MODELING VIRUS TRANSMISSION & EVOLUTION IN MIXED COMMUNITIES

MODELING VIRUS TRANSMISSION & EVOLUTION IN MIXED COMMUNITIES

By MORGAN P. KAIN, B.S.,M.S.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2019) Hamilton, Ontario (Biology)
TITLE: Modeling virus transmission & evolution in mixed communities
AUTHOR: Morgan P. Kain, B.S. (University of Pittsburgh), M.S. (East Carolina University)
SUPERVISOR: Professor Benjamin M. Bolker
NUMBER OF PAGES: ix, Main Text 134, Appendix 227.

Lay Abstract

In their work in the late 1970s and early 1980s, Anderson and May demonstrated that pathogen induced host harm and pathogen transmission ability are intimately linked. This work clearly showed that pathogens maximize their reproductive potential by causing some harm their hosts, contrary to the established belief that pathogens shouldn't harm their hosts at all. I extend this work to study pathogen evolution and transmission in heterogeneous host populations, using two model host-pathogen systems: birds infected with West Nile virus, and European rabbits infected with the myxoma virus, as well as a general model for the evolution of poorly adapted pathogens in small host populations. I show that pathogen transmission in heterogeneous host populations can be estimated using citizen science data, that pathogen transmission is lower in heterogeneous populations, and that pathogens invading naive host populations may experience short-term evolution to higher-than-optimal virulence, increasing infection burden.

Abstract

In the early 1980s Anderson and May showed that parasite virulence (host mortality rate when infected) and parasite transmission are positively correlated because of their joint dependence on host exploitation (e.g. replication rate). This correlation often results in maximum parasite fitness at intermediate virulence, which has important implications for both parasite evolution and transmission. Anderson and May's observation has led to nearly four decades of work on the ecology and evolution of host-parasite interactions, which focuses on making either general predictions for a range of simplified host-parasite systems or detailed predictions for a single host-parasite system. Yet, despite decades of research, we know comparatively little about parasite evolution and transmission in heterogeneous and/or small host populations. Additionally, much previous work has distanced itself from empirical data, either by outpacing the collection of data or under-utilizing available data. My work focuses on the evolution and transmission of parasites in heterogeneous host populations; I rely on tradeoff theory, but adopt a case-study approach to maximize the use of empirical data. Using West Nile virus infections of birds I show that a continent-wide strain displacement event cannot be explained by current data (Chapter 2), and that transmission in heterogeneous host communities can be estimated using data from citizen scientists, laboratory experiments, and phylogenetic comparative analysis (Chapter 3). Using Myxoma virus infection of European rabbits, I show that tradeoff theory can help us to understand parasite evolution in host populations with heterogeneous secondary infection burden (Chapter 4). In Chapter 5 I show that poorly evolved parasites invading new host populations experience transient evolution away from optimal virulence. In addition to my biological focus, I emphasize clarity and rigor in statistical analyses, including the importance of appropriate uncertainty propagation, as well as reproducible science.

Acknowledgements

First and foremost I would like to thank my supervisor Ben Bolker for his support throughout my journey. Thank you for your patience as I entered a new field of study and first attempted to cobble together epidemiological models. Your continued enthusiasm helped to keep me excited about my work and to push through what I found to be the doldrums of model exploration in the absence of any data. I thank my committee member Jonathan Dushoff for his deep insight and for helping me through small roadblocks on nearly a weekly basis. I also thank my committee member Ian Dworkin for a perspective outside of my subfield, which helped me to relay my science to a broader audience. I thank my advisor Ben Bolker and both of my committee members for feedback on each of my thesis chapters, as well as my external examiner for comments and feedback on the complete thesis.

I would also like to thank Jonathan Dushoff and Evan Twomey for inviting me to collaborate on exciting side projects, and my advisor Ben Bolker, my committee members, and the members of the Bolker, Dushoff, Earn, and Kolasa labs for broader scientific discussions. Thank you for sharing both your depth and breadth of knowledge. It is hard to imagine leaving a PhD program with exposure to a greater breadth of scientific ideas than I was exposed to here.

Finally, I would not have been able to complete this thesis without the love and support of my family and Jo Werba. Thank you for keeping me sane!

Contents

List of Figures and Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations and Symbols

α Parasite virulence

β Parasite transmission rate

$\mathcal{R}_0$ Parasite intrinsic reproductive number

AD Adaptive Dynamics

BBS Breeding Bird Survey

CBC Christmas Bird Count

DTS Discrete Time Stochastic

eBird Cornell Laboratory of Ornithology citizen science database

ESS Evolutionary Stable Strategy

GAM Generalized Additive Model

GLMM Generalized Linear Mixed Effects Model

LMM Linear Mixed Effects Model

MYXV Myxoma virus

NEON National Ecological Observation Network

NY02 Lineage I strain of West Nile virus isolated in 1999 in North America

RD Reaction Diffusion

USA United States of America

WN02 Lineage I strain of West Nile virus isolated in 2002 in North America

WNV West Nile virus

Declaration of Academic Achievement

This sandwich thesis contains an introduction (Chapter 1), two published works (Chapter 2, Chapter 4), a manuscript in a third round of review (Chapter 3), a draft of a manuscript in the early stages of preparation for publication (Chapter 5), and a conclusion (Chapter 6). Preambles to Chapters 2-5 describe authors' and committee members' contributions. Chapters 1 and 6 were written by me.

Chapter 1: Introduction

Thesis overview

In this thesis I present a series of works on the ecology and evolution of infectious diseases. The four data chapters in this thesis focus on the evolution of pathogen life-history strategy (Chapters 2, 4, and 5), and/or pathogen transmission in heterogeneous host communities (Chapters 2, 3 and 4). I study these topics using mathematical or statistical models, grounded in biological reality through parameterization with empirical data. I rely primarily on data from two host-pathogen systems: a diversity of bird species infected with West Nile virus (WNV) (Chapters 2 and 3) and European rabbits (*Oryctolagus cuniculus*) infected with the myxoma virus (MYXV) (Chapter 4). The work I present in chapter 5 is primarily conceptual but is motivated in part by the mechanics of parasite exploitation observed in MYXV.

In addition to the biological story that I tell, I work to do my small part in improving reproducibility in the biological sciences. For each of my chapters data and code is published as supplemental files in open-access journals and/or available in GitHub repositories and/or in the form of shiny R (Chang et al., 2018) applications. I also focus on the importance of clarity and rigour in statistical analyses, including an emphasis on appropriate uncertainty propagation. During my work I repeatedly encountered overly narrow uncertainty in published work, the end result of which is overconfidence in published conclusions that do not stand up to scrutiny. In this thesis I emphasize the importance of propagating uncertainty in order to avoid overreaching the biological conclusions provided by our data.

In some of my thesis chapters I use the term “parasite” and in others “pathogen” depending on the venue for that chapter. While these terms do not refer to identical

concepts, the theory I rely upon is applicable to both forms of host exploitation. I use "viral" instead of either of these terms in the title of my thesis because I explicitly focus on a viral pathogen in Chapters 2-4, and the model I present in Chapter 5 is best suited for pathogens with fast generation times. The term "virulence" also has variable definitions in the literature. It can refer generally to pathogen-induced host harm, the mortality rate of an infected host, or the fitness loss experienced by an infected host. I will most often use the host mortality rate definition, though in each chapter's introduction and methods I clarify the definition that I will be using in that chapter.

Background

The word *parasite* comes from the Greek word *parasitos*, which describes a priest's assistant who was invited to share communal meals; as a dinner guest, a respectful *parasitos* would not want to overindulge (Alizon et al., 2009). In accordance with its etymology, (Smith, 1904) concluded that a parasite would evolve to lose their "highly virulent invasive qualities", evolving to complete avirulence (Alizon et al., 2009). Smith attributed any virulence seen in nature to insufficient time for natural selection to remove all virulence following a parasite jump to a new host, or a suboptimal parasite strategy in a non-focal host.

Yet, parasites are organisms that live in or on another organism and which must extract resources from their host in order to reproduce. This necessity begs the question of how complete avirulence could possibly be a parasite's optimal life-history strategy. For many years a number of researchers questioned the conventional wisdom that parasites evolve to complete avirulence (Kostitzin, 1934; Ball, 1943); however, this "avirulence hypothesis" was not abandoned until Anderson and May (1979, 1982)

showed that parasite virulence and parasite transmission rate were positively correlated because of their co-dependence on host exploitation. Using myxoma virus (MYXV) infection in European rabbits *O. cuniculus*, Anderson and May (1982) showed that MYXV attained maximum fitness at intermediate virulence. Since this observation in European rabbits, empiricists have documented trade-offs in λ phage in *E. coli* (Berngruber et al., 2015), HIV in humans (Fraser et al., 2007), *Ophryocystis* in monarch butterflies (De Roode et al., 2008), and Cauliflower Mosaic Virus in *Brassica rapa* (Doumayrou et al., 2013).

The conclusion that parasites evolve to intermediate virulence because of a genetic correlation between virulence and transmission has become known as the "trade-off theory" of virulence evolution, and forms the foundation of most modern research on parasite evolution (Alizon et al., 2009; Cressler et al., 2016). Modeling work has shown that trade-offs between virulence and transmission arise from within-host infection dynamics (Gilchrist and Sasaki, 2002; Alizon and van Baalen, 2005; King et al., 2009), and that trade-off theory can be used to describe optimal parasite strategies in homogeneous host populations constrained by spatial structure (Boots and Mealor, 2007; Best et al., 2008, 2011), as well as in heterogeneous host populations (e.g. Gandon, 2004; Rigaud et al., 2010).

Heterogeneity among hosts to infection influences pathogen evolution, and therefore both the severity of a disease at the host population level and transmission in that community. Host heterogeneity can select for lower virulence (Ebert and Hamilton, 1996; Gandon, 2004; Pugliese, 2011; Osnas and Dobson, 2012), for higher virulence (Gandon et al., 2001; Ganusov et al., 2002; Read et al., 2015), or facilitate the coexistence of two strains with different exploitation strategies (Fleming-Davies et al., 2015). In the presence of variation in non-linear responses among hosts to infection, host heterogeneity can decrease a pathogen's absolute fitness at its evolutionary optimum,

which can decrease the number of hosts that become infected during an epidemic or at endemic equilibrium (Regoes et al., 2000; Gandon, 2004). For a given pathogen strain, host heterogeneous affects not only the number of hosts that become infected, but which hosts become infected. Variation among hosts in susceptibility, transmission potential, and/or contact rate can create different sized epidemics in different host types (Wasserheit and Aral, 1996), affect how pathogens spill over into non target hosts (Schmidt and Ostfeld, 2001; Keesing et al., 2006), alter which control strategies will be successful for eradicating a disease (Bolzoni et al., 2007), and fundamentally alter the total human and/or animal health burden (Gandon and Day, 2007).

Chapter summaries

Forty years of research built upon the tenants of trade-off theory has clarified many of the intricacies generated by interactions among host heterogeneity to infection, host-pathogen dynamics, and evolution. However, the research landscape is beginning to shift once again. While tradeoff theory has unarguably allowed us to think more clearly about parasite evolution, the reduction of a complex and highly variable biological interaction to two dimensions is an over-simplification that can only tell us so much about any specific system. Many authors are beginning to emphasize the importance of keeping trade-off theory in context, and highlight the difficulties in connecting theory to empirical data when the primary drive is to describe universal laws governing host-parasite interactions (Bull and Lauring, 2014; Alizon and Michalakis, 2015; Cressler et al., 2016). Improving the link between theory and empirical data is a critically needed step to advance our understanding of pathogen evolution — arguably the most important challenge currently faced in research on the ecology and evolution of infectious diseases (Alizon and Michalakis, 2015; Cressler et al., 2016).

Given this recent emphasis on marrying theory with data, I see a few promising avenues for advancing our understanding of the ecology and evolution of pathogen dynamics in heterogeneous host communities. In line with the recent suggestions of Bull and Lauring (2014), Alizon and Michalakis (2015), and Cressler et al. (2016), the most fruitful approach may be to return to an all encompassing view of the factors that contribute directly to a parasite's lifetime reproductive potential in specific host-pathogen systems. This strategy will allow us to expand our "reference library" of case studies, at which point common threads among systems may begin to materialize, an approach favored by the ecologist Tony Ives (Halliday, 2013). I use this strategy in Chapters 2 and 3, where I examine the evolution and transmission of WNV in diverse bird communities in North American. In Chapter 2 I examine if empirical data from the literature is sufficient for explaining the continent-wide replacement of the NY99 WNV strain by the WN02 strain. In Chapter 3 I present a mechanistic model for WNV transmission that predicts WNV spread in any bird community in North America by scaling up from the physiological responses of individual birds to transmission at the level of the community.

A second option is to use empirical data from a given system to determine if different aspects of the evolution of pathogen virulence can be explained by tradeoff theory in that system, in a way to determine the extent and coverage of the theory. I use this strategy in Chapter 4, where I examine if tradeoff theory can help to describe the evolution of MYXV virulence in European rabbits co-infected with gastrointestinal macroparasites. In this chapter I present a model for the evolution of pathogen virulence in a heterogeneous host population that is flexible enough to be parameterized with a range of empirical measures of host heterogeneity (such as host age, immune status, genotype, vaccine status, or co-infection with other parasitic species), but which is sufficiently tractable to capture the life history of many natural host-pathogen

systems. Using this model I examine the effects of co-infection of European rabbits with gastrointestinal helminths and MYXV, using data on helminth burden and the virulence of circulating MYXV strains in Australia and the UK.

Finally, it may be fruitful to expand upon the core tenants of tradeoff theory with additional axes of complexity. While this is a common strategy (e.g. using within-host parasite dynamics: Alizon and van Baalen 2005, host evolution: Carval and Ferriere 2010; Best et al. 2014; Papkou et al. 2016, or spatial structure: Lipsitch et al. 1995; Haraguchi and Sasaki 2000; Lion and Gandon 2015), I am unaware of a model that helps to explain parasite evolution both to and on the tradeoff curve by treating the tradeoff curve as a true *optimization frontier*. In Chapter 5 I examine the evolution of parasites whose transmission is constrained by both virulence and investment in compensatory virulence factors that are required for the parasite to realize its transmission potential at given level of virulence. Using this model I examine the ecologically realistic scenario of stochastic evolution in small host populations using a simulation based approach.

Bibliography

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.

Alizon, S. and Y. Michalakis 2015. Adaptive virulence evolution: the good old fitness-based approach. *Trends in ecology & evolution 30*(5), 248–254.

Alizon, S. and M. van Baalen 2005. Emergence of a convex trade-off between transmission and virulence. *The American Naturalist 165*(6), E155–E167.

Anderson, R. M. and R. May 1982. Coevolution of hosts and parasites. *Parasitology 85*(02), 411–426.

Anderson, R. M. and R. M. May 1979. Population biology of infectious diseases: Part i. *Nature 280*(5721), 361.

Ball, G. H. 1943. Parasitism and evolution. *The American Naturalist 77*(771), 345–364.

Berngruber, T. W., S. Lion, and S. Gandon 2015. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS Pathog 11*(4), e1004810.

Best, A., S. Webb, A. White, and M. Boots 2011. Host resistance and coevolution in spatially structured populations. *Proceedings of the Royal Society of London B: Biological Sciences 278*(1715), 2216–2222.

Best, A., A. White, and M. Boots 2008. Maintenance of host variation in tolerance to pathogens and parasites. *Proceedings of the National Academy of Sciences 105*(52), 20786–20791.

Best, A., A. White, and M. Boots 2014. The coevolutionary implications of host tolerance. *Evolution 68*(5), 1426–1435.

Bolzoni, L., L. Real, and G. De Leo 2007. Transmission heterogeneity and control strategies for infectious disease emergence. *PLoS One 2*(8), e747.

Boots, M. and M. Mealor 2007. Local interactions select for lower pathogen infectivity. *Science 315*(5816), 1284–1286.

Bull, J. J. and A. S. Lauring 2014. Theory and empiricism in virulence evolution. *PLoS pathogens 10*(10), e1004387.

Carval, D. and R. Ferriere 2010. A unified model for the coevolution of resistance, tolerance, and virulence. *Evolution 64*(10), 2988–3009.

Chang, W., J. Cheng, J. Allaire, Y. Xie, and J. McPherson 2018. *shiny: Web Application Framework for R*. R package version 1.2.0.

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.

De Roode, J. C., A. J. Yates, and S. Altizer 2008. Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proceedings of the National Academy of Sciences 105*(21), 7489–7494.

Doumayrou, J., A. Avellan, R. Froissart, and Y. Michalakis 2013. An experimental test of the transmission-virulence trade-off hypothesis in a plant virus. *Evolution 67*(2), 477–486.

Ebert, D. and W. D. Hamilton 1996. Sex against virulence: the coevolution of parasitic diseases. *Trends in Ecology & Evolution 11*(2), 79–82.

Fleming-Davies, A. E., V. Dukic, V. Andreasen, and G. Dwyer 2015. Effects of host heterogeneity on pathogen diversity and evolution. *Ecology Letters 18*(11), 1252–1261.

Fraser, C., T. D. Hollingsworth, R. Chapman, F. de Wolf, and W. P. Hanage 2007. Variation in hiv-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proceedings of the National Academy of Sciences 104*(44), 17441–17446.

Gandon, S. 2004. Evolution of multihost parasites. *Evolution 58*(3), 455–469.

Gandon, S. and T. Day 2007. The evolutionary epidemiology of vaccination. *Journal of the Royal Society Interface 4*(16), 803–817.

Gandon, S., M. J. Mackinnon, S. Nee, and A. F. Read 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature 414*(6865), 751–756.

Ganusov, V. V., C. T. Bergstrom, and R. Antia 2002. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution 56*(2), 213–223.

Gilchrist, M. A. and A. Sasaki 2002. Modeling host–parasite coevolution: a nested approach based on mechanistic models. *Journal of Theoretical Biology 218*(3), 289–308.

Halliday, F. 2013, Oct. 29. `http://www.biodiverseperspectives.com/2014/03/11/diverse-introspectives-with-tony-ives/`.

Haraguchi, Y. and A. Sasaki 2000. The evolution of parasite virulence and transmission rate in a spatially structured population. *Journal of Theoretical Biology 203*(2), 85–96.

Keesing, F., R. D. Holt, and R. S. Ostfeld 2006. Effects of species diversity on disease risk. *Ecology letters 9*(4), 485–498.

King, A. A., S. Shrestha, E. T. Harvill, and O. N. Bjørnstad 2009. Evolution of acute infections and the invasion-persistence trade-off. *The American Naturalist 173*(4), 446–455.

Kostitzin, V. 1934. Symbiose, parasitisme et evolution (etude mathématique) hermann.

Lion, S. and S. Gandon 2015. Evolution of spatially structured host–parasite interactions. *Journal of evolutionary biology 28*(1), 10–28.

Lipsitch, M., E. A. Herre, and M. A. Nowak 1995. Host population structure and the evolution of virulence: a "law of diminishing returns". *Evolution 49*(4), 743–748.

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

Papkou, A., C. S. Gokhale, A. Traulsen, and H. Schulenburg 2016. Host–parasite coevolution: why changing population size matters. *Zoology 119*(4), 330–338.

Pugliese, A. 2011. The role of host population heterogeneity in the evolution of virulence. *Journal of biological dynamics 5*(2), 104–119.

Read, A. F., S. J. Baigent, C. Powers, L. B. Kgosana, L. Blackwell, L. P. Smith, D. A. Kennedy, S. W. Walkden-Brown, and V. K. Nair 2015. Imperfect vaccination can enhance the transmission of highly virulent pathogens. *PLoS Biol 13*(7), e1002198.

Regoes, R. R., M. A. Nowak, and S. Bonhoeffer 2000. Evolution of virulence in a heterogeneous host population. *Evolution 54*(1), 64–71.

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 of London B: Biological Sciences 277*(1701), 3693–3702.

Schmidt, K. A. and R. S. Ostfeld 2001. Biodiversity and the dilution effect in disease ecology. *Ecology 82*(3), 609–619.

Smith, T. 1904. Some problems in the life history of pathogenic microorganisms. *Science*, 817–832.

Wasserheit, J. N. and S. O. Aral 1996. The dynamic topology of sexually transmitted disease epidemics: implications for prevention strategies. *Journal of Infectious Diseases 174*(Supplement_2), S201–S213.

Chapter 2: Can existing data on West Nile virus infection in birds and mosquitos explain strain replacement?

In my first of two chapters on West Nile virus (WNV), I show that common explanations presented in the literature for the displacement of the NY99 strain of WNV by the WN02 strain are insufficient to describe the displacement of NY99 when all of the available data on each step of the life cycle of WNV are taken into account. I challenge the notion that a faster incubation rate in mosquitoes is sufficient criteria for WN02's superior transmission capability, and suggest that American Robins may be less important for WNV amplification than is currently believed. This study illustrates the importance of propagating uncertainty in order to avoid over-confident results, which are abundant in the literature on WNV.

Author Contributions

MPK conceived the study with helpful feedback from BMB; MPK collected the data; MPK performed statistical analyses with helpful feedback from BMB; MPK wrote the first draft of manuscript and both authors revised the manuscript.

Acknowledgements

I thank my committee members Jonathan Dushoff and Ian Dworkin, the Bolker, Dushoff, and Earn labs, as well as Jo Werba for helping me refine and present this work.

esa

ECOSPHERE

Can existing data on West Nile virus infection in birds and mosquitos explain strain replacement?

MORGAN P. KAIN[1,†] AND BENJAMIN M. BOLKER[1,2]

[1]*Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1 Canada*
[2]*Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8 Canada*

Citation: Kain, M. P., and B. M. Bolker. 2017. Can existing data on West Nile virus infection in birds and mosquitos explain strain replacement? Ecosphere 8(3):e01684. 10.1002/ecs2.1684

Abstract. Understanding pathogen strain displacement is important for predicting and managing the spread of infectious disease. However, a mechanistic understanding for the outcome of pathogen competition often remains unresolved, as in the case of the displacement of the West Nile virus (WNV) NY99 genotype by the newer WN02 genotype. In this study, we seek to explain the observed displacement of the NY99 genotype by examining evidence for differences between NY99 and WN02 over WNV's entire transmission cycle. We synthesized the available empirical data on key aspects of WNV's transmission cycle including viral titer profiles in birds, survival of birds, bird-to-mosquito transmission probability, and mosquito-to-bird transmission probability for infections with both the NY99 and WN02 genotypes of WNV. Using a Bayesian statistical framework, we combine our literature synthesis on infection dynamics in birds and mosquitos with bird community and mosquito life history parameters to examine fitness differences between NY99 and WN02. We calculate the intrinsic reproduction numbers ($\mathcal{R}_0$) for NY99 and WN02 and assume a larger value of $\mathcal{R}_0$ for WN02 will explain its competitive dominance over NY99. Our analysis of the collective body of experimental infections of birds and mosquitos produced similar $\mathcal{R}_0$ estimates for NY99 and WN02 with wide overlapping credible intervals, which we take as insufficient evidence for greater fitness of WN02. The currently cited explanation for the displacement of WN02 by NY99—the attribution of the displacement of NY99 by WN02 to the latter's more efficient replication in mosquitos—is, at best, weakly supported by the evidence. Further infection studies of American Robins (*Turdus migratorius*) and broader coverage of possibly competent hosts for WN02, as well as more ecological data such as the spatial and temporal differences in vector and bird communities, will be needed to fully understand the community dynamics of WNV and the determinants of higher fitness in the WN02 genotype.

Key words: Bayesian statistics; epidemic; passeriformes; pathogen evolution; viral competition.

Received 4 November 2016; revised 18 December 2016; accepted 19 December 2016. Corresponding Editor: Andrew W. Park.

† **E-mail:** kainm@mcmaster.ca

INTRODUCTION

Pathogens can evolve higher-fitness genotypes during the course of an epidemic through immune escape (e.g., influenza; Bhatt et al. 2013) or by shifting along the transmission–virulence tradeoff curve (e.g., Myxomatosis; Dwyer et al. 1990). In order to understand, predict, and manage epidemics, we need to understand the causes and consequences of such evolutionary change. Sometimes the pathway from genetic change to increased pathogen fitness is apparent, such as the amino acid substitution in chikungunya virus that allowed for efficient vectoring by *Aedes albopictus* (Tsetsarkin et al. 2007, 2014) or the amino acid substitution in West Nile virus (WNV) that increased viral replication in avian hosts (NY99 genotype: Briese et al. 2002,

Lanciotti et al. 1999) and led to the North American epidemic. In other cases, the evidence of genetic substitutions is clear, but the mechanisms by which they lead to a fitness advantage remain unresolved. One such example is a newer amino acid substitution in WNV producing the WN02 genotype, which rapidly displaced the invading NY99 genotype on a landscape scale. Despite statements in the literature about the cause of NY99's displacement (e.g., Moudy et al. 2007, Kilpatrick et al. 2008, Duggal et al. 2014), it remains unclear how the difference in the phenotypes of NY99 and WN02 across different stages of the virus's life history (e.g., mosquito-to-bird transmission rate, mosquito incubation period, bird-to-mosquito transmission rate) resulted in the higher fitness of WN02.

West Nile virus was first detected in North America in 1999. The invading genotype (NY99) differed from closely related strains isolated in Israel in 1997–1998 (Lanciotti et al. 1999, Briese et al. 2002) by a single amino acid substitution in the NS3 helicase. This substitution increases viral replication in avian hosts (Lanciotti et al. 1999) and apparently contributed to a widespread North American epidemic. From 1999 to 2003, NY99 spread rapidly, causing substantial declines in many bird species (Brault et al. 2007, LaDeau et al. 2007, Brault 2009), and numerous spillover infections in humans (Ostroff 2013, Petersen et al. 2013). By 2003, WNV infections had been recorded in birds and/or humans in nearly all contiguous U.S. states, southern Canada, and northern Mexico (Kilpatrick et al. 2007).

During WNV's westward expansion between 1999 and 2003, a single point mutation resulting in an amino acid substitution (Val159Ala) in a viral structural envelope gene (E) produced the WN02 genotype (Beasley et al. 2003, Davis et al. 2005, Di Giallonardo et al. 2016). This mutation fixed so rapidly that a recent analysis (Snapinn et al. 2007) failed to detect any period during which NY99 and WN02 co-occurred, suggestive of a substantial fitness difference between the two genotypes. Earlier and more efficient replication of WN02 in mosquito vectors (Moudy et al. 2007, Kilpatrick et al. 2008) is routinely cited as providing the required fitness difference (as recently as May 2016: Grinev et al. 2016), despite evidence that the incubation rates of NY99 and WN02 in mosquitos do not differ (Anderson et al. 2012, Danforth et al. 2015). Higher competence of House Sparrows for WN02 may also have contributed to NY99's displacement (Duggal et al. 2014); however, many experiments indicate that House Sparrows, among other passerines, exhibit lower titer when infected with WN02. Further complications include evidence for higher survival of birds when infected with WN02, which increases the total number of contacts between infected birds and naive mosquitos (e.g., Brault et al. 2011).

Due to substantial avian mortality and the threat of infection in humans (Ostroff 2013, Petersen et al. 2013), many experimental infection studies in both birds (e.g., Komar et al. 2003) and mosquitos (e.g., Turell et al. 2000, 2001) began after the 1999 outbreak. To determine whether the collective body of empirical data on WNV transmission (based on experimental infections of NY99 and WN02 in birds and mosquitos) can clarify the cause of NY99 displacement, we synthesized the available empirical data on key aspects of WNV's transmission cycle. Beginning with an infected bird, the transmission cycle of WNV proceeds as follows: First, an infected bird infects naive mosquitos. The expected number of naive mosquitos infected by a bird depends on the bird's virus titer, the duration of infection in the bird, and the relationship between the bird's titer and the bird-to-mosquito transmission probability. We gathered data on each of these components: titer profiles in infected birds, survival of infected birds, and titer-dependent transmission probability from an infected bird to a naive mosquito (Fig. 1). Second, infected mosquitos transfer infection to naive birds. Transmission depends on sufficient viral replication and movement of viral particles into the mosquitos' salivary glands; a faster virus replication rate within the mosquito leads to a shorter transmission cycle. The rate of viral replication within mosquitos depends on the viral dose a mosquito receives from the infecting bird (a function of the bird's titer), the mosquito species, and other external variables such as temperature (Fig. 1).

Host titer profiles, host survival, and bird-to-mosquito transmission probability conditioned on host titer are often combined into a single metric called "host competence," which describes bird species' aptitude to transmit WNV to mosquitos (Komar et al. 2003). The probability that a mosquito transmits infection to a naive bird following a blood meal is known as "vector

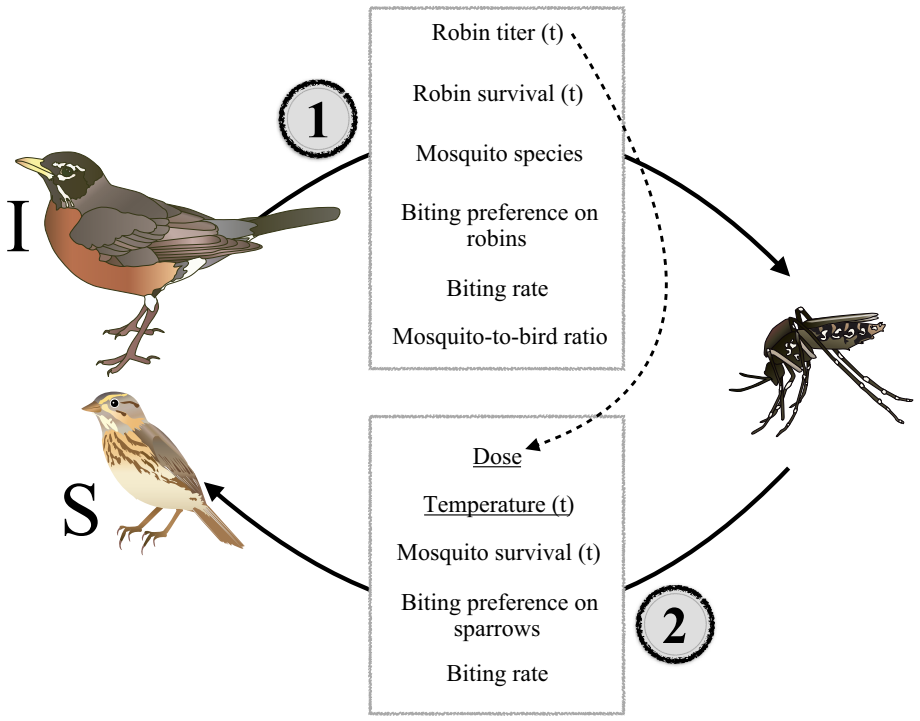

Fig. 1. The transmission process of West Nile virus (WNV), shown from an infected robin to an uninfected, susceptible sparrow. The center column lists the parameters associated with bird-to-mosquito transmission (1) and mosquito-to-bird transmission (2) that are included in our $\mathcal{R}_0$ calculation. Time-dependent parameters (i.e., parameters that vary with the number of days since infection) are denoted by (t). Underlining denotes parameters that may vary among mosquito species. The dashed line from robin titer to dose indicates that the titer of the infected bird at the time of mosquito infection carries forward to affect the probability of mosquito-to-bird infection. Vector images courtesy of the Integration and Application Network, University of Maryland Center for Environmental Science (ian.umces.edu/symbols/).

competence" (Turell et al. 2000). Many studies report aggregated host and vector competence only; here, we retain as much detail as possible about differences between NY99 and WN02 across the virus life cycle. We refer to the four key aspects of transmission (bird titer profiles, bird survival, bird-to-mosquito transmission, and mosquito-to-bird transmission) collectively as "individual-level" transmission dynamics.

To estimate the fitness differences between NY99 and WN02 in an ecological setting, we combined the results of our synthesis on individual-level transmission dynamics with micro-scale ecological dynamics using a case study approach. For the ecological and life history parameters of mosquito biting rate and biting preference, mosquito density, and bird density, we use parameter estimates from suburban Chicago, Illinois, where reasonably complete information on these parameters is available (Hamer et al. 2009, Loss et al. 2009, Ruiz et al. 2010, Newman et al. 2011). We calculate the intrinsic reproduction numbers ($\mathcal{R}_0$) for NY99 and WN02 and assume a larger value of $\mathcal{R}_0$ for WN02 will explain its competitive dominance over NY99, an outcome with strong theoretical support (Heffernan et al. 2005, Keeling and Rohani 2008). We focus on between-genotype differences in $\mathcal{R}_0$ here, rather than the closely related intrinsic growth rate (r), which measures net reproductive rate rather than lifetime parasite fitness. It is possible that a larger intrinsic growth rate for WN02 may have played a role in NY99's displacement, a topic we return to in the *Discussion*. We use a Bayesian statistical framework to combine uncertainty from disparate sources on WNV's complex transmission process to estimate $\mathcal{R}_0$ for NY99 and WN02 (Elderd et al. 2006). This methodology allows for the use of informative

priors when data are scarce and the seamless addition of new data when it becomes available, and could be useful for sparser-data scenarios, such as during emerging epidemics.

Methods

Literature search

We searched the Web of Science for appropriate studies. To search for publications on experimental infections of avian species with WNV, we used: TS = (West AND Nile AND Virus) AND TS = virul* AND TS = infec*. For data on bird-to-mosquito transmission, we combined results from three searches in an attempt to capture variable language used in this literature. We used (from broad to narrow): TS = (West AND Nile AND Virus) AND TS = Culex AND (TS = vector competence); TS = (West AND Nile AND Virus) AND TS = (NY99 OR WN02 OR SW03) AND (TS = vector OR TS = competence); TS = (West AND Nile AND Virus) AND TS = (NY99 OR WN02 OR SW03) AND TS = Culex AND (TS = vector competence OR TS = trans* OR infect*). For mosquito-to-bird transmission (incubation rate of WNV in *Culex* mosquitos), we used: TS = (West AND Nile AND Virus) AND TS = (NY99 OR WN02 OR SW03) AND TS = Culex AND (TS = inc* OR TS = rep*). Finally, for feeding preference of *Culex* mosquitos, we used: TS = Culex AND (TS = bird OR TS = avian) AND TS = (bit* pref* OR feed* pref*). Our searches yielded 332, 35, 384, and 70 results, respectively.

We read the titles and abstracts (when further clarification was needed) of these 821 studies and downloaded 53, 20, 41, and 22 that mentioned experimental infections of birds, experimental infection of mosquitos, or feeding and/or biting preference of mosquitos on birds. We narrowed these 136 studies down to 67 studies from which data were extracted (inclusion criteria and methods of data extraction are described in detail below). The extracted data (Data S1) and a list of all studies, with publication-specific notes (Appendix S2), are available as Supporting Information (also see https://github.com/morgankain/WNV_Synthesis).

Data extraction and curation

Bird titer profiles and survival.—For bird-related phenotypes, we gathered all studies that experimentally infected birds and measured both titer and survival through time. We included studies that presented individual-level data or group means and reported sample size. We discarded studies that took place outside of North America. In total, data were extracted from 29 papers, which included 134 conspecific groups of birds infected with WNV (henceforth referred to as "infection experiments"), resulting in 134 averaged titer profiles. From each paper, we extracted titer, mortality, host species, location and date (when available) of bird collection, WNV strain, and date and location of isolation among other data (see Appendix S2, Data S1).

These 134 averaged titer curves spanned 11 orders and 45 species, with 66 curves from American Crows, House Sparrows, House Finches, or American Robins (for a complete list of all bird species infected, see Data S1). These studies infected birds with over 30 different strains of WNV belonging to either Lineage 1: NY99 genotype, Lineage 1: WN02 genotype, Lineage 1: SW03 genotype, Lineage 2, Lineage 3, or genotypes that were experimentally altered using genome editing techniques (e.g., Langevin et al. 2014). In rare cases in which no lineage was given, we looked up the WNV strain ID (all papers gave strain information) in Davis et al. (2005) or McMullen et al. (2011). Virus doses ranged from 2 $\log_{10}$ titer per mL (Reisen et al. 2005) to 6 $\log_{10}$ titer per mL (Ziegler et al. 2013). Birds were captured from 10 different U.S. states and Mexico between 1999 and 2013.

Because our goal was to determine the difference between NY99 and WN02, we included only experimental infections with Lineage 1 genotypes. Recent genetic work by Di Giallonardo et al. (2016) found that the SW03 group is not distinct from the WN02 group; therefore, we treat all SW03 genotypes as WN02 genotypes. Additionally, what we refer to as the NY99 genotype includes isolates from 1999 and from 2001 and thus corresponds to what Duggal et al. (2014) call the "East Coast genotype." Because 49% of all experimental infections of birds used American Crows, House Finches, House Sparrows, or American Robins as hosts, and because these species are thought to be the most important for WNV amplification (Kilpatrick et al. 2007), we focused on differences in $\mathcal{R}_0$ between NY99 and WN02 in these four focal passerine species. We present analyses

on titer and survival for the other 41 bird species used in experimental infections in Appendix S1. Because only two to five were infected with either NY99 or WN02 for most species, and only four species other than the four focal passerine species have been infected with WN02 (Clay-Colored Thrush, Great-Tailed Grackle, Carolina Wren, Tufted Titmouse), compared to 39 other species infected with NY99, we grouped all of the species apart from the four focal passerine species into a single "other" category for analysis.

Bird-to-mosquito transmission.—We collected data for bird-to-mosquito transmission by extracting the percentage of mosquitos that became infected following a feeding event on infected blood from studies of vector competence. In this literature, assays of mosquito infection routinely take place 14 d after feeding. A subset of studies measured vector infection rate on days 7 and 14, but we focused our efforts on extracting data from day 14. Mosquitos are rarely allowed to feed on a live host (but see Anderson et al. 2012, Turell et al. 2000, 2001), rather they feed on blood packets; this process is assumed to mimic transmission from live birds (Vanlandingham et al. 2004). A small subset of this literature also gave auxiliary information specifying environmental factors such as time of year and classifying infected mosquitos by titer level, age, and number of generations removed from wild-caught individuals. Because of the sparsity of this auxiliary information, our analyses consider only the overall proportion of mosquitos that were infected 14 d after feeding on a source of WNV. Environmental and individual-level variation is likely to increase variability in transmission probability, even among mosquitos of the same species and population (Vaidyanathan and Scott 2006, Vanlandingham et al. 2007, Reisen et al. 2008, Richards et al. 2010). We extracted data from a total of 20 studies including mosquito species, extrinsic temperature, virus genotype, source of blood meal, dose in $\log_{10}$ titer, sample size, and percentage of mosquitos that were infected after 14 d. These data included 10 species of *Culex* mosquitos, a range of $\log_{10}$ dose from 0.22 to 9.5, and temperatures between 26° and 28°C.

Mosquito-to-bird transmission.—For data on WNV incubation rate in mosquitos and transmission probability from mosquitos to birds, we extracted data from studies that infected the avian specialist genus *Culex* and quantified time-dependent probability of transmission and/or time-dependent virus titer. Because Moudy et al. (2007) provided extensive data on both titer and probability of transmission, this data set was used to fit a logistic model to predict transmission from mosquito titer levels (see Appendix S1), so we could use data from two papers that measured titer in mosquitos (Dohm et al. 2002, Johnson et al. 2003, see Appendix S1). Only nine studies from this literature met our criteria for data extraction. Nonetheless, these studies included 45 unique infection experiments that resulted in 34 averaged curves of NY99 incubation in mosquitos and 11 averaged curves of WN02 incubation in mosquitos. These studies included six mosquito species and a temperature range of 14–32°C. The standard procedure in this literature is to collect mosquito salivary excretions using a capillary tube method (Aitken 1977), which is used as a stand-in for successful transmission to a host species. In rare cases, experiments allow mosquitos to feed directly on hosts (e.g., suckling mice; Anderson et al. 2012). We extracted the following data from each paper: location of study, vector species, extrinsic temperature, virus strain, method of infection, receiving dose, and percentage of mosquitos that successfully transmitted virus at a specific number of days post-infection (see Data S1).

Ecological effects.—Finally, for data on mosquito biting preference, we used Hamer et al. (2009) as a case study. Hamer et al. (2009) focused on the spread of WNV at a small spatial scale (suburban Chicago, Illinois, USA) and measured mosquito-to-bird ratios, bird community composition, and mosquito biting preferences. Due to some missing data, we combined these results with data from Loss et al. (2009), Ruiz et al. (2010), and Newman et al. (2011), who also studied transmission in the Chicago area, and one parameter estimated by Simpson et al. (2012) (New Haven, Connecticut, USA; Table 1).

Statistical methods

We analyzed our data using a Bayesian approach with the statistical software Stan (Carpenter et al. 2017), interfaced with the R statistical programming environment (R Development Core Team 2013) using the `rstan` package (Stan Development Team 2016). Our statistical procedure was composed of two steps, model

Table 1. Ecological parameters for WNV transmission in Chicago, Illinois, USA.

Ecological parameter	Value	Source
Mosquito survival	Appendix S1: Fig. S2	Andreadis et al. (2014)
Mosquito daily bite rate	0.14/d	Simpson et al. (2012) from Wonham et al. (2004) and Vinogradova (2000)
Mosquito-to-bird ratio	3:1	Simpson et al. (2012) from Simpson et al. (2009)
Mosquito biting preference	American Crow: 0.54 House Sparrow 0.32 House Finch: 5.69 American Robin: 2.26	Hamer et al. (2009)
Bird density	Total: 9.66 birds/ha American Crow: 0.0007 House Sparrow 4.25 House Finch: 0.0110 American Robin: 2.0	Hamer et al. (2009)
Temperature	22°C	weatherspark.com

fitting to gathered data and calculation of $\mathcal{R}_0$ for each genotype, which combined predictions from models for each step of the transmission cycle. We first describe models for each transmission step and then our method for calculating $\mathcal{R}_0$. Estimates and pointwise 95% credible intervals (CIs) for each model were generated by sampling parameter values from MCMC chains. All models were run with four MCMC chains until all Gelman–Rubin statistic ($\hat{R}$) values were <1.1, following Gelman et al. (2014). Data points in all models were weighted by sample size. Code for all Stan models is available in Data S2.

Titer profiles.—To analyze titer profiles, we used a lognormal linear mixed-effects model using $\log_{10}$ titer as the response variable. We used a quadratic day-by-virus genotype interaction and $\log_{10}$ dose as fixed effects and modeled variation among studies, infection experiments, and bird species using random effects. While titer profiles will vary among individuals within a single bird species—for example, due to differences in sex, age, or nutritional state—the sparsity of data made subdividing data below the level of bird species impractical for the current modeling effort. We provide a description of the age and sex of each group of experimental birds for each publication that reported these data in Appendix S2.

Because of the detection limit for WNV of 1.7 $\log_{10}$ titer, we used a left-censored model for titer reported as 1.7 (in this model, reports of 1.7 $\log_{10}$ titer are modeled as residing between 0.0 and 1.7 $\log_{10}$ titer). When data from each individual were available and the titers of a subset of birds exceeded the detection limit, we averaged across all birds using 0.85 $\log_{10}$ titer units for each bird with non-detectable viremia. If this average was below 1.7 $\log_{10}$ titer, we rounded to 1.7 $\log_{10}$ titer to include it as a censored value. We adopted this approach in an effort to retain as much data as possible while differentiating this case from the case when no birds were above the detection limit. Predicted titer was obtained for each bird species by averaging predicted titers using two values for $\log_{10}$ dose: median inoculates coming from *Culex pipiens* (5.0 $\log_{10}$ dose) and *Culex tarsalis* (6.1 $\log_{10}$ dose) determined by Styer et al. (2007).

Survival.—We modeled infected bird survival on a given day using a mixed-effects model with a binomial error distribution, using proportion of surviving birds as the response variable. We used an interaction of virus genotype with both day and $\log_{10}$ dose as fixed effects and modeled variation among studies, infection experiments, and bird species using random effects. Predicted survival was obtained for each bird species by averaging predictions when using inoculates from *C. pipiens* (5.0) and *C. tarsalis* (6.1).

Bird-to-mosquito transmission.—We modeled bird-to-mosquito transmission using a mixed-effects model with a binomial error distribution, using proportion of mosquitos infected as the response variable. We used an interaction of virus genotype with $\log_{10}$ dose and temperature as fixed effects and modeled variation among studies and mosquito species using random

effects. Predicted bird-to-mosquito transmission was obtained using predicted titer as $\log_{10}$ dose.

