–25: Summary of main model excluding single-host parasites for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

Host taxonomic diversity

Due to incomplete sampling, the host species reported in the GMPD and EID2 databases are unlikely to include the complete set of susceptible hosts for each parasite. As a sensitivity analysis, host species richness (SR_p) was replaced by a measure of taxonomic diversity using data reported by Lefèvre et al. (2010) and the OIE documentation. Host taxonomic diversity varies from 1-6 corresponding to whether parasites infect hosts belonging to a single species (1), genus (2), family (3), order (4), class (5), or multiple classes (6). Just as with host species richness, the ability to infect humans was not included in estimates of taxonomic diversity.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.79	0.90	-2.71	-1.37	-0.75	-0.19	0.87	2261	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	6000	1.00
	Evolutionary Isolation	1.70	0.13	1.45	1.61	1.70	1.78	1.93	5833	1.00
Parasite	Host Taxonomic Diversity	-0.01	0.20	-0.39	-0.15	-0.02	0.12	0.38	2687	1.00
	Vectored	-0.26	0.66	-1.59	-0.69	-0.26	0.18	1.02	5847	1.00
	Reproduction	0.04	0.67	-1.28	-0.40	0.05	0.48	1.36	5347	1.00
	Environmental Resting Stage	0.30	0.98	-1.56	-0.34	0.27	0.90	2.39	6000	1.00
Country	Latitude	-0.20	0.43	-1.05	-0.48	-0.21	0.08	0.65	5261	1.00
	GDP per capita	-0.55	0.37	-1.27	-0.80	-0.55	-0.30	0.18	6000	1.00

Table 6–26: Summary of model with host taxonomic diversity for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

Parasites with avian reservoirs

As an extension of our main model, we include whether or not a parasite uses an avian reservoir (Eastern equine encephalitis, Western equine encephalitis, Venezuelan equine encephalitis, Fowlpox, Newcastle Disease, West Nile Virus, *Pasturella multocida*), as we hypothesize that this might correlate with whether domesticated mammals represent dead-end hosts from which the parasite is not transmitted further, such as is the case for West Nile Virus and other encephalitic viruses that spillover from birds to horses (Weaver and Barrett, 2004). The use of avian species as reservoir hosts were taken from OIE publications (disease summaries from the OIE Terrestrial Manual (OIE, 2012) and OIE technical disease cards), and from Lefèvre et al. (2010), and coded as a binary predictor.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.59	0.72	-2.04	-1.07	-0.59	-0.11	0.80	1568	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.33	-1.33	-1.32	-1.30	5980	1.00
	Evolutionary Isolation	1.69	0.12	1.46	1.61	1.69	1.78	1.93	5935	1.00
Parasite	Host Species Richness	0.90	0.73	-0.45	0.41	0.87	1.36	2.42	3574	1.00
	Vectored	0.28	0.65	-1.60	-0.70	-0.26	0.16	1.01	4318	1.00
	Reproduction	0.14	0.65	-1.14	-0.30	0.13	0.58	1.42	5209	1.00
	Environmental Resting Stage	0.23	0.97	-1.61	-0.39	0.20	0.82	2.29	5934	1.00
	Avian Reservoir	0.18	0.83	-1.47	-0.36	0.17	0.68	1.83	5474	1.00
Country	Latitude	-0.20	0.42	-1.05	-0.48	-0.20	0.09	0.62	3896	1.00
	GDP per capita	-0.55	0.37	-1.29	-0.80	-0.55	-0.31	0.15	5942	1.00

Table 6–27: Summary of model including indicator for avian reservoir for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

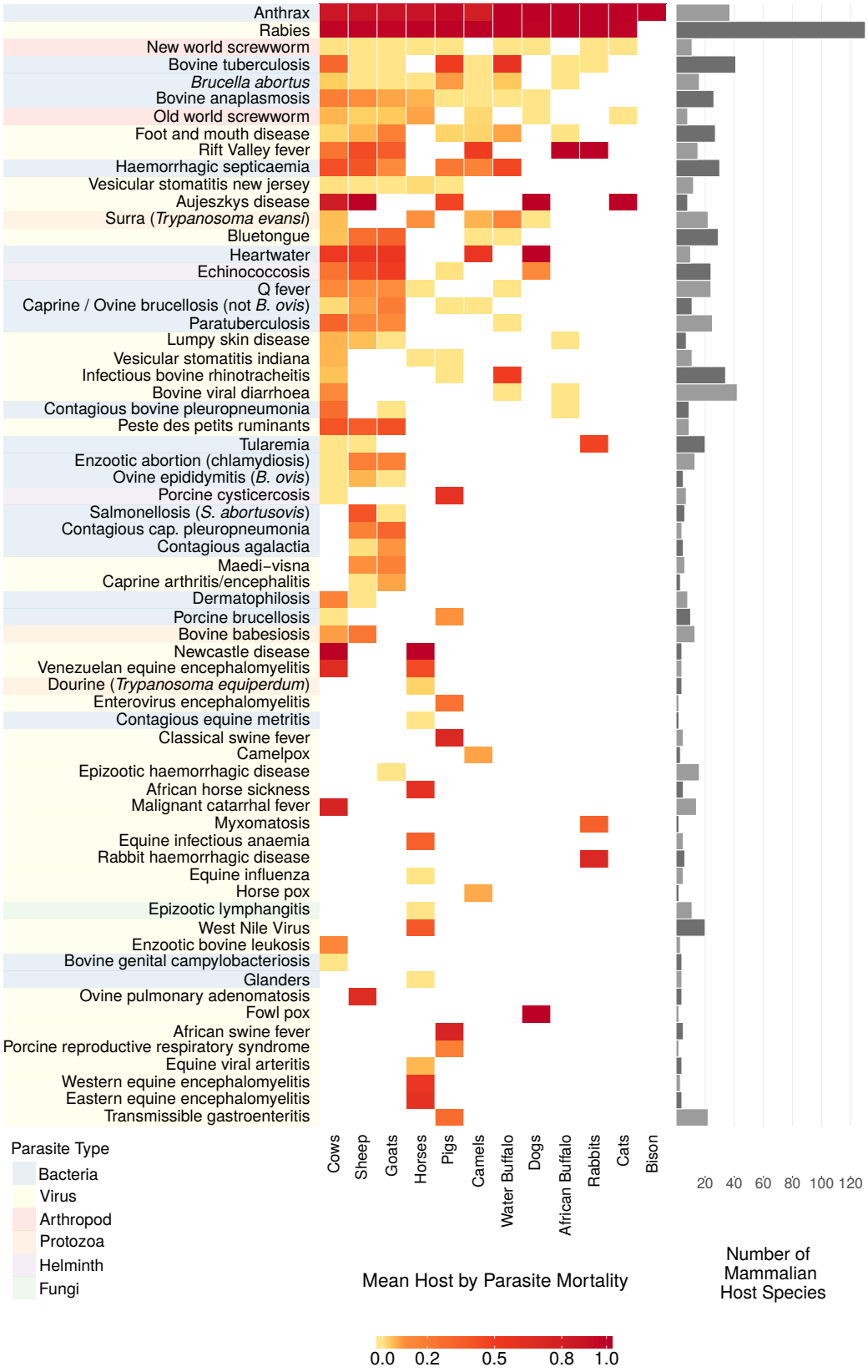


Figure 6–20: Version of Fig. 4–2 including parasite common names. Parasite names are colour coded by parasite type.

6.4 Chapter 5 Supplementary Data and Results

Host Group	Response	N hosts	λ	Phylo df	Predictor	Slope	S.E.	t	p
Ungulates	Proportion Single-Host Parasites	68	0.0	68	Citation Count	0.34	0.083	4.118	<0.0001
					Threatened	1.537	0.447	3.442	0.001
					Adult Body Mass (g)	-0.015	0.108	-0.141	0.888
					GR Area (km^2)	-0.151	0.081	-1.869	0.066
					Population Density (n/km^2)	0.086	0.083	1.033	0.306
					Decreasing Population Trend	0.283	0.328	0.862	0.392
	Richness of Single-Host Parasites	68	0.0	68	Citation Count	0.528	0.094	5.615	<0.0001
					Threatened	0.374	0.425	0.88	0.382
					Adult Body Mass (g)	0.136	0.12	1.131	0.262
					GR Area (km^2)	0.073	0.086	0.846	0.401
					Population Density (n/km^2)	0.122	0.09	1.362	0.178
					Decreasing Population Trend	-0.641	0.374	-1.717	0.091
	Richness of Multi-Host Parasites	68	0.0	68	Citation Count	0.193	0.061	3.14	0.003
					Threatened	-0.964	0.403	-2.392	0.02
					Adult Body Mass (g)	0.18	0.076	2.375	0.021
					GR Area (km^2)	0.166	0.068	2.439	0.018
					Population Density (n/km^2)	0.12	0.063	1.916	0.06
					Decreasing Population Trend	-0.747	0.237	-3.155	0.002
Carnivores	Proportion Single-Host Parasites	64	0.0	64	Citation Count	0.381	0.154	2.468	0.017
					Threatened	-0.648	0.822	-0.789	0.434
					Adult Body Mass (g)	0.195	0.111	1.76	0.084
					GR Area (km^2)	-0.094	0.126	-0.746	0.459
					Population Density (n/km^2)	0.339	0.088	3.848	<0.0001
					Decreasing Population Trend	-0.111	0.348	-0.318	0.752
	Richness of Single-Host Parasites	64	0.0	64	Citation Count	0.833	0.175	4.759	<0.0001
					Threatened	-0.84	1.031	-0.814	0.419
					Adult Body Mass (g)	0.162	0.118	1.372	0.176
					GR Area (km^2)	-0.133	0.161	-0.826	0.412
					Population Density (n/km^2)	0.267	0.09	2.954	0.005
					Decreasing Population Trend	-0.586	0.443	-1.324	0.191
	Richness of Multi-Host Parasites	64	0.178	54.9	Citation Count	0.591	0.101	5.882	<0.0001
					Threatened	-0.352	0.425	-0.83	0.411
					Adult Body Mass (g)	-0.018	0.078	-0.236	0.814
					GR Area (km^2)	-0.077	0.076	-1.015	0.315
					Population Density (n/km^2)	0.044	0.055	0.805	0.425
					Decreasing Population Trend	-0.215	0.24	-0.895	0.375

Table 6–28: Full models

Host Group	Response	N hosts	λ	Phylo df	Predictor	Slope	S.E.	t	p
Ungulates	Proportion Single-Host Parasites	93	0.0	93	Citation Count	0.255	0.065	3.916	<0.0001
					Threatened	1.072	0.029	3.695	<0.0001
					GR Area (km^2)	-0.082	0.063	-1.297	0.198
	Richness of Single-Host Parasites	86	0.0	86	Citation Count	0.613	0.074	8.323	<0.0001
					Threatened	0.506	0.316	1.604	0.113
					Decreasing Population Trend	-0.669	0.345	-1.941	0.056
	Richness of Multi-Host Parasites	68	0.0	68	Citation Count	0.193	0.061	3.14	0.003
					Threatened	-0.964	0.403	-2.392	0.02
					Adult Body Mass (g)	0.18	0.076	2.375	0.021
					GR Area (km^2)	0.166	0.068	2.439	0.018
					Population Density (n/km^2)	0.12	0.063	1.916	0.06
					Decreasing Population Trend	-0.747	0.237	-3.155	0.002
Carnivores	Proportion Single-Host Parasites	85	0.0218	83.5	Citation Count	0.320	0.113	2.834	0.006
					Threatened	0.360	0.368	0.978	0.331
					Adult Body Mass (g)	0.114	0.073	1.567	0.121
					Population Density (n/km^2)	0.246	0.059	4.189	<0.0001
	Richness of Single-Host Parasites	85	0.0	85	Citation Count	0.883	0.133	4.189	<0.0001
					Threatened	-0.008	0.424	-0.02	0.984
					Population Density (n/km^2)	0.185	0.061	3.043	0.003
	Richness of Multi-Host Parasites	140	0.16	121	Citation Count	0.615	0.066	9.338	<0.0001
					Threatened	-0.306	0.201	-1.528	0.129

Table 6–29: Reduced models

Host Group	Response	N hosts	λ	Phylo df	Predictor	Slope	S.E.	t	p
Ungulates	Proportion Single-Host Parasites	95	0.0	95	Citation Count Threatened	0.21 0.956	0.057 0.259	3.717 3.693	<0.0001 <0.0001
	Richness of Single-Host Parasites	95	0.0	95	Citation Count Threatened	0.533 0.51	0.065 0.314	8.205 1.624	<0.0001 0.108
	Richness of Multi-Host Parasites	95	0.0	95	Citation Count Threatened	0.329 -0.570	0.050 0.314	6.55 -1.81	<0.0001 0.073
Carnivores	Proportion Single-Host Parasites	140	0.0281	137	Citation Count Threatened	0.14 0.109	0.08 0.308	1.743 0.353	0.084 0.725
	Richness of Single-Host Parasites	140	0.0	140	Citation Count Threatened	0.742 -0.317	0.107 0.406	6.94 -0.781	<0.0001 0.436
	Richness of Multi-Host Parasites	140	0.16	121	Citation Count Threatened	0.615 -0.306	0.066 0.201	9.338 -1.528	<0.0001 0.129

Table 6–30: Bivariate models

Host Group	Response	N hosts	λ	Phylo df	Predictor	Slope	S.E.	t	p
Ungulates	Proportion Single-Host Micro-Parasites	49	0.0	49	Citation Count	0.366	0.142	2.568	0.0014
					Threatened	2.002	0.61	3.279	0.002
					Adult Body Mass (g)	-0.139	0.18	-0.771	0.445
					GR Area (km^2)	-0.183	0.096	-1.903	0.064
					Population Density (n/km^2)	0.064	0.137	0.468	0.642
					Decreasing Population Trend	0.083	0.578	0.143	0.887
	Proportion Single-Host Macro-Parasites	64	0.0	64	Citation Count	0.367	0.099	3.708	<0.0001
					Threatened	1.234	0.554	2.229	0.030
					Adult Body Mass (g)	0.058	0.117	0.496	0.622
					GR Area (km^2)	-0.182	0.128	-1.426	0.159
					Population Density (n/km^2)	0.071	0.097	0.729	0.469
					Decreasing Population Trend	0.443	0.346	1.279	0.206
Carnivores	Proportion Single-Host Micro-Parasites	53	0.984	13	Citation Count	0.216	0.259	0.831	0.438
					Threatened	-0.524	0.923	-0.568	0.591
					Adult Body Mass (g)	0.042	0.192	0.218	0.835
					GR Area (km^2)	0.035	0.214	0.165	0.875
					Population Density (n/km^2)	0.092	0.015	0.616	0.561
					Decreasing Population Trend	0.906	0.491	1.844	0.115
	Proportion Single-Host Macro-Parasites	60	0.0	60	Citation Count	0.434	0.189	2.291	0.026
					Threatened	-1.14	1.251	-0.911	0.366
					Adult Body Mass (g)	0.206	0.131	1.572	0.122
					GR Area (km^2)	-0.306	0.175	-1.749	0.086
					Population Density (n/km^2)	0.358	0.102	3.516	0.001
					Decreasing Population Trend	-0.599	0.458	-1.307	0.197

Table 6–31: Micro and Macro-Parasite Models

References

Aghokeng, A. F., A. Ayouba, E. Mpoudi-Ngole, S. Loul, F. Liegeois, E. Delaporte, and M. Peeters (2010). Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. *Infection, Genetics and Evolution 10*(3), 386–396.

Agudelo-Romero, P. and S. F. Elena (2008). The degree of plant resilience to infection correlates with virus virulence and host-range. *Spanish Journal of Agricultural Research 6*, 160–169.

Akkari, H., M. Gharbi, and M. a. Darghouth (2012). Dynamics of infestation of tracers lambs by gastrointestinal helminths under a traditional management system in the North of Tunisia. *Parasite 19*(4), 407–15.

Akue, J. P., N. E. Tomo, J. Badiambile, H. Moukana, R. A. Mbou-Mountsimbi, and B. Ngoubangoye (2018). Seroprevalence of Toxoplasma gondii and Neospora caninum in non-human primates at a primate center at Franceville, Gabon. *Journal of Parasitology and Vector Biology 10*(1), 1–7.

Alexander, H. K. and T. Day (2010). Risk factors for the evolutionary emergence of pathogens. *Journal of the Royal Society Interface 7*(51), 1455–1474.

Alizon, S. (2013). Parasite co-transmission and the evolutionary epidemiology of virulence. *Evolution 67*(4), 921–933.

Alizon, S., A. Hurford, N. Mideo, and M. Van Baalen (2009). Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *Journal of Evolutionary Biology 22*(2), 245–259.

Allan, K. J., H. M. Biggs, J. E. Halliday, R. R. Kazwala, V. P. Maro, S. Cleaveland, and J. A. Crump (2015). Epidemiology of leptospirosis in Africa: a systematic review of a neglected zoonosis and a paradigm for One Health in Africa. *PLoS Neglected Tropical Diseases 9*(9), 1–25.

Allen, T., K. A. Murray, C. Zambrana-Torrelio, S. S. Morse, C. Rondinini, M. Di Marco, N. Breit, K. J. Olival, and P. Daszak (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications 8*(1), 1–10.

Allison, A. B., C. E. Harbison, I. Pagan, K. M. Stucker, J. T. Kaelber, J. D. Brown, M. G. Ruder, M. K. Keel, E. J. Dubovi, E. C. Holmes, and C. R. Parrish (2012). Role of multiple hosts in the cross-species transmission and emergence of a pandemic parvovirus. *Journal of virology 86*(2), 865–72.

Altizer, S., R. Bartel, and B. A. Han (2011). Animal migration and infectious disease risk. *Science 331*(6015), 296–302.

Altizer, S., A. Dobson, P. Hosseini, P. Hudson, M. Pascual, and P. Rohani (2006). Seasonality and the dynamics of infectious diseases. *Ecology letters 9*(4), 467–84.

Altizer, S., D. Harvell, and E. Friedle (2003). Rapid evolutionary dynamics and disease threats to biodiversity. *Trends in Ecology & Evolution 18*(11), 589–596.

Altizer, S., C. L. Nunn, and P. Lindenfors (2007). Do threatened hosts have fewer parasites? A comparative study in primates. *The Journal of Animal Ecology 76*(2), 304–314.

Aluoch, A. M., M. Y. Otiende, M. A. Obonyo, P. G. Mungai, D. O. Okun, S. Angelone-Alasaad, and M. J. Jowers (2017). First genetic identification of Pilobolus (Mucoromycotina, Mucorales) from Africa (Nairobi National Park, Kenya). *South African Journal of Botany 111*, 182–188.

Anderson, E. C. and L. W. Rowe (1998). The prevalence of antibody to the viruses of bovine virus diarrhoea, bovine herpes virus 1, rift valley fever, ephemeral fever and bluetongue and to Leptospira sp in free-ranging wildlife in Zimbabwe. *Epidemiology and Infection 121*(2), 441–449.

Anderson, R. M. and R. M. May (1979). Population biology of infectious diseases: Part I. *Nature 280*(2), 361–367.

Anderson, R. M. and R. M. May (1991). A framework for discussing the population biology of infectious diseases. In *Infectious Diseases of Humans: Dynamics and*

Control, Chapter 2, pp. 13–24. Oxford University Press.

Angerbjörn, A. and M. Tannerfeldt (2014). Vulpes lagopus. The IUCN Red List of Threatened Species 2014 .

Anstead, G. M., D. A. Sutton, and J. R. Graybill (2012). Adiaspiromycosis causing respiratory failure and a review of human infections due to Emmonsia and Chrysosporium spp. *Journal of Clinical Microbiology 50*(4), 1346–1354.

Anthony, S. J., J. H. Epstein, K. A. Murray, I. Navarrete-Macias, C. M. Zambrana-Torrelio, A. Solovyov, R. Ojeda-Flores, N. C. Arrigo, A. Islam, S. A. Khan, P. Hosseini, T. L. Bogich, K. J. Olival, M. D. Sanchez-Leon, W. B. Karesh, T. Goldstein, S. P. Luby, S. S. Morse, J. A. Mazet, P. Daszak, and W. I. Lipkin (2013). A strategy to estimate unknown viral diversity in mammals. *mBio 4*(5).

Antonovics, J., M. Boots, D. Ebert, B. Koskella, M. Poss, and B. M. Sadd (2013). The Origin of Specificity By Means of Natural Selection: Evolved and Nonhost Resistance in Host-Pathogen Interactions. *Evolution 67*(1), 1–9.

Arel-Bundock, V. (2013). *WDI: World Development Indicators (World Bank)*. R package version 2.4.

Armelagos, G. J., K. C. Barnes, and J. Lin (1996). Disease in Human Evolution: the Re-Emergence of Infectious disease in the Thrid Epidemiological Transisition. *National Museum of Natural History 18*, 1–6.

Averbeck, C., M. Plath, T. Wronski, and A. Apio (2012). Effect of human nuisance on the social organisation of large mammals: group sizes and compositions of seven ungulate species in Lake Mburo National Park and the adjacent Ankole Ranching Scheme. *Wildlife Biology 18*(2), 180–193.

Bachand, N., A. Ravel, R. Onanga, J. Arsenault, and J.-P. Gonzalez (2012). Public health significance of zoonotic bacterial pathogens from bushmeat sold in urban markets of Gabon, Central Africa. *Journal of Wildlife Diseases 48*(3), 785–9.

Balboni, A., F. Bassi, S. De Arcangeli, R. Zobba, C. Dedola, A. Alberti, and M. Battilani (2018). Molecular analysis of carnivore Protoparvovirus detected in white blood cells of naturally infected cats. *BMC Veterinary Research 14*(1), 1–10.

Ball, S. G., C. Colleoni, D. Kadouche, M. Ducatez, M. C. Arias, and C. Tirtiaux (2015). Toward an understanding of the function of Chlamydiales in plastid endosymbiosis. *Biochimica et Biophysica Acta 1847*(6-7), 495–504.

Barati, S., N. Moori-Bakhtiari, M. G. Najafabadi, H. Momtaz, and L. Shokuhizadeh (2017). The role of zoonotic chlamydial agents in ruminants abortion. *Iranian Journal of Microbiology 9*(5), 288–294.

Barrett, L. G., J. M. Kniskern, N. Bodenhausen, W. Zhang, and J. Bergelson (2009). Continua of specificity and virulence in plant host-pathogen interactions: Causes and consequences. *New Phytologist 183*(3), 513–529.

Bartomeus, I., D. Gravel, J. M. Tylianakis, M. A. Aizen, I. A. Dickie, and M. Bernard-Verdier (2016). A common framework for identifying linkage rules across different types of interactions. *Functional Ecology 30*(12), 1894–1903.

Baselga, A., D. Orme, S. Villeger, J. De Bortoli, and F. Leprieur (2018). *betapart: Partitioning Beta Diversity into Turnover and Nestedness Components*. R package version 1.5.0.

Beineke, A., W. Baumgärtner, and P. Wohlsein (2015). Cross-species transmission of canine distemper virus–an update. *One Health 1*, 49–59.

Belsley, D. A., E. Kuh, and R. E. Welsch (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Bengis, R. G. and J. M. Erasmus (1988). Wildlife diseases in South Africa: a review. *Revue scientifique et technique (OIE) 7*(4), 807–821.

Betancourt-Alonso, M. A., A. Orihuela, V. Aguirre, R. Vázquez, and I. Flores-Pérez (2011). Changes in behavioural and physiological parameters associated with Taenia

pisiformis infection in rabbits (Oryctolagus cuniculus) that may improve early detection of sick rabbits. *World Rabbit Science 19*(1), 21–30.

Betsem, E., R. Rua, P. Tortevoye, A. Froment, and A. Gessain (2011). Frequent and recent human acquisition of simian foamy viruses through apes' bites in central Africa. *PLoS Pathogens 7*(10), 1–15.

Bhola, N., J. O. Ogutu, M. Y. Said, H.-P. Piepho, and H. Olff (2012). The distribution of large herbivore hotspots in relation to environmental and anthropogenic correlates in the Mara region of Kenya. *The Journal of animal ecology 81*(6), 1268–87.

Biek, R. and L. A. Real (2010). The landscape genetics of infectious disease emergence and spread. *Molecular Ecology 19*(17), 3515–3531.

Birhane, M. G., J. M. Cleaton, B. P. Monroe, A. Wadhwa, L. A. Orciari, P. Yager, J. Blanton, A. Velasco-Villa, B. W. Petersen, and R. M. Wallace (2017). Rabies surveillance in the United States during 2015. *Journal of the American Veterinary Medical Association 250*(10), 1117–1130.

Bisson, I. A., B. J. Ssebide, and P. P. Marra (2015, nov). Early Detection of Emerging Zoonotic Diseases with Animal Morbidity and Mortality Monitoring. *EcoHealth 12*(1), 98–103.

Blaxter, M., J. Mann, T. Chapman, F. Thomas, C. Whitton, R. Floyd, and E. Abebe (2005). Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences 360*(1462), 1935–1943.

Blomberg, S. P., T. Garland, and A. R. Ives (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution 57*(4), 717–745.

Bohmann, K., A. Evans, M. T. P. Gilbert, G. R. Carvalho, S. Creer, M. Knapp, D. W. Yu, and M. de Bruyn (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution 29*(6), 358–367.

Bommineni, Y. R., E. J. Dick, J. S. Estep, J. L. Van de Berg, and G. B. Hubbard (2009). Fatal acute Chagas disease in a chimpanzee. *Journal of Medical Primatology 38*(4),

247–251.

Borda, C. E. and M. J. Rea (2006). Intermediate and definitive hosts of Schistosoma mansoni in Corrientes province, Argentina. *Memorias do Instituto Oswaldo Cruz 101*(Suppl. 1), 233–234.

Borer, E. T., J. Antonovics, L. L. Kinkel, P. J. Hudson, P. Daszak, M. J. Ferrari, K. A. Garrett, C. R. Parrish, A. F. Read, and D. M. Rizzo (2011). Bridging taxonomic and disciplinary divides in infectious disease. *EcoHealth 8*(3), 261–267.

Bornstein, S., T. Mörner, and W. M. Samuel (2002). Sarcoptes Scabiei and Sarcoptic Mange. In W. M. Samuel, Margo J. Pybus, and A. A. Kocan (Eds.), *Parasitic Diseases of Wild Mammals* (2nd ed.)., pp. 107–119. Ames, Iowa: Iowa Sate University Press.

Bouchard, C., G. Beauchamp, S. Nguon, L. Trudel, F. Milord, L. R. Lindsay, D. Bélanger, and N. H. Ogden (2011). Associations between Ixodes scapularis ticks and small mammal hosts in a newly endemic zone in southeastern Canada: implications for Borrelia burgdorferi transmission. *Ticks and Tick-borne Diseases 2*(4), 183–90.

Bouskill, N. J., J. Barker-Finkel, T. S. Galloway, R. D. Handy, and T. E. Ford (2010). Temporal bacterial diversity associated with metal-contaminated river sediments. *Ecotoxicology 19*(2), 317–328.

Braga, M. P., E. Razzolini, and W. A. Boeger (2015). Drivers of parasite sharing among Neotropical freshwater fishes. *Journal of Animal Ecology 84*(2), 487–497.

Brown, S. P., M. E. Hochberg, and B. T. Grenfell (2002). Does multiple infection select for raised virulence? *Trends in Microbiology 10*(9), 401–5.

Browne, A. J., C. A. Guerra, R. V. Alves, V. M. Da Costa, A. L. Wilson, D. M. Pigott, S. I. Hay, S. W. Lindsay, N. Golding, and C. L. Moyes (2017). The contemporary distribution of Trypanosoma cruzi infection in humans, alternative hosts and vectors. *Scientific Data 4*, 1–10.

Brownstein, J. S., C. C. Freifeld, B. Y. Reis, and K. D. Mandl (2008, jul). Surveillance sans frontières: Internet-based emerging infectious disease intelligence and the HealthMap

project. *PLoS Medicine 5*(7), 1019–1024.

Bryan, L. K., S. A. Hamer, S. Shaw, R. Curtis-Robles, L. D. Auckland, C. L. Hodo, K. Chaffin, and R. R. Rech (2016). Chagas disease in a Texan horse with neurologic deficits. *Veterinary Parasitology 216*, 13–17.