Mosquito-to-bird transmission.—To analyze mosquito-to-bird transmission, we used a mixed-effects model with a binomial error distribution, using proportion of mosquitos transmitting as the response variable. In our analysis, we separate mosquito-to-bird transmission from bird-to-mosquito transmission by using transmission from infected mosquitos (i.e., conditioning on mosquito infection) to birds. We calculated mosquito-to-bird transmission conditioned on infected mosquitos for all studies that originally present results for transmission from all mosquitos (e.g. Moudy et al. 2007, Kilpatrick et al. 2008), using data on infection prevalence in mosquitos provided in each publication. We used an interaction of virus genotype with $\log_{10}$ dose, day, and temperature as fixed effects and modeled variation among citations, infection experiments, and mosquito species using random effects. To obtain average mosquito lifetime transmission given $\log_{10}$ dose and temperature, we averaged daily mosquito transmission probability weighted by survival using *Culex* survival modeled using data from Andreadis et al. (2014) (see Appendix S1). Data were analyzed with and without weighting by mosquito survival; in the main text, we present transmission including mosquito survival probabilities. In Appendix S1, we present transmission holding mosquito lifespan constant to illustrate the differences in transmission probabilities driven by temperature dependence in infection alone (Appendix S1: Fig. S4.5).

Unfortunately, all of the available mosquito-to-bird transmission studies used a narrow range of $\log_{10}$ dose (6.5–9.5), outside of the range commonly seen in birds (e.g., 4.5–6.5). We adopted two approaches to analyze these data. Our first analysis used only the titer range available in the extracted data. This led to a negative slope across titer at high doses, driven by decreased success in mosquito vector capability at high virulence (see *Discussion*: *Mosquito-to-bird transmission*). When the fitted model was used to predict titer at low values (e.g., 1 or 2 $\log_{10}$ titer), the probability of mosquito-to-bird transmission approached 100% by day 2 or 3, a biological impossibility. We include parameter estimates and figures from this analysis in Appendix S1 (Figs. S4.3, S4.4, S4.6, S5, and S6.9–6.10) and discuss the results below. In the main text, we present results from a second method which incorporates prior information on transmission probabilities at low titer using data extracted from four studies on the transmission of Japanese encephalitis virus (JEV), one of the most phylogenetically similar flaviviruses to WNV (Kuno et al. 1998), by *Culex* mosquitos at lower titer levels (2.5–5 $\log_{10}$ dose). This method, adopted because of the absence of data, assumes that NY99 and WN02 have the same mosquito-to-bird transmission at low $\log_{10}$ dose.

Locally derived mosquito biting parameters and bird community composition.—No models were fit to data for this portion of our analysis. Parameter estimates were obtained directly from Hamer et al. (2009) and others (Table 1). Because our focus was on determining the difference between NY99 and WN02, we do not incorporate the error in the estimates from these papers, instead using presented means. Incorporating variation in these parameter estimates would increase uncertainty in $\mathcal{R}_0$ relative to our estimates, but would be unlikely to affect estimates of the difference between NY99 and WN02 $\mathcal{R}_0$ values (see *Results*: $\mathcal{R}_0$ *calculations*).

$\mathcal{R}_0$ *calculations.*—The intrinsic reproduction number $\mathcal{R}_0$ combines all of the previously estimated quantities: bird titer profiles and survival; bird-to-mosquito and mosquito-to-bird transmission as a function of titer and temperature; and ecological quantities such as biting rate, mosquito-to-bird ratio, and bird community composition. We start by generating bird titer profiles and bird survival from our fitted models. Higher titers increase the instantaneous probability of bird-to-mosquito transmission; a higher titer also increases the dose a mosquito receives from a bite, which increases the probability per unit time that it transfers the infection to a naive bird. Longer durations of viral titer and longer survival times give a broader time window for bird-to-mosquito transmission. We use species-level estimates to predict transmission probabilities to and from individuals of particular bird species. Finally, the ecological parameters of mosquito biting rate, mosquito-to-bird ratio, and bird community composition (Table 1) determine the overall population-level rate of transmission within and among bird species. To calculate $\mathcal{R}_0$ for each bird species for the duration of their infection, we used an extension of the classic Ross–Macdonald model (Smith et al. 2012):

$$\mathcal{R}_{0\{ijCv\}} = \sum_{t_b=1}^{T_b} \Bigg(\mathcal{B}\,\text{Survival}_{t_b v} \times \mathcal{B}\text{-to-}\mathcal{M}\,\text{Trans}_{t_b v} \times \mathcal{M}\,\text{Daily bite rate} \times \sum_{t_m=1}^{T_m} \Big(\mathcal{M}\text{-to-}\mathcal{B}\,\text{Trans}_{t_m t_b v C} \times \mathcal{M}\,\text{Survival}_{C t_m} \times \mathcal{M}\,\text{Daily bite rate} \Big) \Bigg) \times \mathcal{M} : \mathcal{B}\,\text{Ratio} \times \text{Prop'n}\,\mathcal{B}_i \times \mathcal{M}\,\text{Bite Pref on}\,\mathcal{B}_i \times \text{Prop'n}\,\mathcal{B}_j \times \mathcal{M}\,\text{Bite Pref on}\,\mathcal{B}_j \quad (1)$$

This equation gives the $\mathcal{R}_0$ between a transmitting bird species ($\mathcal{B}_i$) and a receiving bird species ($\mathcal{B}_j$) for a given virus genotype (v) at a given temperature (°C; C), with $\mathcal{M}$ = mosquitos; t_b = day of bird infection; T_b = last day of bird infection; t_m = day of mosquito infection; and T_m = median mosquito survival (last day of mosquito infection). "Prop'n $\mathcal{B}_i$" and "$\mathcal{B}_j$" refer to the proportion of the community composed of $\mathcal{B}_i$ or $\mathcal{B}_j$, respectively.

We use a definition of $\mathcal{R}_0$ based on the type number, or the expected number of secondary bird infections caused by a single infected bird in an otherwise susceptible population (Roberts 2007). In this case, $\mathcal{R}_0$ is defined as the product of the host-to-vector and vector-to-host reproductive numbers, in contrast to the definition based on the next-generation-matrix approach that instead uses the geometric mean of this quantity (i.e., the square root of the product; Wonham et al. 2006). The choice between these metrics is largely a matter of taste; they give identical answers about qualitative questions such as the invasion and persistence of disease (which are determined by whether $\mathcal{R}_0$ is >1 or <1; Heffernan et al. 2005, Roberts 2007). More quantitative questions such as the level of control needed to eradicate disease must be answered on a case-by-case basis with careful attention to details of how the control is applied (Roberts 2007). The type-number definition is common in analyses of vector-borne disease (Bailey et al. 1982) because it describes the level of control that would be necessary to achieve disease control by affecting a single stage of the life cycle (e.g., by reducing mosquito densities). However, our type-number results can be converted (approximately) to the geometric mean scale by simply taking the square root of the reported $\mathcal{R}_0$ values.

To calculate community $\mathcal{R}_0$ values, we constructed community WAIFW (Who Acquires Infection from Whom) matrices (Dobson and Foufopoulos 2001) for a given temperature from pairwise species $\mathcal{R}_{0\{ij\}}$ values at a given temperature, where $\mathcal{R}_{0\{ij\}}$ gives the number of secondary infections in bird species j generated by an infected bird of species i (with $i = j$ describing intra-species transmission). The dominant eigenvalue of this matrix, which captures the effects of all transmission pathways within and between bird species, gives the community $\mathcal{R}_0$ (Dobson and Foufopoulos 2001).

Results

First, we present the results of our literature synthesis for each step of the transmission process modeled in Stan. We then present $\mathcal{R}_0$ values for monocultures of each bird species (calculated using the equation for $\mathcal{R}_0$ setting $i = j$ and "Prop'n $\mathcal{B}_i$" = 1); these values are determined only by species-specific titer profiles and survival, independent of other ecological variables. We use the WAIFW matrix to calculate community-level $\mathcal{R}_0$ values for a bird community in suburban Chicago, Illinois, USA, by combining model predictions of individual-level transmission processes with the ecological parameter values presented in Table 1. For visual assessment of the model fit to empirical data, we reduce our predictions to two dimensions for each model by using a single value for some predictors. For example, infection experiments of birds varied in $\log_{10}$ dose. To obtain predicted titers for use in subsequent models, we averaged predictions using inoculates coming from *C. pipiens* and *C. tarsalis* determined by Styer et al. (2007), and plot titer profiles in Fig. 2 using this average. Coefficient plots for all models are available in Appendix S1 (Figs. S6.1–S6.10).

Titer profiles

Our results confirm that American Crows have the highest titers, followed by House Finches (Fig. 2). Median estimates for titer resulting from NY99 infection are higher for American Crows, House Sparrows, and House Finches, while American Robins had marginally higher titer with

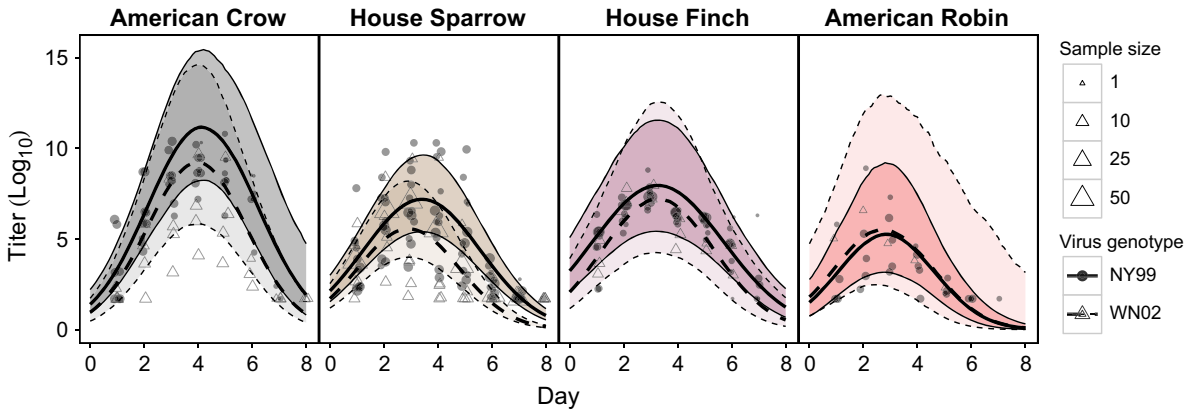

Fig. 2. Titer profiles in four passerine species. Solid lines (and darker shading) show median NY99 infection profile with 95% pointwise CIs and dotted line and lighter shading WN02. The fitted relationship presented here is an average of predictions using 5.0 and 6.1 $\log_{10}$ dose. Circles show data for NY99 and triangles for WN02. Each point is an average from an individual infection experiment on a given day (2–58 birds per experiment).

WN02 infection. However, 95% CIs overlap in all cases (overlap in CIs is a conservative criterion for significant differences between groups; in the cases presented here, there is usually so much overlap that the 95% CIs for each strain include the median estimate of the other strain). Fig. 2 shows titer given a single level of $\log_{10}$ dose equal to the expected dose transferred from a mosquito to a bird. Appendix S1: Fig. S3.1 shows titer profiles in other species. The median NY99 titer profile for these other, non-focal bird species resembles the titer profile of American Robins, and the WN02 profile (for only four total infected bird species) resembles the titer profile of House Finches.

Bird survival

Crows survived worst, followed by House Finches (Fig. 3). Median estimates for survival were higher for all four passerines with WN02 infection; however, 95% CIs for NY99 and WN02 overlapped for all species. Survival apparently decreases with increasing peak titer across bird hosts; American Crows have the highest peak titer and the lowest survival, while American Robins have the lowest titer and highest survival. Appendix S1: Fig. S3.2 shows titer profiles in other species.

Bird-to-mosquito transmission

Transmission from birds to mosquitos was similar for NY99 and WN02 (Fig. 4). For both virus genotypes, the majority of studies infected mosquitos with a $\log_{10}$ dose between 4 and 6, which contains the inflection point of the relationship between $\log_{10}$ dose and transmission probability (Fig. 4). WN02 shows more variation around the fitted relationship than NY99 (Fig. 4), despite the fact that the studies which infected mosquitos with NY99 used more mosquito species overall (9 vs. 5) and collected mosquitos from more states overall (6 vs. 5).

Mosquito-to-bird transmission

Results for mosquito-to-bird transmission presented in the main text include data from JEV (see Appendix S1: Figs. S4.3, S4.4, and S4.6 for analyses excluding JEV and Appendix S1: Fig. S4.7 for "vector competence" at 26°C). Mosquito-to-bird transmission increases with time since initial infection in the mosquito and with increasing temperature (Fig. 5; for coefficient plots, see Appendix S1: Figs. S6.7 and S6.8). CIs for mosquito-to-bird transmission for NY99 and WN02 overlap, with little difference in the medians. Fig. 5a shows transmission of NY99 and WN02 from an infected mosquito to a naive bird for 5.5 $\log_{10}$ titer at 16° and 26°C. Fig. 5b shows mosquito-to-bird transmission at the same two temperatures averaged across the mosquito's lifespan (weighted by survival) using data from Andreadis et al. (2014) (see Appendix S1 for data on mosquito survival and results for mosquito-to-bird transmission without weighting by survival). Here, an increase in temperature decreases lifetime average transmission

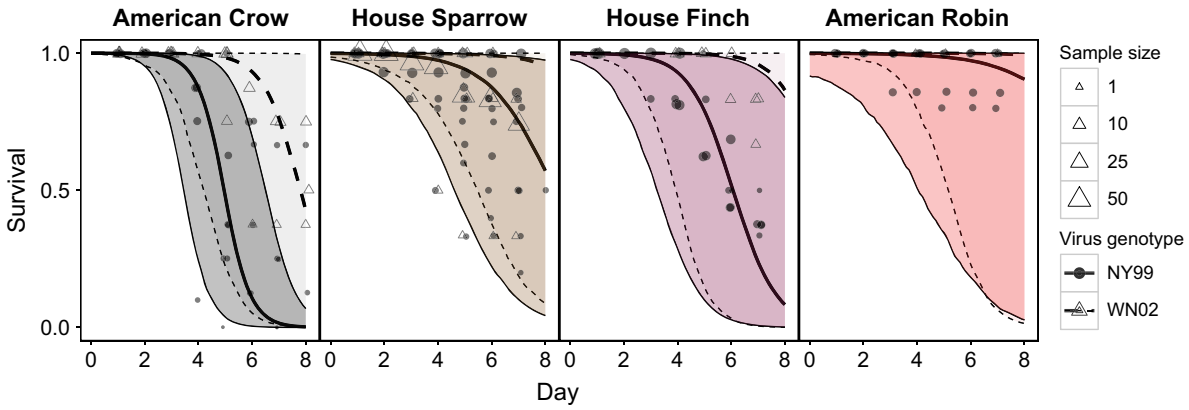

Fig. 3. Proportion of surviving passerines in each infection experiment. Solid lines (and darker shading) show median NY99 infection profile with 95% pointwise CI and dotted line and lighter shading WN02. Circles show data for NY99 and triangles for WN02. Point size corresponds to the total number of infection experiments (2–58 birds per experiment). Fitted relationships here are based on average predictions using 5.0 and 6.1 $\log_{10}$ units.

because the increased rate of mosquito mortality at high temperatures outweighs the gain in incubation rate. Due to the complexity of the raw data (range of 14–32°C, 3–8.5 $\log_{10}$ dose, 40 d of measurements), we present fitted results without raw data in the main text and three-dimensional plots with raw data in Appendix S1: Figs. S4.1–S4.4.

Case studies: $\mathcal{R}_0$ at a local scale

In monoculture, American Crows are predicted to produce the largest $\mathcal{R}_0$ (i.e., if the community were a monoculture of the focal species, and bite rate, bird density, and mosquito density were the same in all communities regardless of bird species). In monoculture American Robins would generate the lowest $\mathcal{R}_0$ because of their low titer (Fig. 2), despite their low mortality rates (Fig. 3).

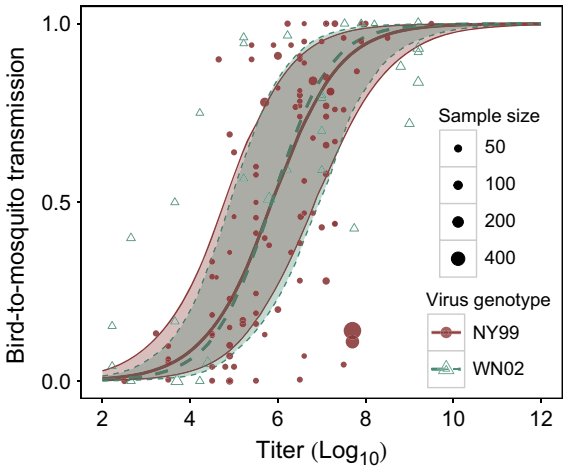

Fig. 4. Bird-to-mosquito transmission for NY99 (solid line, darker red shading) and WN02 (dotted line, lighter blue shading) at 26°C. Points are extracted data.

Monoculture $\mathcal{R}_0$ CIs for NY99 and WN02 are wide and overlap between strains for all species because of substantial uncertainty at each step of the transmission pathway, producing large uncertainty in the sign of the difference in $\mathcal{R}_0$ between NY99 and WN02 (Table 2).

We calculate two values of community $\mathcal{R}_0$, one including and the other excluding all of the non-focal passerine species ("other" birds). When including "other" birds, we assumed that the remaining 35.2% of the Chicago bird community not composed of one of the four focal passerine species were bird species from our "other" category and were bitten at a preference of 1 (biting rate proportional to their density; Hamer et al. 2009). To calculate $\mathcal{R}_0$ for NY99 and WN02 without "other" birds, we scaled the proportions of each of the four passerine species so that the community was composed entirely of the four focal species at the same relative proportions as observed. In the community in Chicago, IL, American Crows and House Finches become less important for the spread of WNV relative to the scenario presented in Fig. 6 because they occur at low densities. Here, American Robins and House Sparrows emerge as important hosts, as

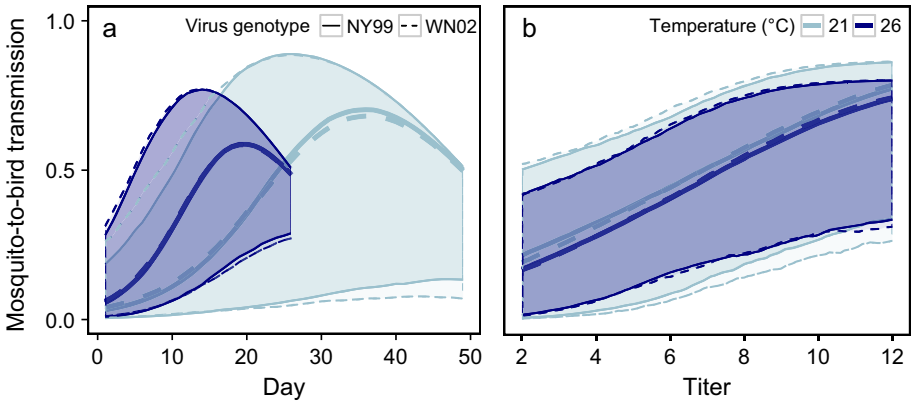

Fig. 5. Mosquito-to-bird transmission for NY99 (solid lines) and WN02 (dotted lines). Panel (a) shows mosquito-to-bird transmission probability as a function of days post-infection at 16° (light blue) and 26°C (dark blue) at 5.5 $\log_{10}$ dose, discounted by mosquito survival. Panel (a) shows transmission until day 50 at 16° and day 26 at 26°C (median mosquito longevity at 16° and 26°C is 89 and 26 d, respectively). Panel (b) shows average daily transmission for a mosquito across its lifespan at different $\log_{10}$ dose at 16° and 26°C.

previously suggested in this and other communities (Kilpatrick et al. 2006, Savage et al. 2007, Hamer et al. 2009, Simpson et al. 2012; Appendix S1: Table S1).

To confirm that our $\mathcal{R}_0$ estimates are consistent with population-level estimates of $\mathcal{R}_0$, we conducted a brief survey of the literature for seroprevalence data (comprising 12 papers; Appendix S2). We extracted fractions of birds seropositive (averaged across bird species) as an estimate of $1 - S^*/N$, or the fraction of birds not immune to WNV from prior infection. Based on this information, $\mathcal{R}_0$ estimates ($\mathcal{R}_0 = 1/(1 - S^*/N)$) ranged from ≈1.1 to 2.1, which overlaps the majority of our $\mathcal{R}_0$ distributions (see *Discussion*: *Local* $\mathcal{R}_0$ *estimates* for a discussion of median $\mathcal{R}_0$ <1). Because there was no temporal trend in seroprevalence, apart from a spike at the epicenter of the epidemic in New York in 1999, we did not calculate separate $\mathcal{R}_0$ estimates for NY99 and WN02 (publications did not report infection with NY99 or WN02).

Table 2. Ratio of $\mathcal{R}$, between NY99 and WN02 on a log scale.

Bird species	Temperature (°C)	Median	CI (2.5%, 97.5%)
American Crow	16	0.89	(−3.06, 4.77)
	26	0.86	(−0.96, 4.17)
House Sparrow	16	1.65	(−3.57, 7.20)
	26	1.19	(−1.08, 4.03)
House Finch	16	0.61	(−4.20, 5.73)
	26	0.31	(−1.92, 3.40)
American Robin	16	−0.07	(−6.87, 7.27)
	26	−0.38	(−4.69, 4.80)
Other	16	−1.60	(−6.58, 3.39)
	26	−1.68	(−3.98, 0.86)

Discussion

WN02's $\mathcal{R}_0$ value could plausibly be larger than NY99's (e.g., at 16°C, without "other" birds, the CI ranges from a WN02's $\mathcal{R}_0$ from −6.26 times proportionately lower on the log scale to 5.23 times higher than NY99's), but clearly cannot support a larger value on the basis of laboratory data alone. Furthermore, the estimated $\mathcal{R}_0$ ratio between NY99 and WN02 (i.e., median of NY99 $\mathcal{R}_0$/WN02 $\mathcal{R}_0$) was positive on the log scale for three of the four focal passerine species at both temperatures, suggesting that NY99 may be a *stronger* competitor than WN02. The apparent advantage of either strain varies among steps in the virus life cycle, as well: Attributing the displacement of NY99 by WN02 to the latter's more efficient replication in mosquitos (Moudy et al. 2007, Kilpatrick et al. 2008) is, at best, weakly supported by the available data.

Titer profiles and survival

In most cases, birds infected with WN02 have lower titer but higher survival than those infected with NY99 (Figs. 2, 3). These opposing

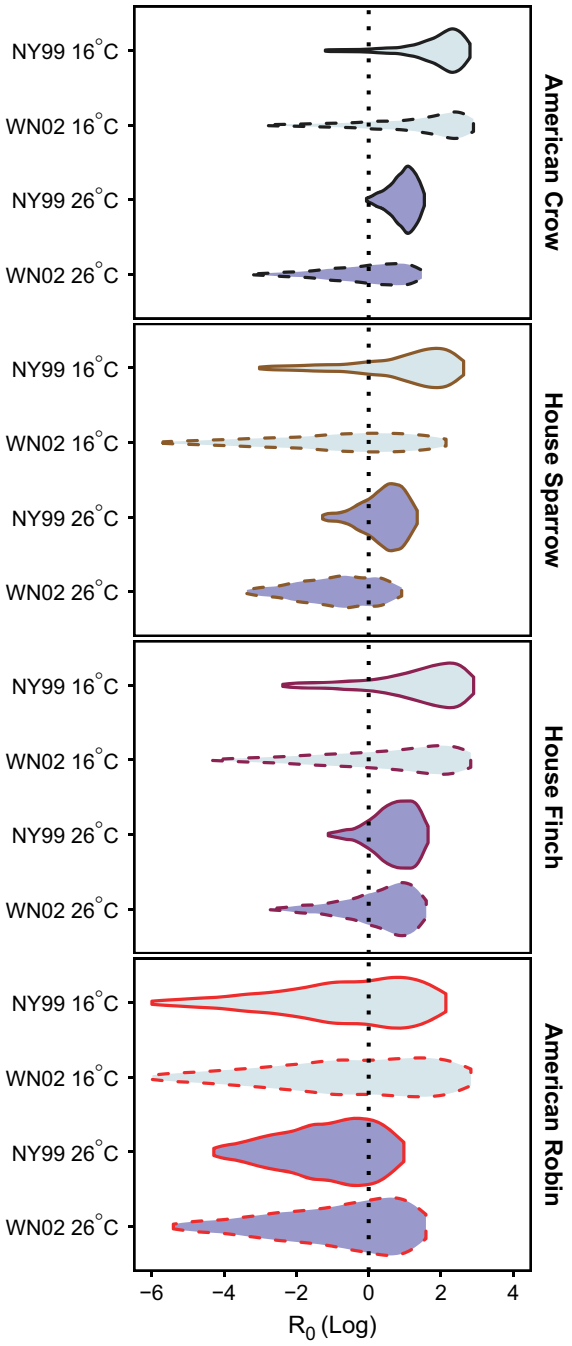

Fig. 6. $\mathcal{R}_t$ distributions for individual birds assuming transmission occurs in monoculture. We use parameter values from Table 1 for biting rate, mosquito-to-bird density, and others. Solid lines show NY99 $\mathcal{R}_t$ and dotted lines WN02 $\mathcal{R}_t$. Here, we present $\mathcal{R}_t$ at 16° (light blue fill) and 26°C (dark blue fill).

attributes of WN02 infection contribute to the ambiguity in fitness differences between genotypes (Table 2). Additionally, given the small number of avian species and families infected with both NY99 and WN02 (five families and six species: American Crow [Corvidae]; House Sparrow [Passeridae]; House Finch [Fringillidae]; American Robin and Clay-Colored Thrush [Turdidae]; and Great-Tailed Grackle [Icteridae]) and the diversity of responses in the four passerines analyzed here (Figs. 2, 3), we cannot really draw broad conclusions about differences between NY99 and WN02.

Despite the importance of American Robins in amplifying WNV in ecological communities, sample sizes for infections of American Robins with both NY99 and WN02 are the smallest of the four passerine species, resulting in large uncertainty in the shape of American Robins' titer profiles (Fig. 2). Additional infections of American Robins with both NY99 and WN02 is an important first step in resolving the fitness differences between NY99 and WN02. Also, because most experimental infections concentrated on four bird species, and WN02 infection sampling outside of these four species was sparse and biased, this relationship might also be resolved by data on WN02 infection from a wider variety of avian species. Specifically, these data raise two important and related questions: First, how well do these results generalize to all birds? Second, are we missing an important host for WN02 amplification?

Great-Tailed Grackles show nearly identical titer profiles and survival when infected with either NY99 or WN02, and infection of Clay-Colored Thrushes (which have the same genus as American Robins: *Turdus*) with WN02 produces higher titer than infection with NY99 (Guerrero-Sánchez et al. 2011). While these disparate examples are based on data from a single study with sparse data, they raise the possibility that an important host for WNV amplification remains unobserved. The search for a host that is more competent for WN02 than for NY99 should concentrate on passerine species (non-passerines have low competence for WNV; Kilpatrick et al. 2007) that have exhibited high titer when infected with WN02 (e.g., Carolina Wren, Tufted Titmouse), are abundant enough, and are bitten

with a high enough biting preference that they can make an impact on a community $\mathcal{R}_0$; further model-based analyses could narrow down the abundances and biting preferences that could allow such a role.

Finally, we clarify that the analysis presented here averages over two additional levels of variation. First, we ignored possible among-individual variation in titer profiles and survival of a given species. Given that our analysis uses nonlinear transformations to compute $\mathcal{R}_0$ from the titer profiles, our use of an average value for each species could cause either over- or underestimation of the pairwise $\mathcal{R}_0$ values (Ruel and Ayres 1999), producing bias in our community-level $\mathcal{R}_0$ values. However, exploring within-species heterogeneity is likely to be less important than finding a species that is more competent for WN02 than for NY99.

Second, we present titer profiles that are calculated using a $\log_{10}$ dose that averages the doses coming from *C. pipiens* and *C. tarsalis*. In the absence of an interaction between $\log_{10}$ dose and virus genotype (i.e., dose effects; the effects of dose transmitted from mosquitos to birds on bird titer are larger in one genotype than in the other), changing the $\log_{10}$ dose would not affect the relative competitive ability of NY99 and WN02. (We omitted this term from our model due to lack of data; a very large difference would likely have shown up in the graphical displays of model fit.) However, a difference in $\log_{10}$ dose will affect the absolute titer produced in a bird. Therefore, predictions of transmission in a specific region of the country where one of these mosquito species is dominant should adjust the $\log_{10}$ dose appropriately.

Bird-to-mosquito transmission

A rich data set comprising 20 publications led to nearly identical point estimates for NY99 and WN02 bird-to-mosquito transmission probability. Thus, additional studies in this area are unlikely to resolve the cause of NY99's displacement. However, the high variance in the relationship between $\log_{10}$ dose and bird-to-mosquito transmission probability for WN02 suggests that we may be missing important predictors of transmission probability for WN02. Previous work has shown highly variable mosquito competence among populations and by time of year (Vaidyanathan and Scott 2006, Vanlandingham et al. 2007, Reisen et al. 2008, Richards et al. 2010), making bird-to-mosquito transmission an important research topic for understanding differences in transmission of WNV at small spatial scales.

Mosquito-to-bird transmission

Mosquito-to-bird transmission for NY99 and WN02 had very similar point estimates and overlapping credible intervals that make it impossible to determine whether NY99 or WN02 has better mosquito-to-bird transmission (Fig. 5). Danforth et al. (2015) emphasize that published studies calculate the percentage of mosquitos transmitting WNV differently, contributing to variation in conclusions among studies. While the calculation method for incubation rate used by Moudy et al. (2007) and Kilpatrick et al. (2008) may contribute to their significant findings (both measure incubation rate using all mosquitos rather than conditioning on infected mosquitos as we do here), the wide distributions for both NY99 and WN02 incubation we derived suggest that the large effects found in these experiments may simply arise from sampling in the tails of these distributions. The small data set, high dimensionality ($\log_{10}$ dose, days post-infection, temperature, mosquito species), heterogeneous results among studies, and variable methods of calculating mosquito-to-bird transmission point to mosquito-to-bird transmission as a focal point of research to resolve the cause of NY99's displacement. Ideally, research should (1) focus on WN02 infection in a variety of mosquito species (Danforth et al. 2015), (2) at a single temperature and value of $\log_{10}$ dose (in order to reduce sources of heterogeneity), and (3) measure transmission conditioned on infected mosquitos (in order to isolate mosquito-to-bird transmission).

How sensitive are our results to the sparse data available for mosquito-to-bird transmission at low titers? In the main text, we conducted our analysis for mosquito-to-bird transmission using data from four papers on incubation rate in Japanese Encephalitis Virus (JEV) at low $\log_{10}$ dose, in an attempt to fill in missing information about mosquito-to-bird transmission at low titers. In Appendix S1, we present all results from analyses without JEV data. These analyses show no qualitative differences in incubation rate in mosquitos through time (Appendix S1: Fig. S4.6) and reveal wide, overlapping CIs for $\mathcal{R}_0$ in NY99 and WN02

Table 3. $\mathcal{R}_0$ estimates for a bird community in Chicago, Illinois, USA.

Temperature (°C)	Bird species included	NY99 median	NY99 CI (2.5%, 97.5%)	WN02 median	WN02 CI (2.5%, 97.5%)
16	All bird species (including "others")	0.42	(0.005, 2.60)	0.97	(0.01, 5.12)
	Four focal passerines (excluding "others")	0.61	(0.08, 4.80)	0.44	(0.002, 8.94)
26	All bird species (including "others")	0.19	(0.04, 0.75)	0.50	(0.08, 1.52)
	Four focal passerines (excluding "others")	0.50	(0.08, 1.50)	0.34	(0.02, 2.60)

(Appendix S1: Fig. S5). With JEV data excluded, the relationship between $\log_{10}$ dose and incubation rate reverses, becoming negative (Appendix S1: Figs. S4.3, S4.4, S4.6, and S6.10). This suggests that mosquitos may become less competent vectors when infected with high viral doses. We emphasize that this result is equivocal, because the upper 95% confidence limit is greater than 0; nevertheless, the pattern is intriguing enough to warrant further investigation. A tradeoff between virulence and transmission is supported by a large body of theory and empirical work (Alizon et al. 2009); however, tradeoffs in arthropod vectors have received relatively little attention (for evidence of a tradeoff between virulence and transmission in mosquitos for WNV, see Ciota et al. 2013). Our results emphasize the potential importance of tradeoffs in disease vectors.

Local $\mathcal{R}_0$ estimates

We presented two scenarios using different sets of ecological parameters to illustrate how our synthesis of the physiological aspects of transmission can produce different $\mathcal{R}_0$ outcomes in different ecological communities. In monoculture, when birds were bitten with the same daily biting parameters used in the case study in Chicago, Illinois (Table 1), American Crows had the highest $\mathcal{R}_0$ resulting from their high titer, followed by House Finches with the second highest titer and moderate survival. However, there is a distinct lack of evidence about the differences in fitness ($\mathcal{R}_0$) between NY99 and WN02. In the Chicago community, including and excluding "other" birds both resulted in overlapping CIs for $\mathcal{R}_0$. Including "other" birds resulted in a larger median estimate for WN02 than for NY99 and excluding "other" birds the reverse (Table 2), due in part to the biased sample of infections in bird species highly susceptible to WN02 (Carolina Wren, Tufted Titmouse).

Using Eq. (1) for $\mathcal{R}_0$ with the ecological parameters in Table 1 produced median estimates for $\mathcal{R}_0$ that were less than 1 for both genotypes at both temperatures (Table 3). While these median estimates taken at face value would suggest that WNV should go extinct, CIs for most estimates do include the values of $\mathcal{R}_0$ derived from serological data. Our low $\mathcal{R}_0$ estimates may be due to using an estimate for mosquito-to-bird ratio from Simpson et al. (2012) who measured this parameter in New Haven, Connecticut, USA, or because we ignored uncertainty in ecological parameter estimates in an effort to focus on individual-level transmission. For example, doubling any single parameter, such as the mosquito-to-bird ratio or mosquito bite rate, would double $\mathcal{R}_0$, producing many estimates above 1 without changing the relationship between NY99 and WN02. A change in temperature would also impact mosquito-to-bird transmission (Fig. 5) and could increase $\mathcal{R}_0$; however, because slope parameter estimates for temperature were similar for NY99 and WN02 (Appendix S1: Fig. S6.7), a different temperature will not change the relationship between NY99 and WN02.

Our calculated $\mathcal{R}_0$ distributions for NY99 and WN02 clearly show that more research is needed before we can reliably estimate the difference in $\mathcal{R}_0$ between these genotypes. However, the similarity in point estimates of $\mathcal{R}_0$ across genotypes means we should also consider the possibility that NY99 and WN02 have similar $\mathcal{R}_0$ values; searching for a greater $\mathcal{R}_0$ in WN02 would be fruitless if something other than a difference in $\mathcal{R}_0$ is responsible for NY99's displacement. Snapinn et al. (2007) find no evidence for population growth of WN02 following NY99's displacement, suggesting that the $\mathcal{R}_0$ values for the two genotypes are similar (because the equilibrium incidence of a pathogen is governed by its $\mathcal{R}_0$ value, at least in simple models). Conceivably, a high

intrinsic growth rate (r) in WN02 could have led to the rapid extirpation of NY99, but the absence of a spike in seroprevalence that would accompany the emergence of a new strain with large r makes this a dubious explanation. Furthermore, WN02 titer profiles in birds and incubation rate in mosquitos provide no evidence for faster viral replication indicative of an increase in r that would result in more rapid spread and increased cases of WN02 relative to NY99.

To explore the possibility of competitive exclusion explained by phenomena that are not captured by differences in $\mathcal{R}_0$, we have begun to examine deterministic between-genotype competition models for WNV in a heterogeneous avian community, parameterized with the results presented here. The model will include seasonal dynamics and among-genotype variation in r and $\mathcal{R}_0$ in heterogeneous communities. We hypothesize that communities with high densities of American Crows and House Finches (the focal species with the largest differences between NY99 and WN02 dynamics) may be able to support the observed displacement of NY99 by WN02 via yearly recruitment of these relatively competent species. Specifically, if mortality is high and recruitment low for these species (e.g., under NY99 infection), low rates of susceptible recruitment may reduce the potential for disease spread, while higher survival under infection with WN02 could lead to endemic disease. This phenomenon is unlikely to play a role in communities where Robins are abundant (Kilpatrick et al. 2006, Savage et al. 2007, Simpson et al. 2012); however, it may affect epidemiological dynamics in communities with a higher proportion of crows.

Conclusions

Combining the available laboratory-derived data on the transmission parameters of WNV suggests that researchers have been overly optimistic in given explanations for the displacement of WN02 by NY99; the reasons for WN02's fitness advantage remain ambiguous. How then do we proceed? Further laboratory studies of individual birds, especially American Robins, and broader coverage of possibly competent hosts for WN02 could uncover the cause of WN02's fitness advantage. Alternatively, we may need to focus on gathering more information and building more realistic models at the level of the community; spatial and temporal differences in vector and bird populations, variation in the effects of environmental covariates of WNV transmission (Ruiz et al. 2010), and uncertainty about the importance of bird density and mosquito preference (e.g., Simpson et al. 2009, Chaves et al. 2010) all contribute to our uncertainty in WNV dynamics. Our work shows the necessity of quantitative synthesis when scaling from individual processes to community-level dynamics of complex host–vector–pathogen systems and highlights the continued challenges of such synthesis.

Acknowledgments

We thank A.M. Kilpatrick and J. Dushoff for helpful comments on the first draft of the manuscript. This work was funded by NSERC Discover Grant 386590-2010.

Literature Cited

Aitken, T. 1977. An in vitro feeding technique for artificially demonstrating virus transmission by mosquitoes. Mosquito News 37:130–133.

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: 245–259.

Anderson, J. F., A. J. Main, G. Cheng, F. J. Ferrandino, and E. Fikrig. 2012. Horizontal and vertical transmission of West Nile Virus genotype NY99 by *Culex salinarius* and genotypes NY99 and WN02 by *Culex tarsalis*. American Journal of Tropical Medicine and Hygiene 86:134–139.

Andreadis, S., O. Dimotsiou, and M. Savopoulou-Soultani. 2014. Variation in adult longevity of *Culex pipiens* f. pipiens, vector of the West Nile Virus. Parisitology Research 113:4315–4319.

Bailey, N. T. J. 1982. The biomathematics of malaria. Oxford University Press, Oxford, UK.

Beasley, D. W., C. T. Davis, H. Guzman, D. L. Vanlandingham, A. P. T. da Rosa, R. E. Parsons, S. Higgs, R. B. Tesh, and A. D. Barrett. 2003. Limited evolution of West Nile Virus has occurred during its southwesterly spread in the United States. Virology 309:190–195.

Bhatt, S., T. Lam, S. Lycett, A. L. Brown, T. Bowden, E. Holmes, Y. Guan, J. Wood, I. Brown, and P. Kellam. 2013. The evolutionary dynamics of influenza A virus adaptation to mammalian hosts.

Philosophical Transactions of the Royal Society B: Biological Sciences 368:20120382.

Brault, A. C. 2009. Changing patterns of West Nile Virus transmission: altered vector competence and host susceptibility. Veterinary Research 40:1–19.

Brault, A. C., C. Y. Huang, S. A. Langevin, R. M. Kinney, R. A. Bowen, W. N. Ramey, N. A. Panella, E. C. Holmes, A. M. Powers, and B. R. Miller. 2007. A single positively selected West Nile viral mutation confers increased virogenesis in American crows. Nature Genetics 39:1162–1166.

Brault, A. C., S. A. Langevin, W. N. Ramey, Y. Fang, D. W. Beasley, C. M. Barker, T. A. Sanders, W. K. Reisen, A. D. Barrett, and R. A. Bowen. 2011. Reduced avian virulence and viremia of West Nile Virus isolates from Mexico and Texas. American Journal of Tropical Medicine and Hygiene 85:758–767.

Briese, T., A. Rambaut, M. Pathmajeyan, J. Bishara, M. Weinberger, S. Pitlik, and W. I. Lipkin. 2002. Phylogenetic analysis of a human isolate from the 2000 Israel West Nile Virus epidemic. Emerging Infectious Diseases 8:528–531.

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

Chaves, L. F., L. C. Harrington, C. L. Keogh, A. M. Nguyen, and U. D. Kitron. 2010. Blood feeding patterns of mosquitoes: Random or structured? Frontiers in Zoology 7:1.

Ciota, A. T., D. J. Ehrbar, A. C. Matacchiero, G. A. Van Slyke, and L. D. Kramer. 2013. The evolution of virulence of West Nile Virus in a mosquito vector: implications for arbovirus adaptation and evolution. BMC Evolutionary Biology 13:1.

Danforth, M. E., W. K. Reisen, and C. M. Barker. 2015. Extrinsic incubation rate is not accelerated in recent California strains of West Nile Virus in *Culex tarsalis* (Diptera: Culicidae). Journal of Medical Entomology 52:1083–1089.

Davis, C. T., G. D. Ebel, R. S. Lanciotti, A. C. Brault, H. Guzman, M. Siirin, A. Lambert, R. E. Parsons, D. W. Beasley, and R. J. Novak. 2005. Phylogenetic analysis of North American West Nile Virus isolates, 2001–2004: evidence for the emergence of a dominant genotype. Virology 342:252–265.

Di Giallonardo, F., J. L. Geoghegan, D. E. Docherty, R. G. McLean, M. C. Zody, J. Qu, X. Yang, B. W. Birren, C. M. Malboeuf, and R. M. Newman. 2016. Fluid spatial dynamics of West Nile Virus in the United States: rapid spread in a permissive host environment. Journal of Virology 90:862–872.

Dobson, A., and J. Foufopoulos. 2001. Emerging infectious pathogens of wildlife. Philosophical Transactions of the Royal Society B: Biological Sciences 356:1001–1012.

Dohm, D. J., M. L. O'Guinn, and M. J. Turell. 2002. Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile Virus. Journal of Medical Entomology 39:221–225.

Duggal, N. K., A. Bosco-Lauth, R. A. Bowen, S. S. Wheeler, W. K. Reisen, T. A. Felix, B. R. Mann, H. Romo, D. M. Swetnam, and A. D. Barrett. 2014. Evidence for co-evolution of West Nile Virus and house sparrows in North America. PLoS Neglected Tropical Diseases 8:e3262.

Dwyer, G., S. A. Levin, and L. Buttel. 1990. A simulation model of the population dynamics and evolution of myxomatosis. Ecological Monographs 60:423–447.

Elderd, B. D., V. M. Dukic, and G. Dwyer. 2006. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. Proceedings of the National Academy of Sciences 103:15693–15697.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2014. Bayesian data analysis. Volume 2. Chapman and Hall/CRC Press, Boca Raton, Florida, USA.

Grinev, A., C. Chancey, E. Volkova, G. Añez, D. A. Heisey, V. Winkelman, G. A. Foster, P. Williamson, S. L. Stramer, and M. Rios. 2016. Genetic variability of West Nile Virus in US blood donors from the 2012 epidemic season. PLoS Neglected Tropical Diseases 10:e0004717.

Guerrero-Sánchez, S., S. Cuevas-Romero, N. M. Nemeth, M. Trujillo-Olivera, G. Worwa, A. Dupuis, A. C. Brault, L. D. Kramer, N. Komar, and J. G. Estrada-Franco. 2011. West Nile Virus infection of birds, Mexico. Emerging Infectious Diseases 17:2245–2252.

Hamer, G. L., U. D. Kitron, T. L. Goldberg, J. D. Brawn, S. R. Loss, M. O. Ruiz, D. B. Hayes, and E. D. Walker. 2009. Host selection by *Culex pipiens* mosquitoes and West Nile Virus amplification. American Journal of Tropical Medicine and Hygiene 80:268–278.

Heffernan, J., R. Smith, and L. Wahl. 2005. Perspectives on the basic reproductive ratio. Journal of the Royal Society Interface 2:281–293.

Johnson, B., T. Chambers, M. Crabtree, J. Arroyo, T. Monath, and B. Miller. 2003. Growth characteristics of the veterinary vaccine candidate ChimeriVax™–West Nile (WN) Virus in Aedes and *Culex* mosquitoes. Medical and Veterinary Entomology 17:235–243.

Keeling, M. J., and P. Rohani. 2008. Modeling infectious diseases in humans and animals. Princeton University Press, Princeton, New Jersey, USA.

Kilpatrick, A. M., P. Daszak, M. J. Jones, P. P. Marra, and L. D. Kramer. 2006. Host heterogeneity dominates West Nile Virus transmission. Proceedings of the Royal Society of London B: Biological Sciences 273:2327–2333.

Kilpatrick, A. M., S. L. LaDeau, and P. P. Marra. 2007. Ecology of West Nile Virus transmission and its impact on birds in the western hemisphere. Auk 124:1121–1136.

Kilpatrick, A. M., M. A. Meola, R. M. Moudy, and L. D. Kramer. 2008. Temperature, viral genetics, and the transmission of West Nile Virus by *Culex pipiens* mosquitoes. PLoS Pathogens 4:e1000092.

Komar, N., S. Langevin, S. Hinten, N. Nemeth, E. Edwards, D. Hettler, B. Davis, R. Bowen, and M. Bunning. 2003. Experimental infection of North American birds with the New York 1999 strain of West Nile Virus. Emerging Infectious Diseases 9:311–322.

Kuno, G., G.-J. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp. 1998. Phylogeny of the genus Flavivirus. Journal of Virology 72:73–83.

LaDeau, S. L., A. M. Kilpatrick, and P. P. Marra. 2007. West Nile Virus emergence and large-scale declines of North American bird populations. Nature 447:710–713.

Lanciotti, R., J. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. Volpe, M. Crabtree, and J. Scherret. 1999. Origin of the West Nile Virus responsible for an outbreak of encephalitis in the northeastern United States. Science 286:2333–2337.

Langevin, S. A., R. A. Bowen, W. K. Reisen, C. C. Andrade, W. N. Ramey, P. D. Maharaj, M. Anishchenko, J. L. Kenney, N. K. Duggal, and H. Romo. 2014. Host competence and helicase activity differences exhibited by West Nile viral variants expressing NS3-249 amino acid polymorphisms. PLoS ONE 9:e100802.

Loss, S. R., G. L. Hamer, E. D. Walker, M. O. Ruiz, T. L. Goldberg, U. D. Kitron, and J. D. Brawn. 2009. Avian host community structure and prevalence of West Nile Virus in Chicago, Illinois. Oecologica 159:415–424.

McMullen, A. R., F. J. May, L. Li, H. Guzman, R. Bueno Jr., J. A. Dennett, R. B. Tesh, and A. D. Barrett. 2011. Evolution of new genotype of West Nile Virus in North America. Emerging Infectious Diseases 5:785–793.

Moudy, R. M., M. A. Meola, L.-L. L. Morin, G. D. Ebel, and L. D. Kramer. 2007. A newly emergent genotype of West Nile Virus is transmitted earlier and more efficiently by *Culex* mosquitoes. American Journal of Tropical Medicine and Hygiene 77: 365–370.

Newman, C. M., F. Cerutti, T. K. Anderson, G. L. Hamer, E. D. Walker, U. D. Kitron, M. O. Ruiz, J. D. Brawn, and T. L. Goldberg. 2011. *Culex flavivirus* and West Nile Virus mosquito coinfection and positive ecological association in Chicago, United States. Vector-Borne and Zoonotic Diseases 11:1099–1105.

Ostroff, S. M. 2013. West Nile Virus: too important to forget. Journal of the American Medical Association 310:267–268.

Petersen, L. R., A. C. Brault, and R. S. Nasci. 2013. West Nile Virus: review of the literature. Journal of the American Medical Association 310:308–315.

R Development Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Reisen, W. K., C. M. Barker, Y. Fang, and V. M. Martinez. 2008. Does variation in *Culex* (Diptera: Culicidae) vector competence enable outbreaks of West Nile Virus in California? Journal of Medical Entomology 45:1126–1138.

Reisen, W., Y. Fang, and V. Martinez. 2005. Avian host and mosquito (Diptera: Culicidae) vector competence determine the efficiency of West Nile and St. Louis encephalitis virus transmission. Journal of Medical Entomology 42:367–375.

Richards, S. L., C. C. Lord, K. N. Pesko, and W. J. Tabachnick. 2010. Environmental and biological factors influencing *Culex pipiens quinquefasciatus* (Diptera: Culicidae) vector competence for West Nile Virus. American Journal of Tropical Medicine and Hygiene 83:126–134.

Roberts, M. 2007. The pluses and minuses of R0. Journal of the Royal Society Interface 4:949–961.

Ruel, J. J., and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. Trends in Ecology and Evolution 14:361–366.

Ruiz, M. O., L. F. Chaves, G. L. Hamer, T. Sun, W. M. Brown, E. D. Walker, L. Haramis, T. L. Goldberg, and U. D. Kitron. 2010. Local impact of temperature and precipitation on West Nile Virus infection in *Culex* species mosquitoes in northeast Illinois, USA. Parasites and Vectors 3:1.

Savage, H. M., D. Aggarwal, C. S. Apperson, C. R. Katholi, E. Gordon, H. K. Hassan, M. Anderson, D. Charnetzky, L. McMillen, and E. A. Unnasch. 2007. Host choice and West Nile Virus infection rates in blood-fed mosquitoes, including members of the *Culex pipiens* complex, from Memphis and Shelby County, Tennessee, 2002–2003. Vector-Borne and Zoonotic Diseases 7:365–386.

Simpson, J. E., C. M. Folsom-O'Keefe, J. E. Childs, L. E. Simons, T. G. Andreadis, and M. A. Diuk-Wasser.

2009. Avian host-selection by *Culex pipiens* in experimental trials. PLoS ONE 4:e7861.

Simpson, J. E., P. J. Hurtado, J. Medlock, G. Molaei, T. G. Andreadis, A. P. Galvani, and M. A. Diuk-Wasser. 2012. Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile Virus as a model system. Proceedings of the Royal Society of London B: Biological Sciences 279:925–933.

Smith, D. L., K. E. Battle, S. I. Hay, C. M. Barker, T. W. Scott, and F. E. McKenzie. 2012. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS Pathogens 8:e1002588.

Snapinn, K. W., E. C. Holmes, D. S. Young, K. A. Bernard, L. D. Kramer, and G. D. Ebel. 2007. Declining growth rate of West Nile Virus in North America. Journal of Virology 81:2531–2534.

Stan Development Team. 2016. Stan: A C++ library for probability and sampling, version 2.10.0. http://mc-stan.org

Styer, L. M., K. A. Kent, R. G. Albright, C. J. Bennett, L. D. Kramer, and K. A. Bernard. 2007. Mosquitoes inoculate high doses of West Nile Virus as they probe and feed on live hosts. PLoS Pathogens 3: e132.

Tsetsarkin, K. A., R. Chen, R. Yun, S. L. Rossi, K. S. Plante, M. Guerbois, N. Forrester, G. C. Perng, E. Sreekumar, and G. Leal. 2014. Multi-peaked adaptive landscape for chikungunya virus evolution predicts continued fitness optimization in *Aedes albopictus* mosquitoes. Nature Communications 3:1895–1906.

Tsetsarkin, K. A., D. L. Vanlandingham, C. E. McGee, and S. Higgs. 2007. A single mutation in chikungunya virus affects vector specificity and epidemic potential. PLoS Pathogens 3:e201.

Turell, M. J., M. L. O'Guinn, D. J. Dohm, and J. W. Jones. 2001. Vector competence of North American mosquitoes (Diptera: Culicidae) for West Nile Virus. Journal of Medical Entomology 38: 130–134.

Turell, M. J., M. O'Guinn, and J. Oliver. 2000. Potential for New York mosquitoes to transmit West Nile Virus. American Journal of Tropical Medicine and Hygiene 62:413–414.

Vaidyanathan, R., and T. W. Scott. 2006. Seasonal variation in susceptibility to West Nile Virus infection in *Culex pipiens pipiens* (L.) (Diptera: Culicidae) from San Joaquin County, California. Journal of Vector Ecology 31:423–425.

Vanlandingham, D. L., C. E. McGee, K. A. Klinger, N. Vessey, C. Fredregillo, and S. Higgs. 2007. Relative susceptibilties of South Texas mosquitoes to infection with West Nile Virus. American Journal of Tropical Medicine and Hygiene 77:925–928.

Vanlandingham, D. L., B. S. Schneider, K. Klingler, J. Fair, D. Beasley, J. Huang, P. Hamilton, and S. Higgs. 2004. Real-time reverse transcriptase–polymerase chain reaction quantification of West Nile Virus transmitted by *Culex pipiens quinquefasciatus*. American Journal of Tropical Medicine and Hygiene 71:120–123.

Vinogradova, E. B. 2000. *Culex pipiens pipiens* mosquitoes: taxonomy, distribution, ecology, physiology, genetics, applied importance and control. Number 2. Pensoft Publishers, Bulgaria.

Wonham, M. J., T. de Camino-Beck, and M. A. Lewis. 2004. An epidemiological model for West Nile Virus: invasion analysis and control applications. Proceedings of the Royal Society of London B: Biological Sciences 271:501–507.

Wonham, M. J., M. A. Lewis, J. Rencławowicz, and P. Van den Driessche. 2006. Transmission assumptions generate conflicting predictions in host–vector disease models: a case study in West Nile Virus. Ecology Letters 9:706–725.

Ziegler, U., J. Angenvoort, D. Fischer, C. Fast, M. Eiden, A. V. Rodriguez, S. Revilla-Fernández, N. Nowotny, J. G. de la Fuente, and M. Lierz. 2013. Pathogenesis of West Nile Virus lineage 1 and 2 in experimentally infected large falcons. Veterinary Microbiology 161:263–273.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1684/full

Chapter 3: Predicting West Nile virus transmission in North American bird communities using phylogenetic mixed effects models and eBird citizen science data

In my second of two chapters on WNV I present the results of a mechanistic model for WNV transmission in diverse North American bird communities. In this chapter I extend my collection of data on WNV transmission from the literature begun in Chapter 2 to include mosquito biting preferences, bird detectabilities, and bird abundance. Using these data and the phylogenetic relationship among these bird species, I expand the scope of my model to predict WNV transmission in any bird community in North America. I focus my analysis on Texas, USA, where I find that Northern Cardinals are the most important hosts for WNV and that increasing species richness decreases the potential for WNV spread.

The text I present here is a submitted second revision of a manuscript in preparation for Parasites & Vectors.

Author Contributions

MPK conceived the study with helpful feedback from BMB; MPK collected the data; MPK performed statistical analyses with helpful feedback from BMB; MPK wrote the first draft of manuscript and both authors revised the manuscript.

Acknowledgements

I thank my committee members Jonathan Dushoff and Ian Dworkin, the Bolker, Dushoff, and Earn labs, as well as Jo Werba for helping me refine and present this work.

Predicting West Nile virus transmission in North American bird communities using phylogenetic mixed effects models and eBird citizen science data

Morgan P Kain[1*] and Benjamin M Bolker[1,2]

*Corresponding author: kainm@mcaster.ca
[1]Department of Biology, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada
[2]Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada

Abstract

Background:
West Nile virus (WNV) is a mosquito-transmitted disease of birds that has caused bird population declines and can spill over into human populations. Previous research has identified bird species that infect a large fraction of the total pool of infected mosquitoes and correlate with human infection risk; however, these analyses cover small spatial regions and cannot be used to predict transmission in bird communities in which these species are rare or absent. Here we present a mechanistic model for WNV transmission that predicts WNV spread (R_0) in any bird community in North America by scaling up from the physiological responses of individual birds to transmission at the level of the community. We predict unmeasured bird species' responses to infection using *phylogenetic imputation*, based on these species' phylogenetic relationship to bird species with measured responses.

Results:
We focus our analysis on Texas, USA, because it is among the states with the highest total incidence of WNV in humans and is well sampled by birders in the eBird database. *Spatio-temporal patterns:* WNV transmission is controlled primarily by temperature variation across time and space, and secondarily by bird community composition. In Texas, we estimate WNV R_0 to be highest in the spring and fall when temperatures maximize the product of mosquito transmission and survival probabilities. In the most favorable months for WNV transmission (April, May, September, and October), we predict R_0 to be the highest in the "Piney Woods" and "Oak Woods & Prairies" ecoregions of Texas, and the lowest in the northern "High Plains" and "South Texas Brush County" ecoregions. *Dilution effect:* More abundant bird species are more competent hosts for WNV, and WNV R_0 decreases with increasing species richness. *Keystone species:* We predict that Northern cardinals (*Cardinalis cardinalis*) are the most important hosts for amplifying WNV and that Mourning doves (*Zenaida macroura*) are the most important sinks of infection across Texas.

Conclusion:
Despite some data limitations, we demonstrate the power of phylogenetic imputation in predicting disease transmission in heterogeneous host communities. Our mechanistic modeling framework shows promise both for assisting future analyses on transmission and spillover in heterogeneous multispecies pathogen systems and for improving model transparency by clarifying assumptions, choices, and shortcomings in complex ecological analyses.

Keywords:
American robin; Dilution effect; Flavivirus; Multiple imputation; Phylogenetic analysis; Zoonotic spillover

Background

West Nile virus (WNV), a mosquito-borne pathogen of birds, is a model system for studying vector-borne disease transmission and virulence evolution [1–6]. West Nile virus caused infrequent outbreaks in Israel, Egypt, India, France, and South Africa from 1937, when it was first isolated in Uganda, until the 1980s [4]. By the mid 1990s WNV had spread across much of Europe; it remains a moderate human and equine health burden in Europe and Africa today [5,7–12]. West Nile virus was first detected in North America in New York, USA in 1999, and by 2003 had spread to all contiguous US states, southern Canada and northern Mexico [1], and has now become the world's most widespread arbovirus [13]. The North American WNV epidemic caused population declines in numerous bird species [1,14,15] and hundreds of thousands of spillover infections in humans [16–18], including 23,000 reported cases of neuroinvasive WNV disease and more than 2,000 deaths between 1999 and 2017 [19].

The life cycle of WNV, which we introduce here from infected mosquito to infected mosquito, is sensitive to abiotic and biotic factors at every stage [4]. First, an infected mosquito infects susceptible birds ("mosquito-to-bird transmission"). Transmission probability during a feeding event depends on the viral load (titer) in a mosquitoes' salivary glands, which is determined by the length of time the mosquito has been infected and the viral replication rate in the mosquito [20]; replication rate is a function of the dose the mosquito received when it became infected, the mosquito species, and environmental variables such as temperature [6]. A mosquito's overall ability to transmit infection to a susceptible host is called "vector competence" [21,22]. Which bird species become infected depends on mosquitoes' biting preferences [23] and on the abundance of each bird species in the community.

In the second step of transmission, infected birds infect susceptible mosquitoes ("bird-to-mosquito transmission"). The probability that a susceptible mosquito becomes infected during a feeding event depends on titer in the bird species, the species of the mosquito, and environmental variables such as temperature [6]. Critically, bird species vary considerably in both their physiological capacity for transmitting infection to mosquito vectors because of differences among species in survival and virus titer (which together comprise "host competence"), and in their relative contribution to the pool of infectious

mosquitoes because of differences in their abundance and attractiveness to mosquitoes [23].

WNV has been intensely studied, including models and/or empirical analysis of: prevention strategies for WNV [24,25]; ecological factors associated with the spread of WNV [26–29]; risk assessment for invasion into new locations [11,30–32]; human infection risk [16,33–36]; and the importance of individual bird species in transmitting WNV [1,6,37,38]. This work has contributed substantially to our understanding of the dynamics of WNV. For example, Wonham et al. [24] and four others reviewed in [39] laid the foundation for WNV transmission models, providing insight into the threshold number of mosquitoes at which WNV $R_0 = 1$ [24], the impact of bird mortality on transmission [40], and the transition from an epidemic to endemic state [41]. However, all of these studies used a differential equation framework that ignores much of the heterogeneity in transmission probabilities over the course of infection and variation among hosts and mosquitoes. Vogels et al. [29] do incorporate transmission probabilities from three vector species at three different temperatures; however, they considered only a single bird species. Kilpatrick et al. [1] and Peterson et al. [37] began to address the abundant variation in competence among bird species, which led to a variety of work on the connection between specific bird species and human infection risk [16,35,36].

Most work neglects much of the heterogeneity in the life cycle of WNV: all of these analyses were focused narrowly on a small subset of the species found in diverse bird communities and/or use a small fraction of the available empirical data. Ideally, predictions for the spread of WNV in diverse communities of birds would be obtained from a mechanistic model that uses as much of the available empirical data as possible on individual-level processes to scale up to transmission at the level of the community while retaining the heterogeneities in WNV transmission. These data include–among many other axes of heterogeneity–the physiological responses of all of the bird species in the community; the biting preferences of mosquitoes on these bird species; and the relative abundance of each bird species. Relative to phenomenological models, mechanistic models are often more powerful because they are better at prediction in conditions beyond those observed [42,43], and help elucidate biological unknowns when they fail [44]. A mechanistic model for WNV would allow for estimation of the force of infection of

WNV in any bird community and help researchers explore causal links between bird community composition and human infection risk.

In North America, there have been over a hundred infection experiments of mosquitoes and birds (see [6] for a synthesis of these data), and extensive studies on mosquito feeding preferences (for a review see [23]; for examples of field observations see [45,46]). Despite this work, bird communities across North America contain hundreds of bird species with unmeasured physiological responses to WNV and unknown mosquito biting preferences. Because of this gap, WNV spread has not yet been predicted mechanistically using full bird communities.

We present a model for predicting WNV R_0 for bird communities in any state or province in North America (aggregated in space and time by county, month, and year) or larger region: `R` code is provided in the online supplemental material. While our model is set up to provide estimates of WNV transmission anywhere in North America, sufficient information about bird species abundance (e.g. from eBird data) may be unavailable in some rural locations in the US and many locations in Canada and Mexico. To get around the problem of unmeasured responses to WNV for many bird species, we estimate missing bird species' responses using these species' phylogenetic relationship to bird species with measured responses, a technique we call "phylogenetic imputation". This is a general method that can be used to model the correlated responses of multiple species and efficiently estimate the response (e.g. traits, response to infection) of species with little or no data (see [47] for a similar method and application). This technique allows us to scale up from the physiological responses of individual birds to disease transmission at the scale of the whole community by considering species-level variation in the physiological response to WNV and the biting preference by mosquito vectors of all of the birds in the community. This allows our model to retain all known heterogeneities in the life cycle of WNV associated with the bird community.

A model allowing for all important WNV transmission heterogeneities would certainly need to allow for spatial and temporal variation in mosquito populations, temperature, and the effects of temperature on transmission probabilities, mosquito survival, and biting rate [12], each of which has large

effects on WNV transmission [29]. While our model considers spatial and temporal variation in temperature and resulting variation in transmission probabilities and mosquito survival, we assume a single homogenous population of mosquitoes because of a lack of data on mosquito populations. Thus, while our model is a step in the right direction, ignoring variation in vector competence among mosquito species is a shortcoming of our approach.

We use a variety of datasets to fit our model including: laboratory infections of birds and mosquitoes (full citations are available in the online supplemental material; further details available in [6]), field data on mosquito biting preferences [45], bird body size data from a searchable database [48], bird detectability from field sampling (citations are listed in the online supplemental material), the comprehensive phylogeny of birds [49,50], and citizen science data on bird abundance from eBird, the Cornell Laboratory of Ornithology citizen science database [51].

We show how our model can be used to predict the intrinsic reproductive number (R_0) of WNV, the expected number of new infections a single infected individual generates in an otherwise susceptible population. We focus our analysis on Texas, USA, because it is among the states with the highest total incidence of WNV in humans [52] (Texas had an estimated total of 534,000 cases between 2003 and 2010 [16], and Dallas county specifically had the highest recorded number of cases anywhere in the US in a 2012 nationwide WNV epidemic [53]), and is well sampled in the eBird database. We use R_0 as a metric to compare transmission potential among bird communities; we do not use R_0 as a metric to predict the exact size of a new epidemic, which would require detailed information on bird seroprevalence. We examine spatio-temporal patterns in WNV R_0 across Texas and determine which bird species in Texas are the best and worst hosts for propagating WNV. For this case study we assume a single mosquito species, which allows us to address our primary focus of variation in the bird community.

Using our imputed responses for full bird communities and R_0 estimates in Texas, we test both an assumption and a prediction of the *dilution effect*, which is the hypothesis that states that increasing biodiversity (in either species richness or evenness) will decrease R_0 or another quantity associated with the spread of disease such as the number of spillover infections into non-target hosts [54–56]. Previous

work in this system has found variable support for the dilution effect [28,33,57,58]. In an attempt to clarify these variable results, we test if more abundant bird species are better hosts for WNV (an assumption of the dilution effect), and whether bird species richness is positively or negatively correlated with WNV R_0 (an amplification or dilution effect respectively).

We structure our paper and supplemental code to serve as a reference for future work analyzing ecological problems that require multi-faceted mechanistic models—mechanistic models that require many (potentially compartmentalized) sub-models, each of which relies on different data sources. We provide a detailed description of each of our sub-models and give reasons for our statistical choices; we emphasize principled ways to estimate missing data, and the importance of propagating uncertainty. The online supplemental material provides extensively commented `R` code and a complete list of all data cleaning and analysis steps required to obtain estimates for the R_0 of WNV in any region in North America using a single compressed eBird data file available upon request from [59].

Methods

Model overview

We introduce our model by working backwards, from the overarching biological questions to the specifics of individual models. We begin by describing our primary model outcomes. We then explain our method for calculating the R_0 of WNV. Finally, we detail how we estimated each parameter in the equation for R_0 using individual sub-models, and how we linked these estimates and propagated uncertainty to calculate R_0. Table 1 describes the components of our overall model and how they fit into our analysis.

Table 1: Sub-model details for our multi-faceted ecological model for WNV R_0. The two transmission steps (Column 2) of WNV's life cycle are: mosquito to bird (M-to-B) transmission—transmission from an infected mosquito to a susceptible bird; bird to mosquito (B-to-M) transmission—transmission from an infected bird to a susceptible mosquito. Citations accompany data available in the online supplemental material; details on data extraction can be found in [6].

Component of community R_0	Transmission step	R_0 equation component (see Eq.2)	Data Soruces(s)	For details see:
Raw eBird counts of bird species *i*	M-to-B	Component of ω_{Si} and $\omega_{\mu i}$	1,437,050 complete lists submitted between 2000-2017 in Texas, USA	*Methods–Model components: Bird community*
Detectability of bird species *i*	M-to-B	Component of ω_{Si}	12 publications, which included estimates for 475 Bird Species	*Methods–Model components: Bird detectability*
Mosquito biting preference on bird species *i*	Both	Component of ω_{Si}	[45] and eBird records for the same spatio-temporal sampling period	*Methods–Model components: Mosquito biting preference*
Mosquito incubation of WNV	M-to-B	Determines P_{MBd}	9 publications, which included 45 infection experiments (see online supplemental material and [6])	*Methods–Community* R_0; model from: [6]
Mosquito survival	M-to-B	S_{Md}	[118]	*Methods–Community* R_0; model from: [6]
Mosquito biting rate	Both	δ	[46] from [24] and [61]	*Methods–Community* R_0
Titer profile of bird species *i*	B-to-M	T_{ij}	30 publications, which included 111 infection experiments of 47 bird species (see online supplemental material and [6])	*Methods–Model components: Bird titer profile and survival*
Survival of bird species *i*	B-to-M	S_{Bij}	30 publications, which included 111 infection experiments of 47 bird species (see online supplemental material and [6])	*Methods–Model components: Bird titer profile and survival*
Bird-to-mosquito transmission probability — titer	B-to-M	P_{BMij}	20 publications (see online supplemental material and [6])	*Methods–Model components: Bird titer profile and survival* ; model from: [6]
Number of mosquitoes per bird	B-to-M	n_{MB}	Based loosely on [46]	*Methods–Community* R_0

Model outcomes

First, we focus on spatial and temporal patterns in R_0 at the level of the community; we calculate WNV R_0 for bird communities between 2000 and 2017 separated spatially by county and temporally by month and year, and then fit a spatio-temporal model to the resulting WNV R_0 estimates which includes 11 ecoregions in Texas, human population density, temperature, and year as predictor variables. Second, we determine which bird species have the largest predicted impact on R_0 in Texas, USA. We quantify the importance of each species within each community by calculating the proportional change in R_0 that would be predicted to occur if that species were removed from the community and replaced by the other species in community in proportion to their relative abundance. We consider species whose removal strongly increases or decreases R_0 as the least or most competent birds for WNV, respectively. In the language of the dilution effect [54– 56], species that increase R_0 when removed can be defined as "diluters", and those that decrease R_0 when removed as "amplifiers". We test if more abundant bird species are more physiologically competent for transmitting WNV and if an increase in species richness is predicted to decrease WNV R_0.

Community R_0

We calculate R_0 as the expected number of mosquitoes that become infected following the introduction of a single infected mosquito into a population of susceptible birds and otherwise uninfected mosquitoes. This calculation assumes that all mosquitoes have identical biting preferences and vector competence. We break R_0 into two transmission steps: mosquito-to-bird transmission, which measures the expected number of each bird of species *i* that would become infected by a single infected mosquito; and bird-to-mosquito transmission, which calculates the expected number of mosquitoes infected by the infected birds of species *i* calculated in the mosquito-to-bird transmission step. Written in this way, the sum of bird-to-mosquito transmission gives the number of new infected mosquitoes resulting from the single infected mosquito, which is the R_0 of WNV.

Mosquito-to-bird transmission is calculated by:

$$\mu_i = \omega_{Si} \sum_{d=1}^{D} P_{MBd} * S_{Md} * \delta, \quad (1)$$

where μ_i is the number of birds of species i that become infected when a single infected mosquito is introduced into a community of susceptible birds. The quantity ω_{Si} is the scaled proportion of susceptible individuals of bird species i, which is given by the observed proportions of species i (determined by eBird data, see *Methods: Bird Community*), weighted by the detectability of species i (see *Methods: Bird Detectability*) and the mosquito biting preference on species i (see *Methods: Mosquito biting preference*). The derivation of ω_{Si} is given in *Methods: Mosquito biting preference*. Total transmission from the infected mosquito to susceptible birds is given by a sum over D, the duration of the mosquito's infectious period. This sum is a measure of vector competence, the total ability of a vector to transmit infection to a susceptible host [21], a key component of which is the transmission probability per feeding event [22]. The probability of transmission per mosquito bite on each day (P_{MBd}) follows a logistic function of titer in the mosquito's salivary glands, which is a function of time since infection, dose received from the infected bird, temperature, mosquito species, and WNV strain (see [6] for a synthesis of these data). Here we assume that the mosquito is introduced into the susceptible population of birds on the first day following infection with the WN02 strain of WNV with a dose of $10^{5.5}$ viral particles. We predict mosquito incubation rate of WNV and mosquito survival (S_{Md}; estimates for mosquito survival are taken from a model for mosquito survival fit in [6]) for each Texas bird community using the average temperature in each Texas county by month and year with temperature data obtained from NOAA [60]. We ignore the effect of mosquito species, which was fitted as a random effect in [6], due to the absence of data. These simplifications do not affect the relative effect of bird species, but will affect overall R_0 values, and could affect spatio-temporal patterns. Finally, δ is mosquito biting rate with units of bites per mosquito per day. We assume a constant mosquito biting rate of 0.14 per mosquito per day (as assumed by [46], taken from [24,61]).

WNV R_0 is calculated using the sum of bird-to-mosquito transmission:

$$\mathcal{R}_0 = \sum_{i=1}^{I} (\omega_{\mu i} \sum_{j=1}^{8} P_{BMij}\,(T_{ij}) * S_{Bij})\, n_{MB} * \delta, \quad (2)$$

Eq.2 gives the expected number of new mosquitoes infected by the expected number of each bird of species i infected by a single mosquito (given by μ_i in Eq.1), weighted by the biting preference of mosquitoes on species i, summed over bird species to obtain overall R_0. The transmission probability from an infected bird of species i to a susceptible mosquito on day j (P_{BMij}) is a function of a bird's titer (T_{ij}). Transmission probability is discounted by the bird's survival probability up to day j (S_{Bij}). We measure bird titer and survival until day 8, which is one day longer than previous measures of host competence [20,62] and long enough to capture all known detectable measures of titer in birds. The inner summation over j captures a quantity commonly called "host competence", which we call "physiological competence" to emphasize that this component is not scaled by mosquito biting preference. Classically, host competence is defined as the daily sum of host-to-vector infection probability over the course of a host's infectious period [20,62], assuming a single mosquito bite per day on an infected bird. Here, when multiplied by $\omega_{\mu i}$, this quantity gives the number of new mosquitoes that infected individuals of species i (μ_i) infect, arising from a single originally infected mosquito (in Eq.2 the entire quantity inside the large parentheses). The R_0 of WNV is given by the sum of this quantity over all infected bird species multiplied by a constant ratio of mosquitoes to birds (n_{MB}) (in the absence of better data we assume a ratio of 3 based approximately on sampling conducted by [46] in New Haven, CT) and the number of bites per mosquito per day (δ).

We focus on estimating the parameters associated with the bird community, which includes ω_{Si}, $\omega_{\mu i}$, P_{BMij}, T_{ij}, and S_{Bij}. For mosquito-to-bird transmission probability (parameters P_{MBd} and S_{Md}) we use estimates from the models fit in [6] and single values from the literature for mosquito biting rate (δ) and the ratio of mosquitoes to birds (n_{MB}). While the mosquito-to-bird ratio and mosquito biting rate will in reality be a function of parameters that vary both spatially and temporally such as ecoregion, season, temperature [29], as well as human population density, we assume a constant mosquito-to-bird ratio here

because of a lack of sufficient data on spatial and seasonal variation in this ratio across Texas and because our primary focus is on estimating R_0 as a function of the bird community. Because we assume no interaction between mosquito species and bird species in the probability of infection, and because the remaining parameters are scalars, differences in these parameters will affect the overall magnitude of R_0 estimates but will not affect qualitative patterns in R_0 due to variation among bird communities in space and time.

In *Methods: Model components* we further unpack Eq.1 and Eq.2 (e.g. ω_{Si}) and describe how we estimated each of the parameters associated with the bird community. The data and models that informed all parameters of both Eq.1 and Eq.2 are described in greater detail in Table 1.

Phylogenetic imputation

The primary difficulty in estimating community competence for a diverse community of birds is that physiological responses to WNV and mosquito biting preferences are unknown for most bird species. Obtaining these data for every species in a diverse community of birds would be infeasible. To address this problem, we use a form of phylogenetic analysis that we call "phylogenetic imputation" in which we fit models using all of the data that is currently available for a given response (e.g. a bird species' titer profile) and estimate the response of species with missing data using the phylogenetic relationship between the missing species and the species for which we have data.

The effects of a predictor variable on the response of multiple species can be modeled using the phylogenetic relationships among the species to estimate the correlation among observations. Classic phylogenetic regression approaches assume a correlated-residual model using phylogenetically independent contrasts (PICs), where the residuals evolve as a Brownian motion process [63]; in other words, residuals are phylogenetically correlated. Many recent approaches, including phylogenetic generalized linear mixed models (PGLMM) [64], Pagel's λ [65], and Blomberg's κ [66], expand upon Felsenstein's PICs by incorporating extra parameters that correct for bias, and by partitioning the phylogenetically correlated residual variation into phylogenetically uncorrelated residual variation

(observation error or tip variation) and phylogenetic signal (biological/evolutionary process error) [67].

Here we use a newly implemented method built on the `lme4` package in `R` that incorporates phylogenetic correlations by modeling them as random effects and allows for random slopes (i.e. phylogenetic signal in response to change in the predictor variable), random interactions, and nested random effect models, and is orders of magnitude faster than alternative methods [68]. Like most previous methods, the evolutionary history for each species is modeled as a sequence of Normal independent errors. Thus, the portion of a species' response attributable to its evolutionary history can be calculated as the sum of the evolutionary change that occurred on each of the internal branches in the phylogeny leading to that species.

We estimate missing values for bird responses (e.g. bird titer) using multiple imputation (where each missing value is replaced by random samples from a distribution of plausible values [69]). To impute, we first fit a phylogenetic mixed model to all of the species for which we have data. Then, for each species without data, we first sum the evolutionary change in the response variable that occurred on all branches of the phylogeny leading to the most recent common ancestor between the species with a missing response and the most similarly related species that has data and was included in the mixed model. This gives the effect on the response variable of the species' shared evolutionary history up to the time when these species diverged. To obtain these values we draw random Normal (multivariate if the mixed model includes multiple correlated species-level random effects) samples for each branch, with means equal to the conditional modes of each branch multiplied by the branch length and variances equal to the conditional variances of each branch multiplied by the square of the branch lengths. Then, the evolutionary change that has occurred since the two species diverged is estimated by drawing random Normal (multivariate Normal if the mixed model includes multiple correlated species-level random effects) samples with a mean of zero (the expected value for each unmeasured species is equal to that of the most closely related measured species because of the assumption of Brownian motion) and standard deviation (sd) equal to the estimated sd of the species-level random effect(s) multiplied by the evolutionary distance (branch length) from the most recent common ancestor of the most closely related

measured species. Together, these estimates give the estimated total effect of a species' evolutionary history on a given response. The remaining portion of a species' response is given by the fixed effects (e.g. body size) and other non-species-level random effects (e.g. variation among infection experiments).

For our analysis we used a bird consensus phylogeny that was calculated using 1000 trees downloaded from [49] (Stage2_MayrPar_Ericson_set1_decisive.tre) [50] using `DendroPy` [70] and methods described in [71].

Phylogenetic imputation validation

We validate our phylogenetic imputation method in two ways: first, we calculate conditional R^2 using the methods outlined in [72,73] for models with and without a species level phylogenetic random effect. We estimate conditional R^2 using code from the `R` package `MuMIn` [74], adapted to accommodate the structure of the phylogenetic mixed model objects. Second, we use blocked leave-one-out cross validation [75] at the level of species for models with and without the species level phylogenetic random effect to assess the effects of phylogenetic imputation on out-of-sample error. We present additional details and results for each of these forms of validation in the online supplement. For a vignette on the phylogenetic models built on `lme4` see [68].

Model components

Bird titer profile and survival

We modeled bird infection profiles and mortality probabilities using data from experimental infections of 47 bird species collected from 30 publications containing 113 individual infection experiments; most of these data have been presented previously [6]. For the bird titer, bird survival, and bird-to-mosquito transmission models in this paper we grouped data from the two primary WNV strains, NY99 and WN02 ([6] were unable to detect a clear difference between the NY99 and WN02 strains).

To model bird titer profiles we used a log-normal mixed effects model; fixed effects included a

Ricker function of day (using day and log(day) as predictors of log-titer; see supplemental material or [76] for more information), infectious dose, bird body size, and the interaction between day and bird body size. We used a random intercept and slope over both day and log(day), which are constrained by the phylogenetic relationship among the species. We also included random intercepts for citation and infection experiment.

To model bird survival we used a generalized linear mixed effects model (GLMM) with a binomial error distribution and complementary log-log link, where the number of birds dying on a given day was taken as the number of "successes" and the number of birds that survived that day as "failures". This model estimates a bird's daily log-hazard [77], which can be back-transformed to estimate daily mortality probability and cumulative survival probability using the cumulative product of the complement of the daily mortality probabilities. We modeled bird survival using the main effects of titer, day, and bird body size as fixed effects; citation, infection experiment, and bird species (phylogenetically constrained) were modeled using random intercepts (due to a lack of data we were unable to estimate species-level variation in sensitivity to titer).

The bird body size data used in both models were obtained from the searchable digital edition of Dunning (2008) [48]. Body size data was averaged if data for a given species was available for both sexes or multiple sub-species. Approximately 7% of the species in the Texas eBird dataset did not have mean body sizes reported in [48] but did have minimum and maximum values reported. The body size for these species was taken as the center of the range. Approximately 0.3% of the species in the Texas eBird dataset were not represented at all in [48]. For these species, the body sizes of all congeners were averaged.

Bird community

We obtained bird abundances data from the Cornell Laboratory of Ornithology citizen science database *eBird* [51,59]. We used all complete checklists [51,80] submitted between January, 2000 and December, 2017. Complete checklists are defined as a report of *all* birds (number of individuals of all species) that

are seen on a given outing. Checklists were aggregated spatially at the level of Texas counties for each month between January, 2000 and December, 2017, which resulted in a total of 30,188 bird communities containing a total of 679 unique species. To match scientific names, which occasionally differed between eBird and the consensus phylogeny, we used an automated lookup procedure to search both the IUCN [81] and Catalogue of Life [82] databases. All unmatched names following the automated lookup were matched by hand using manual searches (< 1% of species).

We focus on results for a reduced eBird dataset that included 2,569 communities and a total of 645 bird species, with a median occurrence (proportion of communities in which bird species *i* was sampled) of 13% (95% of species between .04% and 82%; a total of 167 species were recorded in less than 1% of the communities). In the online supplement we present results for the complete Texas eBird dataset, which included all 30,188 available communities and 679 species. We subset our data for the main analysis because many of the Texas bird communities were under-sampled (e.g. 13,254 communities were sampled with 5 or fewer lists) and therefore these data are unlikely to be a good representation of the true bird community. The 2,569 bird communities were chosen because they were all sampled with a minimum effort of 80 complete checklists. We chose 80 lists in an attempt to maximize the number of communities for our analysis while minimizing the retention of under-sampled communities. To optimize the tradeoff between number of communities and data quality, we resampled 5-120 complete lists from the 46 most sampled communities (communities with greater than 1,300 lists) 100 times. We calculated the proportion of species missing in the subsampled communities as well as the root mean squared error (RMSE) in the relative proportions of all species between the two communities. Using the rate of change in RMSE and species retention (Figures S1, S2), we determined that with fewer than 80 complete lists, the gain in total number of communities was not worth the increased error rate and loss of species representation, while at greater than 80 complete lists the loss in communities was too large for the small decrease in error and species loss. For full simulation results see the online supplement (*Methods: Community resampling,* Figures S1, S2).

Bird detectability

We scaled raw bird counts by the detectability of each bird species to correct for incomplete sampling of bird communities and to control for variation in the quality of eBird records; alternatively or additionally, eBird lists can be weighted by user skill [80,83]. We searched for data on the maximum detection distances of birds using Google scholar with the following search criteria: "X" maximum detection distance, "X" maximum detection radius, "X" effective detection distance, and "X" effective detection radius, where "X" took each of: landbird, land bird, waterbird, water bird, waterfowl, seabird, sea bird, and marsh bird. In all cases the first 60 hits were assessed for relevant information. Because of overlapping results, a total of 1,440 titles, abstracts, and/or entire papers were read for relevant data. In total, we took data from 12 sources which contained maximum detection distances for 469 bird species. However, we failed to find detection distances for waterfowl and shore birds; maximum detection distances (roughly intermediate to values between woodland species and seabirds) were assigned to 21 waterfowl and shore birds based on detection probabilities in the literature and our knowledge of the natural history of these species (*personal birding experience* [84]).

In order to fill in missing information for detection distance, we used the results of our literature search to fit a phylogenetic mixed effects model. Maximum detection distances for species in the Texas eBird data were estimated using a GLMM with a log-normal error distribution. Body size was used as a fixed effect and species was included as a phylogenetic random effect. The eBird counts for each species were then adjusted by multiplying counts by the ratio of the maximum detection distance in the community to the detection distance of each species. Using the square of maximum detection distance to reflect the relative spatial area sampled for each species may also be an appropriate method for adjusting raw eBird counts. We chose linear scaling here because 50% of lists were transects and because squared distance generated unrealistic outliers.

Mosquito biting preference

Finally, bird species proportions were adjusted using the biting preferences of mosquitoes, which scales

true bird proportions to the proportions that mosquitoes "see". Because mosquitoes (*Culex* sp. and others) prefer some hosts to others [20,45,85], this step is required to appropriately translate each bird's physiological response (a bird's mosquito infecting potential) into realized infections of mosquitoes [20]. A mosquito's biting preference on bird species i can be calculated as the rate of mosquito feeding on species i relative to its abundance in the community [20]: $\beta_i = {}^{f_i}/_{a_i}$, where f_i is the fraction of total blood meals from species i, and a_i is the proportion of species i in the community. Experimentally, f_i is determined by sampling mosquitoes and determining the species origin of blood recovered from the mosquitoes; bird surveys are used to determine a_i ([20,45,46]). A value for $\beta_i = 1$ indicates that a bird species is bitten exactly in proportion to its representation in the community. A value of $\beta_i > 1$ or $\beta_i < 1$ indicates a bird species that is preferred or avoided by mosquitoes, respectively. At one extreme, a bird with high infectious potential (high titer and low mortality) may contribute very little to the spread of WNV if it is avoided by mosquitoes. At the other extreme, a bird with low physiological competence (low titer and/or high mortality) may contribute substantially to the spread of WNV if it is among the most preferred species in a community. For example, American robins (*Turdus migratorius*) have been found to infect the largest, or close to the largest, proportion of mosquitoes of any bird species in some bird communities in eastern USA because of their high abundance and mosquito preference [20,45,46,86,87], in spite of their relatively low titer [6].

In previous studies, when the blood of bird species i was recorded in a mosquito, but bird species i was unobserved in the community, the bird was either assigned a proportion corresponding to the rarest bird measured [45], or dropped from the analysis [20]. If bird species i was observed but its blood was not detected in a mosquito, it was assumed that a single mosquito was observed with the blood of bird species i [20,45]. While convenient, the assignment of arbitrary values to missing data here leads to biting preferences spanning three orders of magnitude [20,45], which seems biologically implausible. Alternatively, a Bayesian statistical model can be used to estimate mosquito biting preference (which is not directly observed), when bird species i or its blood is not observed. Here we use a multinomial model

in Stan [78], interfaced with `R` using `rstan` [79]. We model bird proportions using data from [45] and a Dirichlet prior, the conjugate prior to the multinomial distribution [88]. The Dirichlet prior was set proportional to eBird observations for the same location and dates as the sampling originally conducted in [45]; we used all complete checklists in a circle with radius 08' around the focal point of 41°42'N, 87°44'W given as the center of the surveys conducted in [45] for the months of May and October in 2006-2008; this area is shown in Figure S3.

The fraction of total blood meals in mosquitoes was modeled using a Gamma error distribution with data from [45] and a Gamma prior (shape = 0.25, scale = 0.25). This prior distribution has a mean equal to one, median less than one and moderate dispersion, which assumes that birds are preferred in proportion to their abundance on average; the majority of bird species are preferred a bit less than proportional to their relative abundance, while a few bird species are preferred much more than proportional to their relative abundance. This Dirichlet-multinomial model estimates mosquito biting preferences for all of the species recorded on eBird between May and October in 2006-2008 in Cook County, IL.

Estimates of mosquito blood meals from the Dirichlet-multinomial Stan model were then used to impute biting preference on bird species in the Texas dataset by fitting a GLMM with Poisson-distributed error (which includes a species-level phylogenetic random effect) to the biting preferences estimated by the Dirichlet-multinomial model. This step assumes that a mosquito's biting preference on species *i* is the same in Illinois as in Texas; both states share *Cx. tarsalis*, while *Cx. pipiens* is unique to Illinois and *Cx. quinquefasciatus* is unique to Texas [89], making this an unavoidable oversimplification. Biting preference estimates were scaled to a mean of one, and were then used to weight the observed proportions of each bird species. The weighted proportions of each bird species were obtained using:

$$\omega_\iota = \frac{\beta_i \alpha_i \delta_i}{\sum_{i=1}^{I} \beta_i \alpha_i \delta_i}, \tag{3}$$

where ω_i is the adjusted proportion of species i, α_i is the unweighted proportion of each species determined directly from eBird data, β_i is mosquito biting preference on species i, and δ_i is the ratio of

the maximum bird detectability in the community to the detectability of species *i*. The scaling in this equation is equivalent to using a weighted Manly's α index [90].