Buchmann, K. and F. Mehrdana (2016). Effects of anisakid nematodes Anisakis simplex (s.l.), Pseudoterranova decipiens (s.l.) and Contracaecum osculatum (s.l.) on fish and consumer health. *Food and Waterborne Parasitology 4*, 13–22.

Buckland, P. C. and J. P. Sadler (1989). A Biogeography of the Human Flea, Pulex irritans L. (Siphonaptera: Pulicidae). *Journal of Biogeography 16*(2), 115.

Buhnerkempe, M. G., M. G. Roberts, A. P. Dobson, H. Heesterbeek, P. J. Hudson, and J. O. Lloyd-Smith (2015). Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics 10*, 26–30.

Büscher, P., J. M. Bart, M. Boelaert, B. Bucheton, G. Cecchi, N. Chitnis, D. Courtin, L. M. Figueiredo, J. R. Franco, P. Grébaut, E. Hasker, H. Ilboudo, V. Jamonneau, M. Koffi, V. Lejon, A. MacLeod, J. Masumu, E. Matovu, R. Mattioli, H. Noyes, A. Picado, K. S. Rock, B. Rotureau, G. Simo, S. Thévenon, S. Trindade, P. Truc, and N. Van Reet (2018). Do Cryptic Reservoirs Threaten Gambiense-Sleeping Sickness Elimination? *Trends in Parasitology 34*(3), 197–209.

Cabral, J. P. (2010). Water microbiology. Bacterial pathogens and water. *International Journal of Environmental Research and Public Health 7*(10), 3657–3703.

Callahan, B. J., P. J. McMurdie, and S. P. Holmes (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal 11*(12), 2639–2643.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods 13*(7), 581–583.

Callahan, B. J., K. Sankaran, J. A. Fukuyama, P. J. McMurdie, and S. P. Holmes (2017). Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Research 5*(3), 1492.

Campbell, C., A. Borman, C. Linton, P. Bridge, and E. Johnson (2006). Arthroderma olidum, sp. nov. A new addition to the Trichophyton terrestre complex. *Medical Mycology 44*(5), 451–459.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. a. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. a. Lozupone, D. Mcdonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. a. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight (2010). QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nature Methods 7*(5), 335–336.

Cardillo, M., G. M. Mace, J. L. Gittleman, K. E. Jones, J. Bielby, and A. Purvis (2008). The predictability of extinction: biological and external correlates of decline in mammals. *Proceedings of the Royal Society B: Biological Sciences 275*(1641), 1441–1448.

Carey, A. B. and R. G. Mclean (1983). The Ecology of Rabies: Evidence of Co-Adaptation. *Journal of Applied Ecology 20*(3), 777–800.

Carnieli, P., W. d. O. Fahl, J. G. Castilho, R. d. N. Oliveira, C. I. Macedo, E. Durymanova, R. S. Jorge, R. G. Morato, R. O. Spíndola, L. M. Machado, J. E. de Sá, M. L. Carrieri, and I. Kotait (2008). Characterization of Rabies virus isolated from canids and identification of the main wild canid host in Northeastern Brazil. *Virus Research 131*(1), 33–46.

Carpenter, B., A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan: A probabilistic programming language. *Journal of Statistical Software 76*(1), 1–32.

Carroll, D., P. Daszak, N. D. Wolfe, G. F. Gao, C. M. Morel, S. Morzaria, A. Pablos-Méndez, O. Tomori, and J. A. Mazet (2018). The Global Virome Project. *Science 359*(6378), 872–874.

Carvalho, F. M., R. C. Souza, F. G. Barcellos, M. Hungria, and A. T. R. Vasconcelos (2010). Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order Rhizobiales. *BMC Microbiology 10*.

Chan, E. H., T. F. Brewer, L. C. Madoff, M. P. Pollack, A. L. Sonricker, M. Keller, C. C. Freifeld, M. Blench, A. Mawudeku, and J. S. Brownstein (2010). Global capacity for emerging infectious disease detection. *Proceedings of the National Academy of Sciences 107*(50), 21701–6.

Charleston, M. A. and D. L. Robertson (2002). Preferential host switching by primate lentiviruses can account for phylogenetic similarity with the primate phylogeny. *Systematic Biology 51*(3), 528–535.

Chen, C.-C., K. J.-C. Pei, Y.-C. Lai, and J. A. Mortenson (2012). Participatory Epidemiology To Assess Sarcoptic Mange in Serow of Taiwan. *Journal of Wildlife Diseases 48*(4), 869–875.

Chen, J.-C., Y.-J. Tsai, and Y.-L. Wu (2015). Seroprevalence of Toxoplasma gondii antibodies in wild birds in Taiwan. *Research in Veterinary Science 102*, 184–188.

Chilton, N. B., F. Huby-Chilton, A. V. Koehler, R. B. Gasser, and I. Beveridge (2016). Detection of cryptic species of Rugopharynx (Nematoda: Strongylida) from the stomachs of Australian macropodid marsupials. *International Journal for Parasitology: Parasites and Wildlife 5*(2), 124–133.

Chirima, G. J., N. Owen-Smith, and B. F. Erasmus (2012). Changing distributions of larger ungulates in the Kruger National Park from ecological aerial survey data. *Koedoe 54*(1), 1–11.

Cho, S. H., T. S. Kim, Y. Kong, B. K. Na, and W. M. Sohn (2013). Tetrathyridia of Mesocestoides lineatus in Chinese snakes and their adults recovered from experimental

animals. *Korean Journal of Parasitology 51*(5), 531–536.

Chomel, B. B., A. Belotto, and F. X. Meslin (2007). Wildlife, exotic pets, and emerging zoonoses. *Emerging Infectious Diseases 13*(1), 6–11.

Cleaveland, S., M. K. Laurenson, and L. H. Taylor (2001). Diseases of humans and their domestic mammals: Pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences 356*(1411), 991–999.

Clutton-Brock, J. (1999). *A Natural History of Domesticated Mammals* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Colwell, R. K., R. R. Dunn, and N. C. Harris (2012). Coextinction and Persistence of Dependent Species in a Changing World. *Annual Review of Ecology, Evolution, and Systematics 43*(1), 183–203.

Conner, M. M., M. R. Ebinger, J. A. Blanchong, and P. C. Cross (2008). Infectious disease in cervids of North America: data, models, and management challenges. *Annals of the New York Academy of Sciences 1134*, 146–72.

Cooper, N., R. Griffin, M. Franz, M. Omotayo, and C. L. Nunn (2012). Phylogenetic host specificity and understanding parasite sharing in primates. *Ecology Letters 15*(12), 1370 – 1377.

Counotte, M. J., G. Minbaeva, J. Usubalieva, K. Abdykerimov, and P. R. Torgerson (2016). The burden of zoonoses in Kyrgyzstan: a systematic review. *PLoS Neglected Tropical Diseases 10*(7), 1–15.

Cousins, D. V. (2001). Mycobacterium bovis infection and control in domestic livestock. *Revue scientifique et technique (OIE) 20*(1), 71–85.

Craft, M. and D. Caillaud (2011). Network models: an underutilized tool in wildlife epidemiology? *Interdisciplinary Perspectives on Infectious Diseases 2011*, 1–12.

Cressler, C. E., D. V. McLeod, C. Rozins, J. Van den Hoogen, and T. Day (2016). The adaptive evolution of virulence: a review of theoretical predictions and empirical tests.

Parasitology 143(07), 915–930.

Dallas, T. (2016). HelminthR: An R interface to the London Natural History Museum's Host-Parasite Database. *Ecography 39*(4), 391–393.

Dallas, T., S. Huang, C. Nunn, A. W. Park, and J. M. Drake (2017). Estimating parasite host range. *Proceedings of the Royal Society B: Biological Sciences 284*(1861).

Dalponte, J. and O. Courtenay (2008). Lycalopex vetulus. The IUCN Red List of Threatened Species 2008.

Dalponte, J. C. (2009). Lycalopex vetulus (Carnivora: Canidae). *Mammalian Species 847*, 1–7.

Danesi, P., G. da Rold, A. Rizzoli, H. C. Hauffe, S. Marangon, K. Samerpitak, C. Demanche, J. Guillot, G. Capelli, and S. G. de Hoog (2016). Barcoding markers for Pneumocystis species in wildlife. *Fungal Biology 120*(2), 191–206.

Daszak, P., A. A. Cunningham, and A. D. Hyatt (2000). Emerging Infectious Diseases of Wildlife - Threats to Biodiversity and Human Health. *Science 287*, 443–449.

Daszak, P., A. A. Cunningham, and A. D. Hyatt (2001). Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Tropica 78*(2), 103–16.

Daszak, P., C. Zambrana-Torrelio, T. L. Bogich, M. Fernandez, J. H. Epstein, K. A. Murray, and H. Hamilton (2013). Interdisciplinary approaches to understanding disease emergence: The past, present, and future drivers of Nipah virus emergence. *Proceedings of the National Academy of Sciences 110*, 3681–3688.

Davies, T. J. and A. B. Pedersen (2008). Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings of the Royal Society B: Biological Sciences 275*(1643), 1695–1701.

Dayan, T., D. Simberloff, E. Tchernov, and Y. Yom-Tov (1992). Canine carnassials: character displacement in the wolves, jackals and foxes of Israel. *Biological Journal of the Linnean Society 45*(4), 315–331.

de Castro, F. and B. M. Bolker (2005). Parasite establishment and host extinction in model communities. *Oikos 111*(3), 501–513.

de Souza Lima, E., K. E. Dematteo, R. S. P. Jorge, M. L. S. P. Jorge, J. C. Dalponte, H. S. Lima, and S. A. Klorfine (2012). First telemetry study of bush dogs: Home range, activity and habitat selection. *Wildlife Research 39*(6), 512–519.

de Vienne, D. M., M. E. Hood, and T. Giraud (2009). Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology 22*(12), 2532–2541.

Dehove, A., J. Commault, M. Petitclerc, M. Teissier, and J. Macé (2012). Economic analysis and costing of animal health: a literature review of methods and importance. *Revue scientifique et technique (OIE) 31*(2), 605–17, 591–604.

Deiner, K., H. M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology 26*(21), 5872–5895.

Delgado-Baquerizo, M., F. T. Maestre, P. B. Reich, T. C. Jeffries, J. J. Gaitan, D. Encinar, M. Berdugo, C. D. Campbell, and B. K. Singh (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications 7*, 1–8.

DeMatteo, K. E. (2008). Using a survey of carnivore conservationists to gain new insigth into the ecology and conservation status of the bush dog. *Canid News 11*(3), 1–8.

Deredec, A. and F. Courchamp (2003). Extinction thresholds in host-parasite dynamics. *Annales Zoologici Fennici 40*, 115–130.

Diaz, L. A., V. Ré, W. R. Almirón, A. Farías, A. Vázquez, M. P. Sanchez-Seco, J. Aguilar, L. Spinsanti, B. Konigheim, A. Visintin, J. García, M. A. Morales, A. Tenorio, and M. Contigiani (2006). Genotype III Saint Louis encephalitis virus outbreak, Argentina, 2005. *Emerging Infectious Diseases 12*(11), 1752–1754.

Dinnik, J. A. (1951). An itermediate host of the common stomach fluke, Paramphistomum Cervi (Schrank), in Kenya. *The East African Agricultural Journal 16*(3), 124–125.

Dissanayake, D. M. R. B., C. Stephen, S. Daniel, and P. Abeynayake (2012). Gap assessment of animal health legislation in Sri Lanka for emerging infectious disease preparedness. *Outlook on Agriculture 41*(3), 203–208.

Dobson, A. (2004). Population dynamics of pathogens with multiple host species. *The American Naturalist 164 Suppl*, S64–78.

Dobson, a. and J. Foufopoulos (2001). Emerging infectious pathogens of wildlife. *Philosophical Transactions of the Royal Society B: Biological Sciences 356*(1411), 1001–12.

Dobson, A., K. D. Lafferty, A. M. Kuris, R. F. Hechinger, and W. Jetz (2008). Homage to Linnaeus: How many parasites? How many hosts? *Proceedings of the National Academy of Sciences 105*, 11482–11489.

dos Santos, A. P., R. P. dos Santos, A. W. Biondo, J. M. Dora, L. Z. Goldani, S. T. de Oliveira, A. M. de Sá Guimarães, J. Timenetsky, H. A. de Morais, F. H. González, and J. B. Messick (2008). Hemoplasma infection in HIV-positive patient, Brazil. *Emerging Infectious Diseases 14*(12), 1922–1924.

Dubey, J. P., A. Pas, C. Rajendran, O. C. Kwok, L. R. Ferreira, J. Martins, C. Hebel, S. Hammer, and C. Su (2010). Toxoplasmosis in Sand cats (Felis margarita) and other animals in the Breeding Centre for Endangered Arabian Wildlife in the United Arab Emirates and Al Wabra Wildlife Preservation, the State of Qatar. *Veterinary Parasitology 172*(3-4), 195–203.

Dubey, J. P., G. Schares, and L. M. Ortega-Mora (2007). Epidemiology and control of neosporosis and Neospora caninum. *Clinical Microbiology Reviews 20*(2), 323–367.

Dunn, R. R., T. J. Davies, N. C. Harris, and M. C. Gavin (2010). Global drivers of human pathogen richness and prevalence. *Proceedings of the Royal Society B: Biological Sciences 277*(1694), 2587–2595.

Dunn, R. R., N. C. Harris, R. K. Colwell, L. P. Koh, and N. S. Sodhi (2009). The sixth mass coextinction: are most endangered species parasites and mutualists? *Proceedings of the Royal Society B: Biological Sciences 276*(1670), 3037–3045.

Dunne, J. A., K. D. Lafferty, A. P. Dobson, R. F. Hechinger, A. M. Kuris, N. D. Martinez, J. P. McLaughlin, K. N. Mouritsen, R. Poulin, K. Reise, D. B. Stouffer, D. W. Thieltges, R. J. Williams, and C. D. Zander (2013). Parasites affect food web structure primarily through increased diversity and complexity. *PLoS Biology 11*(6), e1001579.