Spatio-temporal patterns in WNV R_0

To determine the spatio-temporal patterns in WNV R_0 we fit a generalized additive model (GAM) using the `mgcv` package in `R`. We use this as a proof of concept example to show how the imputed physiological responses of birds and mosquito biting preferences can be used to predict larger scale patterns. This model included thin plate splines for the log of human population density, temperature, and year. We stress that in the absence of data on mosquito communities on the scale of the bird communities, the R_0 estimates from this model are driven by variation in bird communities and temperature only and cannot be taken at face value as accurate estimates of actual WNV transmission potential.

We first attempted to fit a model using the proportion of each ecoregion in each county, but could not overcome issues of concurvity (analogous to co-linearity in a GAM model [91]) in this model. Instead, we fit a simplified model using a Markov random field to model the effects of ecoregion under the simplified assumption that each county had only a single ecoregion, which we chose as the most abundant ecoregion in each county. We fit a random effect of county to control for repeated measures within counties and to account for spatial variation within ecoregion. Ideally, we would also model fine-scale spatial variation using a thin plate spline over latitude and longitude coordinate pairs; however, models that included this predictor suffered greatly from concurvity problems. We used the inverse of the variance in R_0 estimates as weights.

The 11 major different ecoregions in Texas, population density, and county spatial shape data were obtained from [92]. This model provides estimates of both seasonal and long-term trends in WNV R_0 as the structure of bird communities have changed in the past two decades (due to disturbances such as habitat change [93]; habitat destruction [94]; climate change [95]; and the effect of the WNV epidemic itself [1]) as well as spatial estimates of WNV R_0 by county.

Propagation of uncertainty

Multi-faceted ecological models will underestimate uncertainty (e.g. the width of confidence intervals on estimates of outcomes of interest) if the point estimates from each sub-model are used while neglecting their uncertainty. Point estimates may also differ between models with or without uncertainty because nonlinear transformations of distributions will change the expected value, a phenomenon known as Jensen's inequality [96]. We focus on results from a model with all uncertainty propagated, but briefly discuss the impacts of ignoring uncertainty on both our quantitative and qualitative conclusions (for more detailed results see the online supplemental material). Table 2 gives a list of the sources of uncertainty and how each source was propagated.

Table 2: Details about each source of uncertainty.

Source of Uncertainty	Description	Method of propagation
Fixed effects	Uncertainty in the fixed effects for each sub-model	1000 multivariate (or univariate depending on the model definition) normal samples using the means and vcov matrix of the fixed effects
Phylogenetic random effect	Uncertainty in the amount of evolutionary change in the response variable (e.g. bird titer) that has occurred over each branch of the phylogeny	1000 multivariate (or univariate for models with a single species-level random effect) normal samples for each branch, with means equal to the conditional modes of the species-level random effect for each branch multiplied by the branch lengths and variance equal to the variance of the conditional models of the random effects for each branch multiplied by the squared branch lengths
Phylogenetic tip variation	Evolutionary change that has occurred after the divergence of the species whose response is being imputed from its most closely related species that has an empirically measured (and estimated) response	1000 multivariate (or univariate for models with a single species-level random effect) normal samples with mean 0 (because of the assumption of Brownian motion), and sd equal to the sd of the species-level random effect multiplied by the length of the final (most recent in time) branch leading to the species in question
Other random effects	Uncertainty due to variation among studies and infection experiments	1000 univariate normal samples for each random effect with mean equal to 0 and sd equal to the estimated sd
Stan model overall uncertainty	Summary of the entire uncertainty associated with the three Stan models used in the transmission steps between mosquitoes and birds (bird-to-mosquito transmission probability, mosquito-to-bird transmission probability, and mosquito biting preference)	1000 samples from the posterior distributions for each of the Stan models

We set up our sub-models in the R code provided in the online supplemental material so that each source of uncertainty can be set individually to be either propagated or ignored, which can be used to obtain a first approximation (assuming independence of errors) for the relative effects of uncertainty in each sub-model on uncertainty in R_0 and on spatio-temporal patterns in R_0. We briefly discuss which sources of uncertainty have the largest impact on our conclusions in the online supplemental material.

Results

Community R_0

WNV transmission is controlled primarily by temperature variation across time and space. In Texas, we estimated WNV R_0 to be highest in the spring and fall when temperatures maximize the product of mosquito transmission and survival probabilities (across all ecoregions in April: median R_0 = 2.16, median temperature across all Texas counties = 19°C; May: R_0 = 2.31, 23°C; October: R_0 = 2.27, 19°C) (Figure 1). Within these favorable months, we estimate R_0 to be highest in the "Piney Woods" ecoregion (median R_0 = 2.29) and "Oak Woods & Prairies" (median R_0 = 2.28) ecoregions of Texas, and the lowest in the northern "High Plains" ecoregion (median R_0 = 1.46). Despite these large differences at the larger scale of ecoregions, large uncertainty in the R_0 of individual communities makes it difficult to be certain about the size of the true variation in space and time. For example, despite median estimates of $R_0 > 1$ for 96% of communities in the most favorable months, the 95% CI for all of these communities includes R_0 = 1 (the median across communities of the lower bound of the 95% CI of R_0 is 0.60). In the least favorable months (e.g. December and January), 100% of community median R_0 estimates were less than 1, while 67% of the CI for these communities spanned one (the median of the upper bound of the 95% CI is 2.0).

We decompose the importance of spatial and temporal variation in both temperature and bird community composition by comparing the mean absolute deviation (MAD) in predictions for R_0 between a full model and models with either the bird community or temperature aggregated across space or time (Table 3). Temperature variation across both space and time is more predictive of R_0 than bird community

composition, though ignoring variation in the bird community across space does lead to R_0 estimates that differ from the full model by 0.17 on average (Table 3). Allowing for temporal variation in temperature and bird community composition, the majority of the variation in R_0 within single months is due to spatial variation in temperature; variation in bird community composition is the next most important term (Figure 1).

Table 3: Capability of simplified models to estimate WNV R_0 in Texas. Mean absolute error compares R_0 estimates from a simplified model to the R_0 estimates from a full model for all 2,569 of the bird communities in the reduced eBird dataset. Models are as follows: *Temporally averaged bird community*: each counties' bird community is replaced with the average bird community in that county across all months; *Spatially averaged bird community*: each counties' bird community in each month is replaced with the average bird community across all of Texas in that month; *Spatially averaged temperature*: each counties' temperature in each month is replaced with the average temperature across all of Texas in that month; *Temporally averaged temperature*: each counties' temperature is replaced with the average temperature in that county across all months; *Mean model*: each counties' bird community and temperature is replaced with the average bird community and temperature across all counties and all months.

Model	Mean Absolute Error in R_0 Estimates
Temporally averaged bird community	0.07
Spatially averaged bird community	0.15
Spatially averaged temperature	0.36
Temporally averaged temperature	0.40
Mean model	0.63

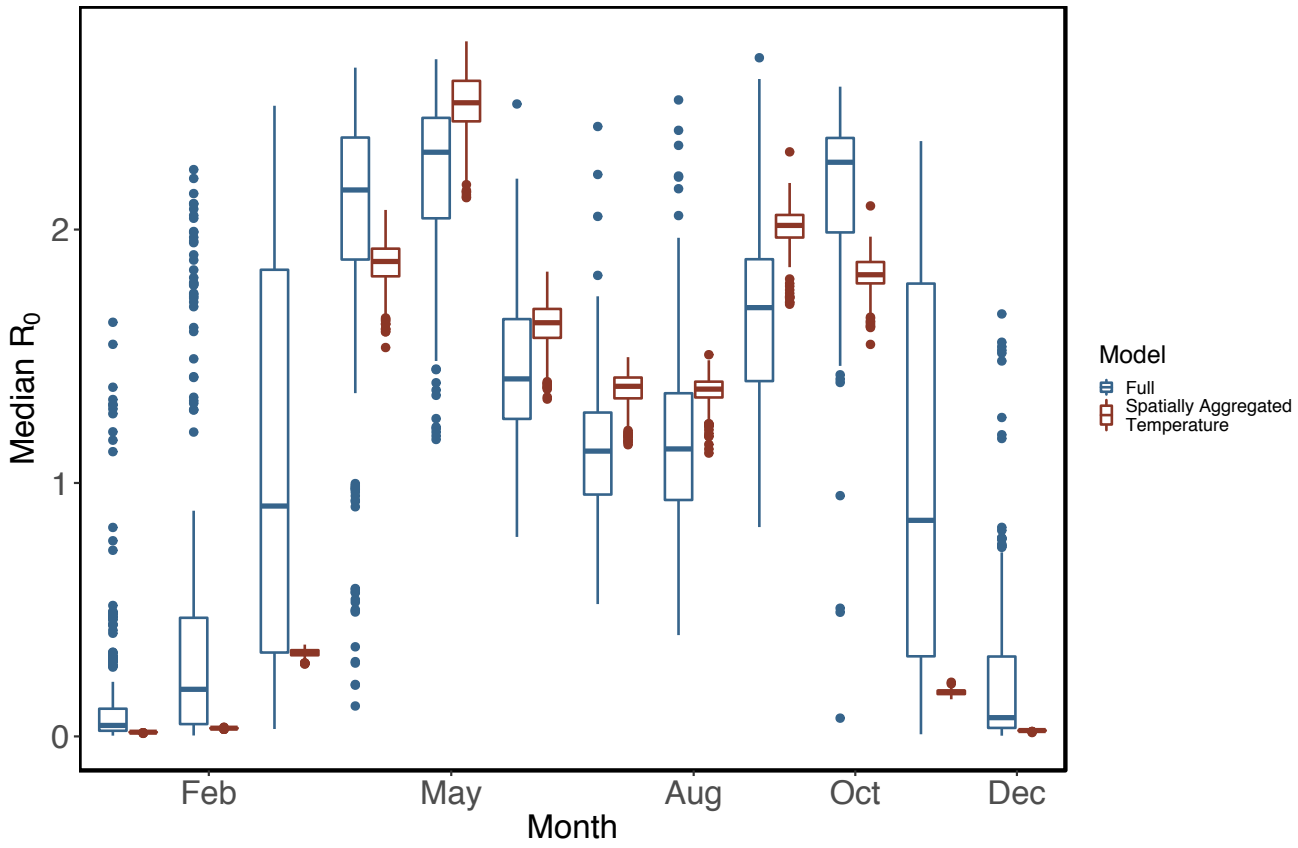

Figure 1: WNV R_0 estimates between months and among Texas counties. Blue boxplots show R_0 estimates across Texas counties within months for a "Full" model, which used the eBird community and NOAA temperature data for each community. Red boxplots show R_0 estimates from a model where each community retained their specific eBird community, but whose temperature was replaced with the average temperature across all of Texas for that month (also see Table 3: *Spatially averaged temperature*). Variation in R_0 within months attributable to variation in the bird communities (red boxplots) is considerably smaller than the variation explained by spatial variation in temperature. Increases or decreases in medians between the models within months is due to the effects of averaging temperature prior to predicting R_0 using the non-linear functions for mosquito-to-bird transmission and mosquito survival across temperature, a manifestation of Jensen's inequality. For example, in November the mean temperature across Texas is 13.6^0C, while the SD among counties is 3.30^0C. We estimate average mosquito-to-bird transmission per bite over the first 30 days of mosquito infection to be 2.5% at 13.6^0C, 8.5% at 16.9^0C (+ 1SD), and is 25% at 20.2^0C (+ 2 SD).

The spatio-temporal GAM model explained 99% of the variation in estimated WNV R_0; results for the spatio-temporal model are presented visually in Figure 2. Most of the variation in WNV R_0 is explained by temperature in the fitted GAM (Figure 2A). Human population density (people/sq.mile) was associated with decreasing R_0, but with small effect and large uncertainty (Figure 2B). Ignoring the

effects of fluctuations in mosquito populations, WNV R_0 was estimated to vary little across years (Figure 2C). Variation due to bird communities among ecoregions after controlling for temperature explained a small fraction of the variation in R_0 among regions (Figure 2D). Our fitted GAM predicts that bird communities in the "High Plains" and "Oak Woods & Prairies" ecoregions are the least favorable for WNV transmission, while bird communities in the "Llano Uplift" are the most favorable (Figure 2D).

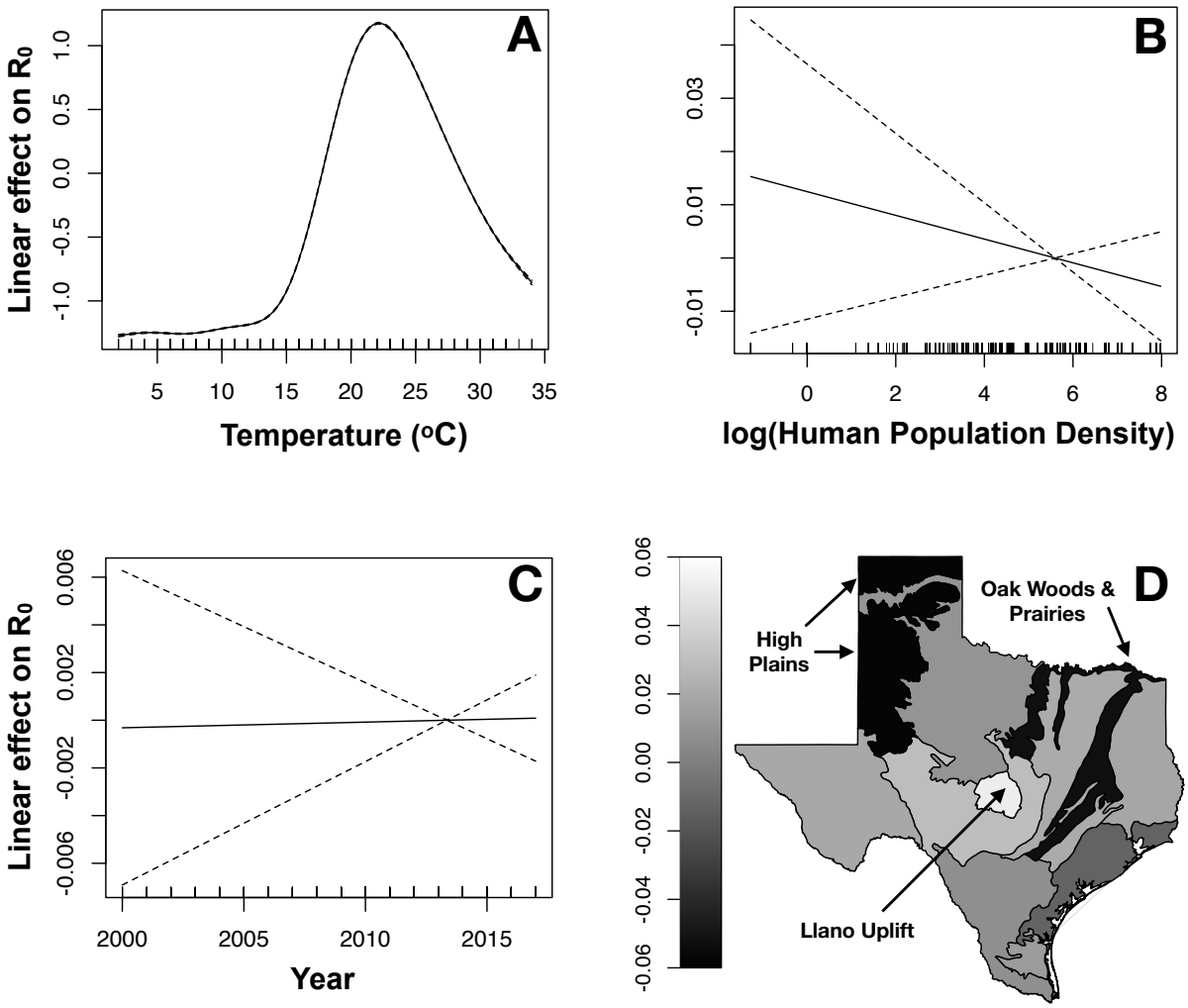

Figure 2: Spatio-temporal GAM model parameter estimates. Y-axes in panels A-C, and the gradient in panel D show the additive effect of centered covariates on R_0. The gradient in panel D shows variation in R_0 among ecoregions explained by variation in bird communities. Dashed lines show 95% CI.

To evaluate the fit of our focal model relative to the model with latitude and longitude coordinate pairs (average estimated concurvities of 0.19 and 0.54 respectively), we used blocked leave-one-out validation [75] at the level of counties. Using this method, RMSE for all estimates for our focal model and the model with latitude and longitude coordinate pairs were 0.07 and 0.08, respectively. This suggests that not including the thin plate spline across coordinate pairs results in little loss in terms of predictive power while also minimizing the possibility of over-fitting by reducing concurvity.

Species-specific contributions to R_0

Across the most sampled bird communities, no single bird species' removal accounted for a median fold decrease in R_0 larger than 0.92 or increase larger than 1.04. Mourning doves (*Zenaida macroura*, recorded in all bird communities) accounted for the largest dilution effect (median: 1.04 fold increase in R_0, CH: 1.01-1.11), while Northern cardinals (*Cardinalis cardinalis*, recorded in 98.7% of the bird communities) accounted for the largest amplification effect (median: 0.92 fold decrease in R_0, CH: 0.83-0.99).

Only two species were estimated to have a median effect greater than a 1.01 fold increase in R_0 (in order of median effect: Mourning dove; White-winged dove: *Zenaida asiatica*), and only five species had a median effect greater than a 0.99 fold decrease in R_0 (in order of median effect: Northern Cardinal; Transvolcanic Jay: *Aphelocoma ultramarina*; Blue Jay: *Cyanocitta cristata*; House Finch; *Carpodacus mexicanus*; Green jay: *Cyanocorax yncas*) (Figure 3). Of the 15 most widespread species (species that appear in at least 95% of communities), the median estimate for five species was of an amplification effect. Eight of the fifteen species act as either diluters or amplifiers in at least 95% of communities, albeit with varying magnitudes. Of the 15 most abundant species (most individuals recorded; recorded in 34-99% of communities), the median effect on R_0 for five species was below a ratio of one. Nine of these fifteen species had an effect in 95% of communities on one side of a ratio of one.

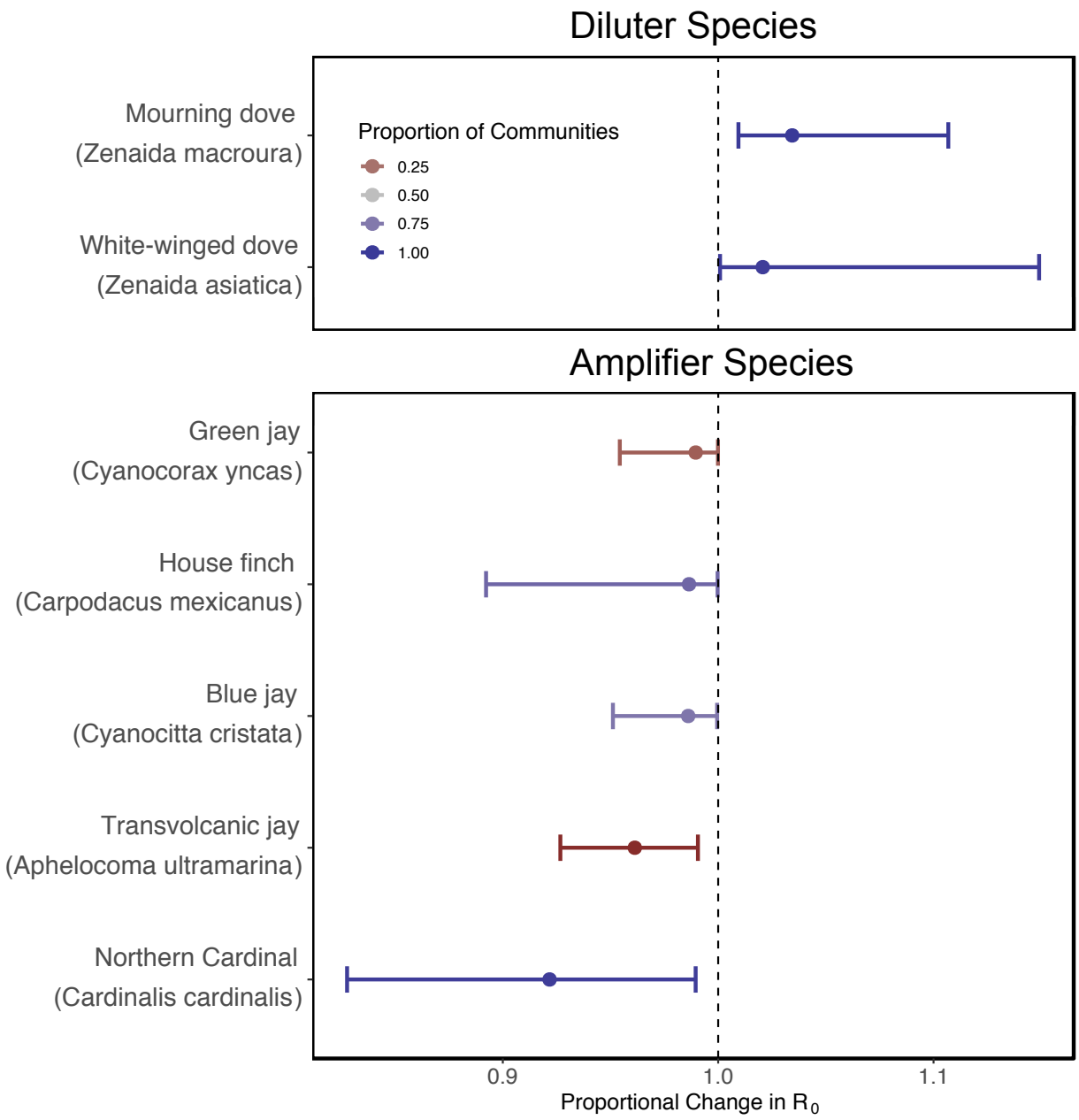

Figure 3: Keystone species. Bird species whose median estimates for their impact on R_0 when they are removed from each community they occupy are greater than a 1.01 (dilution effect—the two species above the plot break in this figure), or less than a 0.99 (amplification effect—the four species below the plot break in this figure) fold change in R_0. Intervals show median effects in 95% of the communities that each bird occupies.

Using a linear model with log of median bird relative abundance as a predictor for species physiological competence, physiological competence was predicted to increase with increasing relative abundance (estimate = 0.05, se = 0.02, t = 3.00, $p < 0.05$). The estimate here refers to the increase in the number of infected mosquitoes with each unit increase of a bird's relative abundance on the log scale (assuming a single mosquito bite per day over the course of a bird's infectious period, which is generally assumed in measures of host competence [20]). We also find evidence for a negative relationship between bird species richness and community R_0 using a linear model with log of species richness and temperature as predictors for median R_0 and variation in R_0 as weights (estimate: -0.15, se = 0.01, t = -10.01, $p < 0.05$).

Propagation of uncertainty

With no uncertainty propagated median WNV R_0 estimates were on average 1.03 times higher throughout the year and 1.06 times higher in the four most favorable months for transmission than in a model with all uncertainty propagated. Ignoring uncertainty had a much larger effect on variation among communities: CV in WNV R_0 estimates were on average 1.32 times higher throughout the year and 1.56 times higher in the four most favorable months for transmission. This increase in magnitude and variation of the R_0 estimates when no uncertainty was propagated is caused by the nonlinear averaging of variation in mosquito-to-bird transmission, mosquito survival, bird-to-mosquito transmission, and bird survival. For example, translating the full distribution for bird's titer profile (uncertainty) instead of a point estimate (no uncertainty), non-linearly, into the probability that a bird transmits infection to a susceptible mosquito given a bite homogenizes birds' responses, decreasing variation among bird communities. This is a manifestation of Jensen's inequality [96].

Species-specific contributions to R_0 also depend on whether uncertainty is propagated. While the most influential bird species (Northern cardinals and Mourning doves) were robust to choices about uncertainty propagation, the ranks and identities of some of the top ten most important amplifier and diluter species changed.

Complete eBird dataset

We present results using the complete eBird data set in the online supplement, but suggest caution when drawing conclusions from these results because many of the estimates were obtained from poorly sampled bird communities. Using the complete eBird data resulted in greater variation in estimates for all outcomes: variation in R_0 among communities increased (Figure S4, Figure S5), variation explained in the spatio-temporal GAM model decreased, and the estimated impacts of individual bird species on R_0 were more extreme.

Discussion

Data limitations

Despite our ability to estimate R_0 in individual bird communities, better data, such as mosquito populations on the same scale as the bird communities, are needed to make reliable quantitative estimates of WNV R_0 across space and time. Given the size of our estimated effect of temperature on WNV transmission and the fact that different mosquito species incubate WNV and feed at different rates across temperatures [6,14], variation in mosquito density and species composition among ecoregions and across seasons are likely the most important missing data needed to predict WNV R_0 reliably. Our WNV R_0 predictions for Texas counties relied on estimates of the mosquito-to-bird ratio and mosquito biting rate based on sparse data from a different geographic region (New Haven, CT) and are assumed to be spatially and temporally homogeneous. While these simplifications let us move forward to explore the variation in R_0 driven by spatial and temporal variation in bird communities and temperature, we emphasize that the magnitude of R_0 values presented here should be taken with a grain of salt. While changes in the mosquito-to-bird ratio and mosquito biting rate for a given species would shift R_0 values by a constant multiple, unknown interactions between mosquito-to-bird ratios (and biting rate), temperature, and mosquito species make it hard to make definitive statements about the robustness of our results.

Higher-resolution mosquito data may be available for some locations outside of Texas. For

example, the NEON (National Ecological Observation Network [97]) database provides mosquito sampling data for many locations across the USA; however, data for Texas was only available for two locations, and estimating mosquito-to-bird ratio from mosquito trapping data would require further simplifying assumptions [98]. In Europe, mosquito data is abundant in at least Italy and Germany (West Nile Disease National Surveillance Plan: [99]). With these data and some additional data on the responses of European birds [38], our model could be extended to predict WNV transmission in Europe where human and equine cases of WNV are increasing [5,8–12]. Alternatively, in the absence of mosquito data for a particular region, information on the ecological drivers of mosquito populations [100,101] might be combined with data on habitat composition to estimate spatio-temporal, multi-species mosquito distributions.

Limited spatial and temporal resolution in the eBird data was another constraint on our analysis. While eBird use is rapidly expanding, it may be worthwhile in the short term to incorporate bird abundance data from additional data sources such as the Breeding Bird Survey (BBS) or Christmas Bird Count (CBC), despite their more restricted seasonal coverage, or to use joint species distribution models to infer local bird community structure from habitat variables.

Community R_0

Though we neglect spatial variation in mosquito-to-bird ratio, mosquito biting rate, and mosquito species, the single values that we use for these parameters result in estimates of WNV R_0 that are similar to those of previous modeling efforts from other regions. For example, most of Hartley et al.'s [102] R_0 estimates for California were between 1.0 and 1.75, while R_0 estimates for New York City were 2.0 and 2.8 assuming mosquito-to-bird ratios of 2 and 4 respectively [103]. Finally, [24] estimate that a mosquito-to-bird ratio of greater than 4.6 would have been required for the epidemic that occurred in New York, USA in 2000 (implying $R_0 = 1$ for $M/B = 4.6$). Using our method, an $R_0 = 1$ is obtained for $M/B = 2.9$ in the median county in July, though a ratio for M/B of only 2.0 is needed in the median county in May.

Bird species-specific contributions to R_0

At the level of individual bird species, some of our conclusions support the results of previous work, while others contradict previous findings. For example, [24] assume a *per capita* mosquito biting rate on American crows of 0.09 per day (CI: 0.03-0.16), which is similar to the biting rate we estimate for crows in our bird communities; our baseline biting rate of 0.14 per day and a median mosquito biting preference on American crows that is ≈ 1.8 times lower than on the average bird gives a biting rate of 0.08. Like previous syntheses [e.g. 1], our model shows that species in the family Corvidae (e.g. Jays, Grackles, and Crows) are highly competent species for WNV. However, our model suggests that no single species ever accounts for more than approximately 30% of WNV R_0, which contrasts with the results of [20] and [86] who found that more than 50% of infectious mosquitoes were infected by American robins, and [87] who found that 96% of mosquitoes were infected by either American robins or House sparrows (*Passer domesticus*). While these studies were conducted over a much smaller and almost entirely urban area with low bird diversity (90% of most of the bird communities sampled were composed of < 6 species), the high proportion of mosquitoes infected by American robins which were present at a relative abundance between approximately 5-20% suggests that either: 1) We are missing an aspect of the interaction between WNV, mosquitoes, and American robins; or 2) Our biological model is adequate and American robins are simply more important in other regions of the country.

To explore these two possibilities, we predicted the proportion of all newly infected mosquitoes attributable to each bird species in the community from [20] and [86] that had the highest proportion of American robins (7.5% of the bird community: Foggy Bottom, District of Columbia, USA). For this community our median estimate for the proportion of all mosquitoes infected by American robins was 18%; however, uncertainty in mosquito biting preferences, bird species physiological competence, and bird species detectability resulted in 95% confidence intervals spanning 3% to 79%. Three conclusions arise from the facts that the composition of mosquito blood meals observed in Foggy Bottom, DC by Kilpatrick et al. [20] is contained within our CI, and that American robins do not show up as one of the most

important hosts in our communities. First, our model estimates are consistent with findings from a very different region of the country (albeit with very large uncertainty arising from propagating the uncertainty across the entire life cycle of WNV). Second, regional differences in bird communities probably cause the differences in the estimated importance of American robins between the current studies and previous studies [86,87]. Finally, regional and seasonal differences in mosquito feeding preferences [20,104] probably also play an important role, reinforcing the need for more data on mosquitoes.

While estimated species-specific seroprevalence rates vary across studies, seroprevalence rates of Northern cardinals are typically among the highest of all birds measured (Tammany Parish, LA [105]; Harris County, TX [106]; Illinois state-wide [107]; Chicago, IL [57]; Atlanta, GA [108]). High seroprevalence in Northern cardinals suggests they may play a critical role in WNV amplification [105,106], as we find here (Figure 3). However, amplification within the bird community may or may not lead to higher human infection risk, and researchers disagree about the effect of Northern cardinals on human infection risk [106,108].

Previous studies have also found high seroprevalence for one of our most effective diluter species, the doves (family *Columbidae*). Rock pigeons (*Columba livia*) had one of the highest antibody prevalence rates in Georgia, USA between 2000 and 2004 [109], while Mourning doves had the highest antibody prevalence rate in Chicago in 2005 and 2006 [57]. Our model shows that these species, which are strongly associated with urban landscapes and which we estimate to be among the least competent species for WNV, could be an important sink protecting human populations from disease. While we found only a small effect of decreasing R_0 with increasing human population density (Figure 2B), these species could potentially drive Nolan et al.'s [16] result that WNV *per capita* risk to humans decreased with increasing human population density.

The dilution effect hypothesis

Studies testing the dilution effect hypothesis for WNV have obtained the full range of possible

results: human cases declined with increasing bird diversity across 742 counties in 38 US states [28]; the proportion of mosquitoes infected with WNV declined with increasing diversity of non-passerine birds in Louisiana, USA [33]; Loss et al. [57] failed to detect a clear effect of species richness on WNV transmission in Chicago, Illinois, USA; Levine et al. [58] detected an amplification effect—overall seroprevalence increased with species diversity in Atlanta, Georgia, USA; in southern France [110] suggest that high bird diversity is a likely explanation of low numbers of horse infections, while [111] suggest that low number of human cases is due to the abundance of horses.

Based on the estimated competence of all 645 species found in the reduced eBird data set and their median abundance in 2,569 bird communities, the host competence (the total number of mosquitoes that would be infected by an infected bird if it was bitten once each day of its infectious period [20,62]) is positively correlated with relative abundance. Additionally, communities with higher species richness had a lower estimated R_0, which is as expected if the most abundant birds are the most competent. These results support both a necessary condition (correlation between abundance and competence) and a primary expectation (correlation between richness and R_0) of the dilution effect. However, we do not know what bird traits (or unobserved underlying ecological covariates) drive these patterns. To put it another way, we expect that R_0 is proximally determined by the composition of the community, which is a function of many environmental covariates, rather than by species richness *per se* [55].

Understanding spillover

Though we do not model human infections directly, we do find variation in WNV R_0 among Texas bird communities that could shed light on patterns of human infection. According to [16] and [19], *per capita* infection risk is highest in northern Texas counties, with maximum risk in Castro, King, and Crosby counties. Two of these counties reside either entirely (Castro) or partially (Crosby) within the "High Plains" ecoregion of Texas, which we estimated to have the smallest R_0 of all 11 ecoregions on average throughout the year. Unfortunately, we cannot validate these estimates in the absence of widespread spatial sampling of infected mosquitoes or birds. However, this apparent failure of our

predictions (we expect human infection risk to be positively correlated with the R_0 of WNV in local bird communities, but have no *a priori* expectation for the strength of this correlation) might be explained by variations in the degree of WNV spillover from birds to humans.

Spillover into human populations varies across microhabitats, seasons, and mosquito communities [16,28,33,86,112]. In Atlanta, Georgia, for example, human infection rates are low despite similar mosquito infection rates and bird seroprevalence to other cities [108]. Levine et al. [108] attribute fewer human infections in Atlanta to high rates of infection in Northern cardinals and Blue jays, which they describe as "supersuppressor" species because they attract mosquito bites but fail to amplify transmission due to low competence. Our results (Figure 3) and others [105] suggest in contrast that Northern cardinals and Blue jays are important *amplifier* species (taking into consideration all experimental infections Northern cardinals and Blue jays are better defined as having moderate competence; their presence increases R_0 within the bird community). Yet, it is still possible that the presence of these species could decrease the number of human cases by drawing mosquito bites, and hence infections, away from humans. Kilpatrick et al. [6] document a related phenomenon, providing correlational evidence to suggest that higher numbers of human cases of WNV could be attributable to an increased number of human bites by *Culex* mosquitoes following seasonal emigration of American robins. Similarly, mosquito feeding on mammals increased in northern California following the fledging of ardeids (heron species) [104].

Our results, combined with the variation in previous results [16,28,33,86,105,108,112], bring into sharp focus how little we really know about the details of human infection risk across space and time in this system. To predict human infection cases for WNV, and for zoonotic diseases with heterogeneous host populations more generally, we envision a fine-scale spatial model that would use a Who Acquires Infection From Whom (WAIFW) matrix approach [113] and explicitly include humans as an additional species in the overall community. This framework would calculate the force of infection between each species pair, and could be used to determine the expected number of human cases during an epidemic. Interspecific contact rates could be parameterized using mosquito biting preferences, natural habitat type,

and land use (urban vs rural) on a very fine spatial scale. While eBird data is currently lacking to estimate the interface of bird communities with humans at a fine spatial scale for most locations, some counties in Texas (and other states) have thousands of complete lists submitted in spring and late summer months that could serve as model locations for analysis.

Propagation of uncertainty

Appropriate uncertainty and point estimates for R_0 are only obtained when uncertainty in every sub-model is considered in calculations of R_0. With the currently available data, we find large uncertainty in most of the models we use in our analysis, which obscures our ability to estimate R_0 with precision in any individual community. While it is a poor practice in general to use median estimates from models instead of all uncertainty, we examined the qualitative and quantitative effects of ignoring uncertainty in order to emphasize the importance of propagating uncertainty (and of reporting the procedures used). Ignoring uncertainty in our analyses would have led us to different quantitative and qualitative conclusions. Ignoring variation in sub-models increased variation in R_0 estimates among communities, for two related reasons: first, large uncertainty in birds' physiological competence and mosquito biting preferences makes it more difficult to differentiate among birds, obscuring differences among communities. Second, birds are further homogenized due to the effects of Jensen's inequality, which occurs when we transform the distribution of titer estimates into the probability that a bird transmits infection to a susceptible mosquito given a bite, which is bounded between zero and one. Jensen's inequality also affects the estimated effects of temperature because of the nonlinear relationship between temperature and mosquito-to-bird transmission and mosquito survival, but has a larger effect when averaging temperature across either space or time (see Figure 1).

In the absence of uncertainty, most bird species have an average titer that results in a bird-to-mosquito transmission probability beneath the inflection point of the logistic relationship between titer and transmission probability. Uncertainty in bird titer results in a non-negligible proportion of the posterior distribution for bird titer that is near or above 10^8, which corresponds to a bird-to-mosquito transmission

probability near one. This decreases variation in physiological competence among birds, which further narrows the variation in estimates among communities. This aspect of Jensen's inequality will increase R_0 estimates because of an increase in bird-to-mosquito transmission; however, increased titer will lead to lower bird survival, counteracting most, but not all, of this increase in R_0 (60% of median estimates for each community were larger when uncertainty was not propagated). Because we were unable to estimate variation among species in mortality probability as a function of titer (that is, species variation in sensitivity to titer), estimated variation among birds is likely to be lower than true variation, further homogenizing birds and estimates among communities.

WNV transmission in Europe

With additional data on the responses of European birds to WNV (e.g. [114–116]) and mosquito biting preferences and code modification, our model could be used to predict WNV transmission in many countries in Europe, with best results in those countries with abundant mosquito surveillance data (e.g. Germany, Italy [99]). Modeling studies on WNV spread in Europe have considered heterogeneities in mosquito transmission due to species [117], and temperature [29], as well as the effects of land cover and type on WNV transmission [110] and human infection risk [12]. However, like their North American counterparts, none of these studies consider full bird communities; our model can provide a method for incorporating heterogeneities in the bird community into spatio-temporal estimates of WNV transmission potential in Europe.

Conclusion

Despite numerous data limitations at the scale we chose for our analyses, WNV remains a promising system for continued study on the mechanisms of vector borne disease spillover on finer spatial scales. Using handpicked locations with sufficient bird community data, mosquito sampling, and temperature variation, our modeling framework can be used as is to predict WNV R_0 incorporating all

known heterogeneities in transmission. With slight modifications, our model could be used to mechanistically estimate human infection probability as a function of bird community composition and other ecological predictors. We emphasize that a critical aspect of multi-faceted ecological analyses, such as modeling human infection risk to WNV, is transparency in model assumptions, choices, and shortcomings; we hope that others will use our structure as a template for future analyses in order to increase model transparency.

Declarations

Acknowledgements

We thank the Dushoff lab and Jo Werba for helpful comments on the first draft of the manuscript.

Funding

This work was funded by NSERC Discovery Grant 386590-2010.

Availability of data and materials

All data and code used in this study are included as additional files.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests

Authors' contributions

MPK and BMB conceived the study; MPK collected the data; MPK and BMB performed statistical analyses; MPK and BMB wrote the manuscript. All authors read and approved the final manuscript.

References

1. Kilpatrick AM, LaDeau SL, Marra PP. Ecology of West Nile virus transmission and its impact on birds in the western hemisphere. Auk. 2007;124:1121–36.

2. Ciota AT, Ehrbar DJ, Matacchiero AC, Van Slyke GA, Kramer LD. The evolution of virulence of West Nile virus in a mosquito vector: implications for arbovirus adaptation and evolution. BMC Evol Biol. 2013;13:71.

3. Hernández-Triana LM, Jeffries CL, Mansfield KL, Carnell G, Fooks AR, Johnson N. Emergence of West Nile virus lineage 2 in Europe: a review on the introduction and spread of a mosquito-borne disease. Front public Heal. 2014;2:271.

4. Chancey C, Grinev A, Volkova E, Rios M. The global ecology and epidemiology of West Nile virus. Biomed Res Int. 2015;2015.

5. Rizzoli A, Jiménez-Clavero MA, Barzon L, Cordioli P, Figuerola J, Koraka P, Martina B, Moreno A, Nowotny N, Pardigon N, Sanders S, Ulbert S, Tenorio, A. The challenge of West Nile virus in Europe: knowledge gaps and research priorities. Eurosurveillance. 2015;20.

6. Kain MP, Bolker BM. Can existing data on West Nile virus infection in birds and mosquitos explain strain replacement? Ecosphere. 2017;8.

7. Venter M, Swanepoel R. West Nile virus lineage 2 as a cause of zoonotic neurological disease in humans and horses in southern Africa. Vector-borne zoonotic Dis. 2010;10:659–64.

8. Reiter P. West Nile virus in Europe: understanding the present to gauge the future. Eurosurveillance. 2010;15:19508.

9. Papa A, Xanthopoulou K, Gewehr S, Mourelatos S. Detection of West Nile virus lineage 2 in mosquitoes during a human outbreak in Greece. Clin Microbiol Infect. 2011;17:1176–80.

10. Roiz D, Vazquez A, Rosà R, Muñoz J, Arnoldi D, Rosso F, Figuerola J, Tenorio A, Rizzoli A. Blood meal analysis, flavivirus screening, and influence of meteorological variables on the dynamics of potential mosquito vectors of West Nile virus in northern Italy. J Vector Ecol. 2012;37:20–8.

11. Fros JJ, Geertsema C, Vogels CB, Roosjen PP, Failloux A-B, Vlak JM, Koenraadt, CJ, Takken W, Pijlman GP. West Nile virus: high transmission rate in north-western European mosquitoes indicates its epidemic potential and warrants increased surveillance. PLoS Negl Trop Dis. 2015;9:e0003956.

12. Marini G, Rosà R, Pugliese A, Rizzoli A, Rizzo C, Russo F, Montarsi F, Capelli G. West Nile virus transmission and human infection risk in Veneto (Italy): a modelling analysis. Sci Rep. Nature Publishing Group; 2018;8:14005.

13. Kramer LD, Styer LM, Ebel GD. A global perspective on the epidemiology of West Nile virus. Annu Rev Entomol. 2008;53:61–81.

14. LaDeau SL, Kilpatrick AM, Marra PP. West Nile virus emergence and large-scale declines of North American bird populations. Nature. 2007;447:710–3.

15. Brault AC. Changing patterns of West Nile virus transmission: altered vector competence and host susceptibility. Vet Res. 2009;40:1–19.

16. Nolan MS, Schuermann J, Murray KO. West Nile virus infection among humans, Texas, USA, 2002--2011. Emerg Infect Dis. 2013;19:137.

17. Ostroff SM. West Nile virus: Too Important to Forget. JAMA. 2013;310:267–8.

18. Petersen LR, Brault AC, Nasci RS. West Nile virus: review of the literature. JAMA. 2013;310:308–15.

19. Center for Disease Control. Final Cumulative Maps and Data, West Nile virus. https://www.cdc.gov/westnile/statsmaps/cumMapsData.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fwestnile%2Fstatsmaps%2Fpreliminarymapsdata2017%2Fdisease-cases-state.html (2017). Accessed 3 Oct 2018.

20. Kilpatrick AM, Daszak P, Jones MJ, Marra PP, Kramer LD. Host heterogeneity dominates West Nile virus transmission. Proc R Soc London B Biol Sci. 2006;273:2327–33.

21. Goddard LB, Roth AE, Reisen WK, Scott TW. Vector competence of California mosquitoes for West Nile virus. Emerg Infect Dis. 2002;8:1385.

22. Kilpatrick AM, Fonseca DM, Ebel GD, Reddy MR, Kramer LD. Spatial and temporal variation in vector competence of *Culex pipiens* and *Cx. restuans* mosquitoes for West Nile virus. Am J Trop Med Hyg. 2010;83:607–13.

23. Takken W, Verhulst NO. Host preferences of blood-feeding mosquitoes. Annu Rev Entomol. 2013;58:433–53.

24. Wonham MJ, de-Camino-Beck T, Lewis MA. An epidemiological model for West Nile virus: invasion analysis and control applications. Proc R Soc London B Biol Sci. 2004;271:501–7.

25. Bowman C, Gumel AB, van den Driessche P, Wu J, Zhu H. A mathematical model for assessing control strategies against West Nile virus. Bull Math Biol. 2005;67:1107–33.

26. Bradley CA, Gibbs SEJ, Altizer S. Urban land use predicts West Nile virus exposure in songbirds. Ecol Appl. 2008;18:1083–92.

27. Brown HE, Childs JE, Diuk-Wasser MA, Fish D. Ecologic factors associated with West Nile virus transmission, northeastern United States. Emerg Infect Dis. 2008;14:1539.

28. Allan BF, Langerhans RB, Ryberg WA, Landesman WJ, Griffin NW, Katz RS, Oberle BJ, Schutzenhofer MR, Smyth KN, Maurice AD, Clark L, Crooks, KR, Hernandez DE, McLean RG, Ostfeld RS, Chase JM. Ecological correlates of risk and incidence of West Nile virus in the United States. Oecologia. 2009;158:699–708.

29. Vogels CBF, Hartemink N, Koenraadt CJM. Modelling West Nile virus transmission risk in Europe: effect of temperature and mosquito biotypes on the basic reproduction number. Sci Rep. 2017;7:5022.

30. Tachiiri K, Klinkenberg B, Mak S, Kazmi J. Predicting outbreaks: a spatial risk assessment of West Nile virus in British Columbia. Int J Health Geogr. 2006;5:21.

31. Douglas KO, Kilpatrick AM, Levett PN, Lavoie MC. A quantitative risk assessment of West Nile virus introduction into Barbados. West Indian Med J. 2007;56:394–7.

32. Chevalier V, Tran A, Durand B. Predictive modeling of West Nile virus transmission risk in the Mediterranean Basin: how far from landing? Int J Environ Res Public Health. 2013;11:67–90.

33. Ezenwa VO, Godsey MS, King RJ, Guptill SC. Avian diversity and West Nile virus: testing associations between biodiversity and infectious disease risk. Proc R Soc London B Biol Sci. 2006;273:109–17.

34. Swaddle JP, Calos SE. Increased avian diversity is associated with lower incidence of human West Nile infection: observation of the dilution effect. PLoS One. 2008;3:e2488.

35. Kilpatrick AM, Pape WJ. Predicting human West Nile virus infections with mosquito surveillance data. Am J Epidemiol. 2013;178:829–35.

36. Levine RS, Mead DG, Kitron UD. Limited spillover to humans from West Nile virus viremic birds in Atlanta, Georgia. Vector-Borne Zoonotic Dis. 2013;13:812–7.

37. Peterson AT, Vieglais DA, Andreasen JK. Migratory birds modeled as critical transport agents for West Nile virus in North America. Vector-Borne Zoonotic Dis. 2003;3:27–37.

38. Koraka P, Barzon L, Martina BEE. West Nile virus infections in (European) birds. J Neuroinfect Dis. 2016;7:226.

39. Wonham MJ, Lewis MA, Rencławowicz J, van den Driessche P. Transmission assumptions generate conflicting predictions in host--vector disease models: a case study in West Nile virus. Ecol Lett. 2006;9:706–25.

40. Lord CC, Day JF. Simulation studies of St. Louis encephalitis and West Nile viruses: the impact of bird mortality. Vector Borne Zoonotic Dis. 2001;1:317–29.

41. Cruz-Pacheco G, Esteva L, Montaño-Hirose JA, Vargas C. Modelling the dynamics of West Nile virus. Bull Math Biol. 2005;67:1157–72.

42. Rastetter EB, Aber JD, Peters DPC, Ojima DS, Burke IC. Using mechanistic models to scale ecological processes across space and time. AIBS Bull. 2003;53:68–76.

43. Gustafson EJ. When relationships estimated in the past cannot be used to predict the future: using mechanistic models to predict landscape ecological dynamics in a changing world. Landsc Ecol. 2013;28:1429–37.

44. Epstein JM. Why model? J Artif Soc Soc Simul. 2008;11:12.

45. Hamer GL, Kitron UD, Goldberg TL, Brawn JD, Loss SR, Ruiz MO, Hayes DB, Walker ED. Host selection by *Culex pipiens* mosquitoes and West Nile virus amplification. Am J Trop Med Hyg. 2009;80:268–78.

46. Simpson JE, Hurtado PJ, Medlock J, Molaei G, Andreadis TG, Galvani AP, Diuk-Wasser MA. Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile virus as a model system. Proc R Soc London B Biol Sci. 2012;279:925–33.

47. Ogle K, Pathikonda S, Sartor K, Lichstein JW, Osnas JLD, Pacala SW. A model-based meta-analysis for estimating species-specific wood density and identifying potential sources of variation. J Ecol. 2014;102:194–208.

48. Dunning JBJ. CRC Handbook of Avian Body Masses, Second Edition. Boca Raton, FL: CRC press; 2007.

49. A global phylogeny of birds. BirdTree.org (2014) Accessed 14 Nov 2017.

50. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. A comprehensive phylogeny of birds *Aves* using targeted next-generation DNA sequencing. Nature. 2015;526:569–73.

51. Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D, Kelling S. eBird: A citizen-based bird observation network in the biological sciences. Biol Conserv. 2009;142:2282–92.

52. Lindsey NP, Staples JE, Lehman JA, Fischer M. Surveillance for human West Nile virus disease—United States, 1999--2008. MMWR Surveill Summ. 2010;59:1–17.

53. Chung WM, Buseman CM, Joyner SN, Hughes SM, Fomby TB, Luby JP, Haley RW. The 2012 West Nile encephalitis epidemic in Dallas, Texas. JAMA. 2013;310:297–307.

54. Schmidt KA, Ostfeld RS. Biodiversity and the dilution effect in disease ecology. Ecology. 2001;82:609–19.

55. Salkeld DJ, Padgett KA, Jones JH. A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. Ecol Lett. 2013;16:679–86.

56. Civitello DJ, Cohen J, Fatima H, Halstead NT, Liriano J, McMahon TA, Ortega CN, Sauer EL, Sehgal T, Young S, Rohr JR. Biodiversity inhibits parasites: broad evidence for the dilution effect. Proc Natl Acad Sci. 2015;112:8667–71.

57. Loss SR, Hamer GL, Walker ED, Ruiz MO, Goldberg TL, Kitron UD Brawn JD. Avian host community structure and prevalence of West Nile virus in Chicago, Illinois. Oecologica. 2009;159:415–24.

58. Levine RS, Hedeen DL, Hedeen MW, Hamer GL, Mead DG, Kitron UD. Avian species diversity and transmission of West Nile virus in Atlanta, Georgia. Parasit Vectors. 2017;10:62.

59. Cornell Lab of Ornithology. eBird request data. https://eBird.org/data/request Accessed 15 Sep 2017.

60. NOAA National Centers for Environmental Information, Data Access. https://www.ncdc.noaa.gov/data-access (2018) Accessed March 21 2018.

61. Vinogradova EB. *Culex pipiens pipiens* mosquitoes: taxonomy, distribution, ecology, physiology, genetics, applied importance and control. Sofia, Bulgaria: Pensoft Publishers; 2000.

62. Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, Hettler D, Davis B, Bowen R, Bunning M. Experimental infection of North American birds with the New York 1999 strain of West Nile virus. Emerg Infect Dis. 2003;9:311–22.

63. Felsenstein J. Phylogenies and the comparative method. Am Nat. 1985;125:1–15.

64. Ives AR, Helmus MR. Generalized linear mixed models for phylogenetic analyses of community structure. Ecol Monogr. 2011;81:511–25.

65. Pagel M. Inferring the historical patterns of biological evolution. Nature. 1999;401:877.

66. Blomberg SP, Garland Jr T, Ives AR. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution. 2003;57:717–45.

67. Hansen TF, Bartoszek K. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Syst Biol. 2012;61:413–25.

68. Li M, Bolker BM. wzmli/phyloglmm: First release of phylogenetic comparative analysis in lme4-verse (Version v1.0.0). Zenodo. 2019. http://doi.org/10.5281/zenodo.2639887

69. Rubin DB. Multiple imputation for nonresponse in surveys. Hoboken, NJ: John Wiley & Sons; 2004.

70. Sukumaran J, Holder MT. DendroPy: a Python library for phylogenetic computing. Bioinformatics. 2010;26:1569–71.

71. Rubolini D, Liker A, Garamszegi LZ, Møller AP, Saino N. Using the BirdTree. org website to obtain robust phylogenies for avian comparative studies: A primer. Curr Zool. 2015;61:959–65.

72. Nakagawa S, Schielzeth H. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol Evol. 2013;4:133–42.

73. Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J R Soc Interface. 2017;14:20170213.

74. Bartoń K. MuMIn: Multi-Model Inference. R package version 1.42.1. 2018.

75. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography. 2017;40:913–29.

76. Bolker BM. Ecological models and data in R. Princeton, NJ: Princeton University Press; 2008.

77. Jenkins SP. Survival analysis. Unpubl manuscript, Inst Soc Econ Res Univ Essex, Colchester, UK. 2005;42:54–6. https://www.iser.essex.ac.uk/files/teaching/stephenj/ec968/pdfs/ec968lnotesv6.pdf Accessed 22 July 2018.

78. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Marcus B, Jiqiang G, Peter L, Allen R. Stan: A probabilistic programming language. J Stat Softw. 2017;76.

79. Stan Development Team. RStan: the R interface to Stan. R package version 2.16.2. 2017

80. Kosmala M, Wiggins A, Swanson A, Simmons B. Assessing data quality in citizen science. Front Ecol Environ. 2016;14:551–60.

81. IUCN 2017. ISSN 2307-8235. The IUCN Red List of Threatened Species. http://www.iucnredlist.org (2017) Accessed 2 Nov 2017.

82. Roskov Y, Abucay L, Orrell T, Nicolson D, Bailly N, Kirk PM, Bourgoin T, DeWalt RE, Decock W, De Wever A, Nieukerken E van, Zarucchi J, Penev L, eds. Species 2000 & ITIS Catalogue of Life, 2017 Annual Checklist. ISSN 2405-884X. www.catalogueoflife.org/annual-checklist/2017. (2017). Species 2000: Naturalis, Leiden, the Netherlands. Accessed 2 Nov 2017.

83. Yu J, Wong W-K, Kelling S. Clustering Species Accumulation Curves to Identify Skill Levels of Citizen Scientists Participating in the eBird Project. AAAI. 2014. p. 3017–23.

84. eBird user Morgan Kain. https://eBird.org/profile/Nzk4Mzg1/world (2016) Accessed May 22 2018.

85. Edman JD, Taylor DJ. *Culex nigripalpus*: seasonal shift in the bird-mammal feeding ratio in a mosquito vector of human encephalitis. Science. 1968;161:67–8.

86. Kilpatrick AM, Kramer LD, Jones MJ, Marra PP, Daszak P. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. PLoS Biol. 4:e82.

87. Hamer GL, Chaves LF, Anderson TK, Kitron UD, Brawn JD, Ruiz MO, Loss SR, Walker ED, Goldberg TL. Fine-scale variation in vector host use and force of infection drive localized patterns of West Nile virus transmission. PLoS One. 2011;6:e23767.

88. Tu S. The Dirichlet-multinomial and Dirichlet-categorical models for Bayesian inference. Comput Sci Div UC Berkeley. Tech. Rep., 2014.

89. Evans M V, Dallas TA, Han BA, Murdock CC, Drake JM. Data-driven identification of potential Zika virus vectors. Elife. 2017;6:e22053.

90. Manly BFJ. A model for certain types of selection experiments. Biometrics. 1974;281–94.

91. Amodio S, Aria M, D'Ambrosio A. On concurvity in nonlinear and nonparametric regression models. Statistica. 2014;74:85–98.

92. Texas Natural Resources Information System: TNRIS. Texas data search and download. https://tnris.org/data-download/#!/statewide Accessed Apr 14 2018.

93. Flanders AA, Kuvlesky Jr WP, Ruthven III DC, Zaiglin RE, Bingham RL, Fulbright TE, Hernández F, Brennan LA. Effects of invasive exotic grasses on south Texas rangeland breeding birds. Auk. 2006;123:171–82.

94. Brennan LA, Kuvlesky Jr. WP. North American grassland birds: an unfolding conservation crisis? J Wildl Manage. 2005;69:1–13.

95. Böhning-Gaese K, Lemoine N. Importance of climate change for the ranges, communities and conservation of birds. Adv Ecol Res. 2004;35:211–36.

96. Ruel JJ, Ayres MP. Jensen's inequality predicts effects of environmental variation. Trends Ecol Evol. 1999;14:361–6.

97. National Ecological Observatory Network data portal. http://data.neonscience.org (2018) Accessed August 11 2018.

98. Rund SSC, Braak K, Cator L, Copas K, Emrich SJ, Giraldo-Calderón GI, Johansson MA, Heydari N, Hobern D, Kelly SA, Lawson D, Lord C, MacCallum RM, Roche DG, Ryan SJ, Schigel D, Vandegrift K, Watts M, Zaspel JM, Pawar S. MIReAD, a minimum information standard for reporting arthropod abundance data. Sci Data. 2019;6:40.

99. Engler O, Savini G, Papa A, Figuerola J, Groschup M, Kampen H, Medlock J, Vaux A, Wilson AJ, Werner D, Jöst H, Goffredo M, Capelli G, Federici V, Tonolla M, Patocchi N, Flacio E, Portmann J, Rossi-Pedruzzi A, Mourelatos S, Ruiz S, Vázquez, Calzolari M, Bonilauri P, Dottori M, Schaffner F, Mathis A, Johnson N. European surveillance for West Nile virus in mosquito populations. Int J Environ Res Public Health. 2013;10:4869–95.

100. Bisanzio D, Giacobini M, Bertolotti L, Mosca A, Balbo L, Kitron U, Vazquez-Prokopec GM. Spatio-temporal patterns of distribution of West Nile virus vectors in eastern Piedmont Region, Italy. Parasit Vectors. 2011;4:230.

101. Sallam MF, Xue R-D, Pereira RM, Koehler PG. Ecological niche modeling of mosquito vectors of West Nile virus in St. John's County, Florida, USA. Parasit Vectors. 2016;9:371.

102. Hartley DM, Barker CM, Le Menach A, Niu T, Gaff HD, Reisen WK. Effects of temperature on emergence and seasonality of West Nile virus in California. Am J Trop Med Hyg. 2012;86:884–94.

103. Magori K, Bajwa WI, Bowden S, Drake JM. Decelerating spread of West Nile virus by percolation in a heterogeneous urban landscape. PLoS Comput Biol. 2011;7:e1002104.

104. Thiemann TC, Wheeler SS, Barker CM, Reisen WK. Mosquito host selection varies seasonally with host availability and mosquito density. PLoS Negl Trop Dis. 2011;5:e1452.

105. Komar N, Panella NA, Langevin SA, Brault AC, Amador M, Edwards E, Owen JC. Avian hosts for West Nile virus in St. Tammany Parish, Louisiana, 2002. Am J Trop Med Hyg. 2005;73:1031–7.

106. Dennett JA, Bala A, Wuithiranyagool T, Randle Y, Sargent CB, Guzman H, Siirin M, Hassan HK, Reyna-Nava M, Unnasch TR, Tesh RB, Parsons RE, Bueno Jr R. Associations between two mosquito populations and West Nile virus in Harris County, Texas, 2003-06. J Am Mosq Control Assoc. 2007;23:264.

107. Beveroth TA, Ward MP, Lampman RL, Ringia AM, Novak RJ. Changes in seroprevalence of West Nile virus across Illinois in free-ranging birds from 2001 through 2004. Am J Trop Med Hyg. 2006;74:174–9.

108. Levine RS, Mead DG, Hamer GL, Brosi BJ, Hedeen DL, Hedeen MW, McMillan JR, Bisanzio D, Kitron UD. Supersuppression: Reservoir competency and timing of mosquito host shifts combine to reduce spillover of West Nile virus. Am J Trop Med Hyg. 2016;95:1174–84.

109. Gibbs SE, Allison AB, Yabsley MJ, Mead DG, Wilcox BR, Stallknecht DE. West Nile virus antibodies in avian species of Georgia, USA: 2000-2004. Vector Borne Zoonotic Dis. 2006;6,57-72.

110. Pradier S, Leblond A, Durand B. Land cover, landscape structure, and West Nile virus circulation in southern France. Vector-Borne Zoonotic Dis. 2008;8:253–64.

111. Durand B, Chevalier V, Pouillot R, Labie J, Marendat I, Murgue B, Zeller H, Zientara S. West Nile virus outbreak in horses, southern France, 2000: results of a serosurvey. Emerg Infect Dis. 2002;8:777.

112. Liu A, Lee V, Galusha D, Slade MD, Diuk-Wasser M, Andreadis T, Scotch M, Rabinowitz PM. Risk factors for human infection with West Nile Virus in Connecticut: a multi-year analysis. Int J Health Geogr. 2009;8:67.

113. Dobson A, Foufopoulos J. Emerging infectious pathogens of wildlife. Philos Trans R Soc B Biol Sci 2001;356:1001–12.

114. Dridi M, Vangeluwe D, Lecollinet S, van den Berg T, Lambrecht B. Experimental infection of Carrion crows (*Corvus corone*) with two European West Nile virus (WNV) strains. Vet Microbiol. 2013;165:160–6.

115. Lim SM, Brault AC, van Amerongen G, Sewbalaksing VD, Osterhaus ADME, Martina BEE, Koraka P. Susceptibility of European jackdaws (*Corvus monedula*) to experimental infection with lineage 1 and 2 West Nile viruses. J Gen Virol. 2014;95:1320–9.

116. Gamino V, Höfle U. Pathology and tissue tropism of natural West Nile virus infection in birds: a review. Vet Res. 2013;44:39.

117. Vogels CBF, Göertz GP, Pijlman GP, Koenraadt CJM. Vector competence of northern and southern European *Culex pipiens pipiens* mosquitoes for West Nile virus across a gradient of temperatures. Med Vet Entomol. 2017;31:358–64.

118. Andreadis SS, Dimotsiou OC, Savopoulou-Soultani M. Variation in adult longevity of *Culex pipiens f. pipiens*, vector of the West Nile virus. Parisitology Res. 2014;113:4315–9.

Additional Files

Additional file 1 – Supplemental Methods and Results

Appendix C in this thesis.

Additional file 2 – ReadMe for Code

Appendix D in this thesis.
Accompanies code at: https://github.com/morgankain/WNV_Mechanistic_Model.

Chapter 4: The evolutionary response of virulence to host heterogeneity: a general model with application to myxomatosis in rabbits co-infected with intestinal helminths

In this chapter I present a model for the evolution of parasite virulence in a heterogeneous host population that is designed to be simple enough to be parameterized with empirical data from a specific host-pathogen system obtainable from two laboratory experiments and minimal data from the field. As a case study, I focus on the evolution of virulence in MYXV in populations of European rabbits in Australia and the UK that vary in their exposure to gastrointestinal helminth (worm) parasites. I show that spatial variation in secondary infection by helminth infection is unlikely to explain differences in the average virulence of strains circulating in the UK vs those circulating in Australia; however, my results show that heterogeneity across Australia in mean helminth burden may help to explain the presence of high variation in virulence among Australian strains of MYXV.

Author Contributions

MPK conceived the study with helpful feedback from IMC and BMB; IMC provided valuable knowledge about the biological system; MPK performed statistical analyses with helpful feedback from BMB; MPK wrote the manuscript and all authors revised the manuscript.

Acknowledgements

I thank my committee members Jonathan Dushoff and Ian Dworkin, the Bolker, Dushoff, and Earn labs, as well as Jo Werba for helping me refine and present this work.

Evolutionary Ecology Research, 2018, **19**: 257–278

The evolutionary response of virulence to host heterogeneity: a general model with application to myxomatosis in rabbits co-infected with intestinal helminths

Morgan P. Kain[1], Isabella M. Cattadori[2,3] and Benjamin M. Bolker[1,4]

[1]*Department of Biology, McMaster University, Hamilton, Ontario, Canada,* [2]*Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, Pennsylvania, USA,* [3]*Department of Biology, Pennsylvania State University, University Park, Pennsylvania, USA and* [4]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada*

ABSTRACT

Background: Changes in the mean and variance of traits in a host population modify selection pressures on pathogen virulence; increasing heterogeneity (variance) leads to the over- or under-exploitation of a subset of hosts and thus decreases the pathogen's ability to spread in the population.

Objective: To improve the links between theory and data, we develop a model of pathogen evolution in heterogeneous host populations that can be parameterized with a range of empirical measures of host heterogeneity to infection, and which is sufficiently flexible to capture the life history of many natural host–pathogen systems. We use this model to determine whether rabbits co-infected with gastrointestinal helminths could have contributed to the attenuation of the myxoma virus, and might explain differences in the virulence of strains more recently circulating in Australia and Scotland.

Methods: We constructed a deterministic model of pathogen transmission and solved it numerically to determine evolutionarily stable strategies with respect to transmission and virulence. Using this model and empirical data from the rabbit–myxoma virus system, we examine how host heterogeneity in co-infection with gastrointestinal helminths affects the severity of an evolved pathogen in a given host type, and to what degree host heterogeneity affects a pathogen's ability to spread in the host population.

Results: Host heterogeneity to infection always decreases pathogen spread and often leads to the under-exploitation of a typical host. However, the specifics are sensitive to: the shape of the distribution describing host heterogeneity, the relationship between virulence and transmission, and the relationship between the heterogeneous host trait and pathogen virulence. In the rabbit–myxoma virus system, gastrointestinal helminths plausibly contributed to the attenuation of the myxoma virus but are unlikely to have contributed to the higher virulence of circulating strains in Scotland relative to those in Australia in the years following the release of the virus.

Keywords: evolutionarily stable strategy, MYXV, trade-off, vector-borne.

Correspondence: M.P. Kain, Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. email: kainm@mcmaster.ca *or* morganpkain@gmail.com

INTRODUCTION

Heterogeneity among hosts to infection influences host–pathogen interactions and, in doing so, pathogen evolution. For example, host heterogeneity can affect the evolution of pathogen virulence, defined broadly as the harm caused by pathogens to hosts, and thus the severity of a disease at the host population level. Host heterogeneity may select for lower virulence (Ebert and Hamilton, 1996; Gandon, 2004; Pugliese, 2011; Osnas and Dobson, 2012), for higher virulence (Gandon *et al.*, 2001; Ganusov *et al.*, 2002; Read *et al.*, 2015), or facilitate the co-existence of two strains with different exploitation strategies (Fleming-Davies *et al.*, 2015). Pathogen evolution in a heterogeneous host population can also disproportionately harm a subset of the population, where disease severity can increase in the most susceptible hosts, such as unvaccinated individuals (Gandon *et al.*, 2001). Moreover, host heterogeneity can decrease a pathogen's absolute fitness at its evolutionary optimum. For example, heterogeneity can reduce the intrinsic reproduction number (R_0) of a pathogen that cannot adjust its exploitation strategy to its current host (Regoes *et al.*, 2000; Gandon, 2004).

Previous theoretical studies have clarified many of the intricacies generated by the interactions among host heterogeneity to infection, host–pathogen dynamics, and evolution. However, the complexity of the models used, and the variation in model structures and assumptions across studies, make it difficult to calibrate models from empirical data and to predict outcomes for the pathogen and for hosts. Improving the link between theory and empirical data is a critical step in advancing our understanding of pathogen evolution – arguably the most important challenge currently faced in research on the ecology and evolution of infectious diseases (Alizon and Michalakis, 2015; Cressler *et al.*, 2016). While the field has made substantial progress, for example by parameterizing virulence–transmission trade-off curves for a range of pathogens (Fraser *et al.*, 2007; De Roode *et al.*, 2008; Doumayrou *et al.*, 2013; Berngruber *et al.*, 2015), and by quantifying the relationship between pathogen virulence and transmission across the full life cycle of the pathogen (De Roode *et al.*, 2008; Kain and Bolker, 2017), connections between theory and data that account for the role of host heterogeneity in pathogen evolution remain rare.

Here we present a model for the evolution of pathogen virulence in a heterogeneous host population that is flexible enough to be parameterized with a range of empirical measures of host heterogeneity (such as host age, immune status, genotype, vaccine status, or co-infection with other parasitic species), and which is sufficiently tractable to capture the life history of many natural host–pathogen systems. We parameterize our model with empirical data from the *Oryctolagus cuniculus*–myxoma virus (MYXV) system, the canonical example of host–pathogen evolution following pathogen invasion (Fenner and Marshall, 1957; Dwyer *et al.*, 1990; Fenner and Fantini, 1999). We examine variation among rabbits in gastrointestinal helminth burden, which is known to interact with the outcome of MYXV infection (Cattadori *et al.*, 2007). The effects of co-infection by multiple pathogen species on the evolution of virulence have been previously examined (Choisy and de Roode, 2010; Restif and Graham, 2015); however, despite the important role of co-infection in infection and transmission dynamics (Cattadori *et al.*, 2007; Graham *et al.*, 2007; Fenton, 2008; Thakar *et al.*, 2012; Cattadori *et al.*, 2014), the effects of heterogeneity in co-infection on the evolution of a focal pathogen have rarely been considered. We explore the impact of a heterogeneous helminth burden on the severity of an evolved pathogen in a given host type and on the pathogen's ability to spread in the host population. We also evaluate whether or not the observed differences in the virulence of circulating MYXV strains in Scotland and Australia can plausibly be explained by differences in the prevalence of gastrointestinal helminths.

METHODS

Model structure

Our model examines the evolutionarily stable strategy (ESS) of a pathogen exploiting a heterogeneous host population. We define ESS virulence as the virulence that maximizes a pathogen's intrinsic reproductive number R_0, the expected number of new infections a single infected individual generates in an otherwise susceptible population [this criterion, derived from a more general criterion of non-invasibility, holds for a broad range of epidemiological models (Alizon *et al.*, 2009)]. Given that we assume a relationship between transmission rate and virulence (Alizon *et al.*, 2009; Cressler *et al.*, 2016), ESS virulence determines both overall pathogen severity at the population level and pathogen severity in a subset of the population, and indirectly affects the pathogen's ability to spread. Virulence is not fundamentally an intrinsic trait of the pathogen, but rather the result of a complex interaction between host and pathogen. Virulence can be defined as the disease-induced mortality rate of a host within a homogeneous population, or as the mortality rate of an infected *reference host* in a heterogeneous population. In the *O. cuniculus*–MYXV system, the virulence of a circulating MYXV strain is defined as the mortality rate of an infected laboratory breed of rabbit with no shared evolutionary history with the virus (Fenner and Marshall, 1957). In a heterogeneous host population, infection with a single pathogen strain will cause a range of mortality rates. For consistency with the *O. cuniculus*–MYXV system literature, we define ESS virulence in the heterogeneous population as the disease-induced mortality rate of a reference rabbit with no helminths. Alternatively, the reference host could be defined as a well-nourished host, a fully immunocompetent or immunodeficient host, or a host with a reference genotype.

Our model comprises three flexible components: (1) a trade-off curve that relates pathogen virulence to transmission (Alizon *et al.*, 2009); (2) the distribution of *observable heterogeneity* in the host population, which we modelled as a continuous distribution, in contrast to some previous studies that used discrete groups; and (3) the *heterogeneity map*, i.e. a function that translates observable heterogeneity to the impacts on the host interaction with the pathogen, such as host mortality rate. Mathematically, the observable heterogeneity distribution and the heterogeneity map are redundant – we need only the distribution of pathogen-induced host mortality or pathogen clearance rate to predict ESS virulence. However, we include both components in our model to facilitate empirical parameterization and testing of the model. Once the heterogeneity map has been parameterized using laboratory experiments or field data, for example, only the observable heterogeneity, such as secondary infection severity among hosts, is required to determine pathogen ESS virulence. Ultimately, this framework should allow easier model calibration and prediction of ESS virulence in a variety of systems.

Trade-off curve

Due to the trade-off between virulence and transmission (Anderson and May, 1982; Ewald, 1983), transmission (β) is assumed to be an increasing function of virulence (α). When β is a decelerating function of α, a single value of α maximizes R_0 (Alizon *et al.*, 2009). We examine both power-law and sigmoidal functions for the relationship between α and β; these functions are commonly used to model directly transmitted and vector-borne diseases, respectively (Alizon and van Baalen, 2005). Both transmission rate (β) and instantaneous rate of

disease-induced host mortality (virulence, α) can be modelled as a function of within-host pathogen exploitation strategy [e.g. replication rate, titre, or set-point viral load (Fraser *et al.*, 2007)], though this relationship does not need to be explicitly modelled and can be ignored without loss of generality. Here we model pathogen exploitation strategy (σ) as titre load and assume a linear relationship between φ and α ($\alpha = b\varphi$), which has previously been assumed in the *O. cuniculus*–MYXV system (Anderson and May, 1982; Dwyer *et al.*, 1990).

A power-law relationship is given by:

$$\beta(\alpha) = c\alpha^{\frac{1}{\gamma}},$$

with $\alpha = b\varphi$, where b is the slope of the linear relationship between replication rate and mortality rate. For the remainder of the paper we focus on α, using $b = 0.1$ in all simulations. The shape of the curve is controlled by γ (larger γ increases curvature and decreases optimal virulence), while c is a scaling factor. The value for R_0 is given by $(\beta(\alpha))/(\mu + \alpha)$, or for the power law:

$$R_0 = \frac{c\alpha^{\frac{1}{\gamma}}}{\mu + \alpha},$$

where μ is the background rate of host mortality. In this case, the α that maximizes R_0 is given by

$$\alpha^* = \frac{\mu}{\gamma - 1}.$$

When the trade-off curve follows a sigmoidal curve, for example the Hill function (Tjørve, 2003),

$$\beta(\alpha) = \frac{c\alpha^n}{\alpha + \alpha^n},$$

then

$$R_0 = \frac{c(b\alpha)^n}{(a + (b\alpha)^n)(\mu + b\alpha)},$$

where n and a jointly control the location of the inflection point and slope of the curve at the inflection point, and c is a scaling parameter. The α that optimizes R_0 for the Hill function has no closed-form solution but can be found numerically. Figure 1 shows examples of a power-law and sigmoidal trade-off curve.

Host heterogeneity

We assume hosts possess a single trait that determines their reaction to pathogen infection (e.g. host age, immune status, genotype, vaccine status, or co-infection with another pathogen species). We use a Gamma distribution to model the among-host variation (a continuous analog of the negative binomial distribution more commonly used to quantify macro-parasite load), exploring a range of scenarios by adjusting the Gamma shape and scale parameters. Unlike Gandon (2004), we assume that all pathogen strains are perfectly implastic and can only adopt a single strategy across all hosts. This assumption is supported

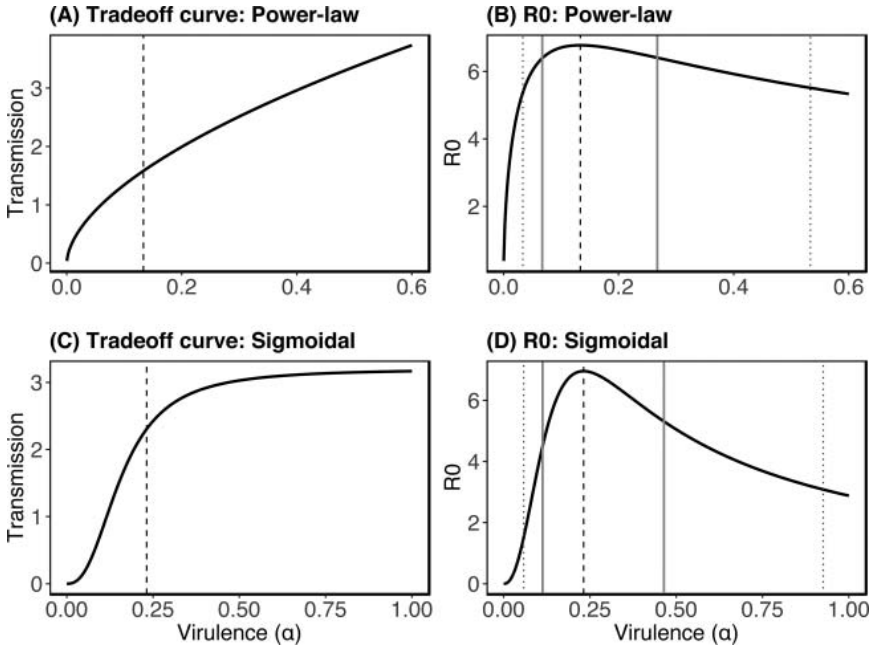

Fig. 1. Panels A and B show an example of a power-law trade-off curve and the R_0 curve that results from this power-law trade-off curve, respectively. Panels C and D show an example of a sigmoidal trade-off curve and the R_0 curve that results from this sigmoidal trade-off curve, respectively. The vertical dashed black lines show optimum α values. The vertical dotted black lines and solid grey lines show the R_0 for a 2× and 4× higher or lower α than optimum, respectively. The parameter values used here were specifically chosen to accentuate the differences between the power-law and sigmoidal trade-off curves (power-law: $c = 5$, $\gamma = 1.75$, $m = 0.2$, $\mu = 0.1$; sigmoidal: $a = 0.01$, $b = 0.1$, $n = 2.5$, $\mu = 0.01$, $c = 3.2$). Figure S1 shows power-law and sigmoidal trade-off curves with the parameter values used in the simulation for Figs. 2–4 (with $\gamma = 1.05$ for the power-law relationship and $c = 1.5$ for the sigmoidal relationship).

for the *O. cuniculus*–MYXV system (Best and Kerr, 2000; Kerr *et al.*, 2017); however, evidence for pathogen plasticity is evident in other systems [e.g. *P. aeruginosa* infection in mammals (Furukawa *et al.*, 2006)].

Heterogeneity map

The distribution of host heterogeneity is translated into a distribution of virulence using a heterogeneity map. Our first choice of heterogeneity function describes a monomolecular, or saturating exponential, relationship between a host trait and pathogen virulence:

$$\alpha_h = \alpha(z(1 - e^{-rx}) + 1),$$

where z controls the maximum of the curve, r controls the rate of approach to z, and α_h is the mortality rate of a host with a trait value (e.g. helminth burden) equal to x. A saturating function is a plausible first approximation of the relationship between a host trait, such as nutrient status or immunocompetence, and pathogen virulence (Bedhomme *et al.*, 2004). In

the MYXV–helminth case, this function describes the increasing, but saturating effect of helminth burden on MYXV virulence, reflecting a type-1/type-2 immune response trade-off (for details, see 'The *O. cuniculus*–myxoma virus case study', p. 264).

Alternatively, we consider a heterogeneity map that allows for a reduction in pathogen virulence at intermediate values of the heterogeneous trait. To model this situation, we use a piecewise function that combines a quadratic function with the saturating exponential function. This function could be used, for example, to capture variation in the immune response, where a smaller immune upregulation reduces pathogen virulence but a larger upregulation increases virulence because of an increase in pathogen exploitation in an effort to evade the stronger immune constraints or immunopathology. For the *O. cuniculus*–myxoma virus case study (see below), we use this piecewise function to model a primed immune response, where low to intermediate helminth burdens reduce MYXV virulence.

$$\alpha_h = \begin{cases} \left(\dfrac{x}{h} - 1\right)^2 (1-k) + k, & x \le h \\ z(1 - e^{-r(x-h)}) + k, & x > h \end{cases}$$

Here h determines the location of minimum pathogen virulence, which defines the helminth burden that provides a rabbit with the largest decrease in MYXV virulence (i.e. the lowest possible MYXV mortality rate: α_h). This function starts at 1 and decreases to a helminth burden of h following a quadratic relationship, where relative virulence at h is given by k ($0 < k < 1$). At trait values greater than h, virulence increases to $z + 1$ following a saturating exponential function. The complete model is illustrated in Fig. 2.

Model outcomes

From a host-centric view, this model can be used to examine the relationship between the parameters of the observed heterogeneity distribution and pathogen ESS virulence, such as the median trait or variation in that trait. For example, a change in ESS virulence attributable to a change in the median trait alone (a pure location shift) can be examined by comparing two homogeneous populations with a different trait value, while a change in ESS virulence attributable to just a change in variance can be quantified by comparing a homogeneous population with a given trait value to a heterogeneous population with the same median. The latter scenario can be used to examine what we term *median-exploitation*, which quantifies the amount of over- or under-exploitation experienced by a *typical host* (i.e. a host with a trait value at the median of the heterogeneity distribution) that is directly attributable to host heterogeneity. A direct comparison of ESS virulence in a heterogeneous population to a reference population with trait value equal to zero involves changes in both median and variance, and therefore the predicted change in ESS virulence cannot be attributed to host heterogeneity. Nonetheless, this comparison is useful for predicting ESS virulence in rabbit populations from empirical data, without an explicit decomposition of the relative effect of a change in median and change in variance.

From a pathogen-centric view, our model can be used to determine how host heterogeneity affects a pathogen's ability to spread in the host population. This can be quantified as the fold (multiplicative) change in pathogen R_0 that results from exploitation of the heterogeneous population, a quantity we refer to as pathogen *efficiency*. Efficiency quantifies the cost to the pathogen for exploiting a heterogeneous host population. In our case

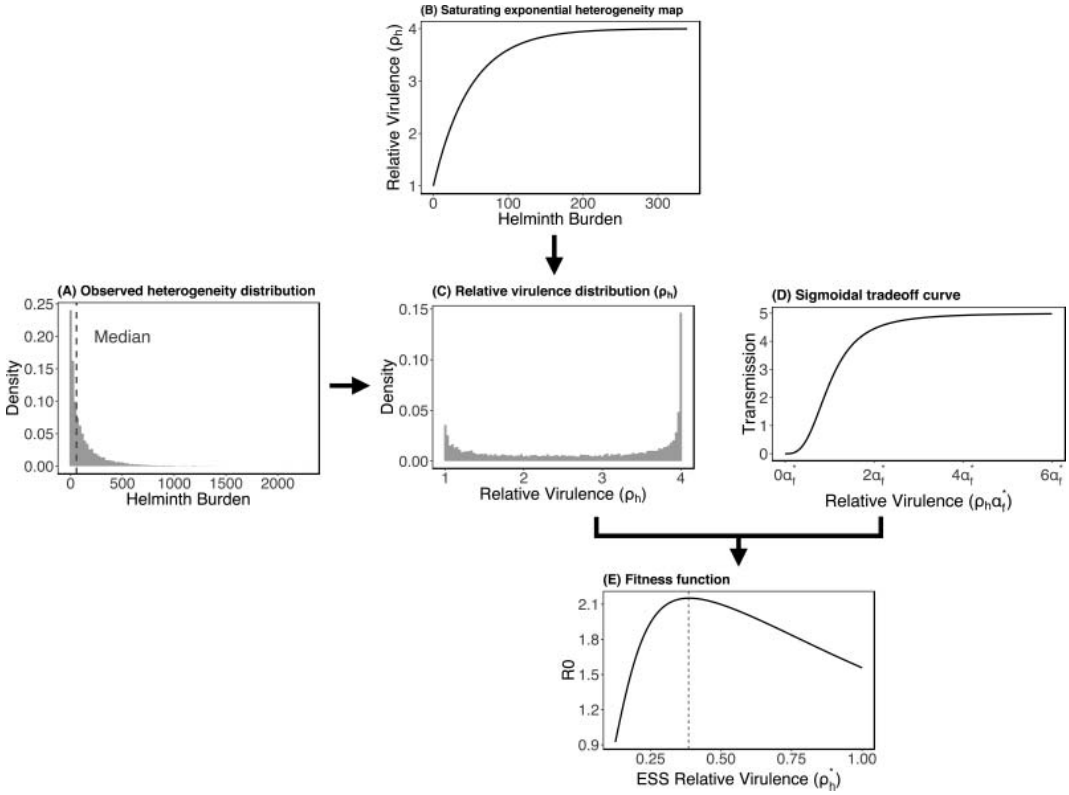

Fig. 2. Model schematic. Panel A shows the population distribution of helminth burden, our measure of host heterogeneity. Panel B shows a saturating-exponential heterogeneity map. Panel C shows the population distribution of relative virulence (ρ_h: the fold-increase in virulence relative to a helminth-free rabbit) that results from translating the distribution in A through the heterogeneity map in B. Together the ρ_h distribution (panel C) and the trade-off curve (panel D: a sigmoidal curve in this example) determine the R_0 value at different levels of pathogen relative virulence (panel E). The horizontal axis in panel E shows the virulence of a given pathogen strain in the heterogeneous population relative to ESS virulence in the homogeneous population, with the dotted line showing ESS relative virulence in the heterogeneous population (ρ_h^*).

study, efficiency is defined as the R_0 of the ESS strain in the heterogeneous population divided by the R_0 of the ESS strain in a homogeneous host population. Because we assume the pathogen always reaches its ESS virulence, R_0 and thus efficiency is unaffected by the median trait value, and therefore it is just a function of host heterogeneity.

Patterns in ESS virulence, over- or under-exploitation of a typical host, and efficiency can be explored generally using the *shiny* R applications (Chang *et al.*, 2017) in the Appendix (evolutionary-ecology.com/data/3118Appendix.pdf). These applications allow users to select parameter values and visualize model outcomes in a layout similar to Fig. 2. Users can explore these applications on their own or paired with the general model exploration presented in the Appendix. The material there describes in detail the qualitative patterns we found most interesting while exploring the model, including details about qualitative differences in results between the power-law and sigmoidal trade-off curve, and between the

saturating exponential and piecewise heterogeneity map. Parameter values for the results presented in the Appendix are available in Table S1.

In the main text we instead focus more narrowly on applying our model to the *O. cuniculus*–myxoma virus system.

The *O. cuniculus*–myxoma virus case study

The initial release of MYXV in Australia and Europe in the 1950s caused unprecedented rabbit mortality: case mortality for the Standard Laboratory Strain (SLS) was ~99%, with an infected rabbit's average lifespan under 13 days (Fenner, 1953). However, within three years of its release, MYXV evolved to intermediate virulence on both continents (Kerr *et al.*, 2012). It is well established that attenuation of the myxoma virus occurred due to a combination of increased resistance in rabbits and higher fitness of lower virulence myxoma strains (Kerr *et al.*, 2015). Yet, many fundamental aspects in the evolution of MYXV virulence remain unresolved, including the impact of heterogeneity in host susceptibility to MYXV infection. In natural settings, rabbits are commonly infected with a community of gastrointestinal helminths that interact directly or indirectly within the host (Lello *et al.*, 2004; Cattadori *et al.*, 2008, 2014; Murphy *et al.*, 2013) and can affect the virulence and transmission of MYXV (Kerr *et al.*, 2004).

Detailed information on the interaction of MYXV and helminths, such as the effects of helminth burden on MYXV infection severity and colonization within the host, is currently unavailable. However, it is known that helminths, including the common gastrointestinal species found in rabbits, cause an upregulation of the type-2 immune response (Cattadori *et al.*, 2007, 2016; Murphy *et al.*, 2011, 2013; Thakar *et al.*, 2012), which has antagonistic components to the type-1 reaction developed against viruses and bacteria, including MYXV (Nash *et al.*, 1999; Kerr *et al.*, 2004; Cattadori *et al.*, 2007). These contrasting immune responses impair the host's ability to successfully control both infections (Kerr *et al.*, 2004; Cattadori *et al.*, 2007). Therefore, we would expect infection by helminths to increase the mortality of rabbits infected with MYXV. Under this scenario, gastrointestinal helminths might have contributed to the rapid attenuation of MXYV virulence by causing the accelerated mortality of co-infected rabbits and facilitating the selection for less virulent MYXV strains that allow longer host survival. We assess this strict type-1/type-2 immune response trade-off (Cattadori *et al.*, 2007), dependent on the total helminth burden (but irrespective of helminth species) using the saturating heterogeneity map.

The alternative scenario is where low (but non-zero) helminth burdens may increase the ability of rabbits to cope with MYXV infection. Rabbits that have co-evolved with helminths for centuries (Audebert and Durette-Desset, 2007), and with MYXV for decades (Fenner and Fantini, 1999), have developed some resistance to the virus, and can manage co-infections with both when helminth intensity is not too high or MYXV is not too virulent (Cattadori *et al.*, 2007, 2008; Kerr *et al.*, 2017). Indeed, observations of wild rabbits from recent populations show that rabbits can cope with both infections (Cattadori *et al.*, 2007). These observations suggest that low to intermediate helminth burdens may train (prime) the host immune response to deal with both infections to the point of being beneficial to the host in coping with the MYXV infection (Kemp and Björkstén, 2003; Okada *et al.*, 2010; Graham *et al.*, 2011). At intermediate to high helminth burdens, the effect of the type-1/type-2 immune response trade-off is assumed to outweigh the benefit of a primed immune reaction, leading to an increase in MYXV virulence in rabbits with intermediate to high helminth burdens. We model this scenario

using the non-monotonic piecewise heterogeneity map, which allows for rabbits with low to intermediate helminth burdens to have *reduced* MYXV virulence relative to helminth-free rabbits but retains the saturating exponential relationship between helminth burden and MYXV virulence at high helminth burdens, which captures the type-1 and type-2 response trade-off. Under this general scenario, a helminth burden distribution with a low median would select for higher MYXV virulence to allow the virus to escape the immune response and achieve maximum transmission.