Durbin, L., A. Venkataraman, S. Hedges, and W. Duckworth (2005). South Asia South of the Himalaya (Oriental). In C. Sillero-Zubiri, M. Hoffmann, and D. Macdonald (Eds.), *Canids: Foxes, Wolves, Jackals and Dogs*, Chapter 8, pp. 210–219. Island Press.

Ebert, D. (1998). Experimental Evolution of Parasites. *Science 282*(5393), 1432–1436.

Eidson, M., S. D. Matthews, A. L. Willsey, B. Cherry, R. J. Rudd, and C. V. Trimarchi (2005). Rabies virus infection in a pet guinea pig and seven pet rabbits. *Journal of the American Veterinary Medical Association 227*(6), 932–935.

Eisenberg, J. N., M. A. Desai, K. Levy, S. J. Bates, S. Liang, K. Naumoff, and J. C. Scott (2007, aug). Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. *Environmental Health Perspectives 115*(8), 1216–1223.

Eklöf, A., U. Jacob, J. Kopp, J. Bosch, R. Castro-Urgal, N. P. Chacoff, B. Dalsgaard, C. de Sassi, M. Galetti, P. R. Guimarães, S. B. Lomáscolo, A. M. Martín González, M. A. Pizo, R. Rader, A. Rodrigo, J. M. Tylianakis, D. P. Vázquez, and S. Allesina (2013). The dimensionality of ecological networks. *Ecology Letters 16*(5), 577–583.

Elmasri, M., M. Farrell, and D. A. Stephens (2017). A hierarchical Bayesian model for predicting ecological interactions using evolutionary relationships. *arXiv preprint: 1707.08354*, 1–30.

Elmore, S. A., J. L. Jones, P. A. Conrad, S. Patton, D. S. Lindsay, and J. P. Dubey (2010). Toxoplasma gondii: epidemiology, feline clinical aspects, and prevention. *Trends in*

Parasitology 26(4), 190–196.

Ewald, P. (1983). Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics 14*(1983), 465–485.

Ezenwa, V. O. (2003). Habitat overlap and gastrointestinal parasitism in sympatric African bovids. *Parasitology 126*(4), 379–388.

Ezenwa, V. O., S. A. Price, S. Altizer, N. D. Vitone, and K. C. Cook (2006). Host traits and parasite species richness in even and odd-toed hoofed mammals, Artiodactyla and Perissodactyla. *Oikos 115*(3), 526–536.

Farrell, M. J., L. Berrang-Ford, and T. J. Davies (2013). The study of parasite sharing for surveillance of zoonotic diseases. *Environmental Research Letters 8*(1), 015036.

Faust, C. L., H. I. McCallum, L. S. Bloomfield, N. L. Gottdenker, T. R. Gillespie, C. J. Torney, A. P. Dobson, and R. K. Plowright (2018). Pathogen spillover during land conversion. *Ecology Letters 21*(4), 471–483.

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist 125*(1), 1–15.

Fera, M. T., T. L. Maugeri, C. Gugliandolo, C. Beninati, E. L. Camera, M. Carbone, and M. Giannone (2004). Detection of Arcobacter spp. in the coastal environment of the Mediterranean Sea. *Applied and Environmental Microbiology 70*(3), 1271–1276.

Fèvre, E. M., B. M. C. Bronsvoort, K. A. Hamilton, and S. Cleaveland (2006). Animal movements and the spread of infectious diseases. *Trends in Microbiology 14*(3), 125–131.

Figuerola, J. and A. Green (2000). Haematozoan parasites and migratory behaviour in waterfowl. *Evolutionary Ecology 14*(2), 143–153.

Fish, R., Z. Austin, R. Christley, P. M. Haygarth, L. A. Heathwaite, S. Latham, W. Medd, M. Mort, D. M. Oliver, R. Pickup, J. M. Wastling, and B. Wynne (2011). Uncertainties in the governance of animal disease: An interdisciplinary framework for analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences 366*(1573),

2023–2034.

Fisher, B. and T. Christopher (2007). Poverty and biodiversity: Measuring the overlap of human poverty and the biodiversity hotspots. *Ecological Economics 62*(1), 93–101.

Fisher, J. C., R. J. Newton, D. K. Dila, and S. L. McLellan (2015). Urban microbial ecology of a freshwater estuary of Lake Michigan. *Elementa: Science of the Anthropocene 3*(000064), 1–14.

Fitzpatrick, J. L., J. L. Dyer, J. D. Blanton, I. V. Kuzmin, and C. E. Rupprecht (2014). Rabies in rodents and lagomorphs in the United States, 1995-2010. *Journal of the American Veterinary Medical Association 245*(3), 333–337.

Flanagan, M. L., C. R. Parrish, S. Cobey, G. E. Glass, R. M. Bush, and T. J. Leighton (2012). Anticipating the species jump: surveillance for emerging viral threats. *Zoonoses and public health 59*(3), 155–63.

Fleming, P. J., H. Nolan, S. M. Jackson, G. A. Ballard, A. Bengsen, W. Y. Brown, P. D. Meek, G. Mifsud, S. K. Pal, and J. Sparkes (2017). Roles for the Canidae in food webs reviewed: Where do they fit? *Food Webs 12*, 14–34.

Fonjungo, P. N., G. A. Alemnji, Y. Kebede, A. Opio, C. Mwangi, T. J. Spira, R. S. Beard, and J. N. Nkengasong (2017). Combatting Global Infectious Diseases: A Network Effect of Specimen Referral Systems. *Clinical Infectious Diseases 64*(6), 796–803.

Ford, T. E., R. R. Colwell, J. B. Rose, S. S. Morse, D. J. Rogers, and T. L. Yates (2009). Using satellite images of environmental changes to predict infectious disease outbreaks. *Emerging Infectious Diseases 15*(9), 1341–6.

Fowler, M. (1996). Husbandry and diseases of camelids. *Revue scientifique et technique (OIE) 15*(1), 155–169.

Fritz, S. A., O. R. P. Bininda-Emonds, and A. Purvis (2009). Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters 12*(6), 538–549.

Fuehrer, H. P. (2014). An overview of the host spectrum and distribution of Calodium hepaticum (syn. Capillaria hepatica): Part 2 - Mammalia (excluding Muroidea). *Parasitology Research 113*(2), 641–651.

Gandon, S. (2004). Evolution of Multihost Parasites. *Evolution 58*(3), 455.

Ganz, H. H., U. Karaoz, W. M. Getz, W. Versfeld, and E. L. Brodie (2012). Diversity and structure of soil bacterial communities associated with vultures in an African savanna. *Ecosphere 3*(6), 1–18.

Garamszegi, L. Z. (2006). The evolution of virulence and host specialization in malaria parasites of primates. *Ecology Letters 9*(8), 933–940.

Gardy, J. L. and N. J. Loman (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics 19*(1), 9–20.

Gautret, P., J. Blanton, L. Dacheux, F. Ribadeau-Dumas, P. Brouqui, P. Parola, D. H. Esposito, and H. Bourhy (2014). Rabies in nonhuman primates and potential for transmission to humans: a literature review and examination of selected French national data. *PLoS Neglected Tropical Diseases 8*(5), 1–7.

Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics 2*(4), 1360–1383.

Geoghegan, J. L. and E. C. Holmes (2017). Predicting virus emergence amid evolutionary noise. *Open Biology 7*(10), 170189.

Ghai, R. R., C. A. Chapman, P. A. Omeja, T. J. Davies, and T. L. Goldberg (2014). Nodule worm infection in humans and wild primates in Uganda: cryptic species in a newly identified region of human transmission. *PLoS Neglected Tropical Diseases 8*(1), 39.

Ghosh, J., Y. Li, and R. Mitra (2015). On the use of Cauchy prior distributions for Bayesian logistic regression. *arXiv preprint 1507.07170*, 1–59.

Giangaspero, M. (2017). Trypanosoma cruzi and domestic animals. *Clinical Microbiology: Open Access 06*(03), 20–22.

Gibson, D. I., R. A. Bray, and E. A. C. Harris (2005). Host-Parasite Database of the Natural History Museum, London.

Gilbert, G. S. and C. O. Webb (2007). Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Sciences 104*(12), 4979–4983.

Gittleman, J. L. (1989). Carnivore group living: comparative trends. In J. Gittleman (Ed.), *Carnivore behavior, ecology and evolution* (1 ed.)., pp. 183–207. Ithaca: Cornell University Press.

Gleason, F. H. and A. V. Marano (2011). The effects of antifungal substances on some zoosporic fungi (Kingdom Fungi). *Hydrobiologia 659*(1), 81–92.

Goldberg, T. L., T. R. Gillespie, I. B. Rwego, E. Wheeler, E. L. Estoff, and C. a. Chapman (2007). Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation 135*(4), 511–517.

Gómez, A. and A. A. Aguirre (2008). Infectious diseases and the illegal wildlife trade. *Annals of the New York Academy of Sciences 1149*, 16–19.

Gómez, A. and E. Nichols (2013). Neglected wild life: Parasitic biodiversity as a conservation target. *International Journal for Parasitology: Parasites and Wildlife 2*, 222–227.

Gomez, J. M., C. L. Nunn, and M. Verdu (2013). Centrality in primate-parasite networks reveals the potential for the transmission of emerging infectious diseases to humans. *Proceedings of the National Academy of Sciences 110*(19), 7738–7741.

Gonçalves, R. (2007). Does Mycoplasma mycoides subsp. mycoides LC cause human infection? Technical report, Instituto Nacional de Investigação Agrária e Veterinária (INIAV).

Grace, D., J. Gilbert, T. Randolph, and E. Kang'ethe (2012). The multiple burdens of zoonotic disease and an ecohealth approach to their assessment. *Tropical Animal Health and Production 44*, 67–73.

Graham, A. L., J. E. Allen, and A. F. Read (2005). Evolutionary Causes and Consequences of Immunopathology. *Annual Review of Ecology, Evolution, and Systematics 36*(1), 373–397.

Gravel, D., T. Poisot, C. Albouy, L. Velez, and D. Mouillot (2013). Inferring food web structure from predator-prey body size relationships. *Methods in Ecology and Evolution 4*(11), 1083–1090.

Greger, M. (2007). The human/animal interface: emergence and resurgence of zoonotic infectious diseases. *Critical reviews in microbiology 33*(4), 243–99.

Gregory, R. (1990). Parasites and host geographic range as illustrated by waterfowl. *Functional Ecology 4*(5), 645–654.

Gubernot, D. M., B. L. Boyer, and M. S. Moses (2008). Animals as early detectors of bioevents: veterinary tools and a framework for animal-human integrated zoonotic disease surveillance. *Public health reports 123*(3), 300–15.

Guk, S.-M., J.-L. Kim, J.-H. Park, and J.-Y. Chai (2007). A Human Case of Plagiorchis vespertilionis (Digenea: Plagiorchiidae) Infection in the Republic of Korea. *Journal of Parasitology 93*(5), 1225–1227.

Gürtler, R. E. and M. V. Cardinal (2015). Reservoir host competence and the role of domestic and commensal hosts in the transmission of Trypanosoma cruzi. *Acta Tropica 151*(1), 32–50.

Haller, S. L., C. Peng, G. McFadden, and S. Rothenburg (2014). Poxviruses and the evolution of host range and virulence. *Infection, Genetics and Evolution 21*, 15–40.

Halliday, J., C. Daborn, H. Auty, Z. Mtema, T. Lembo, B. M. C. Bronsvoort, I. Handel, D. Knobel, K. Hampson, and S. Cleaveland (2012). Bringing together emerging and endemic zoonoses surveillance: Shared challenges and a common solution. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*(1604), 2872–2880.

Halliday, J. E., A. L. Meredith, D. L. Knobel, D. J. Shaw, B. M. C. Bronsvoort, and S. Cleaveland (2007). A framework for evaluating animals as sentinels for infectious

disease surveillance. *Journal of the Royal Society Interface 4*(16), 973–984.

Halliday, J. E. B., K. Hampson, N. Hanley, T. Lembo, J. P. Sharp, D. T. Haydon, and S. Cleaveland (2017). Relevant Capacity Strengthening. *Science 357*, 146–148.

Hampson, K., L. Coudeville, T. Lembo, M. Sambo, A. Kieffer, M. Attlan, J. Barrat, J. D. Blanton, D. J. Briggs, S. Cleaveland, P. Costa, C. M. Freuling, E. Hiby, L. Knopf, F. Leanes, F. X. Meslin, A. Metlin, M. E. Miranda, T. Müller, L. H. Nel, S. Recuenco, C. E. Rupprecht, C. Schumacher, L. Taylor, M. A. N. Vigilato, J. Zinsstag, and J. Dushoff (2015). Estimating the Global Burden of Endemic Canine Rabies. *PLoS Neglected Tropical Diseases 9*(4), 1–20.

Han, B. A., J. P. Schmidt, L. W. Alexander, S. E. Bowden, D. T. Hayman, and J. M. Drake (2016). Undiscovered bat hosts of filoviruses. *PLoS Neglected Tropical Diseases 10*(7), 1–10.

Han, B. A., J. P. Schmidt, S. E. Bowden, and J. M. Drake (2015). Rodent reservoirs of future zoonotic diseases. *Proceedings of the National Academy of Sciences 112*(22), 7039–7044.

Harmon, L., J. Weir, C. Brock, R. Glor, and W. Challenger (2008). geiger: investigating evolutionary radiations. *Bioinformatics 24*, 129–131.

Harmon, L. J., J. B. Losos, T. Jonathan Davies, R. G. Gillespie, J. L. Gittleman, W. Bryan Jennings, K. H. Kozak, M. A. McPeek, F. Moreno-Roark, T. J. Near, A. Purvis, R. E. Ricklefs, D. Schluter, J. A. Schulte, O. Seehausen, B. L. Sidlauskas, O. Torres-Carvajal, J. T. Weir, and A. T. Mooers (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution 64*(8), 2385–2396.

Harris, N. C. and R. R. Dunn (2010). Using host associations to predict spatial patterns in the species richness of the parasites of North American carnivores. *Ecology letters 13*(11), 1411–8.

Harvey, P. H. and M. D. Pagel (1991). *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.

Hayward, M. W. (2009). The need to rationalize and prioritize threatening processes used to determine threat status in the IUCN Red List. *Conservation Biology 23*(6), 1568–1576.

Heard, M. J., K. F. Smith, K. J. Ripp, M. Berger, J. Chen, J. Dittmeier, M. Goter, S. T. McGarvey, and E. Ryan (2013). The threat of disease increases as species move toward extinction. *Conservation Biology 27*(6), 1378–1388.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. DeWaard (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences 270*(1512), 313–321.

Herrero, M., P. K. Thornton, P. Gerber, and R. S. Reid (2009). Livestock, livelihoods and the environment: understanding the trade-offs. *Current Opinion in Environmental Sustainability 1*(2), 111–120.

Højsgaard, S., U. Halekoh, and J. Yan (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software 15*(2), 1–11.

Hopkins, M. E. and C. L. Nunn (2007). A global gap analysis of infectious agents in wild primates. *Diversity and Distributions 13*(5), 561–572.

Hotez, P. J., M. Alvarado, M. G. Basáñez, I. Bolliger, R. Bourne, M. Boussinesq, S. J. Brooker, A. S. Brown, G. Buckle, C. M. Budke, H. Carabin, L. E. Coffeng, E. M. Fèvre, T. Fürst, Y. A. Halasa, R. Jasrasaria, N. E. Johns, J. Keiser, C. H. King, R. Lozano, M. E. Murdoch, S. O'Hanlon, S. D. Pion, R. L. Pullan, K. D. Ramaiah, T. Roberts, D. S. Shepard, J. L. Smith, W. A. Stolk, E. A. Undurraga, J. Utzinger, M. Wang, C. J. Murray, and M. Naghavi (2014). The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases. *PLoS Neglected Tropical Diseases 8*(7).

Hotez, P. J., L. Savioli, and A. Fenwick (2012). Neglected tropical diseases of the Middle East and North Africa: review of their prevalence, distribution, and opportunities for control. *PLoS Neglected Tropical Diseases 6*(2).

Huang, S., O. R. P. Bininda-Emonds, P. R. Stephens, J. L. Gittleman, and S. Altizer (2014). Phylogenetically related and ecologically similar carnivores harbour similar parasite assemblages. *Journal of Animal Ecology 83*(3), 671–680.

Huang, S., J. M. Drake, J. L. Gittleman, and S. Altizer (2015). Parasite diversity declines with host evolutionary distinctiveness: A global analysis of carnivores. *Evolution 69*(3), 621–630.

Hudson, P. J., A. P. Dobson, and K. D. Lafferty (2006). Is a healthy ecosystem one that is rich in parasites? *Trends in Ecology & Evolution 21*(7), 381–385.

International Livestock Research Institute (2012). Mapping of poverty and likely zoonoses hotspots. *Zoonoses Report 4*, 1–119.

International Union for the Conservation of Nature (2014). IUCN Red List of Threatened Species Version 2014.3.

Ito, A. and C. M. Budke (2014). Culinary delights and travel? A review of zoonotic cestodiases and metacestodiases. *Travel Medicine and Infectious Disease 12*(6), 582–591.

IUCN SSC Antelope Specialist Group (2016). Kobus vardonii, The IUCN Red List of Threatened Species 2016.

Jenkins, D. J. and C. N. Macpherson (2003). Transmission ecology of Echinococcus in wild-life in Australia and Africa. *Parasitology 127*, S63–S72.

Johnson, P. (2013). Host and parasite diversity jointly control disease risk in complex communities. *Proceedings of the National Academy of Sciences 110*(42), 16916–16921.

Jones, B. A., D. Grace, R. Kock, S. Alonso, J. Rushton, M. Y. Said, D. McKeever, F. Mutua, J. Young, J. McDermott, and D. U. Pfeiffer (2013, may). Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences 110*(21), 8399–8404.

Jones, K., J. Bielby, M. Cardillo, and S. Fritz (2009). PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals.

Ecology 90(9), 2648.

Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and P. Daszak (2008, feb). Global trends in emerging infectious diseases. *Nature 451*(7181), 990–993.

Jongejan, F. and G. Uilenberg (1994). Ticks and control methods. *Revue scientifique et technique (OIE) 13*(4), 1201–1226.

Jordaan, K. and C. C. Bezuidenhout (2013). The impact of physico-chemical water quality parameters on bacterial diversity in the Vaal River, South Africa. *Water SA 39*(3), 385–396.

Jordaan, K. and C. C. Bezuidenhout (2016). Bacterial community composition of an urban river in the North West Province, South Africa, in relation to physico-chemical water quality. *Environmental Science and Pollution Research 23*(6), 5868–5880.

Jorge, R., E. Lima, and L. Lucarts (2008). Sarna sarcóptica ameaçando cachorros-vinagres (Speothos venaticus) de vida livre em Nova XavantinaMT. In *XXXIII Congresso Anual da Sociedade de Zoológicos do Brasil, At Sorocaba, SP, Brasil*.

Jorge, R. S. P., M. S. Pereira, R. G. Morato, K. C. Scheffer, P. Carnieli, F. Ferreira, M. M. Furtado, C. K. Kashivakura, L. Silveira, A. T. A. Jacomo, E. S. Lima, R. C. de Paula, and J. A. May-Junior (2010). Detection of rabies virus antibodies in Brazilian free-ranging wild carnivores. *Journal of Wildlife Diseases 46*(4), 1310–1315.

Joshi, D. (1982). Yak and Chauri Husbandry in Nepal. *H.M. Government Press, XVII*, 145.

Juan, C. Durn. R., M. F. C. (1989). Rabia en zorro gris (pseudalopex griseus) patagnico, magallanes, chile. *Avances en Ciencias Veterinarias 4*(2).

Kagira, J. M., P. N. Kanyari, N. Maingi, S. M. Githigia, C. Ng'ang'a, and J. Gachohi (2013). Relationship between the prevalence of ectoparasites and associated risk factors in free-range pigs in Kenya. *ISRN Veterinary Science*, 650890.

Kamins, A. O., O. Restif, Y. Ntiamoa-Baidu, R. Suu-Ire, D. T. Hayman, A. A. Cunningham, J. L. Wood, and J. M. Rowcliffe (2011, dec). Uncovering the fruit bat bushmeat

commodity chain and the true extent of fruit bat hunting in Ghana, West Africa. *Biological Conservation 144*(12), 3000–3008.

Kamiya, T., K. O'Dwyer, S. Nakagawa, and R. Poulin (2014). What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts. *Biological Reviews of the Cambridge Philosophical Society 89*, 123–134.

Kembel, S., P. Cowan, M. Helmus, W. Cornwell, H. Morlon, D. Ackerly, S. Blomberg, and C. Webb (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics 26*, 1463–1464.

Koh, L. P., R. R. Dunn, N. S. Sodhi, R. K. Colwell, H. C. Proctor, and V. S. Smith (2004). Species coextinctions and the biodiversity crisis. *Science 305*(5690), 1632–1634.

Kopp, A., T. R. Gillespie, D. Hobelsberger, A. Estrada, J. M. Harper, R. A. Miller, I. Eckerle, M. A. Müller, L. Podsiadlowski, F. H. Leendertz, C. Drosten, and S. Junglen (2013). Provenance and geographic spread of St. Louis encephalitis virus. *mBio 4*(3), 1–11.

Kruse, H., A. M. Kirkemo, and K. Handeland (2004). Wildlife as source of zoonotic infections. *Emerging Infectious Diseases 10*(12), 2067–2072.

Laakkonen, J. (1998). Pneumocystis carinii in wildlife. *International Journal for Parasitology 28*(2), 241–252.

Lafferty, K. (2009). The ecology of climate change and infectious diseases. *Ecology 90*(4), 888–900.

Lafferty, K. D. (2012). Biodiversity loss decreases parasite diversity: Theory and patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*(1604), 2814–2827.

Lafferty, K. D. and A. M. Kuris (2002). Trophical strategies, animal diversity and body size. *Trends in Ecology & Evolution 17*(11), 507–513.

Laforest-Lapointe, I., A. Paquette, C. Messier, and S. W. Kembel (2017). Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature 546*(7656),

145–147.

Lahaye, R., M. van der Bank, D. Bogarin, J. Warner, F. Pupulin, G. Gigot, O. Maurin, S. Duthoit, T. G. Barraclough, and V. Savolainen (2008). DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences 105*(8), 2923–2928.

Laidemitt, M. R., E. T. Zawadzki, S. V. Brant, M. W. Mutuku, G. M. Mkoji, and E. S. Loker (2017). Loads of trematodes: Discovering hidden diversity of paramphistomoids in Kenyan ruminants. *Parasitology 144*(2), 131–147.

Lebov, J., K. Grieger, D. Womack, D. Zaccaro, N. Whitehead, B. Kowalcyk, and P. D. MacDonald (2017). A framework for One Health research. *One Health 3*, 44–50.

Lee, L. M., R. S. Wallace, V. L. Clyde, A. Gendron-fitzpatrick, S. D. Sibley, M. Stuchin, M. Lauck, D. H. O. Connor, M. Nakao, A. Lavikainen, E. P. Hoberg, and T. L. Goldberg (2016). Definitive hosts of Versteria Tapeworms (Cestoda: Taeniidae) causing fatal infection in North America. *Emerging Infectious Diseases 22*(4), 707–710.

Lefèvre, P.-C., J. Blancou, R. Chermette, and G. Uilenberg (2010). *Infectious and parasitic diseases of livestock*. Cachan, France: Lavoisier - Editions Médicales Internationales.

Lefkowitz, E. J., D. M. Dempsey, R. C. Hendrickson, R. J. Orton, S. G. Siddell, and D. B. Smith (2018). Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research 46*(D1), D708–D717.

Leggett, H. C., A. Buckling, G. H. Long, and M. Boots (2013). Generalism and the evolution of parasite virulence. *Trends in Ecology & Evolution 28*(10), 592–596.

Lerner, J., V. Brumberger, and V. Preac-Mursic (1994). Severe diarrhea associated with Arcobacter butzleri. *European Journal of Clinical Microbiology & Infectious Diseases 13*(8), 660–662.

Ley, R. E., M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight, and J. I. Gordon (2008). Evolution of mammals and their gut microbes. *Science 320*(5883), 1647–1651.

Li, K., J. Gao, M. Shahzad, Z. Han, F. Nabi, M. Liu, D. Zhang, and J. Li (2014). Seroprevalence of Toxoplasma gondii infection in yaks (Bos grunniens) on the Qinghai-Tibetan Plateau of China. *Veterinary Parasitology 205*(1-2), 354–356.

Lindenfors, P., C. L. Nunn, K. E. Jones, A. A. Cunningham, W. Sechrest, and J. L. Gittleman (2007). Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density. *Global Ecology and Biogeography 16*(4), 496–509.

Liu, G., C. M. Tang, and R. M. Exley (2015). Non-pathogenic Neisseria: members of an abundant, multi-habitat, diverse genus. *Microbiology 161*(7), 1297–1312.

Lloyd-Smith, J. (2013). Vacated niches, competitive release and the community ecology of pathogen eradication. *Philosophical Transactions of the Royal Society B: Biological Sciences 368*, 20120150.

Locatelli, S. and M. Peeters (2012). Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *AIDS 26*(6), 659–73.

Lodge, D., C. Turner, and C. Jerde (2012). Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. *Molecular Ecology 21*(11), 2555–8.

Longdon, B., M. A. Brockhurst, C. A. Russell, J. J. Welch, and F. M. Jiggins (2014). The evolution and genetics of virus host shifts. *PLoS Pathogens 10*(11), e1004395.

Longdon, B., J. D. Hadfield, J. P. Day, S. C. L. Smith, J. E. McGonigle, R. Cogni, C. Cao, and F. M. Jiggins (2015). The causes and consequences of changes in virulence following pathogen host shifts. *PLoS Pathogens 11*(3), e1004728.

Longdon, B., J. D. Hadfield, C. L. Webster, D. J. Obbard, and F. M. Jiggins (2011, sep). Host phylogeny determines viral persistence and replication in novel hosts. *PLoS Pathogens 7*(9), e1002260.

Loveridge, A. J. and D. W. Macdonald (2003). Niche separation in sympatric jackals (Canis mesomelas and Canis adustus). *Journal of Zoology 259*(2), 143–153.

Loy, A., C. Schulz, S. Lücker, A. Schöpfer-Wendels, K. Stoecker, C. Baranyi, A. Lehner, and M. Wagner (2005). 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order Rhodocyclales. *Applied and Environmental Microbiology 71*(3), 1373–1386.

Lozupone, C. A. and R. Knight (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences 104*(27), 11436–11440.

Lucherini, M. and E. M. Luengos Vidal (2008). Lycalopex gymnocercus (Carnivora: Canidae). *Mammalian Species 820*(820), 1–9.

Luis, A. D., T. J. O'Shea, D. T. Hayman, J. L. Wood, A. A. Cunningham, A. T. Gilbert, J. N. Mills, and C. T. Webb (2015). Network analysis of host-virus communities in bats and rodents reveals determinants of cross-species transmission. *Ecology Letters 18*(11), 1153–1162.

Luque, J. a. D., H. Müller, L. González, and I. Berkunsky (2014). Clinical signs suggestive of mange infestation in a free-ranging maned wolf (Chrysocyon brachyurus) in the Moxos Savannhs of Beni, Bolivia. *Mastozoologia Neotropical 21*(1), 135–138.

Lysnyansky, I., S. Levisohn, M. Bernstein, G. Kenigswald, S. Blum, I. Gerchman, D. Elad, and J. Brenner (2007). Diagnosis of Mycoplasma conjunctivae and Chlamydophila spp in an episode of conjunctivitis/keratoconjunctivitis in lambs. *Israel Journal of Veterinary Medicine 62*(3-4), 79–82.

Ma, M., X. Jiang, Q. Wang, M. Ongena, D. Wei, J. Ding, D. Guan, F. Cao, B. Zhao, and J. Li (2018). Responses of fungal community composition to long-term chemical and organic fertilization strategies in Chinese Mollisols. *MicrobiologyOpen e597*, 1–12.

Mace, G. M., N. J. Collar, K. J. Gaston, C. Hilton-Taylor, H. R. Akçakaya, N. Leader-Williams, E. J. Milner-Gulland, and S. N. Stuart (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology 22*(6),

1424–1442.