To explore the effects of these biological scenarios, we first examine qualitative patterns in MYXV ESS virulence under both scenarios across a broad parameter set. We focus on MYXV ESS virulence as a function of the two heterogeneity maps and helminth burden distributions and examine an isolated change in both median and variance in helminth burden. Here we consider only a sigmoidal trade-off relationship between virulence and transmission of MYXV fit to data on the relationship between virulence, transmission, and recovery rate; more cases are considered in the Appendix. The quantitative relationship is taken from Fenner and Marshall (1957; presented in Dwyer *et al.*, 1990) and estimated using non-linear least squares [with the nlxb function in the package nlmrt (Nash, 2016)].

Second, we examine the possibility that a non-monotonic relationship between helminth burden and MYXV virulence could explain current-day local MYXV strain diversity and historical differences in virulence between the UK and Australia (Appendix Fig. S2). Phylogenetic studies from recent rabbit populations of Australia and the UK have found that highly virulent strains are currently co-circulating with attenuated strains (Kerr *et al.*, 2012, 2015, 2017). There is also evidence that the average virulence of strains has been historically higher in the UK than in Australia (Fig. S2). An exploratory statistical analysis over years 4–30 after release, ignoring the first 4 years following release (a poorly sampled 4-year transient period), shows that the average case mortality of UK MYXV strains has been higher than Australian strains on average between 1955 and 1985 (Australia 95% CI: 0.73–0.77, UK 95% CI: 0.82–0.89; $P < 0.05$; generalized linear model with binomial error distribution and numbers of strains sampled as weights). Differences in the gastrointestinal helminth burden may help to explain these results. We examine what combinations of parameters for the location of the piecewise minimum (h) and the maximum reduction of MYXV virulence (k) must be assumed for heterogeneity in gastrointestinal helminth burden to generate higher ESS virulence. Here, h determines the helminth burden at which relative MYXV virulence is the lowest, where this minimum virulence is given by $k(0 < k < 1)$. Beyond this helminth burden, relative MYXV virulence follows the saturating exponential curve. We use helminth burden estimates from three sites in Australia (Dunsmore, 1966) and one site in Scotland (Cattadori *et al.*, 2007). We note that the historic trend of higher MYXV virulence in the UK (Fig. S2) is not directly associated with the Scotland site, and that the helminth burdens measured at the Scotland site may not be representative of helminth burdens across the UK.

Computational methods

We write ESS virulence in a homogeneous reference population as α_f^* and in a heterogeneous population as α_h^*. We define α_h^* as $\alpha_f^* \rho_h^*$, where we call ρ_h^* the ESS relative virulence. In the reference population, optimal virulence α_f^* is calculated by setting the derivative of the R_0 expression to 0 and solving for α; we denote the corresponding R_0 as R_{0f}. To determine α_h^*, we first translate the distribution of helminth burden into a distribution of pathogen

virulence among hosts using the given heterogeneity map. With our assumption of a perfectly implastic pathogen, an intrinsic pathogen exploitation strategy results in a different realized virulence in each individual in the heterogeneous population. We refer to pathogen virulence in a given individual as α_h, and to the distribution of α_h values as the *realized virulence distribution*. We define α_h as $\alpha_f^* \rho_h$, where ρ_h is relative virulence in a particular host. We refer to the distribution of ρ_h values as the *relative virulence distribution*. A given pathogen exploitation strategy will result in a population-level R_0 that depends on the distribution of $\alpha_h(\rho_h)$. A pathogen with virulence α_f^* in a reference host has an R_0 in a heterogeneous population given by the following integral:

$$R_0 = \int_0^\infty \frac{\beta(\alpha_f^* \rho_h)}{\mu + \alpha_f^* \rho_h} P(\rho_h)\, \mathrm{d}\rho_h,$$

where $P(\rho_h)$ is the distribution of relative virulence in the heterogeneous population, and $\beta(\alpha_f^* \rho_h)$ is given by the trade-off curve.

We solve this integral numerically in R and find the level of pathogen virulence that optimizes R_0 in the heterogeneous population ($\alpha_h^* = \alpha_f^* \rho_h^*$) using the optimize function. Like previous models, this method for determining pathogen ESS virulence results in a pathogen that adopts a life-history strategy tailored to hosts that contribute the most to the pathogen's reproductive potential – called 'high quality' hosts by Gandon (2004) and 'prime hosts' by Pugliese (2011) – weighted by host abundance. We refer to the R_0 of the pathogen in the heterogeneous population with virulence α_h^* as R_{0h}. In this notation, relative virulence is expressed as $\rho_h^* = \alpha_h^*/\alpha_f^*$, median-exploitation is defined by ρ_h^* in the presence of changing heterogeneity without a change in the median, and efficiency is expressed as R_{0h}/R_{0f}. Table S1 shows parameter values used in the general model exploration presented in the Appendix as a companion to the *shiny* R application, as well as parameters used in our examination of the *O. cuniculus*–myxoma virus case study.

RESULTS

General model exploration: qualitative patterns

In our general model exploration, we focus on the effects of the shape of the trade-off curve, heterogeneity map, and host trait distribution on qualitative patterns in ESS virulence and efficiency, as well as robust patterns in the sensitivity of these model outcomes to changes in the parameter values. In brief, we find that an increase in host heterogeneity to infection (variance) always decreases the ability for a pathogen to spread in the host population, but that the amount of pathogen efficiency loss depends strongly on the shape of both the trade-off curve and the heterogeneity map. An increase in the slope of either trade-off curve decreases parasite R_0, although the change in R_0 is larger and more sensitive to changes in the slope of the sigmoidal trade-off curve. Both ESS virulence and the amount of over- or under-exploitation of a typical host also depend on the shape of the trade-off curve. For example, for the saturating exponential heterogeneity map and for a given host heterogeneity distribution, increasing the slope/curvature (γ) of the power-law trade-off curve decreases ESS relative virulence, while an increase in the slope of the sigmoidal trade-off curve at its inflection point increases ESS relative virulence and over-exploitation of a typical host (Fig. S3).

These patterns as well as the effects of the shape of the heterogeneity map and observable heterogeneity distribution are presented in the text of the Appendix, and can be explored visually using the online *shiny* R applications presented there.

The *O. cuniculus*–myxoma virus case study

Might co-infection with gastrointestinal helminths have contributed to the observed post-introduction attenuation of MYXV virulence?

Any saturating heterogeneity map capturing the type-1/type-2 immune response trade-off always predicts lower ESS virulence (i.e. MYXV attenuation) in a population of rabbits with a non-zero helminth burden, regardless of the shape of the helminth burden distribution. However, the relative importance of changes both in the median and variance of helminth burden depends on the shape of the heterogeneity distribution. For example, with the saturating heterogeneity map shown in Fig. 2 adding a helminth burden homogeneous to all rabbits of 56.9 parasites (this value is the median of the Gamma distribution pictured in Fig. 2) leads to a 3.05-fold decrease in MYXV virulence. Adding a heterogeneous worm burden following the Gamma distribution pictured reduces ESS virulence by a factor of 2.50. Thus, the effect of heterogeneity is to moderately *increase* MYXV virulence. Similar but variable increases in ESS virulence attributable to heterogeneity are seen with different Gamma distributions and can be examined using the online *shiny* R applications. Illustrations of the relative effects of changing the median and the variance are available in Figs. S3 and S4. The Gamma distribution of helminth burden does not allow for a median equal to zero and positive variance, so we cannot compare a homogeneous population with a median equal to 0 and a heterogeneous population with a median equal to 0.

In contrast, a non-monotonic heterogeneity map that models a decrease in MYXV virulence at low to intermediate helminth burden greatly reduces the parameter space leading to MYXV attenuation. Instead, it more commonly selects for an *increase* in MYXV ESS virulence relative to a homogeneous rabbit population without helminth parasites. For example, with the helminth burden distribution for Urana, NSW, and the parameters for the saturating exponential proportion of the piecewise heterogeneity map used in Fig. 2 ($z = 3$, $r = 0.020$), the addition of the quadratic portion of the piecewise heterogeneity map leads to an increase in MYXV ESS virulence at most parameter values for k and h (Fig. 3). Alternative values for z and r can be explored using the *shiny* R application in the Appendix.

When partitioning the relative impact of a change in the median and a change in the variance of the helminth burden distribution on ESS virulence using a piecewise heterogeneity map, we find that a change in variance has the largest effect when the median of the heterogeneous host population is near the minimum of the quadratic portion of the piecewise heterogeneity map. In this case a homogeneous population with helminth burden near the quadratic vertex experiences a MYXV strain with higher virulence than a population with zero helminth burden. Here, an increase in the variance of the helminth burden without a change in the median increases the proportion of the rabbit population with high helminth burdens who select for a decrease in MYXV virulence; a large increase in the variance can switch ESS virulence from >1 (higher virulence) to <1 (lower virulence, i.e. attenuation). For example, with a piecewise heterogeneity map with parameters $z = 3$, $r = 0.02$, $h = 21$, $k = 0.6$, a homogeneous rabbit population with a helminth burden of 20.33

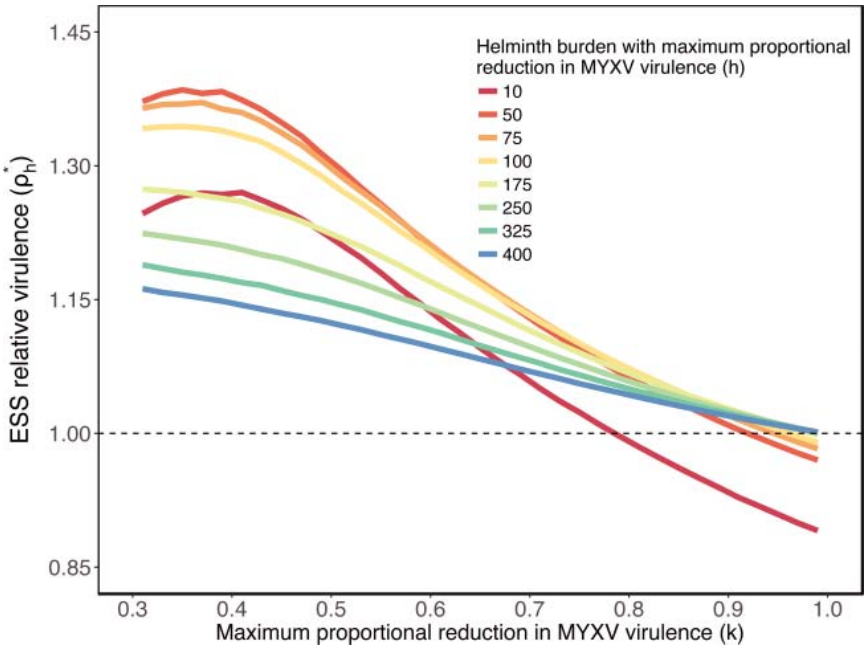

Fig. 3. ESS relative virulence, ρ_h^*, given by the ratio of ESS virulence in the heterogeneous rabbit population (α_h^*) to ESS virulence in a homogeneous helminth-free rabbit population (α_f^*). Results shown here use the helminth burden distribution fit to Urana, NSW and a piecewise heterogeneity map (saturating exponential portion: $z = 3$, $r = 0.02$ as pictured in Fig. 2; quadratic portion: $k = 0.31$–0.99, $h = 10$–400). The x-axis shows a range for k, which is the proportional reduction in MYXV virulence at the quadratic vertex. Colours show k, which is the helminth burden at the vertex of the quadratic portion of the piecewise heterogeneity map.

parasites results in a 2.5× *higher* ESS virulence of MYXV than in a homogeneous rabbit population without helminths. However, in a rabbit population with a helminth burden following a Gamma distribution with a coefficient of variation of 1.83 and a median of 20.33 parasites (scale = 278, shape = 0.3), MYXV ESS virulence is 1.11-fold *lower* than in a homogeneous population with no helminths. This pattern is presented visually in Fig. S4.

Under what conditions can co-infection with helminths facilitate higher ESS MYXV virulence?

Given that a non-monotonic relationship between helminth burden and MYXV virulence can increase MYXV ESS virulence in a rabbit population with helminths, the difference in virulence in MYXV strains circulating in Scotland and Australia could be explained by the differences in helminth burdens in these two countries. In this section, we examine the parameters for the non-monotonic heterogeneity map that could support differences in ESS virulence of the magnitude seen in Australia and Scotland.

Only an intermediate reduction in virulence (e.g. magnitude of MYXV virulence reduction $k = 0.4$) at a high helminth burden (e.g. helminth burden at virulence, minimum $h = 300$) selects for higher ESS virulence in Scotland than in Snowy Plains, NSW (with a peak of ~1.10× higher virulence) (Fig. 4). At all other parameter combinations, virulence is

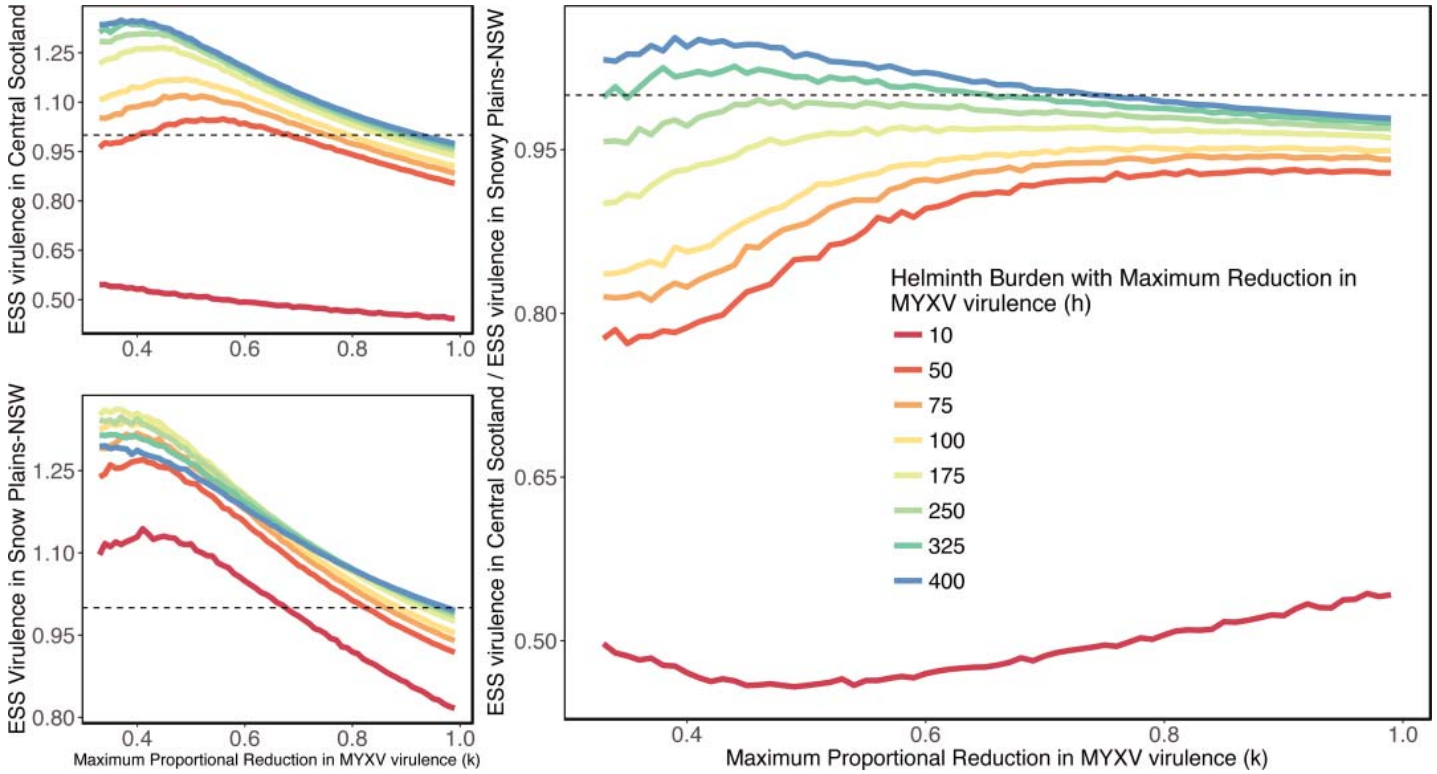

Fig. 4. Panels A and B show the fold change in MYXV virulence in Central Scotland and Snowy Plains, NSW relative to MYXV virulence in a homogeneous helminth-free rabbit population. In panels A and B, regions above and below the dashed reference line show parameter combinations that lead to a higher or lower ESS MYXV virulence in the wild rabbit populations, respectively. Panel C shows the ratio of ESS virulence in Scotland relative to ESS virulence in Snowy Plains, NSW. Here, regions above and below the dashed reference line show parameter combinations that lead to higher or lower virulence in Scotland, respectively. Parameter values for the piecewise heterogeneity map are the same as those used in Fig. 3 ($z = 3$, $r = 0.02$, $k = 0.31$–0.99, $h = 10$–400).

higher in Snowy Plains, NSW. A rapid change in the ratio of ESS virulence between Scotland and Australia occurs between a piecewise minimum (h) of 10 and 50, respectively. This range of precipitous change occurs because of a high variance in the relative virulence distribution in the Snowy Plains, NSW population. Similar qualitative results occur using different parameter values for the piecewise heterogeneity map (Fig. S5).

Comparing the rabbit population from Scotland and rabbit populations in other locations in Australia (either Urana, NSW or Mitchell, Queensland), the parameter space for h and k that leads to higher ESS virulence in Scotland is slightly larger than for Snowy Plains, NSW (Fig. S6). The larger difference between the Urana, NSW or Mitchell, Queensland populations and Scotland (Snowy Plains has the most similar helminth burden to Scotland of the three Australian populations), supports a larger range of parameters for the heterogeneity map that selects for higher ESS virulence in Scotland (Fig. S6). As the difference in median helminth burden between the two countries increases, a smaller proportion of the two helminth burden distributions overlap the same region of the heterogeneity map, which leads to divergent responses in the populations from the two countries. This allows for more heterogeneity maps that have a quadratic portion spanning the range of helminth burdens found in the Australian but not in the Scotland population, which selects for higher virulence in Scotland.

DISCUSSION

We have presented a general model on the evolution of virulence in heterogeneous host populations that retains the fundamental structure and emergent properties of previous models (e.g. Gandon *et al.*, 2001; Pugliese, 2011) but has been structured to be more easily parameterized with empirical data (Alizon and Michalakis, 2015; Cressler *et al.*, 2016). Specifically, experimental or observational data are needed to parameterize the distribution of the focal host trait that influences infection (host heterogeneity), the heterogeneity map between pathogen virulence and the host trait, and the transmission–virulence trade-off curve. In the *O. cuniculus*–MYXV system, the distributions of gastrointestinal helminths in wild rabbit populations were calculated from sampling wild animals (Dunsmore, 1966; Cattadori *et al.*, 2007). Experiments to parameterize the heterogeneity map remain incomplete. In the end, two separate experiments will be needed in the *O. cuniculus*–MYXV system to parameterize the heterogeneity distribution and heterogeneity map. However, with an *a priori* goal of model parameterization, the distribution of host heterogeneity could be collected in conjunction with an experiment designed to parameterize the heterogeneity map. The sigmoidal trade-off curve in the *O. cuniculus*–MYXV system was parameterized in two steps using laboratory experiments performed previously. First, viral loads and the lifespan of infected rabbits were determined by examining the survival of laboratory animals infected with a diversity of MYXV strains (Fenner and Marshall, 1957). Second, the relationship between viral load and vector transmission was determined by allowing mosquitoes to feed on a rabbit at the site of the primary skin infection and calculating the probability of MYXV transmission to an uninfected rabbit (Fenner *et al.*, 1956). In the case of a directly transmitted disease, both transmission and host mortality could be measured in the same experiment (De Roode *et al.*, 2008).

Since no host–pathogen system currently provides sufficient data to parameterize the entire model, including the *O. cuniculus*–MYXV system, we first explored the qualitative effects of different shapes of the heterogeneity map and the virulence–transmission trade-off curve on ESS virulence, median-exploitation, and pathogen efficiency. Second, we

used our model to examine the effects of host heterogeneity to helminth infection on viral evolution in the *O. cuniculus*–MYXV system. We examined whether helminths could have been important drivers of the initial rapid attenuation of MYXV virulence in Australia and Europe (Kerr *et al.*, 2004; Cattadori *et al.*, 2007), and/or of the prolonged difference in MYXV virulence between these two continents (Fig. S2) (Kerr *et al.*, 2012, 2015, 2017). Because the exact effects of helminth co-infection on MYXV virulence are unknown, we investigated two possibilities. First, based on the general concept of the dichotomy between the type-1 and type-2 immune response trade-off to micro- and macro-parasites, we examined a monotonically increasing but saturating relationship between helminth burden and MYXV virulence. Second, motivated by observations of rabbits in the field (Cattadori *et al.*, 2007), we examined a non-monotonic relationship between helminth burden and MYXV virulence that modelled lower realized MYXV virulence in rabbits co-infected with low to intermediate helminth burdens, where the immune system has already been primed.

Model construction and assumptions

The heterogeneity map (which describes the relationship between a measurable host trait and pathogen virulence) and the distribution of host heterogeneity for a focal trait are key drivers of changes in relative virulence, median-exploitation, and efficiency (i.e. ability to transmit). These two model components affect results directly through their joint control of the relative virulence distribution (ρ_h, Fig. 2), and indirectly by regulating the impact of the functional form (power-law or sigmoidal) and slope of the trade-off curve. The shape of the trade-off curve has the largest impact on relative virulence, median-host exploitation, and efficiency (Fig. S3) when variance in the relative virulence distribution is maximized, which occurs when the relative virulence distribution is bimodal (Fig. 2). This can occur, for example, when the heterogeneity map approaches its asymptote near the mean of the host heterogeneity distribution (Fig. 2). With a bimodal relative virulence distribution, any change in either the measured host heterogeneity distribution or the heterogeneity map causes a shift in which mode of the bimodal relative virulence distribution has higher density. For example, in the extreme of a relative virulence distribution composed of half low-virulence and half high-virulence extremes (e.g. a population with 50% helminth-free rabbits and 50% rabbits with heavy helminth burden), a change to a distribution of either 51%:49% or 49%:51% causes a shift in the preferred host type and a large difference in model outcomes. That is, a bimodal relative virulence distribution will result in an implastic pathogen specializing on the marginally more abundant host type, leading to large efficiency losses because of extreme over- or under-exploitation of the other host types. Although we do not explicitly model the case of two discrete host types, a bimodal relative virulence distribution that approximates a host population with two host types has been previously examined (e.g. Gandon *et al.*, 2001; Gandon, 2004). Additionally, while our model is limited to a single monomorphic strain, a bimodal relative virulence distribution favours the evolution of a specialist pathogen on a subset of the host population (Regoes *et al.*, 2000). In the opposite extreme of a uniform relative virulence distribution, the ESS pathogen strategy is to specialize on hosts with a mean trait value. In this scenario, changes in the distribution of helminths or the shape of the trade-off curve have little effect on model outcomes.

Because of the dependence of the relative virulence distribution on the shape of the heterogeneity map, our main conclusions depend strongly on our assumption of a saturating heterogeneity map (which holds for both the saturating exponential and the piecewise

maps used here). Specifically, a saturating function increases the range of population heterogeneity distributions that create a bimodal distribution of relative virulence. This not only increases the sensitivity of model outcomes to changes in population heterogeneity, but also increases the sensitivity of results to the shape of the trade-off curve. In our case study, the relationship between helminth burden and MYXV virulence is likely not perfectly saturating (i.e. the virus–helminth interaction does not produce a true ceiling on MYXV virulence), though a saturating function is a plausible first approximation of the effects of co-infection and other sources of host heterogeneity. Alternative scenarios can be smoothly incorporated into the model framework. For example, it is possible to specify a power-law curve that decelerates (negative second derivative) but increases indefinitely (no asymptote).

At one extreme, if the full range of the heterogeneity distribution spans only the approximately linear portion of the heterogeneity map, the relative virulence distribution will resemble a scaled heterogeneity distribution, reducing the importance of the shape of the trade-off curve. At the other extreme, when the heterogeneity distribution spans the full range of the heterogeneity map such that an appreciable portion of its density (e.g. 30%) is near the asymptote of the heterogeneity map, variance in the relative virulence distribution will be maximized (see Fig. 2), and the shape of the trade-off curve increases in importance. Together, these results imply that while a decelerating curve has the potential to create a bimodal relative virulence distribution, it will greatly depend on the slope of the curve within the range of the heterogeneity distribution.

Previous studies have gathered data on individual heterogeneity in traits similar to those needed to parameterize the heterogeneity map, including host size (Cable and Van Oosterhout, 2007), nutritional status (Bedhomme *et al.*, 2004), and maternal environment (Stjernman and Little, 2011; Garbutt *et al.*, 2014). However, the data from most of these studies are insufficient for our model because they are based on experiments that used two discrete host groups (ANOVA design) instead of a treatment range [regression design (Inouye, 2001; Cottingham *et al.*, 2005)]. Despite noisy data, there is some preliminary empirical evidence for a saturating heterogeneity map (Cable and Van Oosterhout, 2007), but also for a linear relationship between food availability and virulence, at least within the range examined (Bedhomme *et al.*, 2004). Regardless, more empirical work that follows a regression design is needed for parameterization of the heterogeneity map. One promising case is the *Daphnia magna–Pasteuria ramosa* system, in which virulence differs between low and high rearing temperature (Garbutt *et al.*, 2014) and where a trade-off curve has been quantified (Jensen *et al.*, 2006).

In an effort to keep our model simple enough to be readily parameterized with experimental data, and given the ecological specifics of the *O. cuniculus*–MYXV system, we ignore the evolution of pathogen specialists (Regoes *et al.*, 2000), as well as complexities such as pathogen plasticity (Gandon, 2004; Fleming-Davies *et al.*, 2015), host–pathogen co-evolution (Pugliese, 2011), or the emergence of individual heterogeneities during the course of the infection (Osnas and Dobson, 2012). Our assumption that the focal pathogen is perfectly implastic causes heterogeneity to maximally increase over- and under-exploitation in a subset of hosts; a perfectly implastic pathogen will experience the highest possible efficiency loss in a heterogeneous population. Alternatively, a completely plastic pathogen will change its virulence to match any value of the relative virulence distribution, preventing efficiency loss in a heterogeneous population. While we did not examine pathogen plasticity, it has recently been shown that pathogen plasticity, in the presence of host heterogeneity, can lead to the co-existence of high and low virulence strains (Fleming-Davies *et al.*, 2015). The lack of plasticity shown by MYXV makes this an implausible explanation for the variation in virulence of circulating

MYXV strains, but may be important to consider in other systems. The model framework can incorporate other assumptions for pathogen plasticity by allowing the pathogen's intrinsic exploitation strategy to vary depending on the host type. For example, a single additional scaling parameter could be added to the model that pulls a host's effective virulence (α_h) towards the intrinsic exploitation strategy of the pathogen, which would reduce the impact of heterogeneity.

In the *O. cuniculus*–MYXV system, the trade-off between transmission and virulence is the dominant constraint affecting pathogen evolution, though the original formulation of the trade-off using this system actually used the trade-off between transmission and recovery (Anderson and May, 1982). This trade-off may be more important in other systems than the trade-off we model here (Alizon, 2008; Cressler *et al.*, 2016). While the fundamental structure of our model does not need to change to incorporate a transmission–recovery trade-off, adjustments to each model component are needed. The equation for R_0 would need to be rewritten to focus on recovery, and both the heterogeneity distribution and the heterogeneity map would need to focus on traits associated with recovery rather than virulence. The computational methods for finding the optimum pathogen exploitation strategy would be applied as described here.

Finally, we ignore the possibility that changes in MYXV virulence, a vector-transmitted disease, affect both transmission from an infected rabbit to a susceptible mosquito/flea and transmission from an infected mosquito/flea to a susceptible rabbit. Given that mosquitoes or fleas act as mechanical vectors for MYXV, changes in virulence are not expected to affect the vector in this system. However, in other systems pathogen virulence in the host and pathogen virulence in the vector could be coupled to some degree. Ignoring this coupling could potentially lead to an incorrect estimate of the optimal pathogen strategy.

The *O. cuniculus*–myxoma virus case study

This study was in part motivated by our interest in the role of helminth infections in the attenuation of MYXV following its initial release in Australia and Europe in the 1950s. It is well understood that MYXV evolved towards a more attenuated virulence on both continents (Kerr *et al.*, 2012) because of selection against both highly virulent strains that killed rabbits too quickly and low virulence strains that produced too little virus for transmission by vectors (Fenner and Marshall, 1957; Dwyer *et al.*, 1990). The parallel increased resistance of rabbits to MYXV was an additional cause of reduced virus transmission (Kerr *et al.*, 2015). We argue that helminths might have contributed to the initial attenuation of MYXV by having susceptible rabbits with an already impaired immune response to the virus experience faster mortality. To capture this immune alteration, specifically the type-1/type-2 antagonistic immune reaction between MYXV and helminths (Nash *et al.*, 1999; Cattadori *et al.*, 2007), we used a saturating function to describe the heterogeneity map. A saturating heterogeneity map always predicts lower ESS virulence in the presence of helminths regardless of the shape of the heterogeneity distribution and trade-off curve. Critically, partitioning the effects of a change in the median helminth burden from a change in the variance of helminth burden, under a strict type-1/type-2 immune response trade-off, we found that an increase in the heterogeneity of helminth burden without a change in the median tends to decrease ESS virulence. Thus, any effect of MYXV attenuation attributable to helminths under the scenario of a strict type-1/type-2 trade-off is caused by an increase in the overall helminth burden in the population (measured here using the median).

Under the assumption of reduced MYXV virulence in rabbits infected with a range of helminth burdens, which we modelled using a piecewise heterogeneity mapping function, ESS virulence can be higher or lower than in a homogeneous population (Figs. 3 and S4) depending on the shape of the heterogeneity map. Here, an increase in variation without a change in the median can decrease ESS virulence substantially, potentially reversing the overall change in virulence from an increase (comparing between a homogeneous population with no helminths and a homogeneous population with helminths) to a decrease (comparing between a homogeneous population with helminths to a heterogeneous population with median helminth burden equal to the helminth burden in the homogeneous population).

Together, these results suggest that little can be said definitively. There is an urgent need for experiments that quantify the shape of the heterogeneity mapping function in order to provide a conclusive answer to the possibilities presented here. For example, laboratory experiments of MYXV–helminth co-infections can quantify the MYXV replication rate and virulence in rabbits infected with different helminth burdens. With the ability to parameterize the heterogeneity map, this simple heterogeneity model could be combined with a more detailed epidemiological model (e.g. Dwyer *et al.*, 1990) for MYXV transmission to capture more realistically the trajectory and causes of the attenuation of MYXV following the initial introduction as well as the current patterns of circulating strain diversity. Tentatively, we argue that the most realistic scenario is that co-infections, in concert with an increase in rabbit resistance, would have contributed to the decrease of MYXV virulence in wild rabbits following the release of the virus, but further work is needed to support this conclusion.

We also examined whether lower MYXV virulence in rabbits with intermediate helminth burden relative to helminth-free rabbits could explain the approximately 1.2 times higher virulence of circulating MYXV strains in Scotland than in Australia in the decades following the release of the virus in both countries (Fig. S2). Using a piecewise function for the heterogeneity mapping function, we show that there is a relatively restricted and moderately extreme parameter set that selects for higher virulence in a population of rabbits with higher mean helminth burden. For example, minimum MYXV virulence has to occur above 300 helminths, which is much higher than the median of 138 helminths estimated from the Gamma distribution fit to the helminth burdens measured in Scotland (Cattadori *et al.*, 2007). These 300 helminths must also be accompanied by a reduction in MYXV virulence of ~50% (Fig. 4). Given the prevalence of both MYXV and helminths in Scotland and the much larger effect MYXV has on rabbit mortality relative to helminths, including the evolution of increased host resistance to MYXV on a time frame of decades, it is unlikely that a reduction of 50% in virulence would be maintained in the host population without evolution towards increased susceptibility to helminths. Smaller reductions in MYXV virulence (k) can also lead to greater ESS virulence in Scotland, but only at even larger helminth burdens (Fig. 4). For example, with $h = 400$, virulence in Scotland is only a maximum of ~1.1 times that in Snowy Plains, NSW (Fig. 4), providing further support against helminths as a causal factor in selecting for higher MYXV virulence.

A comparison between Scotland and Urana, NSW (a population with lower helminth burden than the Snowy Plains, NSW population), a larger parameter space supports higher virulence in the Scotland population (Fig. S6). If the Urana, NSW population was isolated from the rest of Australia, a mediated helminth-controlled reduction in relative MYXV virulence could be a plausible explanation for differences in MYXV virulence between the

populations. However, with the mixing of rabbit populations and the equally regional spread of MYXV by mosquitoes and fleas, it is unlikely that the shape of the heterogeneity map varies by population. Therefore, while a given heterogeneity map could plausibly explain why virulence is higher in Scotland than a single region of Australia, this explanation is brittle in the face of a well-mixed host population on the continental scale. Yet, this does not rule out the possibility that some of the spatial diversity observed in MYXV virulence (Kerr *et al.*, 2010, 2015) is driven by interpopulation variation in helminth burden. Little is known about the rate of strain emergence or extinction, or the scale at which strains circulate, though recent work indicates new strains are gained and lost each year (Kerr *et al.*, 2017). Our simulations presented in Figs. 3 and 4 suggest that differences in the distribution of helminth burden can select for large differences in ESS virulence and could in part be driving the patterns of the high spatial diversity in MYXV virulence (Kerr *et al.*, 2010, 2015).

To improve our model, the within-host dynamics of MYXV in rabbits free of helminths can provide a baseline of the fundamental processes of virus evolution. The empirical trade-off curve used here, derived from laboratory data (Fenner and Marshall, 1957), does not capture the important biological MYXV–rabbit interactions that directly affect virulence. Indeed, the use of viral titre measured at the site of the mosquito skin lesion [fibroma (Dwyer *et al.*, 1990)] leads to a range of values for peak titre or area under the curve [two standard methods of measuring transmission potential (Handel and Rohani, 2015)], resulting in two possible values of both case mortality and transmission (i.e. two values from a non 1 : 1 function). Ultimately, more details on the dynamics of infection within rabbits and the role of secondary skin lesion in MYXV transmission are needed if we are to develop more realistic evolutionary models of infection.

CONCLUSION

The mapping function of host heterogeneity to pathogen virulence has a marked influence on specific outcomes of interest. In the absence of empirical data, we caution that assumptions about its characteristics, and to a lesser extent, the shape of the transmission–virulence trade-off curve, can have large effects on model outcomes. While our simulations, like others, are largely driven by model assumptions (Pugliese, 2011), the ease of substitution of functional forms in our model will help bridge the gap between theory and data in this field. As Alizon and Michalakis (2015) and Cressler *et al.* (2016) indicate, the field drastically needs more studies that incorporate knowledge of the full life cycle of both host and pathogen. Infection with a second pathogen species is likely to be a large part of the host's life cycle as well as modulate the life history and traits of the focal pathogen (Graham *et al.*, 2007; Fenton, 2008; Thakar *et al.*, 2012; Cattadori *et al.*, 2014) and ultimately pathogen evolution.

ACKNOWLEDGEMENTS

We thank the Dushoff lab and Jo Werba for helpful comments on the first draft of the manuscript. This work was funded by NSERC Discovery Grant 386590-2010.

REFERENCES

Alizon, S. 2008. Transmission–recovery trade-offs to study parasite evolution. *Am. Nat.*, **172**: 113–121.

Alizon, S. and Michalakis, Y. 2015. Adaptive virulence evolution: the good old fitness-based approach. *Trends. Ecol. Evol.*, **30**: 248–254.

Alizon, S. and van Baalen, M. 2005. Emergence of a convex trade-off between transmission and virulence. *Am. Nat.*, **165**: 155–167.

Alizon, S., Hurford, A., Mideo, N. and Van Baalen, M. 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.*, **22**: 245–259.

Anderson, R.M. and May, R. 1982. Coevolution of hosts and parasites. *Parasitology*, **85**: 411–426.

Audebert, F. and Durette-Desset, M. 2007. Do lagomorphs play a relay role in the evolution of the trichostrongylina nematodes? *Parasite*, **14**: 183–197.

Bedhomme, S., Agnew, P., Sidobre, C. and Michalakis, Y. 2004. Virulence reaction norms across a food gradient. *Proc. R. Soc. Lond. B: Biol. Sci.*, **271**: 739–744.

Berngruber, T.W., Lion, S. and Gandon, S. 2015. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS Pathog.*, **11**: e1004810.

Best, S.M. and Kerr, P.J. 2000. Coevolution of host and virus: the pathogenesis of virulent and attenuated strains of myxoma virus in resistant and susceptible European rabbits. *Virology*, **267**: 36–48.

Cable, J. and Van Oosterhout, C. 2007. The impact of parasites on the life history evolution of guppies (*Poecilia reticulata*): the effects of host size on parasite virulence. *Int. J. Parasitol.*, **37**: 1449–1458.

Cattadori, I.M., Albert, R. and Boag, B. 2007. Variation in host susceptibility and infectiousness generated by co-infection: the myxoma–*Trichostrongylus retortaeformis* case in wild rabbits. *J. R. Soc. Interface*, **4**: 831–840.

Cattadori, I., Boag, B. and Hudson, P. 2008. Parasite co-infection and interaction as drivers of host heterogeneity. *Int. J. Parasitol.*, **38**: 371–380.

Cattadori, I.M., Wagner, B.R., Wodzinski, L.A., Pathak, A.K., Poole, A. and Boag, B. 2014. Infections do not predict shedding in co-infections with two helminths from a natural system. *Ecology*, **95**: 1684–1692.

Cattadori, I.M., Sebastian, A., Hao, H., Katani, R., Albert, I., Eilertson, K.E. *et al.* 2016. Impact of helminth infections and nutritional constraints on the small intestine microbiota. *PLoS One*, **11**: e0159770.

Chang, W., Cheng, J., Allaire, J., Xie, Y. and McPherson, J. 2017. *shiny: Web Application Framework for R*. R package version 1.0.0 [https://cran.r-project.org/web/packages/shiny/index.html].

Choisy, M. and de Roode, J.C. 2010. Mixed infections and the evolution of virulence: effects of resource competition, parasite plasticity, and impaired host immunity. *Am. Nat.*, **175**: 105–118.

Cottingham, K.L., Lennon, J.T. and Brown, B.L. 2005. Knowing when to draw the line: designing more informative ecological experiments. *Front. Ecol. Environ.*, **3**: 145–152.

Cressler, C.E., McLeod, D.V., Rozins, C., Van Den Hoogen, J. and Day, T. 2016. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology*, **143**: 915–930.

De Roode, J.C., Yates, A.J. and Altizer, S. 2008. Virulence–transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proc. Natl. Acad. Sci. USA*, **105**: 7489–7494.

Doumayrou, J., Avellan, A., Froissart, R. and Michalakis, Y. 2013. An experimental test of the transmission–virulence trade-off hypothesis in a plant virus. *Evolution*, **67**: 477–486.

Dunsmore, J. 1966. Nematode parasites of free-living rabbits, *Oryctolagus cuniculus* (L.) in eastern Australia. *Aust. Zool.*, **14**: 185–199.

Dwyer, G., Levin, S.A. and Buttel, L. 1990. A simulation model of the population dynamics and evolution of myxomatosis. *Ecol. Monogr.*, **60**: 423–447.

Ebert, D. and Hamilton, W.D. 1996. Sex against virulence: the coevolution of parasitic diseases. *Trends Ecol. Evol.*, **11**: 79–82.

Ewald, P.W. 1983. Host–parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Evol. Syst.*, **14**: 465–485.

Fenner, F. 1953. Changes in the mortality-rate due to myxomatosis in the Australian wild rabbit. *Nature*, **172**: 228–230.

Fenner, F. and Fantini, B. 1999. *Biological Control of Vertebrate Pests: The History of Myxomatosis – An Experiment in Evolution.* Wallingford, UK: CABI Publishing.

Fenner, F. and Marshall, I. 1957. A comparison of the virulence for European rabbits (*Oryctolagus cuniculus*) of strains of myxoma virus recovered in the field in Australia, Europe and America. *Epidemiol. Infect.*, **55**: 149–191.

Fenner, F., Day, M. and Woodroofe, G.M. 1956. Epidemiological consequences of the mechanical transmission of myxomatosis by mosquitoes. *Epidemiol. Infect.*, **54**: 284–303.

Fenton, A. 2008. Worms and germs: the population dynamic consequences of microparasite–microparasite co-infection. *Parasitology*, **135**: 1545–1560.

Fleming-Davies, A.E., Dukic, V., Andreasen, V. and Dwyer, G. 2015. Effects of host heterogeneity on pathogen diversity and evolution. *Ecol. Lett.*, **18**: 1252–1261.

Fraser, C., Hollingsworth, T.D., Chapman, R., de Wolf, F. and Hanage, W.P. 2007. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. USA*, **104**: 17441–17446.

Furukawa, S., Kuchma, S. and O'Toole, G. 2006. Keeping their options open: acute versus persistent infections. *J. Bacteriol.*, **188**: 1211–1217

Gandon, S. 2004. Evolution of multihost parasites. *Evolution*, **58**: 455–469.

Gandon, S., Mackinnon, M.J., Nee, S. and Read, A.F. 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature*, **414**: 751–756.

Ganusov, V.V., Bergstrom, C.T. and Antia, R. 2002. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution*, **56**: 213–223.

Garbutt, J.S., Scholefield, J.A., Vale, P.F. and Little, T.J. 2014. Elevated maternal temperature enhances offspring disease resistance in *Daphnia magna. Funct. Ecol.*, **28**: 424–431.

Graham, A.L., Cattadori, I.M., Lloyd-Smith, J.O., Ferrari, M.J. and Bjørnstad, O.N. 2007. Transmission consequences of coinfection: cytokines writ large? *Trends Parasitol.*, **23**: 284–291.

Graham, A.L., Shuker, D.M., Pollitt, L.C., Auld, S.K., Wilson, A.J. and Little, T.J. 2011. Fitness consequences of immune responses: strengthening the empirical framework for ecoimmunology. *Funct. Ecol.*, **25**: 5–17.

Handel, A. and Rohani, P. 2015. Crossing the scale from within-host infection dynamics to between-host transmission fitness: a discussion of current assumptions and knowledge. *Phil. Trans. R. Soc. Lond. B: Biol. Sci.*, **370**: 20140302.

Inouye, B.D. 2001. Response surface experimental designs for investigating interspecific competition. *Ecology*, **82**: 2696–2706.

Jensen, K.H., Little, T., Skorping, A. and Ebert, D. 2006. Empirical support for optimal virulence in a castrating parasite. *PLoS Biol.*, **4**: e197.

Kain, M.P. and Bolker, B.M. 2017. Can existing data on West Nile virus infection in birds and mosquitos explain strain replacement? *Ecosphere*, **8**: e01684.

Kemp, A. and Björkstén, B. 2003. Immune deviation and the hygiene hypothesis: a review of the epidemiological evidence. *Pediatr. Allergy Immunol.*, **14**: 74–80.

Kerr, P.J., Perkins, H., Inglis, B., Stagg, R., McLaughlin, E., Collins, S. *et al.* 2004. Expression of rabbit IL-4 by recombinant myxoma viruses enhances virulence and overcomes genetic resistance to myxomatosis. *Virology,* **324**: 117–128.

Kerr, P.J., Hone, J., Perrin, L., French, N. and Williams, C. 2010. Molecular and serological analysis of the epidemiology of myxoma virus in rabbits. *Vet. Microbiol.*, **143**: 167–178.

Kerr, P.J., Ghedin, E., DePasse, J.V., Fitch, A., Cattadori, I.M., Hudson, P.J. *et al.* 2012. Evolutionary history and attenuation of myxoma virus on two continents. *PLoS Pathog.*, **8**: e1002950.

Kerr, P.J., Liu, J., Cattadori, I.M., Ghedin, E., Read, A.F. and Holmes, E.C. 2015. Myxoma virus and the leporipoxviruses: an evolutionary paradigm. *Viruses*, **7**: 1020–1061.