Mariano, V., C. M. E. McCrindle, B. Cenci-Goga, and J. A. Picard (2009). Case-control study to determine whether river water can spread tetracycline resistance to unexposed impala (Aepyceros melampus) in Kruger National Park (South Africa). *Applied and Environmental Microbiology 75*(1), 113–118.

Marinkelle, C. J. (1982). The prevalence of Trypanosoma (Schizotrypanum) cruzi cruzi infection in Colombian monkeys and marmosets. *Annals of Tropical Medicine & Parasitology 76*(2), 121–124. PMID: 6807227.

Martin, C., P.-P. Pastoret, B. Brochier, M.-F. Humblet, and C. Saegerman (2011). A survey of the transmission of infectious diseases/infections between wild and domestic ungulates in Europe. *Veterinary research 42*(1), 70.

Massolo, A., S. Liccioli, C. Budke, and C. Klein (2014). Echinococcus multilocularis in North America: the great unknown. *Parasite 21*, 73.

Maudlin, I., M. C. Eisler, and S. C. Welburn (2009). Neglected and endemic zoonoses. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*(1530), 2777–87.

May, R. M. and R. M. Anderson (1979). Population biology of infectious diseases: Part II. *Nature 280*(9), 455–416.

Mazel, F., T. J. Davies, L. Gallien, J. Renaud, M. Groussin, T. Münkemüller, and W. Thuiller (2016). Influence of tree shape and evolutionary time-scale on phylogenetic diversity metrics. *Ecography 39*(10), 913–920.

McAllister, C. T., C. R. Bursey, and J. F. Roberts (2004). Physocephalus sexalatus (Nematoda: Spirurida: Spirocercidae) in three species of rattlesnakes, Crotalus atrox, Crotalus lepidus, and Crotalus scutulatus, from southwestern Texas. *Journal of Herpetological Medicine and Surgery 14*(3), 10–12.

McCallum, H., N. Barlow, and J. Hone (2001). How should pathogen transmission be modelled? *Trends in Ecology & Evolution 16*(6), 295–300.

McDermott, J. J., D. Grace, and J. Zinsstag (2013). Economics of brucellosis impact and control in low-income countries. *Revue technique et scientifique (OIE) 32*(1), 249–261.

Mcgarry, J. W., A. Higgins, N. G. White, K. C. Pounder, and U. Hetzel (2015). Zoonotic helminths of urban brown rats (Rattus norvegicus) in the UK: Neglected public health considerations? *Zoonoses and Public Health 62*(1), 44–52.

McMurdie, P. J. and S. Holmes (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE 8*(4).

Mei, C., H. Wang, W. Zhu, H. Wang, G. Cheng, K. Qu, X. Guang, A. Li, C. Zhao, W. Yang, C. Wang, Y. Xin, and L. Zan (2016). Whole-genome sequencing of the endangered bovine species Gayal (Bos frontalis) provides new insights into its genetic features. *Scientific Reports 6*, 1–8.

Mennerat, A., F. Nilsen, D. Ebert, and A. Skorping (2010). Intensive farming: Evolutionary implications for parasites and pathogens. *Evolutionary Biology 37*(2-3), 59–67.

Mercier, A., D. Ajzenberg, S. Devillard, M. P. Demar, B. de Thoisy, H. Bonnabau, F. Collinet, R. Boukhari, D. Blanchet, S. Simon, B. Carme, and M. L. Dardé (2011). Human impact on genetic diversity of Toxoplasma gondii: Example of the anthropized environment from French Guiana. *Infection, Genetics and Evolution 11*(6), 1378–1387.

Michel, A. L. and R. G. Bengis (2012). The African buffalo: A villain for inter-species spread of infectious diseases in southern Africa. *Onderstepoort Journal of Veterinary Research 79*, 1–5.

Michel, A. L., L.-M. de Klerk, N. van Pittius, R. M. Warren, and P. D. van Helden (2007). Bovine tuberculosis in African buffaloes: observations regarding Mycobacterium bovis shedding into water and exposure to environmental mycobacteria. *BMC Veterinary Research 3*(1), 23.

Mihalca, A. D., C. M. Gherman, and V. Cozma (2011). Coendangered hard-ticks: threatened or threatening? *Parasites & Vectors 4*(71), 1–7.

Miller, R. S., M. L. Farnsworth, and J. L. Malmberg (2013). Diseases at the livestock-wildlife interface: Status, challenges, and opportunities in the United States. *Preventive Veterinary Medicine 110*(2), 119–132.

Moir, M. L., P. A. Vesk, K. E. C. Brennan, D. A. Keith, L. Hughes, and M. A. McCarthy (2010). Current constraints and future directions in estimating coextinction. *Conservation Biology 24*(3), 682–690.

Moir, M. L., P. a. Vesk, K. E. C. Brennan, R. Poulin, L. Hughes, D. a. Keith, M. a. McCarthy, and D. J. Coates (2012). Considering extinction of dependent species during translocation, ex situ conservation, and assisted migration of threatened hosts. *Conservation Biology 26*(2), 199–207.

Mollentze, N., R. Biek, and D. G. Streicker (2014). The role of viral evolution in rabies host shifts and emergence. *Current Opinion in Virology 8*, 68–72.

Molyneux, D., Z. Hallaj, G. T. Keusch, D. P. McManus, H. Ngowi, S. Cleaveland, P. Ramos-Jimenez, E. Gotuzzo, K. Kar, A. Sanchez, A. Garba, H. Carabin, A. Bassili, C. L. Chaignat, F. X. Meslin, H. M. Abushama, A. L. Willingham, and D. Kioy (2011, jan). Zoonoses and marginalised infectious diseases of poverty: Where do we stand? *Parasites and Vectors 4*(1), 106.

Molyneux, D. and D. P. Sankara (2017). Guinea worm eradication: Progress and challenges should we beware of the dog? *PLoS Neglected Tropical Diseases 11*(4), 4–7.

Morales-Castilla, I., M. G. Matias, D. Gravel, and M. B. Araújo (2015). Inferring biotic interactions from proxies. *Trends in Ecology & Evolution 30*(6), 347–356.

Morens, D. M., G. K. Folkers, and A. S. Fauci (2004). The challenge of emerging and re-emerging infectious diseases. *Nature 430*(6996), 242–249.

Morse, S. S., J. a. K. Mazet, M. Woolhouse, C. R. Parrish, D. Carroll, W. B. Karesh, C. Zambrana-Torrelio, W. I. Lipkin, and P. Daszak (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet 380*(9857), 1956–65.

Mouillot, D., B. R. Krasnov, G. I. Shenbrot, K. J. Gaston, and R. Poulin (2006). Conservatism of host specificity in parasites. *Ecography 29*(4), 596–602.

Murphy, T. M., J. O'Connell, M. Berzano, C. Dold, J. D. Keegan, A. McCann, D. Murphy, and N. M. Holden (2012). The prevalence and distribution of Alaria alata, a potential zoonotic parasite, in foxes in Ireland. *Parasitology Research 111*(1), 283–290.

Mwirichia, R., S. Cousin, A. W. Muigai, and E. Boga, H. I. Stackebrandt (2011). Bacterial diversity in the haloalkaline Lake Elmenteita, Kenya. *Current Microbiology 62*(1), 209–221.

Nakao, M., A. Lavikainen, T. Iwaki, V. Haukisalmi, S. Konyaev, Y. Oku, M. Okamoto, and A. Ito (2013). Molecular phylogeny of the genus Taenia (Cestoda: Taeniidae): Proposals for the resurrection of Hydatigera Lamarck, 1816 and the creation of a new genus Versteria. *International Journal for Parasitology 43*(6), 427–437.

Nemeth, N. M., M. G. Ruder, R. W. Gerhold, J. D. Brown, B. A. Munk, P. T. Oesterle, S. V. Kubiski, and M. K. Keel (2014). Demodectic mange, dermatophilosis, and other parasitic and bacterial dermatologic diseases in free-ranging white-tailed deer (Odocoileus virginianus) in the United States from 1975 to 2012. *Veterinary Pathology 51*(3), 633–640.

Nganga, V. G., F. W. Kariuki, and K. Kotut (2012). A comparison of the physico-chemical and bacteriological quality of greywater from water deficient households in Homabay town and Githurai estates in Kenya. *Open Environmental Engineering Journal 5*, 110–118.

Nieuwenhuijs, J., J. Haagsma, and P. Lina (1992). Epidemiology and control of rabies in bats in The Netherlands. *Revue scientifique et technique (OIE) 11*(4), 1155–1161.

Nunn, C. L., S. Altizer, K. E. Jones, and W. Sechrest (2003). Comparative tests of parasite species richness in primates. *The American Naturalist 162*(5), 597–614.

Nunn, C. L. and S. M. Altizer (2005). The global mammal parasite database: An online resource for infectious disease records in wild primates. *Evolutionary Anthropology 14*(1),

1–2.

Nyakatura, K. and O. R. P. Bininda-Emonds (2012). Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biology 10*(12), 1–31.

OIE (2005). *World Animal Health in 2005*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2006). *World Animal Health in 2006*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2007). *World Animal Health in 2007*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2008). *World Animal Health in 2008*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2009). *World Animal Health in 2009*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2010). *World Animal Health in 2010*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2011). *World Animal Health in 2011*. 12, rue de Prony - 75017 Paris, France: World Organisation for Animal Health (OIE).

OIE (2012). *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (7th Edition ed.).

OIE (2015). 21st Conference of the OIE Regional Commission for Africa. Technical Report February 16-20, OIE, Rabat, Morocco.

OIE (2016). Terrestrial Animal Health Code: General provisions. Technical report, World Organization for Animal Health, Paris, France.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner (2018). *vegan: Community Ecology Package*. R package version 2.4-6.

Oksi, J., S. Rantala, S. Kilpinen, R. Silvennoinen, M. Vornanen, V. Veikkolainen, E. Eerola, and A. T. Pulliainen (2013). Cat scratch disease caused by Bartonella grahamii in an immunocompromised patient. *Journal of Clinical Microbiology 51*(8), 2781–2784.

Olival, K. J., P. R. Hosseini, C. Zambrana-Torrelio, N. Ross, T. L. Bogich, and P. Daszak (2017). Host and viral traits predict zoonotic spillover from mammals. *Nature 546*(7660), 646–650.

Olson, S. H., C. M. Benedum, S. R. Mekaru, N. D. Preston, J. A. Mazet, D. O. Joly, and J. S. Brownstein (2015). Drivers of emerging infectious disease events as a framework for digital detection. *Emerging Infectious Diseases 21*(8), 1285–1292.

Osnas, E. E. and A. P. Dobson (2012). Evolution of virulence in heterogeneous host communities under multiple trade-offs. *Evolution 66*(2), 391–401.

Otero-Abad, B. and P. R. Torgerson (2013). A systematic review of the epidemiology of echinococcosis in domestic and wild animals. *PLoS Neglected Tropical Diseases 7*(6).

Ovaskainen, O., G. Tikhonov, A. Norberg, F. Guillaume Blanchet, L. Duan, D. Dunson, T. Roslin, and N. Abrego (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters 20*(5), 561–576.

Owen-Smith, N. (1996). Ecological guidelines for waterpoints. *South African Journal of Wildlife Research 26*(4), 107–112.

Page, R. D. (1993). Parasites, phylogeny and cospeciation. *International Journal for Parasitology 23*(4), 499–506.

Palmer, R. J., O. P. Settnes, J. Lodal, and A. E. Wakefield (2000). Population structure of rat-derived Pneumocystis carinii in Danish wild rats. *Applied and Environmental Microbiology 66*(11), 4954–4961.

Paradis, E. and J. Claude (2002). Analysis of comparative data using generalized estimating equations. *Journal of Theoretical Biology 218*, 175–185.

Paradis, E., J. Claude, and K. Strimmer (2004). ape: analyses of phylogenetics and evolution in R language. *Bioinformatics 20*(2), 289–290.

Park, A. W., M. J. Farrell, J. P. Schmidt, S. Huang, T. A. Dallas, P. Pappalardo, J. M. Drake, P. R. Stephens, R. Poulin, C. L. Nunn, and T. J. Davies (2018). Characterizing the phylogenetic specialism-generalism spectrum of mammal parasites. *Proceedings of the Royal Society B: Biological Sciences 285*(1874).

Parker, I. M., M. Saunders, M. Bontrager, A. P. Weitz, R. Hendricks, R. Magarey, K. Suiter, and G. S. Gilbert (2015). Phylogenetic structure and host abundance drive disease pressure in communities. *Nature 520*(7548), 542–544.

Parrish, C. R., E. C. Holmes, D. M. Morens, E.-C. Park, D. S. Burke, C. H. Calisher, C. A. Laughlin, L. J. Saif, and P. Daszak (2008, sep). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews 72*(3), 457–470.

Patz, J. A., P. Daszak, G. M. Tabor, and A. A. Aguirre (2004). Unhealthy landscapes - policy recommendations on land use change and infectious disease emergence. *Environmental Health Perspectives 112*(10), 1092 – 1098.

Pearce-Duvet, J. M. C. (2006). The origin of human pathogens: evaluating the role of agriculture and domestic animals in the evolution of human disease. *Biological reviews of the Cambridge Philosophical Society 81*(3), 369–82.

Pedersen, A. B., S. Altizer, M. Poss, A. A. Cunningham, and C. L. Nunn (2005, may). Patterns of host specificity and transmission among parasites of wild primates. *International Journal for Parasitology 35*(6), 647–657.

Pedersen, A. B. and T. J. Davies (2009). Cross-species pathogen transmission and disease emergence in primates. *EcoHealth 6*(4), 496–508.

Pedersen, A. B., K. E. Jones, C. L. Nunn, and S. Altizer (2007). Infectious diseases and extinction risk in wild mammals. *Conservation Biology 21*(5), 1269–79.

Peltroche-Llacsahuanga, H., B. Frye, and G. Haase (2012). Isolation of Streptococcus urinalis from a human blood culture. *Journal of Medical Microbiology 61*(5), 740–742.

Pence, D. and E. Ueckermann (2002). Sarcoptic mange in wildlife. *Revue scientifique et technique (OIE) 21*(2), 385–398.

Penzhorn, B. L., E. C. Netherlands, C. A. Cook, N. J. Smit, I. Vorster, R. F. Harrison-White, and M. C. Oosthuizen (2018). Occurrence of Hepatozoon canis (Adeleorina: Hepatozoidae) and Anaplasma spp. (Rickettsiales: Anaplasmataceae) in black-backed jackals (Canis mesomelas) in South Africa. *Parasites and Vectors 11*(1), 1–7.

Pérez, J. M., I. Sánchez, and R. L. Palma (2013). The dilemma of conserving parasites: the case of Felicola (Lorisicola) isidoroi (Phthiraptera: Trichodectidae) and its host, the endangered Iberian lynx (Lynx pardinus). *Insect Conservation and Diversity 6*(6), 680–686.

Perlman, S. and J. Jaenike (2003). Infection success in novel hosts: an experimental and phylogenetic study of Drosophila-parasitic nematodes. *Evolution 57*(3), 544–557.

Perry, B. and D. Grace (2009). The impacts of livestock diseases and their control on growth and development processes that are pro-poor. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*(1530), 2643–55.

Perry, B. D., D. Grace, and K. Sones (2011). Current drivers and future directions of global livestock disease dynamics. *Proceedings of the National Academy of Sciences 110*(52), 20871–7.

Peterson, A. T., V. Sánchez-cordero, C. B. Beard, and J. M. Ramsey (2002). Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerging Infectious Diseases 8*(7), 662–667.

Picard-Meyer, E., E. Robardet, L. Arthur, G. Larcher, C. Harbusch, A. Servat, and F. Cliquet (2014). Bat rabies in France: A 24-year retrospective epidemiological study. *PLoS ONE 9*(6).

Pilosof, S., S. Morand, B. R. Krasnov, and C. L. Nunn (2015). Potential parasite transmission in multi-host networks based on parasite sharing. *PLoS ONE 10*(3).

Plourde, B. T., T. L. Burgess, E. A. Eskew, T. M. Roth, N. Stephenson, and J. E. Foley (2017). Are disease reservoirs special? Taxonomic and life history characteristics. *PLoS ONE 12*(7), 1–23.

Plowright, R. K., S. H. Sokolow, M. E. Gorman, P. Daszak, and J. E. Foley (2008). Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Frontiers in Ecology and the Environment 6*(8), 420–429.

Poirotte, C., P. M. Kappeler, B. Ngoubangoye, S. Bourgeois, M. Moussodji, and M. J. Charpentier (2016). Morbid attraction to leopard urine in toxoplasma-infected chimpanzees. *Current Biology 26*(3), R98–R99.

Poulin, R. (2007). *Evolutionary Ecology of Parasites* (2 ed.). Princeton, New Jersey: Princeton University Press.

Poulin, R. and C. Combes (1999, dec). The concept of virulence: Interpretations and implications. *Parasitology Today 15*(12), 474–475.

Poulin, R., B. R. Krasnov, and D. Mouillot (2011). Host specificity in phylogenetic and geographic space. *Trends in Parasitology 27*(8), 355–361.

Poulin, R. and D. Mouillot (2003). Parasite specialization from a phylogenetic perspective: a new index of host specificity. *Parasitology 126*(5), 473–480.

Poulin, R. and D. Mouillot (2005). Combining phylogenetic and ecological information into a new index of host specificity. *Journal of Parasitology 91*(3), 511–514.

Powell, F. A. (2011). Can early loss of affiliates explain the coextinction paradox? An example from Acacia-inhabiting psyllids (Hemiptera: Psylloidea). *Biodiversity and Conservation 20*(7), 1533–1544.

Price, M. N., P. S. Dehal, and A. P. Arkin (2010). FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE 5*(3).

Price, S. A. and J. L. Gittleman (2007). Hunting to extinction: biology and regional economy influence extinction risk and the impact of hunting in artiodactyls. *Proceedings of the Royal Society B: Biological Sciences 274*(1620), 1845–1851.

Pusterla, N. and C. Gebhart (2009). Equine proliferative enteropathy caused by Lawsonia intracellularis. *Equine Veterinary Education 21*(8), 415–419.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research 41*(D1), 590–596.

R Core Team (2014). R: A Language and Environment for Statistical Computing.

R Core Team (2015). R: A Language and Environment for Statistical Computing.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rappé, M. S. and S. J. Giovannoni (2003). The uncultured microbial majority. *Annual Review of Microbiology 57*(1), 369–394.

Rar, V. and I. Golovljova (2011). Anaplasma, Ehrlichia, and Candidatus Neoehrlichia bacteria: Pathogenicity, biodiversity, and molecular genetic characteristics, a review. *Infection, Genetics and Evolution 11*(8), 1842–1861.

Ratnasingham, S. and P. D. N. Hebert (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PLoS ONE 8*(7).

Redfern, J. V., C. C. Grant, A. Gaylard, and W. M. Getz (2005). Surface water availability and the management of herbivore distributions in an African savanna ecosystem. *Journal of Arid Environments 63*(2), 406–424.

Rees, H. C., B. C. Maddison, D. J. Middleditch, J. R. Patmore, and K. C. Gough (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology 51*(5), 1450–1459.

Reluga, T., R. Meza, D. B. Walton, and A. P. Galvani (2007). Reservoir interactions and disease emergence. *Theoretical Population Biology 72*(3), 400–408.

Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution 1*(4), 319–329.

Ricci, F., L. Rokach, and B. Shapira (2011). *Introduction to Recommender Systems Handbook*. Springer.

Rigaud, T., M. J. Perrot-Minnot, and M. J. Brown (2010). Parasite and host assemblages: Embracing the reality will improve our knowledge of parasite transmission and virulence. *Proceedings of the Royal Society B: Biological Sciences 277*(1701), 3693–3702.

Ripple, W. J., J. a. Estes, R. L. Beschta, C. C. Wilmers, E. G. Ritchie, M. Hebblewhite, J. Berger, B. Elmhagen, M. Letnic, M. P. Nelson, O. J. Schmitz, D. W. Smith, A. D. Wallach, and A. J. Wirsing (2014). Status and ecological effects of the world's largest carnivores. *Science 343*(6167), 1241484.

Roeder, P., J. Mariner, and R. Kock (2013). Rinderpest: The veterinary perspective on eradication. *Philosophical Transactions of the Royal Society B: Biological Sciences 368*(1623).

Rwego, I. B., G. Isabirye-Basuta, T. R. Gillespie, and T. L. Goldberg (2008). Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology 22*(6), 1600–7.

Said, Y., M. Gharbi, M. Mhadhbi, M. Dhibi, and S. Lahmar (2018). Molecular identification of parasitic nematodes (Nematoda: Strongylida) in feces of wild ruminants from Tunisia. *Parasitology 145*(7), 901–911.

Sakae, C. and T. Ishida (2012). Direct evidence for Toxoplasma gondii infection in a wild serow (Capricornis crispus) from mainland Japan. *Journal of Parasitology 98*(1), 224–225.

Schaller, G. B. and A. Kang (2008). Status of Marco Polo sheep Ovis ammon polii in China and adjacent countries: Conservation of a vulnerable subspecies. *Oryx 42*(1), 100–106.

Schatz, J., A. R. Fooks, L. Mcelhinney, D. Horton, J. Echevarria, S. Vázquez-Moron, E. A. Kooi, T. B. Rasmussen, T. Müller, and C. M. Freuling (2013). Bat rabies surveillance in Europe. *Zoonoses and Public Health 60*(1), 22–34.

Schemske, D. W., G. G. Mittelbach, H. V. Cornell, J. M. Sobel, and K. Roy (2009). Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology, Evolution, and Systematics 40*(1), 245–269.

Schindler, E. (1957). Comparative investigations on the behaviour of mice and continental field voles (Microtus arvalis) after infection with rabies virus. *Zeitschrift fur Tropenmedizin und Parasitologie 8*(1/2), 233–241.

Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology 75*(23), 7537–7541.

Schlundt, J., H. Totfuku, J. Fisher, M. Artois, T. Morner, and C. Tate (2004). The role of wildlife in emerging and re-emerging zoonoses. *Revue scientifique et technique (OIE) 23*(2), 485–496.

Sempere, J., V. E. Sokolov, and A. A. Danilkin (1996). Capreolus capreolus. *Mammalian Species* (538), 1–9.

Settnes, O. and S. A. Henriksen (1989). Pneumocystis carinii in large domestic animals in denmark. a preliminary report. *Acta veterinaria Scandinavica 30*(4), 437–440.

Shaw, D. J. and A. P. Dobson (1995). Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. *Parasitology 111*, S111–S133.

Shokralla, S., J. L. Spall, J. F. Gibson, and M. Hajibabaei (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular ecology 21*(8), 1794–805.

Shwab, E. K., P. Saraf, X.-Q. Zhu, D.-H. Zhou, B. M. McFerrin, D. Ajzenberg, G. Schares, K. Hammond-Aryee, P. van Helden, S. A. Higgins, R. W. Gerhold, B. M. Rosenthal, X. Zhao, J. P. Dubey, and C. Su (2018). Human impact on the diversity and virulence of the ubiquitous zoonotic parasite Toxoplasma gondii. *Proceedings of the National Academy of Sciences 115*(29), 201722202.

Siembieda, J. L., R. A. Kock, T. A. McCracken, and S. H. Newman (2011). The role of wildlife in transboundary animal diseases. *Animal health research reviews / Conference of Research Workers in Animal Diseases 12*(1), 95–111.

Silva-Rodríguez, E., A. Farias, D. Moreira-Arce, J. Cabello, M. Hidalgo-Hermoso, E. Lucherini, and J. Jiménez (2016). Lycalopex fulvipes. The IUCN Red List of Threatened Species 2016.

Šimić, I., N. Krešić, I. Lojkić, and T. Bedeković (2017). The first serological confirmation of bats rabies in croatia. In *Rabies Serology Meeting*.

Smit, I. and C. Grant (2009). Managing surface-water in a large semi-arid savanna park: Effects on grazer distribution patterns. *Journal for Nature Conservation 17*(2), 61–71.

Smit, I. P., C. C. Grant, and B. J. Devereux (2007). Do artificial waterholes influence the way herbivores use the landscape? Herbivore distribution patterns around rivers and artificial surface water sources in a large African savanna park. *Biological Conservation 136*(1), 85–99.

Smit, I. P., E. S. Riddell, C. Cullum, and R. Petersen (2013). Kruger National Park research supersites: Establishing long-term research sites for cross-disciplinary, multiscaled learning. *Koedoe 55*(1), 1–7.

Smith, C. S., J. H. Epstein, A. C. Breed, R. K. Plowright, K. J. Olival, C. de Jong, P. Daszak, and H. E. Field (2011). Satellite telemetry and long-range bat movements. *PLoS ONE 6*(2), e14696.

Smith, K. F., K. Acevedo-Whitehouse, and A. B. Pedersen (2009). The role of infectious diseases in biological conservation. *Animal Conservation 12*(1), 1–12.

Smith, K. L., V. DeVos, H. Bryden, L. B. Price, M. E. Hugh-Jones, and P. Keim (2000). Bacillus anthracis diversity in Kruger National Park. *Journal of Clinical Microbiology 38*(10), 3780–3784.

Smith, N. H., R. G. Hewinson, K. Kremer, R. Brosch, and S. V. Gordon (2009). Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology 7*(7), 537–44.

Solden, L., K. Lloyd, and K. Wrighton (2016). The bright side of microbial dark matter: Lessons learned from the uncultivated majority. *Current Opinion in Microbiology 31*, 217–226.

Solórzano-García, B. and G. Pérez-Ponce de León (2018). Parasites of Neotropical Primates: A Review. *International Journal of Primatology 39*(2), 155–182.

Sousa, O. E. (1972). Anotaciones sobre la enfermedad de Chagas en Panamá. Frecuencia y distribución de Trypanosoma cruzi y Trypanosoma rangeli. *Revista de biologia tropical 20*(2), 167–179.

Spengler, J. R., É. Bergeron, and P. E. Rollin (2016). Seroepidemiological studies of Crimean-Congo hemorrhagic fever virus in domestic and wild animals. *PLoS Neglected Tropical Diseases 10*(1), 1–28.

Spickler, A. (2015). Bluetongue. Technical report, Institute for International Cooperation in Animal Biologies.

Stamm, I., M. Hailer, B. Depner, P. A. Kopp, and J. Rau (2013). Yersinia enterocolitica in diagnostic fecal samples from European dogs and cats: Identification by Fourier transform infrared spectroscopy and matrix-Assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Clinical Microbiology 51*(3), 887–893.

Stan Development Team (2017a). RStan: the R interface to Stan. R package version 2.17.2.

Stan Development Team (2017b). shinystan: Interactive visual and numerical diagnostics and posterior analysis for bayesian models. R package version 2.4.0.

Stan Development Team (2017c). The Stan Core Library, Version 2.17.0.

Standley, W. and P. McCue. Blood characteristics of San Joaquin kit fox (Vulpes velox macrotis) at Camp Roberts Army National Guard Training Site, California. *Department of Defense Techincal Report*.

Steck, F. and A. Wandeler (1980). The epidemiology of fox rabies in Europe. *Epidemiologic reviews 2*, 71–96.

Stephens, P. R., S. Altizer, K. F. Smith, A. Alonso Aguirre, J. H. Brown, S. A. Budischak, J. E. Byers, T. A. Dallas, T. Jonathan Davies, J. M. Drake, V. O. Ezenwa, M. J. Farrell, J. L. Gittleman, B. A. Han, S. Huang, R. A. Hutchinson, P. Johnson, C. L. Nunn, D. Onstad, A. Park, G. M. Vazquez-Prokopec, J. P. Schmidt, and R. Poulin (2016). The macroecology of infectious diseases: a new perspective on global-scale drivers of pathogen distributions and impacts. *Ecology letters 19*(9), 1159–1171.

Stephens, P. R., P. Pappalardo, S. Huang, J. E. Byers, R. Critchlow, M. J. Farrell, A. Gehman, R. R. Ghai, S. Haas, B. A. Han, A. W. Park, J. P. Schmidt, S. Altizer, V. O. Ezenwa, and C. L. Nunn (2017). Global Mammal Parasite Database version 2.0. *Ecology* (1), 42–49.