Kerr, P.J., Cattadori, I.M., Rogers, M.B., Fitch, A., Geber, A., Liu, J. *et al.* 2017. Genomic and phenotypic characterization of myxoma virus from Great Britain reveals multiple evolutionary pathways distinct from those in Australia. *PLoS Pathog.*, **13**: e1006252.

Lello, J., Boag, B., Fenton, A., Stevenson, I.R. and Hudson, P.J. 2004. Competition and mutualism among the gut helminths of a mammalian host. *Nature*, **428**: 840–844.

Murphy, L., Nalpas, N., Stear, M. and Cattadori, I.M. 2011. Explaining patterns of infection in free-living populations using laboratory immune experiments. *Parasite Immunol.*, **33**: 287–302.

Murphy, L., Pathak, A. and Cattadori, I.M. 2013. A co-infection with two gastrointestinal nematodes alters host immune responses and only partially parasite dynamics. *Parasite Immunol.*, **35**: 421–432.

Nash, J.C. 2016. *nlmrt: Functions for Nonlinear Least Squares Solutions*. R package version 2016. 3.2 [https://cran.r-project.org/web/packages/nlmrt/nlmrt.pdf].

Nash, P., Barrett, J., Cao, J.X., Hota-Mitchell, S., Lalani, A.S., Everett, H. *et al.* 1999. Immunomodulation by viruses: the myxoma virus story. *Immunol. Rev.*, **168**: 103–120.

Okada, H., Kuhn, C., Feillet, H. and Bach, J.F. 2010. The 'hygiene hypothesis' for autoimmune and allergic diseases: an update. *Am. J. Clin. Exp. Immunol.*, **160**: 1–9.

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

Pugliese, A. 2011. The role of host population heterogeneity in the evolution of virulence. *J. Biol. Dyn.*, **5**: 104–119.

Read, A.F., Baigent, S.J., Powers, C., Kgosana, L.B., Blackwell, L., Smith, L.P. *et al.* 2015. Imperfect vaccination can enhance the transmission of highly virulent pathogens. *PLoS Biol.*, **13**: e1002198.

Regoes, R.R., Nowak, M.A. and Bonhoeffer, S. 2000. Evolution of virulence in a heterogeneous host population. *Evolution*, **54**: 64–71.

Restif, O. and Graham, A.L. 2015. Within-host dynamics of infection: from ecological insights to evolutionary predictions. *Proc. R. Soc. Lond. B: Biol. Sci.*, **370**: e1002345.

Stjernman, M. and Little, T. 2011. Genetic variation for maternal effects on parasite susceptibility. *J. Evol. Biol.*, **24**: 2357–2363.

Thakar, J., Pathak, A.K., Murphy, L., Albert, R. and Cattadori, I.M. 2012. Network model of immune responses reveals key effectors to single and co-infection dynamics by a respiratory bacterium and a gastrointestinal helminth. *PLoS Comput. Biol.*, **8**: e1002345.

Tjørve, E. 2003. Shapes and functions of species–area curves: a review of possible models. *J. Biogeogr.*, **30**: 827–835.

Chapter 5: The evolution of parasites constrained by a virulence-transmission tradeoff and compensatory virulence factor "tuning"

In this chapter I examine the evolution of a parasite following its introduction into a new host population. I model parasites that evolve in replication rate, which I model as the underlying cause of virulence, which in turn is correlated with transmission following the tradeoff theory of virulence evolution. My model extends classic virulence theory by adding a second quantitative trait that represents investment in "compensatory" traits (e.g. secondary virulence factors), which is required for parasites to effectively translate virulence into transmission. I use a stochastic, discrete time model as well as reaction-diffusion and adaptive dynamics deterministic models to understand the effects that random mutations in finite populations have on parasite evolution. I find that all models agree that parasites evolve to their global fitness maximum, but that models that the discrete time stochastic and reaction diffusion models that include eco-evolutionary dynamics predict greater transient virulence than the adaptive dynamics model. Smaller populations and higher mutation rates cause parasites to depart further and more frequently from optimal virulence.

The text I present here is an early draft of a manuscript planned for publication.

Author Contributions

MPK and BMB conceived the study, wrote code for the project, and analyzed preliminary results; MPK analyzed the results presented with helpful feedback from BMB; MPK wrote the early draft of the manuscript presented here.

Acknowledgements

I thank Jo Werba for help on the first draft of this manuscript.

The evolution of parasites constrained by a virulence-transmission tradeoff and compensatory virulence factor "tuning"

Morgan P Kain[1]* and Benjamin M Bolker[1,2]

*Corresponding author: kainm@mcaster.ca
[1]Department of Biology, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada
[2]Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada

Abstract

Parasite exploitation after the emergence of parasite in a new host population leads to parasite (and host) evolution on ecological time scales. Yet, most models for parasite evolution are either restricted to long term evolution or assume that parasite fitness is defined only by virulence and transmission, which means that parasites always reside on their optimization frontier. This is an unrealistic assumption for any organism in a new environment. We present a model for the evolution of a parasite that is initially beneath its tradeoff frontier using a second quantitative trait that represents investment in "compensatory" traits (e.g. secondary virulence factors), which is required for parasites to effectively translate virulence into transmission. We analyze our model using three different approaches: a finite population, discrete time stochastic simulation-based model (DTS), a deterministic reaction-diffusion (differential equation based) model (RD), and an adaptive dynamics model (AD). Using these three models we aim to understand the effects of small populations and stochasticity on transient evolution, time to the adaptive peak, and periodic departures of the population from its adaptive peak. We find that the RD and DTS models show qualitatively similar patterns across a range of parameter values and starting conditions, though these patterns begin to diverge as host populations become very small. We show that with an adaptive landscape with a single maximum, an AD also provide an adequate description qualitative description of parasite evolutionary dynamics. In all three models parasites evolve higher than optimal virulence in the short term because of the negative fitness effects of a mismatch between replication rate and compensatory virulence factors. Eco-evolutionary dynamics also contribute to higher transient virulence in the RD and DTS models. In simple evolutionary models we find that a DTS model may be unnecessary; simple RD or AD models are sufficient to understand parasite qualitative evolutionary dynamics. However, if rare evolutionary events in small populations are of primary interest, only the DTS model is appropriate.

Keywords:

Adaptive Dynamics; Parasite Exploitation; Reaction Diffusion; Stochastic Simulation; Tradeoff theory

Introduction

Parasite exploitation of their host exerts strong selection pressure on both parasite exploitation strategy and host defense strategy (Schneider and Ayres, 2008, Ayres and Schneider, 2012, Kutzer and Armitage, 2016). The resulting evolution often occurs at ecological time scales (Day and Proulx, 2004, Bolker et al., 2010, Lion, 2018), and sometimes in a predictable way (Ebert, 1998, Pugliese, 2002, Laine and Tellier, 2008, Frickel et al., 2016). Often, rapid evolution occurs after a parasite jumps species or emerges in a new host populations; examples of rapid parasite evolution following invasion include: WNV in North American birds (Beasley et al., 2003), MYXV in Eurpoean rabbits (Fenner and Marshall, 1957), *Plasmodium falciparum* in House finches (Fleming-Davies et al., 2018), HIV in Humans (Fraser et al., 2007), λ phage in *E. coli.* (Berngruber et al., 2015), *phycodnaviridae* family dsDNA viruses in algae (Frickel et al., 2016, 2018), and *Pasteuria ramosa* in *Daphnia sp.* (Duffy et al., 2009).

In the past four decades, models for parasite evolution informed by these examples have revealed generalities in parasite evolution and helped to explain parasite exploitation strategy in specific systems (for a review see Cressler et al., 2016). However, most of these models are restricted to long term evolution despite common rapid evolution (but see Bolker et al., 2010, Lion, 2018, Parsons et al., 2018). Models often also collapse high-dimensional host-parasite interactions to two axes: parasite virulence (the rate of host mortality) and parasite transmission rate, which jointly determine parasite fitness (Anderson and May, 1982, Ewald, 1983, Alizon and Michalakis, 2015, Cressler et al., 2016). These choices result in a model where parasites move only *on* their optimization frontier, which is unrealistic for any organism in a new environment. In this paradigm it is impossible to understand how a parasite evolves *to* this frontier because parasite traits apart from virulence and transmission are taken as a "black box" (Alizon et al., 2009, Bull and Lauring, 2014, Alizon and Michalakis, 2015).

While it is relatively common for models to extend tradeoff theory using additional axes of complexity such as within-host parasite dynamics (Alizon and van Baalen, 2005, Day et al., 2011, Mideo et al., 2011), host evolution (Carval and Ferriere, 2010, Best et al., 2014, Papkou et al., 2016), or spatial structure (Lipsitch et al., 1995, Haraguchi and Sasaki, 2000, Lion and Gandon, 2015), we are unaware of a model that helps to explain parasite evolution both to and on the tradeoff curve by treating the tradeoff curve as a true *optimization frontier*. Alizon et al. (2009) mention in passing that parasites beneath their frontier will first evolve to their frontier and then to the tradeoff optimum; however, no additional discussion is given here, or to our knowledge elsewhere, on the evolutionary trajectory a parasite takes to reach its global optimum while under the constraint of tradeoff between virulence and transmission.

Here we present a model for the evolution of a parasite that is initially beneath its tradeoff frontier when

it invades a new host population. Our model introduces a second trait axis, in addition to the commonly used virulence, on which parasites evolve: a quantitative trait representation of investment in "compensatory" traits (e.g. secondary virulence factors) that affect the ability of a parasite to efficiently translate virulence into transmission. In classic tradeoff theory, a change in parasite virulence results in a correlated change in transmission following an assumed tradeoff curve function (often either a power-law or sigmoidal function: Alizon and van Baalen 2005, Bolker et al. 2010, Kain et al. 2018). However, an increase in parasite replication rate may only lead to an increase in transmission rate following changes in other genes. For example, changes to the MYXV virulence factor M148R, which helps to downregulate the host immune system, are needed for an increase in MYXV replication rate to increase transmission, because a secondary side effect of an increased replication rate is a more rapid upregulation of the host immune system (Blanié et al., 2009). Even in the absence of host immune regulation, it is not difficult to imagine that an increase in investment in replication rate, for example, leads to a decrease in *per capita* infection efficiency, enough that *total* efficiency decreases, because of a quantity vs quality life-history tradeoff. To recover this *per capita* efficiency and further increase total efficiency, adjustments in secondary traits are needed. While it is conceivably possible to model the evolutionary dynamics of multiple individual virulence factors, to our knowledge no system is simple enough nor is enough known currently about the function of all of the protein products of a specific parasite for this approach to be fruitful.

Models of parasite evolution often also assume both deterministic evolution in an infinite population and a separation of ecological and evolutionary time scales, despite strong evidence that parasite exploitation affects host population size which influences evolution on ecological time scales (Frickel et al., 2016, Papkou et al., 2016, Frickel et al., 2018). These assumptions make it impossible to study rapid evolution following parasite emergence, or as a result of environmental fluctuations such as seasonality (however, see Day and Proulx 2004, Bolker et al. 2010, Lion 2018, and Parsons et al. 2018). These simplifications are often made in order to obtain analytic solutions, which are usually more easily generalizable; solutions from stochastic simulations are likely to be tied more closely to specific systems and can potentially lead to less powerful conclusions. However, because many researchers have recently been calling for more of a case-study approach (Bull and Lauring, 2014, Alizon and Michalakis, 2015, Cressler et al., 2016), a reduced ability to make widely generalizable conclusions may be a minimal drawback. Work on evolution in finite populations is also restricted by a lack of mathematical theory for deriving analytic solutions for evolution in finite populations on ecological time scales (Lion, 2018), though Parsons et al. (2018) have recently derived analytic solutions for parasite evolution in finite populations using *Stochastic Adaptive Dynamics*. While this model more closely resembles parasite evolution in practice, it assumes rare and small-effect mutations which are unlikely for at least viral pathogens (e.g. MYXV, WNV, *phycodnaviridae* dsDNA viruses).

We analyze our model for the mechanics of parasite exploitation using three different modeling approaches: a biologically realistic scenario of parasite evolution in finite populations using a discrete time stochastic simulation-based model, as well as a deterministic reaction-diffusion (differential equation based) model (RD) that captures both selection (advection) and mutation (diffusion) and an adaptive dynamics representation of the problem (AD). These deterministic approaches serve as two forms of "null-models" against which we compare the results of our stochastic model. Using these three models we aim to understand the effects of small populations and stochasticity on transient evolution, time to the adaptive peak, and periodic departures of the population from its adaptive peak. We examine when and what interesting biological patterns emerge as a function of increasing model complexity; that is, what we can learn about parasite evolution using a stochastic model that is missed when using a simpler framework.

Methods

We examine a parasite that is constrained by a tradeoff between virulence and transmission (Anderson and May, 1982, Ewald, 1983), where transmission rate (β) is an increasing function of host mortality rate (virulence: α). When β is a convex function of α ($\frac{\partial^2 \beta}{\partial \alpha^2} < 0$), a single intermediate value of α maximizes $\mathcal{R}_0$ (Alizon et al., 2009). Here we use a power-law function to model the relationship between α and β (Alizon and van Baalen, 2005, Bolker et al., 2010):

$$\beta(\alpha) = c\alpha^{\frac{1}{\rho}},$$

where ρ controls the function's curvature and c is a scaling factor. For a power-law function, $\mathcal{R}_0$ is given by:

$$\mathcal{R}_0 = \frac{c\alpha^{\frac{1}{\rho}}}{\mu + \alpha + \gamma}$$

where μ is the background rate of host mortality and γ is the rate of host recovery. The $\mathcal{R}_0$ of a parasite constrained by only a power-law tradeoff is obtained at $\alpha^* = \frac{\mu+\gamma}{(\rho-1)}$ (Bolker et al., 2010, Kain et al., 2018).

Here we model parasite transmission that is controlled by both replication rate and a trait that we call "tuning", which we take as a collective representation of the status of a collection of "compensatory" secondary virulence factors. For example, an increase in investment in replication rate could lead to a decrease in parasite *per capita* efficiency; to recover and/or increase *total* efficiency, adjustments in secondary virulence factors are needed. In general, we imagine that an appropriate level of investment in compensatory traits is what allows a parasite to realize maximum transmission at a given replication rate.

We define β given by the power-law tradeoff as the frontier (i.e. maximum achievable) transmission rate (β_f) that is obtained when parasite tuning is perfectly matched to replication rate. The actual transmission

rate of a parasite, realized transmission rate β_r, is a decreasing function of the mismatch between replication rate and tuning. We calculate β_r as β_f multiplied by parasite efficiency (ω), which is given by:

$$\omega = e^{\frac{-(\phi-\theta)^2}{\delta}}, \tag{1}$$

where log(replication rate) is given by ϕ, tuning by θ, and δ is a scaling factor (which we call the efficiency scale) that determines the cost of a mismatch between ϕ and θ. Thus, $\beta_r = \beta_f * \omega$, where $\beta_r = \beta_f$ occurs only when $\phi = \theta$ (i.e. a perfectly tuned parasite). Because $\omega = 1$ when $\phi = \theta$, δ does not change the ϕ that optimizes $\mathcal{R}_0$, but does change $\mathcal{R}_0$ when $\phi \neq \theta$. In the limit of $\delta \to \infty$, the impact of θ on parasite fitness decreases to zero and the model collapses to classic tradeoff theory. In this model, parasite $\mathcal{R}_0$ is given by:

$$\omega \frac{\beta(\alpha)}{\alpha + \gamma + \mu}. \tag{2}$$

While parasite fitness ($\mathcal{R}_0$) is defined by virulence (α) and transmission (β), replication rate (ϕ) and tuning (θ) are the evolving traits. We model ϕ as the log or replication rate, and α as the inverse-logit of ϕ ($\alpha = \frac{e^{\phi}}{e^{\phi}+1}$). We use a logistic scale for two reasons: First, the non-linear logistic scale is a convenient scale on which to model mutations that have little effect when a parasite has either a low or high trait value; second, the logistic scale maps the unbounded scale on which mutation occurs to (0, 1), which allows us to translate parasite replication rate and tuning to probabilities of transmission and host mortality for the DTS model.

In reality, parasite virulence and transmission are a joint function of parasite exploitation strategy and host defense strategy; following parasite invasion, parasite exploitation and parasite defense are known to co-evolve (Fenner and Marshall, 1957, Carval and Ferriere, 2010, Best et al., 2014). To simplify the problem, we make the unrealistic assumption of no host evolution, and instead assume a static level of total host defensive capabilities. While we do not model host defenses, they are in a sense intrinsically captured in our model through parasite tuning. That is, a change in parasite replication rate can be imagined to lead to an initial decline in transmission because of increased host immune efficiency, at which parasites must "retune" to maximize transmission (similar to what has been seen in MYXV).

Using these definitions for parasite exploitation and fitness, we examine parasite evolution using three common modeling techniques: a discrete time, stochastic, finite population model (DTS), a reaction-diffusion, differential equation model (RD) and an adaptive dynamics (AD) model. Our primary interest is on stochastic evolutionary dynamics in small populations; however, results from stochastic simulation can be difficult to interpret on their own. We use results from the RD model, which is deterministic and assumes an infinite population, but which allows for eco-evolutionary feedbacks, as well as the AD model which assumes a

separation ecological and evolutionary time scales, and small-effect mutations, to examine the effects of small populations and stochasticity on transient evolution, time to the adaptive peak, and periodic departures of the population from its adaptive peak. We begin with a description of the DTS model and then describe the simplified RD and AD models.

Discrete time, stochastic, finite population model (DTS)

We model parasite transmission using an SISD model, defined by three host classes: Susceptible (S), infected (I), and dead (D). Infected hosts either recover from infection, at which point they become susceptible, or die and become removed from the population. We assume a homogeneous host population. A unique parasite strain is defined as a strain with a different value for replication rate and/or tuning, and thus virulence and transmission.

In each time step of the DTS model the following events occur: S and I hosts die with background mortality probability μ; I hosts die with probability equal to the virulence of the strain they are infected with; S and I hosts reproduce in a density dependent manner, where the probability of reproduction is equal to one minus the ratio of the sum of living S and I hosts to the starting population size if the population size is less than the initial population size and zero otherwise; Infected hosts recover with probability γ (hosts were assumed to have the same recovery rate from infection for all parasite strains; for the evolution of recovery see Anderson and May 1982, Alizon and van Baalen 2005); and S hosts become infected. A proportion of S hosts *escape* infection with probability equal to:

$$\prod_i (1-\beta_i),$$

while the susceptible hosts that do become infected are infected with strain i with probability *approximately* equal to:

$$p_i = \frac{n_i R_{0i}}{\sum_j n_j R_{0j}}.$$

Using probabilities instead of rates for transmission, death, and recovery leads to a slightly modified equation for $\mathcal{R}_0$ of:

$$\mathcal{R}_0 = \omega \frac{c\alpha^{\frac{1}{\rho}}}{1-((1-\mu)(1-\alpha)(1-\gamma))}. \tag{3}$$

This equation leads to a slightly higher level of optimal virulence than for Eq.2 when using the same parameter values for μ and γ.

We assume a mutation occurs in parasite strain i during transmission with probability π, and that mutations have an additive effect on the logistic scale. Mutational effect sizes for both traits are drawn from a multivariate Normal distribution with a mean of zero, positive standard deviation (the same value for σ was used for both traits), and zero covariance. That is:

$$\left.\begin{aligned}\phi_i &\rightarrow \phi_i + MVN(0, \Sigma_{1,1}) \\ \theta_i &\rightarrow \theta_i + MVN(0, \Sigma_{2,2})\end{aligned}\right\} | B(1, \mu) = 1$$

Newly birthed and recovered hosts are not available to be infected until the next time step and newly infected hosts cannot die, recover, or infect S hosts until the next time step. This process proceeds for a specified length of time or until all hosts or all parasites go extinct.

Reaction diffusion model

The model described above can be written using a system of differential equations, using rates instead of probabilities:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= bS - \mu S - S\int_0^\infty \int_0^\infty \beta(\alpha(\phi,\theta,t)))i(\phi,\theta,t)d\phi d\theta + \gamma \int_0^\infty \int_0^\infty i(\phi,\theta,t)d\phi d\theta \\ \frac{\partial i}{\partial t} &= [S\beta(\alpha(\phi,\theta,t)) - (\alpha+\mu+\gamma)]i(\phi,\theta,t) + D\frac{\partial^2 i}{\partial \phi^2} + D\frac{\partial^2 i}{\partial \theta^2},\end{aligned}$$

where the second derivative (diffusion) terms model the mutational process in replication rate and tuning. To remove the integrals in these equations to efficiently solve this system, we discretized replication rate and tuning into an *m x m* matrix. Each cell of this matrix represents a combination of trait values, and the number in each cell representing the number (proportion) of individuals infected with a strain with that combination of trait values. Diffusion on this matrix represents mutation from one trait value to another, using constant step sizes on the logistic scale. We solved this system of differential equations using the `ode` function in the `deSolve` package (Soetaert et al., 2010) in `R` (R Core Team, 2019); diffusion was implemented using the `tran.2d` function in the `ReacTran` (Soetaert and Meysman, 2012) package using a zero-flux boundary condition (no loss of infected individuals due to mutation).

Adaptive dynamics model

When ecological dynamics are assumed to be fast relative to evolutionary dynamics (a new parasite mutant only invades at ecological equilibrium), and mutational effect size is small, it has been shown that parasites

evolve to optimize $\mathcal{R}_0$ (Dieckmann, 2002, Keeling and Rohani, 2008, Lion, 2018). By definition, once at ecological equilibrium a resident strain has an $\mathcal{R}_0 = 1$. A new mutant is able to invade if it has an $\mathcal{R}_E > 1$ (a single host infected with a mutant strain infects greater than one new host at the equilibrium number of S set by the resident strain). In simple epidemiological models (e.g. SIS, SIR) (Dieckmann, 2002, Keeling and Rohani, 2008, Lion, 2018), an $\mathcal{R}_E > 1$ corresponds to an $\mathcal{R}_0 > 1$. Starting with an arbitrary resident stain, we allow mutant strains with $\mathcal{R}_E = 1 + \epsilon$ to invade and fixate until no new mutant strains have $\mathcal{R}_E > 1$. This process can be used to find the optimal parasite strategy, though it is not guaranteed that the optimum found is the global optimum (Dieckmann, 2002).

Our model generates a fitness landscape (we use this term to describe the relationship between parasite tuning, replication rate and $\mathcal{R}_0$) with a single global maximum and no local maxima for all parameter values. We implement AD by relying on the criteria that a mutant strain with $\mathcal{R}_0$ greater than the resident strain will displace the resident strain. Using mutations with the small additive effect of 0.005 on the logistic scale ($\phi_{t+1} = \phi_t \pm 0.005$; $\theta_{t+1} = \theta_t \pm 0.005$), we allow mutant strains to replace the resident strain until all mutants have an $\mathcal{R}_E < 1$. We consider parasite evolution when a parasite can have a simultaneous mutation in both replication rate and tuning (which we assume in both the DTS and RD models), or a mutation in only a single trait at a time. For the model with mutation in both traits it is possible that a number of mutational changes (e.g. an increase in replication rate and tuning vs an increase in replication rate but a decrease in tuning) can lead to a higher $\mathcal{R}_0$; we assume that the mutant parasite with the larger $\mathcal{R}_0$ invades, which removes any stochasticity from this model. That is, the invading strain, within the mutational range of $\pm\epsilon$ *for both traits*, with the highest $\mathcal{R}_0$ displaces the resident (for a stochastic version of adaptive dynamics see Parsons et al. 2018).

Model outcomes

Our primary aim is to understand the dynamics of the DTS model; we use the results from the RD and AD models to help us clarify the effects of stochasticity and small populations. We focus primarily on three outcomes: transient patterns in parasite virulence, the time it takes a parasite to reach its global optimum, and the size of the long-term virulence distribution (due to mutation/selection balance and/or stochastic departures from the global optimum). We ran all three models across a range of parameter values for μ, σ, N, δ, and initial tuning and replication rate (Table 1). We explore only a single power-law tradeoff. For our DTS model we used 250 stochastic simulations for each parameter combination.

Table 1: Parameter values used in the DTS (discrete time, stochastic, finite population) RD (reaction-diffusion), and AD (adaptive dynamics) models. Parameter values that are not used in a given model are written as "–". Starting values for ϕ and θ for the RD model are closer to 0 on the logit scale because of numeric instability at more extreme starting conditions. Values for the power-law scaling parameter (c) was increased in the RD model to insure an $\mathcal{R}_0 > 1$ at extreme combinations of ϕ and θ with a δ of 10. An increase in c increases $\mathcal{R}_0$ but does not affect the shape of the fitness surface.

Parameter	**DTS model**	**RD model**	**AD model**
δ	10, 30, 50	10, 30, 50	10, 30, 50
μ	0.001, 0.005, 0.025	–	1 or 2 mutations in each new mutant, occurs after ecological equilibrium is reached
Mutational Σ	0.05, 0.01, 0.25	$0.025 * 2.5^{0, 1, 2, 3}$	Constant mutational size of 0.005
Population Size	200, 600, 1800	–	–
Starting [ϕ, θ]	[logit(0.03), logit(0.97)], [logit(0.97), logit(0.03)], [logit(0.03), logit(0.03)]	[logit(0.80), logit(0.20)], [logit(0.20), logit(0.80)], [logit(0.20), logit(0.20)]	[logit(0.03), logit(0.97)], [logit(0.97), logit(0.03)], [logit(0.03), logit(0.03)]
Power-law c	0.75	2-20	0.75
Power-law ρ	2	2	2

Results

We first compare qualitative patterns in parasite evolution among the DTS, RD, and AD models for single parameter sets. We then describe the effects of each parameter on each of our metrics of interest in the DTS model. In all cases we focus on a scenario where a parasite evolves from an initial position of low replication rate and low efficiency (high tuning), which we use to represent a parasite that has recently jumped species and is poorly suited to its new host. In the online supplement we present results when a parasite evolves from high replication rate and low efficiency (another possible scenario following a species jump) (Figure S1) as well as low replication rate, high efficiency (Figure S2). We focus on patterns in parasite virulence and transmission, which are the observable traits from the point of view of an exploited host population. We focus less explicitly on tuning, which effects transmission and constrains virulence evolution but serves as a "silent" trait that is a step removed from epidemiology/ecology.

Climbing the fitness landscape: qualitative patterns among models

The RD and DTS models show qualitatively similar patterns when using similar parameter values for mutational variance in the DTS model and the strength of diffusion in the RD model (σ in both models) (Figure 1). The medians of the RD and DTS models show similar patterns for a parasites' evolution in tuning and replication rate, despite marginal differences in the curvature of the fitness landscape, which occurs because of the difference in the equation for $\mathcal{R}_0$ when using rates or probabilities (see Figure S3 for a comparison of the landscapes). With a $\rho = 2$ in the power-law tradeoff curve, parasites maximize $\mathcal{R}_0$ at a virulence of 0.21 and 0.26 in the RD and DTS models respectively (Figure 1). Variation among stochastic runs also roughly match the size of the variation in the DTS model at these parameter values; however, these patterns describe somewhat different processes. The RD model describes variation within a population of parasites, while the DTS model variation pictured in Figure 1 describes variation among stochastic realizations (we discuss variation within a single stochastic run in the following results section). Even so, Figure 1 shows that the simpler RD model provides an adequate qualitative match to the DTS model dynamics.

The AD model results show that even though parasites overshoot their optimal virulence in the absence of ecological dynamics, parasites evolve to their optimum phenotype in a more direct route than in either the RD and DTS models (Figure 1 bold black line). Eco-evolutionary feedbacks in the RD and DTS models select for increased β in the short term (Day and Proulx, 2004, Bolker et al., 2010), which results in higher transient virulence than in the AD model (Figure 1, top left panel). Selection for higher β also leads to a positive feedback for higher β (and therefore α), because strains with higher β mutate faster than strains with lower β. In AD, when mutations only occur in a single trait at a time (Figure 1 bold black dashed line),

a parasites evolves to its optimum level of virulence first, and then to its optimum level of tuning. This pattern is seen across a wide range of parameter values for δ and the shape of the tradeoff curve, including the parameters used in Figure 1 (e.g. Figures S 1, S 2); a parasite evolving according to AD with mutations in single traits will evolve tuning first only with a small δ and flat power-law tradeoff curve (small ρ) (not pictured).

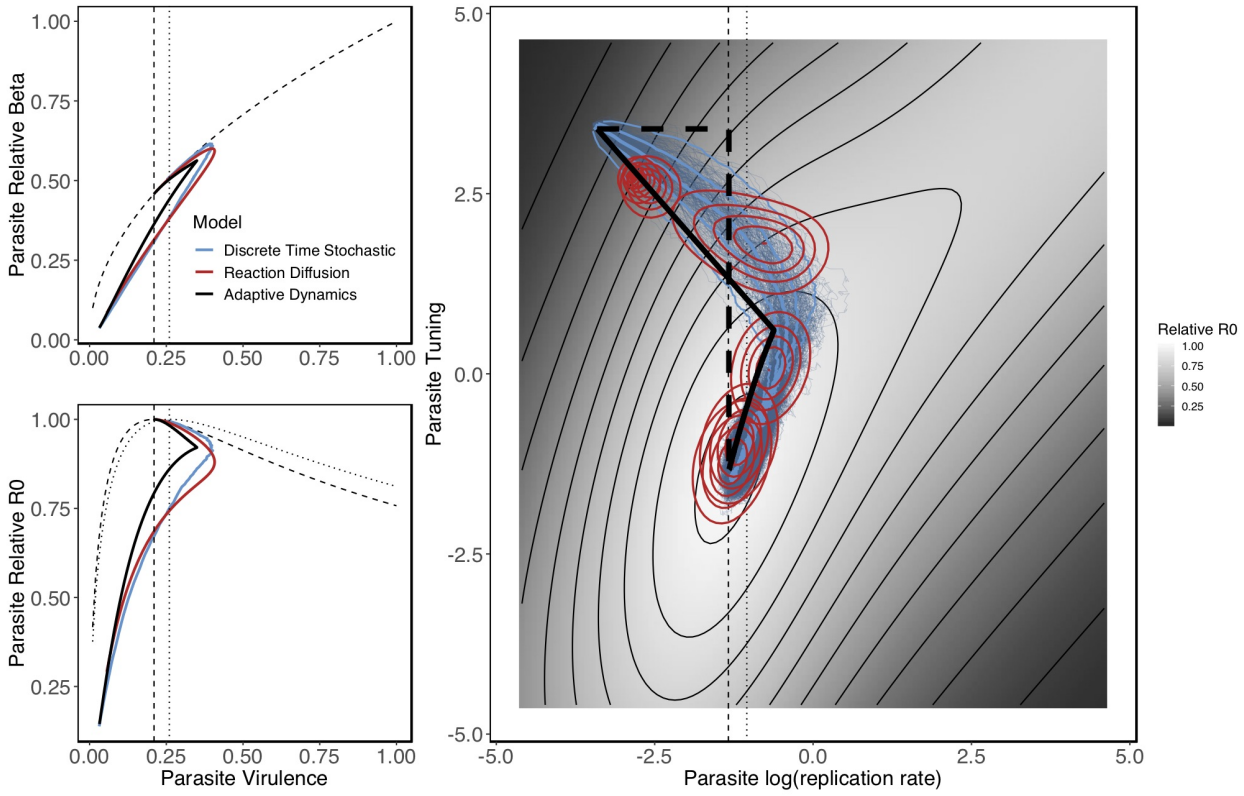

Figure 1: The large right panel shows the evolutionary trajectory of a parasite on its fitness landscape in the DTS model (thin blue lines show 250 stochastic simulations, while the solid blue lines show the median, 50%, and 95% quantities of these simulations), RD model (red ellipses show 95% of the distribution of parasite strains), and AD model (bold black solid line shows simultaneous evolution in two traits; dashed lines show evolution in a single trait per time step), using the following parameter values: DTS: $\mu = 0.005$, $\sigma = 0.05$, N = 600; RD: $\sigma = 0.0625$; All models: $\delta = 30$. The surface pictured is calculated using the $\mathcal{R}_0$ equation for rates (Eq.2), and has been scaled so that the adaptive peak has a value of 1 (shown in light grey). Due to the slight difference in $\mathcal{R}_0$ for the RD and DTS models (Eq.2 vs Eq.3), the DTS surface is marginally different (see Figure S3 for a comparison of these surfaces). The thin dashed vertical line shows the location of the adaptive peak for this surface; the thin dotted vertical line shows the location of the peak for the DTS model. The top left and bottom left panels show a parasite evolving in log(replication rate) and tuning, according to the right panel, translated into virulence and transmission following a power-law tradeoff (top) and virulence and $\mathcal{R}_0$ (bottom). The dashed curve in the bottom left panel shows parasite $\mathcal{R}_0$ for the RD and AD models; the dotted curves shows parasite $\mathcal{R}_0$ for the DTS model. Maximum $\mathcal{R}_0$ is designated by the vertical dashed and dotted lines respectively. All models have the same power-law tradeoff curve (top). A numeric summary of the DTS model for these parameter values is available in Table 2 row 2.

A smaller efficiency scale (δ) value increases the fitness cost of a mismatch between replication rate and tuning (Eq.1), which leads to higher parasite virulence in the short term in all models (Figure 2) (a larger δ decreases the cost, and leads to lower transient parasite virulence: Figure 3). Qualitative differences between the AD and RD/DTS models are similar across the range of values for δ used here (Figures 1-3). In the limit of $\delta \to \infty$, a parasite's fitness landscape collapses to two dimensions and a parasite evolving from lower-than-optimum virulence evolves monotonically to optimum virulence in AD, RD, and DTS models (not pictured).

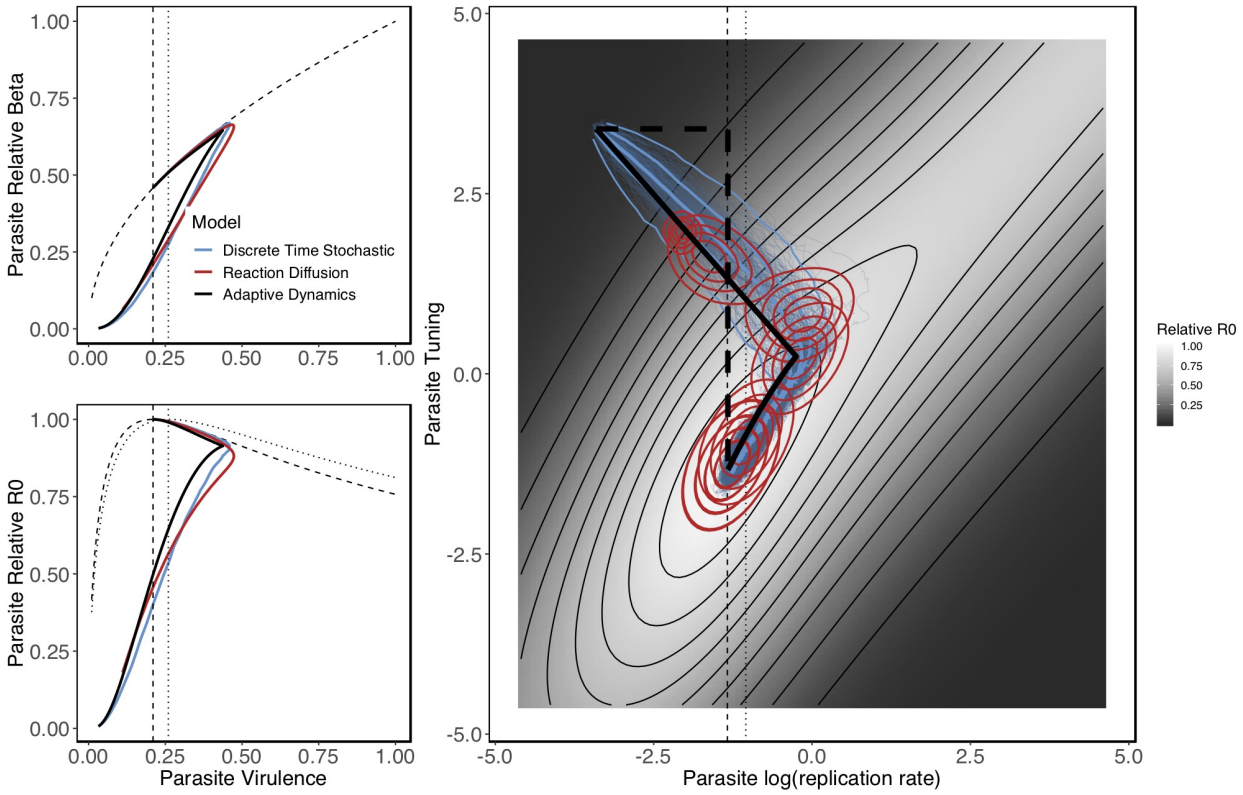

Figure 2: For a complete figure caption see Figure 1. Panels show the evolutionary trajectory of a parasite in the DTS (blue lines), RD (red ellipses), and AD (black solid and dashed lines) models, using the same parameter values as Figure 1, except with $\delta = 10$. A numeric summary of the DTS model for these parameter values is available in Table 2 row 8.

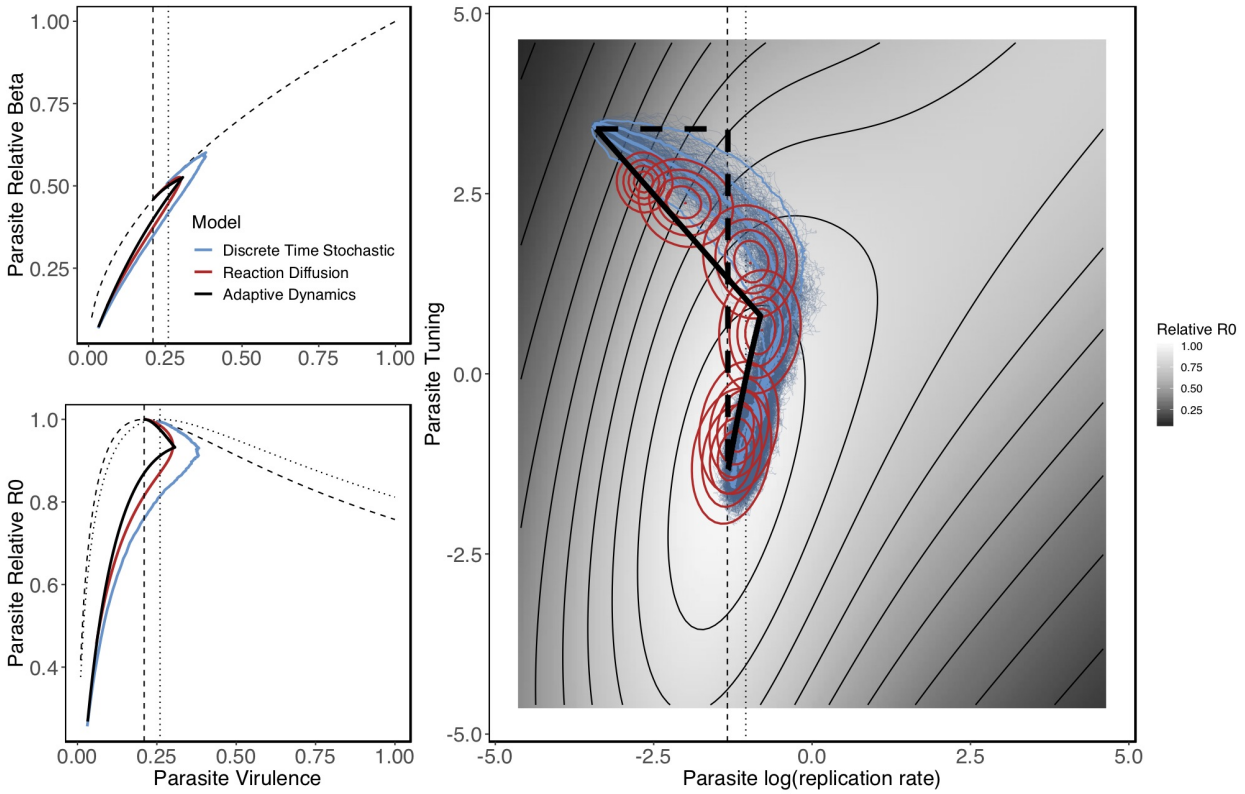

Figure 3: For a complete figure caption see Figure 1. Panels show the evolutionary trajectory of a parasite in the DTS (blue lines), RD (red ellipses), and AD (black solid and dashed lines) models, using the same parameter values as Figure 1, except with $\delta = 50$. A numeric summary of the DTS model for these parameter values is available in Table 2 row 9.

Parameter dependence: focus on the DTS model

For each of our metrics of interest we find that a single parameter is the dominant source of variation; other parameters contribute, but to a lesser degree. Mutational probability (μ) (followed by population size, N) has the strongest effect on the number of circulating strains (results for each of our model outcomes are shown in Table 2). The time it takes for a parasite population to reach its adaptive peak is a decreasing function of both mutation probability and population size, and to a lesser degree mutational standard deviation (σ). Maximum transient virulence is very consistent across parameters except for efficiency scale (δ); decreasing δ results in lower transient virulence (Figures 1-3). The total time a parasite circulates while above optimum virulence is primarily a function of time to optimum virulence, and therefore is the largest for low mutation probability, low mutational standard deviation, and a smaller population size. Finally, departure from optimum virulence is driven almost entirely by population size. A smaller population size leads to a larger stochastic departure from the adaptive peak. For example, in Table 2, last column, row 3, the value of 0.17 means that $\sim$ 95% of the 39 circulating strains (column 1 of "Model Outcomes") had a virulence between 0.20 and 0.33.

All of the results presented in Table 2 can be visualized in Figure 4 except for the number of circulating strains. For parasite evolution from high replication rate and low efficiency see Figure S1; for evolution from low replication rate, high efficiency see Figure S2.

Table 2: Results for each of our model outcome for a subset of the parameter values presented in Table 1. The right half of the table lists results for the parameter values listed on the left. To improve presentation, parameter values are listed sparsely: for a given focal parameter, parameter values for all other parameters are constant. For example, in rows 1-3, parameter values for σ, N, and δ follow row 1. Model outcome columns, in order from left to right, are as follows: 1) Median number of circulating strains over the whole simulation; 2) Time (steps) to first reach optimum virulence, which can be roughly translated to a number of host generations by dividing by 100 (average life span of an uninfected host); 3) Maximum virulence reached during the evolution of the parasite to its adaptive peak; 4) Total time period where the parasite has a higher-than-optimum virulence; 5) Departure from optimum virulence lists the standard deviation in log(replication rate) (virulence on the logit scale) among circulating strains (median among all stochastic simulations) after optimum virulence is reached.