Stojmanovski, Z., M. Zdravkovska, V. Taleski, S. Jovevska, and V. Markovski (2014). Human brucellosis in the republic of Macedonia by regions depending on vaccination procedures in sheep and goats. *Macedonian Journal of Medical Sciences 7*(1), 135–140.

Stoltenow, C. L., K. Solemsass, M. Niezgoda, P. Yager, and C. E. Rupprecht (2000). Rabies in an American Bison from North Dakota. *Journal of Wildlife Diseases 36*(1), 169–171.

Strait, K., J. G. Else, and M. L. Eberhard (2012). *Parasitic Diseases of Nonhuman Primates* (2nd ed.). Elsevier.

Strauss, S. Y., C. O. Webb, and N. Salamin (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences 103*(15), 5841–5845.

Streicker, D. G., A. S. Turmelle, M. J. Vonhof, I. V. Kuzmin, G. F. McCracken, and C. E. Rupprecht (2010). Host phylogeny constrains cross-species emergence and

establishment of rabies virus in bats. *Science 329*(5992), 676–9.

Strona, G., M. L. D. Palomares, N. Bailly, P. Galli, and K. D. Lafferty (2013). Host range, host ecology, and distribution of more than 11800 fish parasite species. *Ecology 94*(2), 2013.

Su, W. F., J. K. Lin, and S. Nieh (1997). Entomophthora coronata infection of the paranasal sinuses: Report of a case. *Journal of Oral and Maxillofacial Surgery 55*(10), 1158–1161.

Sundquist, A. (2007). Bacterial flora-typing with targeted, chip-based Pyrosequencing. *BMC Microbiology 7*(1), 108.

Sweet, A. D., S. E. Bush, D. R. Gustafsson, J. M. Allen, E. DiBlasi, H. R. Skeen, J. D. Weckstein, and K. P. Johnson (2018). Host and parasite morphology influence congruence between host and parasite phylogenies. *International Journal for Parasitology 48*(8), 641–648.

Taberlet, P., E. Coissac, M. Hajibabaei, and L. H. Rieseberg (2012). Environmental DNA. *Molecular ecology 21*(8), 1789–93.

Tao, X. Y., Z. Y. Guo, H. Li, W. T. Jiao, X. X. Shen, W. Y. Zhu, S. Rayner, and Q. Tang (2015). Rabies Cases in the West of China Have Two Distinct Origins. *PLoS Neglected Tropical Diseases 9*(10), 1–8.

Tatem, A. J., D. J. Rogers, and S. I. Hay (2006). Global transport networks and infectious disease spread. *Advances in parasitology 62*, 293 – 343.

Taylor, L. H., S. M. Latham, and M. E. Woolhouse (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences 356*(1411), 983–989.

Tekere, M., A. Lötter, J. Olivier, N. Jonker, and S. Venter (2011). Metagenomic analysis of bacterial diversity of Siloam hot water spring, Limpopo, South Africa. *African Journal of Biotechnology 10*(78), 18005–18012.

Tekere, M., A. Prinsloo, J. Olivier, N. Jonker, and S. Venter (2012). An evaluation of the bacterial diversity at Tshipise, Mphephu and Sagole hot water springs, Limpopo

Province, South Africa. *African Journal of Microbiology Research 6*(23), 4993–5004.

Tenaillon, O., D. Skurnik, B. Picard, and E. Denamur (2010). The population genetics of commensal Escherichia coli. *Nature Reviews Microbiology 8*(3), 207–217.

Thornton, P. K. (2010). Livestock production: recent trends, future prospects. *Philosophical Transactions of the Royal Society B: Biological Sciences 365*(1554), 2853–67.

Tomley, F. M. and M. W. Shirley (2009). Livestock infectious diseases and zoonoses. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*(1530), 2637–42.

Turmelle, A. S. and K. J. Olival (2009). Correlates of viral richness in bats (order Chiroptera). *EcoHealth 6*(4), 522–539.

Välimäki, P., A. Kaitala, K. Madslien, L. Härkönen, G. Várkonyi, J. Heikkilä, M. Jaakola, H. Ylönen, R. Kortet, and B. Ytrehus (2011). Geographical variation in host use of a blood-feeding ectoparasitic fly: implications for population invasiveness. *Oecologia 166*(4), 985–995.

van Baarlen, P., A. van Belkum, R. C. Summerbell, P. W. Crous, and B. P. H. J. Thomma (2007). Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? *FEMS microbiology reviews 31*(3), 239–77.

van der Hoeden, J. (1955). Leptospira Canicola in Cattle. *Journal of Comparative Pathology and Therapeutics 65*, 278–283.

Van Heerden, J., M. G. Mills, M. J. Van Vuuren, P. J. Kelly, and M. J. Dreyer (1995). An investigation into the health status and diseases of wild dogs (Lycaon pictus) in the Kruger National Park. *Journal of the South African Veterinary Association 66*(1), 18–27.

Van Heuverswyn, F., Y. Li, C. Neel, E. Bailes, B. F. Keele, W. Liu, S. Loul, C. Butel, F. Liegeois, Y. Bienvenue, E. M. Ngolle, P. M. Sharp, G. M. Shaw, E. Delaporte, B. H. Hahn, and M. Peeters (2006). Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature 444*(7116), 164.

Van Wilgen, B. W., H. C. Biggs, S. P. O'Regan, and N. Mare (2000). A fire history of the savanna ecosystems in the Kruger National Park, South Africa, between 1941 and 1996.

Van Wyk, L. (2011). Review of the ecological implications of artificial waterhole closures in the Kruger National Park and the effect thereof on tourism. *University of Johannesburg MSc Thesis*, 1–80.

Venter, F. J., R. J. Naiman, H. C. Biggs, and D. J. Pienaar (2008). The evolution of conservation management philosophy: Science, environmental change and social adjustments in Kruger National Park. *Ecosystems 11*(2), 173–192.

Viana, M., R. Mancy, R. Biek, S. Cleaveland, P. C. Cross, J. O. Lloyd-Smith, and D. T. Haydon (2014). Assembling evidence for identifying reservoirs of infection. *Trends in Ecology & Evolution 29*(5), 270–279.

Wagner, D. C. and M. S. Edwards (1984). Hand-rearing Black and White Rhinoceroses: A Comparison. *San Diego Animal Park*, 1–10.

Walz, P. H., D. L. Grooms, T. Passler, J. F. Ridpath, R. Tremblay, D. L. Step, R. J. Callan, and M. D. Givens (2010). Control of bovine viral diarrhea virus in ruminants. *Journal of Veterinary Internal Medicine 24*(3), 476–486.

Wang, L., Q. Tang, and G. Liang (2014). Rabies and rabies virus in wildlife in mainland China, 1990-2013. *International Journal of Infectious Diseases 25*, 122–129.

Wang, L., Y. Wang, S. Jin, Z. Wu, D. P. Chin, J. P. Koplan, and M. E. Wilson (2008). Emergence and control of infectious diseases in China. *The Lancet 372*(9649), 1598–1605.

Wang, Z., X. Wang, and Q. Lu (2007). Selection of land cover by the Tibetan fox Vulpes ferrilata on the eastern Tibetan Plateau, western Sichuan Province, China. *Acta Theriologica 52*(2), 215–223.

Wardeh, M., C. Risley, M. K. Mcintyre, C. Setzkorn, and M. Baylis (2015). Database of host-pathogen and related species interactions, and their global distribution. *Scientific Data 2*, 150049.

Warnecke, L., J. M. Turner, T. K. Bollinger, J. M. Lorch, V. Misra, P. M. Cryan, G. Wibbelt, D. S. Blehert, and C. K. R. Willis (2012). Inoculation of bats with European Geomyces destructans supports the novel pathogen hypothesis for the origin of white-nose syndrome. *Proceedings of the National Academy of Sciences 109*(18), 6999–7003.

Weaver, S. C. and A. D. T. Barrett (2004). Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature Reviews Microbiology 2*(10), 789–801.

Weiss, R. A. and A. J. McMichael (2004). Social and environmental risk factors in the emergence of infectious diseases. *Nature Medicine 10*(12S), S70–S76.

White, P. J., M. T. Hanson, W. H. Berry, and J. J. Eliason (2000). Catastrophic decrease in an isolated population of kit foxes. *Southwestern Naturalist 45*(2), 204–211.

Wiens, J. J., D. D. Ackerly, A. P. Allen, B. L. Anacker, L. B. Buckley, H. V. Cornell, E. I. Damschen, T. Jonathan Davies, J. A. Grytnes, S. P. Harrison, B. A. Hawkins, R. D. Holt, C. M. McCain, and P. R. Stephens (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters 13*(10), 1310–1324.

Wiethoelter, A. K., D. Beltrán-Alcrudo, R. Kock, and S. M. Mor (2015). Global trends in infectious diseases at the wildlife-livestock interface. *Proceedings of the National Academy of Sciences 112*(31).

Wilkinson, K., W. P. Grant, L. E. Green, S. Hunter, M. J. Jeger, P. Lowe, G. F. Medley, P. Mills, J. Phillipson, G. M. Poppy, and J. Waage (2011). Infectious diseases of animals and plants: an interdisciplinary approach. *Philosophical Transactions of the Royal Society B: Biological Sciences 366*(1573), 1933–42.

Wilson, D. E. and D. M. Reeder (2005). *Mammal Species of the World: A Taxonomic and Geographic Reference* (3 ed.). Baltimore, Maryland: Johns Hopkins University Press.

Wolfe, N. D., P. Daszak, A. M. Kilpatrick, and D. S. Burke (2005). Bushmeat Hunting, Deforestation , and Prediction of Zoonoses Emergence. *Emerging Infectious Diseases 11*(12), 1882–1827.

Wolfe, N. D., C. P. Dunavan, and J. Diamond (2007). Origins of major human infectious diseases. *Nature 447*(7142), 279–83.

Wolfe, N. D., W. M. Switzer, J. K. Carr, V. B. Bhullar, V. Shanmugam, U. Tamoufe, a. T. Prosser, J. N. Torimiro, A. Wright, E. Mpoudi-Ngole, F. E. McCutchan, D. L. Birx, T. M. Folks, D. S. Burke, and W. Heneine (2004). Naturally acquired simian retrovirus infections in central African hunters. *Lancet 363*(9413), 932–7.

Wood, C. L., J. E. Byers, K. L. Cottingham, I. Altman, M. J. Donahue, and A. M. H. Blakeslee (2007). Parasites alter community structure. *Proceedings of the National Academy of Sciences 104*(22), 9335–9339.

Wood, J. L. N., M. Leach, L. Waldman, H. MacGregor, a. R. Fooks, K. E. Jones, O. Restif, D. Dechmann, D. T. S. Hayman, K. S. Baker, a. J. Peel, a. O. Kamins, J. Fahr, Y. Ntiamoa-Baidu, R. Suu-Ire, R. F. Breiman, J. H. Epstein, H. E. Field, and a. a. Cunningham (2012). A framework for the study of zoonotic disease emergence and its drivers: spillover of bat pathogens as a case study. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*(1604), 2881–2892.

Woods, C. and J. Versalovic (1993). Whole-cell repetitive element sequence-based polymerase chain reaction allows rapid assessment of clonal relationships of bacterial isolates. *Journal of Clinical Microbiology 31*(7), 1927 – 1931.

Woolhouse, M. (2011). How to make predictions about future infectious disease risks. *Philosophical Transactions of the Royal Society B: Biological Sciences 366*(1573), 2045–54.

Woolhouse, M. and E. Gaunt (2007). Ecological origins of novel human pathogens. *Critical Reviews in Microbiology 33*(4), 231–242.

Woolhouse, M., F. Scott, Z. Hudson, R. Howey, and M. Chase-Topping (2012). Human viruses: discovery and emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences 367*(1604), 2864–2871.

Woolhouse, M. E. and S. Gowtage-Sequeria (2005). Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases 11*(12), 1842–1847.

Woolhouse, M. E. J. (2002). Population biology of emerging and re-emerging pathogens. *Trends in Microbiology 10*(10 Suppl), S3–7.

Woolhouse, M. E. J., L. H. Taylor, and D. T. Haydon (2001). Population Biology of Multihost Pathogens. *Science 292*(5519), 1109–1112.

World Health Organization (1983). Rabies bulletin europe. *Rabies Bulletin Europe 7*(2), 1–34.

Yang, J., Z. Liu, Q. Niu, J. Liu, R. Han, G. Guan, M. A. Hassan, G. Liu, J. Luo, and H. Yin (2017). A novel zoonotic Anaplasma species is prevalent in small ruminants: potential public health implications. *Parasites and Vectors 10*(1).

Young, E. and S. P. Kruger (1967). Trichinella spiralis (Owen, 1835) Railliet, 1895 infestation of wild carnivores and rodents in South Africa. *Journal of the South African Veterinary Association 38*(4), 441–443.

Zapata-Ríos, G. and L. C. Branch (2016). Altered activity patterns and reduced abundance of native mammals in sites with feral dogs in the high Andes. *Biological Conservation 193*, 9–16.

Zeglin, L. H. (2015). Stream microbial diversity in response to environmental changes: Review and synthesis of existing research. *Frontiers in Microbiology 6*, 1–15.

Zhang, S. Y., M. X. Wei, Z. Y. Zhou, J. Y. Yu, and X. Q. Shi (2000). Prevalence of antibodies to Toxoplasma gondii in the sera of rare wildlife in the Shanghai Zoological Garden, People's Republic of China. *Parasitology International 49*(2), 171–174.

Zhou, D. H., W. B. Zheng, J. L. Hou, J. G. Ma, X. X. Zhang, X. Q. Zhu, and W. Cong (2017). Molecular detection and genetic characterization of Toxoplasma gondii in farmed raccoon dogs (Nyctereutes procyonoides) in Shandong province, eastern China. *Acta Tropica 172*(May), 143–146.

Zhu, H., X. Chen, X. Shao, H. Ba, F. Wang, H. Wang, Y. Yang, N. Sun, J. Ren, S. Cheng, and Y. Wen (2015). Characterization of a virulent dog-originated rabies virus affecting more than twenty fallow deer (Dama dama) in Inner Mongolia, China. *Infection, Genetics and Evolution 31*, 127–134.