Parameters					Model Outcomes				
Focal Parameter	**Mutation Probability: μ**	**Mutational SD: σ**	**Population Size: N**	**Efficiency Scale: δ**	**Median number of circulating strains:**	**Time to optimum virulence:**	**Maximum transient virulence:**	**Time with greater than optimum virulence:**	**Departure from optimum virulence:**
	0.001	0.05	600	30	3	$> 5\text{x}10^5$	0.41	$4.6\text{x}10^5$	–
μ	0.005				11	$1.1\text{x}10^5$	0.40	$1.0\text{x}10^5$	0.14
	0.025				39	$3.4\text{x}10^4$	0.40	$3.0\text{x}10^4$	0.17
σ	0.005	0.01	600	30	11	$> 5\text{x}10^5$	0.41	$3.1\text{x}10^5$	–
		0.25			11	$1.1\text{x}10^4$	0.40	$1.0\text{x}10^4$	0.21
N	0.005	0.05	200	30	4	$3.6\text{x}10^5$	0.41	$3.3\text{x}10^5$	0.23
			1800		33	$6.7\text{x}10^4$	0.41	$6.0\text{x}10^4$	0.11
δ	0.005	0.05	600	10	11	$1.7\text{x}10^5$	0.47	$1.7\text{x}10^5$	0.09
				50	11	$1.3\text{x}10^5$	0.38	$1.2\text{x}10^5$	0.18

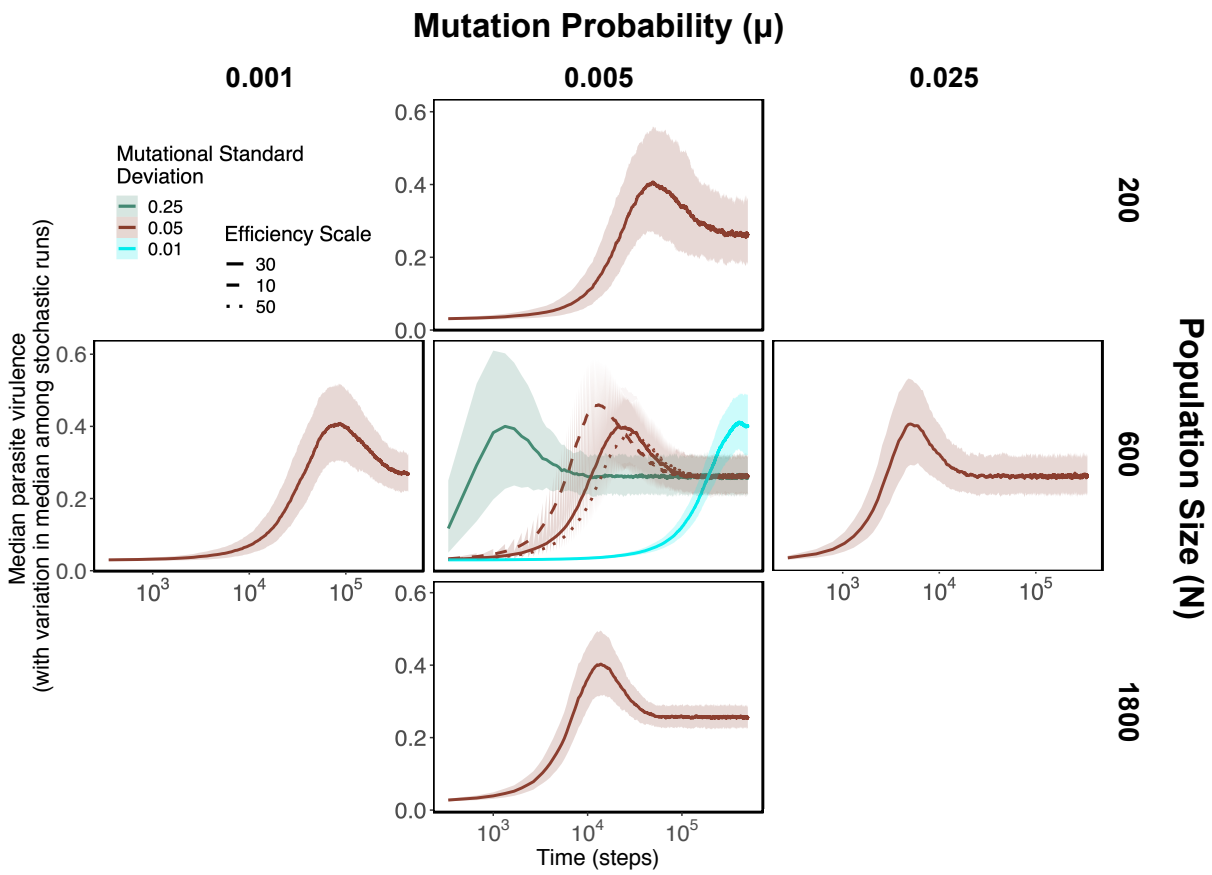

Figure 4: Parasite virulence for all parameter combinations presented in Table 2. Bold lines represent medians across 250 stochastic simulations; shaded envelopes show 95% quantiles. The correspondence between the model outcomes presented in Table 2, from left to right, are as follows: 1) Median number of circulating strains: not pictured; 2) Time to optimum virulence: Time at which median virulence first reaches 0.26; 3) Maximum transient virulence: maximum of each curve; 4) Time with greater than optimum virulence: width of the peak; 5) Departure from optimum virulence: not *directly* pictured. The width of the shaded envelope shows the variation in median virulence among stochastic simulations, not the median among all stochastic simulations for standard deviation in virulence among circulating strains within each simulation, which is what Table 2 shows. However, because all simulations converge on a virulence of 0.26, there is a strong correlation between these two metrics; the width of the shaded envelope after the parasite population reaches a median of 0.26 follows the same *pattern* as the numbers presented in the Table 2, but are not quantitatively identical.

We present an expanded center panel from Figure 4 in Figure 5. This plot shows each of the 250 stochastic simulations for the first three rows of Table 2; the center red curve is the result presented in Figure 1 (only an efficiency scale of 30 is pictured in Figure 5). This plot also shows the median and 95% density for the virulence of the parasite strains in the RD solution, which is similar to the stochastic simulation result. The optimum virulence of 0.21 for the RD model and 0.26 for the DTS model are seen here.

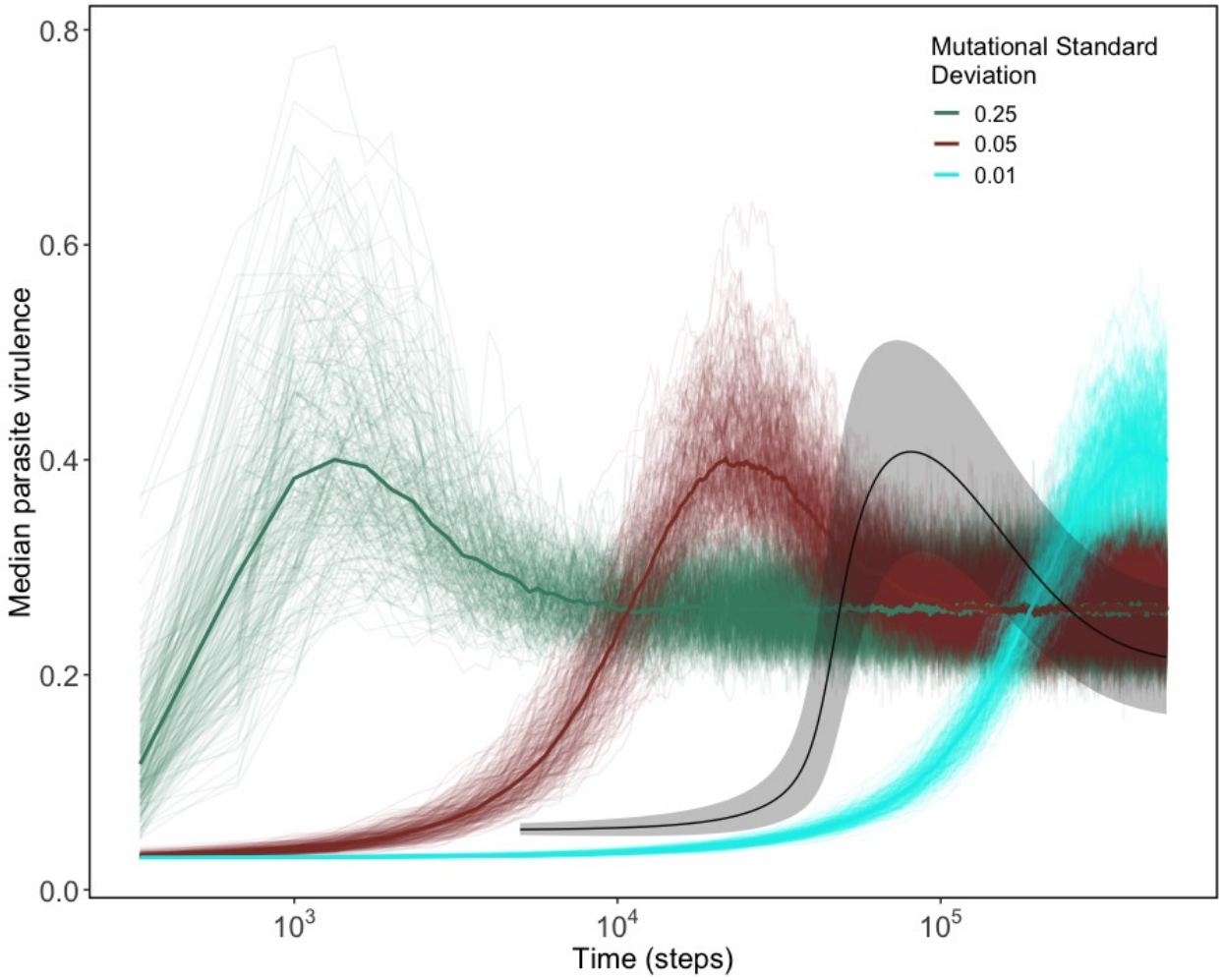

Figure 5: Parasite median virulence and virulence across all stochastic simulations for the center panel of Figure 4 with only $\delta = 30$ shown (these parameter combinations are also shown in Figure 1 and in the first three rows of Table 2). The solid black curve and shaded grey envelope show the median and 95% density for the virulence of the parasite strains in the RD solution.

Discussion

Previous work assuming a tradeoff between virulence and transmission has shown that evolution on ecological timescales, for example during an epidemic, selects for strains with higher instantaneous growth (little r) more strongly than for strains with higher $\mathcal{R}_0$ (Day and Proulx, 2004, Bolker et al., 2010). Eco-evolutionary feedbacks can be seen here in the comparison between AD and RD/DTS solutions in Figure 1 and Figure S2, though our model also provides a new mechanism in addition to eco-evo dynamics for why a pathogen may evolve higher transient virulence: parasite tuning selects for higher virulence in the short term when a parasite evolves from an initial position of low virulence and low efficiency. A larger cost of a mismatch between virulence tuning (which lowers β at a given level of α) leads to very high transient virulence (Figure 2). When scaled to host lifespan in the absence of infection, the transient virulence period pictured in Figure 2 represents 1,600 host lifespans. While this is unrealistic for any vertebrate host, it may be an appropriate time scale for viral parasites of algae (Frickel et al., 2016, 2018) or bacteria Berngruber et al. (2015).

When a parasite evolves from high virulence and low efficiency, another plausible scenario for an emerging pathogen, all three models agree that a parasite evolves monotonically to its optimum. However, the length of time that it takes for a parasite to reach its optimum virulence increases as the fitness impact of tuning increases (e.g. as δ decreases). The evolutionary distance covered by a parasite evolving directly to its optimal virulence reflects evolution only *on* the tradeoff curve; this distance is shown in the horizontal portion of the bold dashed line in Figure S1. Collectively, these results show that in the simple case of a fitness landscape with a single global maximum AD, RD, and DTS models agree on the general trajectory of parasite evolution. With a multi-peaked surface, however, AD solutions may not agree with the RD or DTS solutions (Dieckmann, 2002). For example, when a local maximum exists with a steeper gradient, AD models can show that a parasite resides at this local maximum indefinitely. Diffusion in RD and DTS models allow populations to evolve through fitness valleys, which has been documented empirically (Jain and Krug, 2007, Gokhale et al., 2009, Weissman et al., 2010). This scenario requires more detailed analyses than those presented here.

Despite the similarities between the RD and DTS model solutions, the dynamics we plot reflect slightly different processes. The RD model reveals variation within a population of parasites, while the DTS model solution represents variation among stochastic realizations. While we do not show them here (and only report variation after equilibrium is reached–Table 2 last column), simulation results show that with similar parameter values for σ in the RD and DTS model, population level variation in the RD model is about an order of magnitude larger than in the DTS model. Approximately an order of magnitude smaller value for

σ in the RD model is needed for roughly similar estimates of within-population variation.

If general qualitative patterns are of primary interest, a RD model is likely preferable to a DTS model because it is easier to implement and faster to solve. However, while RD and DTS models (and under many starting conditions the AD model) broadly agree on qualitative dynamics, it is important to keep in mind that these models generate slightly different landscapes shapes for the same numeric values for host recovery and host background death rate. The even simpler AD model used here provides a description of gradient ascent (tracking the steepest path up the landscape), which can provide a description of the final evolutionary state of the system (as long as careful attention is paid to the muli-peakedness of the landscape), but may underestimate transient dynamics. If rare evolutionary events in small populations are of primary interest, only the DTS model can provide an adequate estimate. It is also important to keep in mind however, that with any more complicated of a model than is used here, for example with host evolution in defense traits, the number of differential equations in the RD model becomes unwieldy and solving this system can take longer than solving the DTS model.

For the DTS model specifically, results were as expected. An increase in population size minimized the effects of stochasticity, decreased the time to optimal virulence, and decreased a populations departure from optimum virulence. Both a larger mutation probability and mutational standard deviation decrease time to equilibrium and increase transient departure from equilibrium. The transient virulence peak is determined mostly by the shape of the fitness landscape, with an increase in the cost of a mismatch between replication rate and tuning increasing transient virulence.

Caveats and model extensions

Despite our breadth of coverage of methods for solving our model, we considered only an SIS type epidemiological model, which does not allow for recovered hosts that are permanently immune to infection (e.g. an SIR model), despite lifetime immunity in many wildlife diseases (e.g. WNV, MYXV). We also ignore variation in host recovery across parasite strains, even though tradeoff between host recovery and parasite virulence can also constrain parasite exploitation strategy (Alizon, 2008). It would be interesting to extend our model an SIR model with recovered hosts of incomplete and waning immunity as seen in *Mycoplasma gallisepticum* infection of House Finches (Fleming-Davies et al., 2018). In this system incomplete and waning immunity favors the evolution of more virulent strains which are able to overwhelm a hosts incomplete immunity. However, these evolutionary patterns are not yet fully understood. With a small extension to explicitly consider host defense, our model for secondary trait evolution could help to explain the evolution of virulence in this system.

We also define host-parasite interactions using only a quantitative genetic model, though many host-parasite systems (e.g. flax-flax rust: Islam and Shepherd 1991, human/bird-influenza: Smith et al. 2004, Poullain and Nuismer 2012, *Pseudomonas fluorescens* and its DNA phage: Poullain et al. 2008) can arguably be better understood using a more explicit gene based approach. A matching allele model (MA), for example, assumes that a perfect match between host and parasite genotype is required for infection, while a gene-for-gene (GFG) model relaxes these assumptions slightly, allowing for a decay in infection success as a function of increasing genetic distance (Agrawal and Lively, 2002, Nuismer and Otto, 2005, Thrall et al., 2016). Often these models are usually analyzed in a co-evolutionary framework, and have shown to lead to evolutionary cycles such as Red-Queen dynamics (Decaestecker et al., 2007). In the absence of host evolution, our quantitative genetic model has a clear parallel to a GFG model, with tuning serving as a form of genetic distance, and transmission as a measure of infection success. However, to fully realize the connection between these models it is necessary to extend our model to allow for host evolution.

For the model that we do examine, a few of our assumptions could be explored/relaxed in order to cover a wider range of possible dynamics. First, we assume a very strong form of density dependent birth that greatly reduces the probability of population extinction (e.g. with N = 200 none of populations in the 250 simulations went extinct, though with populations of 100, extinction probability was roughly 10-20%). Given our focus here on a parasites' evolutionary path to their global optimum and the size of the departure from optimum virulence, we chose population sizes to minimize extinction. It will be important to examine the starting conditions that make it unlikely a parasite would escape extinction; these are likely dead end hosts for the parasite. Given our strong assumption for density dependent birth, it is possible that the results we present here are modeling an unrealistic scenario of a pathogen evolving to exploit what is in reality an unsuitable host.

Second, while our assumption of mutational effects on a logistic scale is logical, mutation on one scale and selection on another has the potential to produce very odd dynamics. For example, mutation on the log scale and selection on the linear scale can leads to exponentially increasing fitness, albeit in the absence of tradeoffs. Regardless, it may be fruitful to examine mutation and selection on the same scale in order to understand what portion of the dynamics we see are due simply to our modeling choices.

Finally, as mentioned previously we assume a static, homogeneous host population, which is an unrealistic assumption for any host population. A vast literature exists for a static, heterogeneous host population which shows that host heterogeneity has a large influence on parasite optimum virulence (e.g. (Ebert and Hamilton, 1996, Gandon et al., 2001, Ganusov et al., 2002, Gandon, 2004, Kain et al., 2018)). When host populations evolve with parasite populations, evolutionary cycles and stable polymorphisms can occur (Carval and Ferriere, 2010, Best et al., 2014, Athanasiadou et al., 2015). In order to make our model more realistic it

will be important to extend our model to consider the evolution of host resistance (reducing parasite load) and tolerance (decreasing fitness costs for a given pathogen load) on ecological time scales (for reviews on resistance and tolerance see Schneider and Ayres 2008, Ayres and Schneider 2012, Kutzer and Armitage 2016).

References

Agrawal, A. and C. M. Lively 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research 4*(1), 91–107.

Alizon, S. 2008. Transmission-recovery trade-offs to study parasite evolution. *The American Naturalist 172*(3), E113–E121.

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.

Alizon, S. and Y. Michalakis 2015. Adaptive virulence evolution: the good old fitness-based approach. *Trends in ecology & evolution 30*(5), 248–254.

Alizon, S. and M. van Baalen 2005. Emergence of a convex trade-off between transmission and virulence. *The American Naturalist 165*(6), E155–E167.

Anderson, R. M. and R. May 1982. Coevolution of hosts and parasites. *Parasitology 85*(02), 411–426.

Athanasiadou, S., K. Tolossa, E. Debela, A. Tolera, and J. G. Houdijk 2015. Tolerance and resistance to a nematode challenge are not always mutually exclusive. *International journal for parasitology 45*(4), 277–282.

Ayres, J. S. and D. S. Schneider 2012. Tolerance of infections. *Annual review of immunology 30*, 271–294.

Beasley, D. W., C. T. Davis, H. Guzman, D. L. Vanlandingham, A. P. T. da Rosa, R. E. Parsons, S. Higgs, R. B. Tesh, and A. D. Barrett 2003. Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. *Virology 309*(2), 190–195.

Berngruber, T. W., S. Lion, and S. Gandon 2015. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS Pathog 11*(4), e1004810.

Best, A., A. White, and M. Boots 2014. The coevolutionary implications of host tolerance. *Evolution 68*(5), 1426–1435.

Blanié, S., J. Mortier, M. Delverdier, S. Bertagnoli, and C. Camus-Bouclainville 2009. M148r and m149r are two virulence factors for myxoma virus pathogenesis in the european rabbit. *Veterinary research 40*(1), 1.

Bolker, B. M., A. Nanda, and D. Shah 2010. Transient virulence of emerging pathogens. *Journal of the Royal Society Interface 7*(46), 811–822.

Bull, J. J. and A. S. Lauring 2014. Theory and empiricism in virulence evolution. *PLoS pathogens 10*(10), e1004387.

Carval, D. and R. Ferriere 2010. A unified model for the coevolution of resistance, tolerance, and virulence. *Evolution 64*(10), 2988–3009.

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.

Day, T., S. Alizon, and N. Mideo 2011. Bridging scales in the evolution of infectious disease life histories: theory. *Evolution: International Journal of Organic Evolution 65*(12), 3448–3461.

Day, T. and S. R. Proulx 2004. A general theory for the evolutionary dynamics of virulence. *The American Naturalist 163*(4), E40–E63.

Decaestecker, E., S. Gaba, J. A. Raeymaekers, R. Stoks, L. Van Kerckhoven, D. Ebert, and L. De Meester 2007. Host–parasite 'red queen'dynamics archived in pond sediment. *Nature 450*(7171), 870.

Dieckmann, U. 2002. Adaptive dynamics of pathogen-host interactions.

Duffy, M. A., S. R. Hall, C. E. Cáceres, and A. R. Ives 2009. Rapid evolution, seasonality, and the termination of parasite epidemics. *Ecology 90*(6), 1441–1448.

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

Ebert, D. and W. D. Hamilton 1996. Sex against virulence: the coevolution of parasitic diseases. *Trends in Ecology & Evolution 11*(2), 79–82.

Ewald, P. W. 1983. Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics 14*(1), 465–485.

Fenner, F. and I. Marshall 1957. A comparison of the virulence for European rabbits (*Oryctolagus cuniculus*) of strains of myxoma virus recovered in the field in australia, europe and america. *Journal of Hygiene 55*(02), 149–191.

Fleming-Davies, A. E., P. D. Williams, A. A. Dhondt, A. P. Dobson, W. M. Hochachka, A. E. Leon, D. H. Ley, E. E. Osnas, and D. M. Hawley 2018. Incomplete host immunity favors the evolution of virulence in an emergent pathogen. *Science 359*(6379), 1030–1033.

Fraser, C., T. D. Hollingsworth, R. Chapman, F. de Wolf, and W. P. Hanage 2007. Variation in hiv-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proceedings of the National Academy of Sciences 104*(44), 17441–17446.

Frickel, J., P. G. Feulner, E. Karakoc, and L. Becks 2018. Population size changes and selection drive patterns of parallel evolution in a host–virus system. *Nature communications 9*.

Frickel, J., M. Sieber, and L. Becks 2016. Eco-evolutionary dynamics in a coevolving host–virus system. *Ecology letters 19*(4), 450–459.

Gandon, S. 2004. Evolution of multihost parasites. *Evolution 58*(3), 455–469.

Gandon, S., M. J. Mackinnon, S. Nee, and A. F. Read 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature 414*(6865), 751–756.

Ganusov, V. V., C. T. Bergstrom, and R. Antia 2002. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution 56*(2), 213–223.

Gokhale, C. S., Y. Iwasa, M. A. Nowak, and A. Traulsen 2009. The pace of evolution across fitness valleys. *Journal of Theoretical Biology 259*(3), 613–620.

Haraguchi, Y. and A. Sasaki 2000. The evolution of parasite virulence and transmission rate in a spatially structured population. *Journal of Theoretical Biology 203*(2), 85–96.

Islam, M. and K. Shepherd 1991. Present status of genetics of rust resistance in flax. *Euphytica 55*(3), 255–267.

Jain, K. and J. Krug 2007. Deterministic and stochastic regimes of asexual evolution on rugged fitness landscapes. *Genetics 175*(3), 1275–1288.

Kain, M. P., I. M. Cattadori, and B. M. Bolker 2018. The evolutionary response of virulence to host heterogeneity: a general model with application to myxomatosis in rabbits co-infected with intestinal helminths. *Evolutionary Ecology Research 19*(3), 257–278.

Keeling, M. J. and P. Rohani 2008. *Modeling infectious diseases in humans and animals*. Princeton University Press.

Kutzer, M. A. and S. A. Armitage 2016. Maximising fitness in the face of parasites: a review of host tolerance. *Zoology 119*(4), 281–289.

Laine, A.-L. and A. Tellier 2008. Heterogeneous selection promotes maintenance of polymorphism in host–parasite interactions. *Oikos 117*(9), 1281–1288.

Lion, S. 2018. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *The American Naturalist 191*(1), 21–44.

Lion, S. and S. Gandon 2015. Evolution of spatially structured host–parasite interactions. *Journal of evolutionary biology 28*(1), 10–28.

Lipsitch, M., E. A. Herre, and M. A. Nowak 1995. Host population structure and the evolution of virulence: a "law of diminishing returns". *Evolution 49*(4), 743–748.

Mideo, N., W. A. Nelson, S. E. Reece, A. S. Bell, A. F. Read, and T. Day 2011. Bridging scales in the evolution of infectious disease life histories: application. *Evolution: International Journal of Organic Evolution 65*(11), 3298–3310.

Nuismer, S. L. and S. P. Otto 2005. Host–parasite interactions and the evolution of gene expression. *PLoS biology 3*(7), e203.

Papkou, A., C. S. Gokhale, A. Traulsen, and H. Schulenburg 2016. Host–parasite coevolution: why changing population size matters. *Zoology 119*(4), 330–338.

Parsons, T. L., A. Lambert, T. Day, and S. Gandon 2018. Pathogen evolution in finite populations: slow and steady spreads the best. *Journal of The Royal Society Interface 15*(147), 20180135.

Poullain, V., S. Gandon, M. A. Brockhurst, A. Buckling, and M. E. Hochberg 2008. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution: International Journal of Organic Evolution 62*(1), 1–11.

Poullain, V. and S. L. Nuismer 2012. Infection genetics and the likelihood of host shifts in coevolving host-parasite interactions. *The American Naturalist 180*(5), 618–628.

Pugliese, A. 2002. On the evolutionary coexistence of parasite strains. *Mathematical biosciences 177*, 355–375.

R Core Team 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schneider, D. S. and J. S. Ayres 2008. Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. *Nature Reviews Immunology 8*(11), 889.

Smith, D. J., A. S. Lapedes, J. C. de Jong, T. M. Bestebroer, G. F. Rimmelzwaan, A. D. Osterhaus, and R. A. Fouchier 2004. Mapping the antigenic and genetic evolution of influenza virus. *Science 305*(5682), 371–376.

Soetaert, K. and F. Meysman 2012. Reactive transport in aquatic ecosystems: Rapid model prototyping in the open source software r. *Environmental Modelling Software 32*, 49–60.

Soetaert, K., T. Petzoldt, and R. W. Setzer 2010. Solving differential equations in r: Package desolve. *Journal of Statistical Software 33*(9), 1–25.

Thrall, P. H., L. G. Barrett, P. N. Dodds, and J. J. Burdon 2016. Epidemiological and evolutionary outcomes in gene-for-gene and matching allele models. *Frontiers in plant science 6*, 1084.

Weissman, D. B., M. W. Feldman, and D. S. Fisher 2010. The rate of fitness-valley crossing in sexual populations. *Genetics 186*(4), 1389–1410.

Chapter 6: Concluding Remarks

Principal Findings and Contribution to the Field

In Chapters 2 and 3 I show that when the collective data for WNV is used in an epidemiological model, little clarity remains in our understanding of WNV transmission. Though I am able to show that temperature has a large impact on WNV transmission, I find that there is insufficient evidence to describe why the WN02 lineage is currently the dominant lineage in North America, and that our ability to estimate WNV transmission potential as a function of variation in bird communities across space and time is poor. Bird community composition can, in theory, have a large impact on WNV $\mathcal{R}_0$ because of known differences in the competence of different bird species; however, large uncertainty in the response to infection of many bird species makes it difficult to say how differences among diverse bird communities contribute to WNV transmission. When using median estimates for each bird species' response, for example, I find that some Texas bird counties in the early spring and late fall can support a WNV epidemic in spite of a temperature that would be unable to support an epidemic in the average bird community. However, because of large uncertainty in these species' responses, appropriately using the full distribution for birds' responses homogenizes bird species, reducing the influence of variation in bird communities across space and time on WNV transmission. Despite this uncertainty, I find that WNV transmission potential declines with increasing bird species richness, and that Northern Cardinals specifically, and species in the family Corvidae more broadly, are the most important species for WNV amplification, while dove species (family Columbidae) are the most important species for dampening WNV transmission. Together, these species can help identify communities where epidemics may be more or less common.

In Chapter 4 I show that a small extension to tradeoff theory can help predict the evolution of virulence, even while retaining a simple model designed to be parameterized with empirical data obtainable form a small number of experiments. Using this model I show that heterogeneity among rabbits generated from secondary infection with gastrointestinal helminth parasites is an unlikely explanation for differences in MYXV virulence between Scotland and Australia, but could help explain spatial variation in MYXV virulence across Australia.

Finally, in Chapter 5, I describe an additional mechanism apart from eco-evolutionary feedbacks that could help explain transient patterns in parasite virulence. This model is primarily conceptual, and designed in part as a thought piece for future work exploring more deeply how parasites evolve to the virulence-transmission evolutionary frontier. While this extension may have scientific merit, I believe this chapter may be better suited for a chapter on disease modeling in an ecological modeling textbook. As an entering PhD student, I would have found the exploration of the qualitative dynamics of two deterministic models and a discrete time stochastic model on a problem in disease ecology a useful starting point for my studies.

Future Directions

Empirical work

My work is primarily stymied because of insufficient data. Though my WNV work was completed from an ecological viewpoint, it has the potential to influence human health management decisions. However, this work would only make an impact given both more accurate and precise $\mathcal{R}_0$ estimates across time and space. To improve $\mathcal{R}_0$ estimates variation in mosquito density and species composition among ecoregions and across seasons are likely the most important missing data needed to predict WNV

$\mathcal{R}_0$ accurately. Secondarily, more infection experiments of common, but currently underrepresented families of bird species, such as many non-passerine families, are needed. Despite the simplicity of my model for MYXV evolution, I still had to make an educated guess for the shape of the relationship between worm burden and MYXV infection intensity because of a lack of data. Finally, while the model I present in Chapter 5 was conceptual, a single quantitative estimate for the evolutionary history of the replication rate of a parasite in a new host would have helped me to parameterize my model. Ultimately, these missing data diluted my message and decreased the impact of this work.

The importance of data is widely understood in the field of disease ecology as a major roadblock in advancing our understanding (Alizon and Michalakis, 2015; Cressler et al., 2016). Not only would the utility of my work increase with more data, but more data would also have opened up additional modeling avenues that were currently closed to me. For example, I was forced to abandon a detailed within-host model for MYXV dissemination in rabbits because of a lack of data. I believe one of the biggest contributions I could have made to the field would have been to generate new data. Though I did not present it here, I designed an experiment (including full laboratory protocols, detailed treatments, and power analyses) for the evolution of *Pasteuria ramosa* mortality virulence and castration in populations of *Daphnia magna* constrained by space. While I was unable to run this experiment during my PhD work, I hope to play a part in a related experiment in the future.

Host evolution

A shortcoming of my models is the lack of attention paid to the host component of antagonistic host-parasite interactions. While I consider host variation, I assume that this variation is static despite much empirical evidence for evolving host populations

(European rabbits Fenner and Marshall 1957, *Daphnia* Ebert et al. 1998, and *C. elegans* (Irazoqui et al., 2010) to name but a few). The literature contains a large theoretical and mathematical foundation for me to draw upon to extend my models, including models for the evolution of host resistance (reducing pathogen load) (Boots and Haraguchi, 1999; Gandon et al., 2002; Laine, 2005; Carval and Ferriere, 2010) and tolerance (reducing the negative fitness impact of infection) (Roy and Kirchner, 2000; Best et al., 2010, 2014). Although these results suggest a complicated biological picture, it is apparent that resistance and tolerance are both viable outcomes under a variety of scenarios. However, the majority of evidence of host defense evolution in both natural and laboratory systems involve host resistance evolution, for example, European rabbits to the myxoma virus (MYXV) (Kerr et al., 2015), House finches to mycoplasma (Bonneaud et al., 2012), *E. coli* to phage T4 (Bohannan and Lenski, 2000), and algae to *phycodnaviridae* dsDNA family viruses ((Frickel et al., 2016, 2018). Evidence for the evolution of tolerance is comparatively rare (I am only aware of empirical or observational evidence for tolerance evolution in blue crabs to *Mytilicola intestinalis* Feis et al. 2016).

It remains an open question why the evolution of resistance appears to be much more common than the evolution of tolerance. It is possible that we simply haven't been looking for tolerance for as long, that it may be harder to detect, or that standing genetic variation for tolerance is lower than for virulence. However, it may be that it is more difficult for hosts to evolve tolerance, potentially because tolerance is more costly or less beneficial than resistance. Soares et al. (2017) suggest that whether tolerance or resistance is likely to evolve will be heavily influenced by a parasite's traits. The number of intricate immune mechanisms Soares et al. (2017) propose could be factors would benefit from some reductionist modeling with the goal of illuminating broad classes of parasite strategies that favor host evolution of either resistance or tolerance.

By extending the model I present in Chapter 5 to consider both tolerance and resistance evolution simultaneously, I would be able to begin to address these questions by examining resistance and tolerance evolution as a function of the functional form of the interactions between host and parasite. A powerful modelling approach to begin to tease apart evolutionary change and empirical sampling would be to pair a stochastic simulation model with a power analysis to determine how difficult it is to detect a given evolved change in a natural population.

Spillover and the Dilution Effect

Currently, support for the dilution effect is unequivocally mixed (Salkeld et al., 2013; Civitello et al., 2015), with disagreement over its ubiquity leading to an "acerbic debate" (McCallum, 2015) that has plagued the literature in recent years. While a series of papers since 2013 on *when* the dilution effect occurs has helped this debate (e.g., Joseph et al., 2013; Mihaljevic et al., 2014; Dobson and Auld, 2016; Levi et al., 2016), detailed analysis of the mechanisms driving *how* the dilution effect operates remain rare. I found that WNV $\mathcal{R}_0$ decreases as a function of increasing species richness, which supports the dilution effect hypothesis (Schmidt and Ostfeld, 2001; Keesing et al., 2006) only in the broadest sense. The most immediate and likely most impactful extension of my WNV work would be to use my mechanistic WNV model to estimate human infection risk in well sampled bird communities to begin to unravel the mechanistic basis for human disease risk. While data does not exist to test the dilution effect on a large scale, this would still be a significant contribution. To predict human infection cases for WNV my approach would be to extend my model to a finer spatial scale, and use a Who Acquires Infection From Whom (WAIFW) matrix approach (Dobson and Foufopoulos, 2001) to explicitly include humans as an additional species in the overall community.

Bibliography

Alizon, S. and Y. Michalakis 2015. Adaptive virulence evolution: the good old fitness-based approach. *Trends in ecology & evolution 30*(5), 248–254.

Best, A., A. White, and M. Boots 2010. Resistance is futile but tolerance can explain why parasites do not always castrate their hosts. *Evolution: International Journal of Organic Evolution 64*(2), 348–357.

Best, A., A. White, and M. Boots 2014. The coevolutionary implications of host tolerance. *Evolution 68*(5), 1426–1435.

Bohannan, B. J. and R. E. Lenski 2000. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecology letters 3*(4), 362–377.

Bonneaud, C., S. L. Balenger, J. Zhang, S. V. Edwards, and G. E. Hill 2012. Innate immunity and the evolution of resistance to an emerging infectious disease in a wild bird. *Molecular ecology 21*(11), 2628–2639.

Boots, M. and Y. Haraguchi 1999. The evolution of costly resistance in host-parasite systems. *The american naturalist 153*(4), 359–370.

Carval, D. and R. Ferriere 2010. A unified model for the coevolution of resistance, tolerance, and virulence. *Evolution 64*(10), 2988–3009.

Civitello, D. J., J. Cohen, H. Fatima, N. T. Halstead, J. Liriano, T. A. McMahon, C. N. Ortega, E. L. Sauer, T. Sehgal, S. Young, et al. 2015. Biodiversity inhibits parasites: broad evidence for the dilution effect. *Proceedings of the National Academy of Sciences 112*(28), 8667–8671.

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.

Dobson, A. and J. Foufopoulos 2001. Emerging infectious pathogens of wildlife. *Philosophical Transactions of the Royal Society B: Biological Sciences 356*(1411), 1001–1012.

Dobson, A. D. and S. K. Auld 2016. Epidemiological implications of host biodiversity and vector biology: key insights from simple models. *The American Naturalist 187*(4), 405–422.

Ebert, D., C. D. Zschokke-Rohringer, and H. J. Carius 1998. Within–and between–population variation for resistance of daphnia magna to the bacterial endoparasite pasteuria ramosa. *Proceedings of the Royal Society of London. Series B: Biological Sciences 265*(1410), 2127–2134.

Feis, M. E., M. A. Goedknegt, D. W. Thieltges, C. Buschbaum, and K. M. Wegner 2016. Biological invasions and host–parasite coevolution: different coevolutionary trajectories along separate parasite invasion fronts. *Zoology 119*(4), 366–374.

Fenner, F. and I. Marshall 1957. A comparison of the virulence for european rabbits (oryctolagus cuniculus) of strains of myxoma virus recovered in the field in australia, europe and america. *Journal of Hygiene 55*(02), 149–191.

Frickel, J., P. G. Feulner, E. Karakoc, and L. Becks 2018. Population size changes and selection drive patterns of parallel evolution in a host–virus system. *Nature communications 9*.

Frickel, J., M. Sieber, and L. Becks 2016. Eco-evolutionary dynamics in a coevolving host–virus system. *Ecology letters 19*(4), 450–459.

Gandon, S., M. van Baalen, and V. A. Jansen 2002. The evolution of parasite virulence, superinfection, and host resistance. *The American Naturalist 159*(6), 658–669.

Irazoqui, J. E., J. M. Urbach, and F. M. Ausubel 2010. Evolution of host innate defence: insights from caenorhabditis elegans and primitive invertebrates. *Nature Reviews Immunology 10*(1), 47.

Joseph, M. B., J. R. Mihaljevic, S. A. Orlofske, and S. H. Paull 2013. Does life history mediate changing disease risk when communities disassemble? *Ecology letters 16*(11), 1405–1412.

Keesing, F., R. D. Holt, and R. S. Ostfeld 2006. Effects of species diversity on disease risk. *Ecology letters 9*(4), 485–498.

Kerr, P. J., J. Liu, I. Cattadori, E. Ghedin, A. F. Read, and E. C. Holmes 2015. Myxoma virus and the leporipoxviruses: an evolutionary paradigm. *Viruses 7*(3), 1020–1061.

Laine, A.-L. 2005. Evolution of host resistance: looking for coevolutionary hotspots at small spatial scales. *Proceedings of the Royal Society B: Biological Sciences 273*(1584), 267–273.

Levi, T., F. Keesing, R. D. Holt, M. Barfield, and R. S. Ostfeld 2016. Quantifying dilution and amplification in a community of hosts for tick-borne pathogens. *Ecological applications 26*(2), 484–498.

McCallum, H. I. 2015. Lose biodiversity, gain disease. *Proceedings of the National Academy of Sciences 112*(28), 8523–8524.

Mihaljevic, J. R., M. B. Joseph, S. A. Orlofske, and S. H. Paull 2014. The scaling of host density with richness affects the direction, shape, and detectability of diversity-disease relationships. *PloS one 9*(5), e97812.

Roy, B. and J. Kirchner 2000. Evolutionary dynamics of pathogen resistance and tolerance. *Evolution 54*(1), 51–63.

Salkeld, D. J., K. A. Padgett, and J. H. Jones 2013. A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. *Ecology letters 16*(5), 679–686.

Schmidt, K. A. and R. S. Ostfeld 2001. Biodiversity and the dilution effect in disease ecology. *Ecology 82*(3), 609–619.

Soares, M. P., L. Teixeira, and L. F. Moita 2017. Disease tolerance and immunity in host protection against infection. *Nature Reviews Immunology 17*(2), 83.

Appendix : Additional figures for Chapter 2

Supplemental figures, originally presented as a supplemental component of the published work presented as Chapter 2.

Supplemental Figures for: Can existing data on WNV infection in birds and mosquitos explain strain replacement?

Morgan P. Kain[1†], Benjamin M. Bolker[1,2]

[1]Department of Biology, McMaster University, 1280 Main St. West, Hamilton, ON, Canada, L8S 4K1

[2]Department of Mathematics and Statistics, McMaster University, 1280 Main St. West, Hamilton, ON, Canada, L8S 4L8

[†]Correspondence author.
LSB-215
1280 Main St. West
Hamilton, ON
L8S 4K1
E-mail: kainm@mcmaster.ca

Table of Contents

(1) Model for mosquito titer to mosquito transmission

Due to a lack of publications that explicitly measured transmission from mosquitos after X days following infection with WNV, we fit a model to transmission with titer as a predictor used data from (Moudy et al., 2007) to obtain data from papers that only measured titer temporally and not transmission. To do so we fit a parameterization of a logistic cdf to transmission using non-linear least squares with the nlmrt package (Nash, 2012).

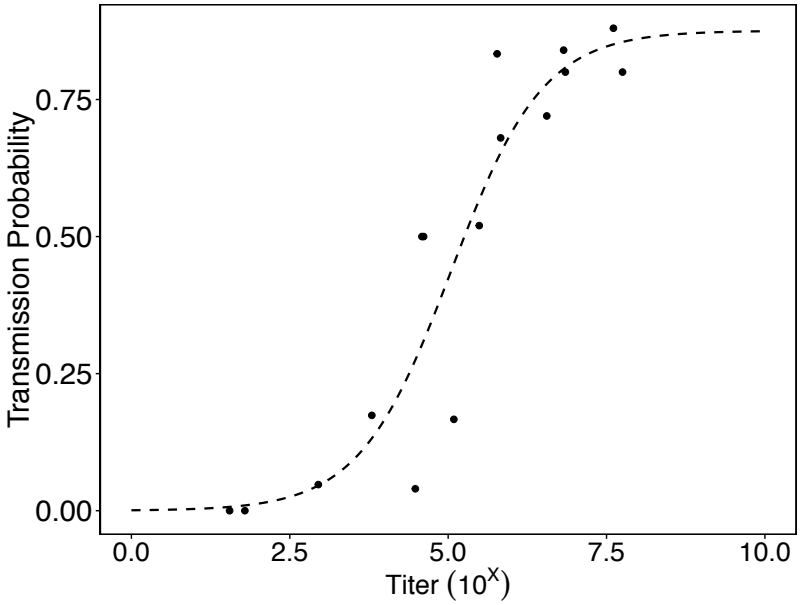

Figure S1: Relationship between Titer and Transmission from Moudy et al. 2007.

(2) Model for mosquito survival

We used data from (Andreadis et al., 2014) to fit a logistic model to mosqutio longevity. In our data analysis we used median survival at each temperature to calculate R_0. We appreciate that this study took place in Greece, far from the transmission events we are interested in, but it includes the most complete data on temperature dependent *Culex* survival that we could find.

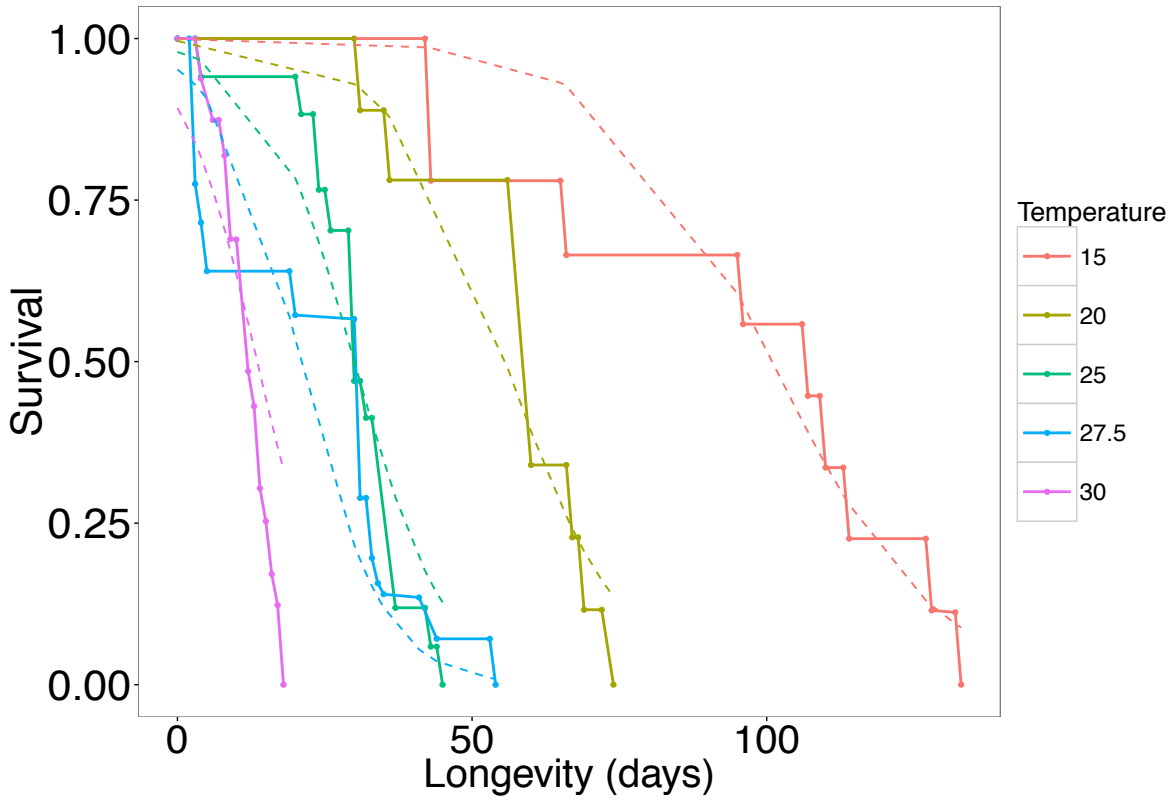

Figure S2: Relationship between Mosqutio Survival Probability and Days. Colored solid lines are extracted data from Andreadis et al. 2014 Figure 1C. Colored dashed lines are model estimates.

(3) Titer and Survival for "Other" Birds

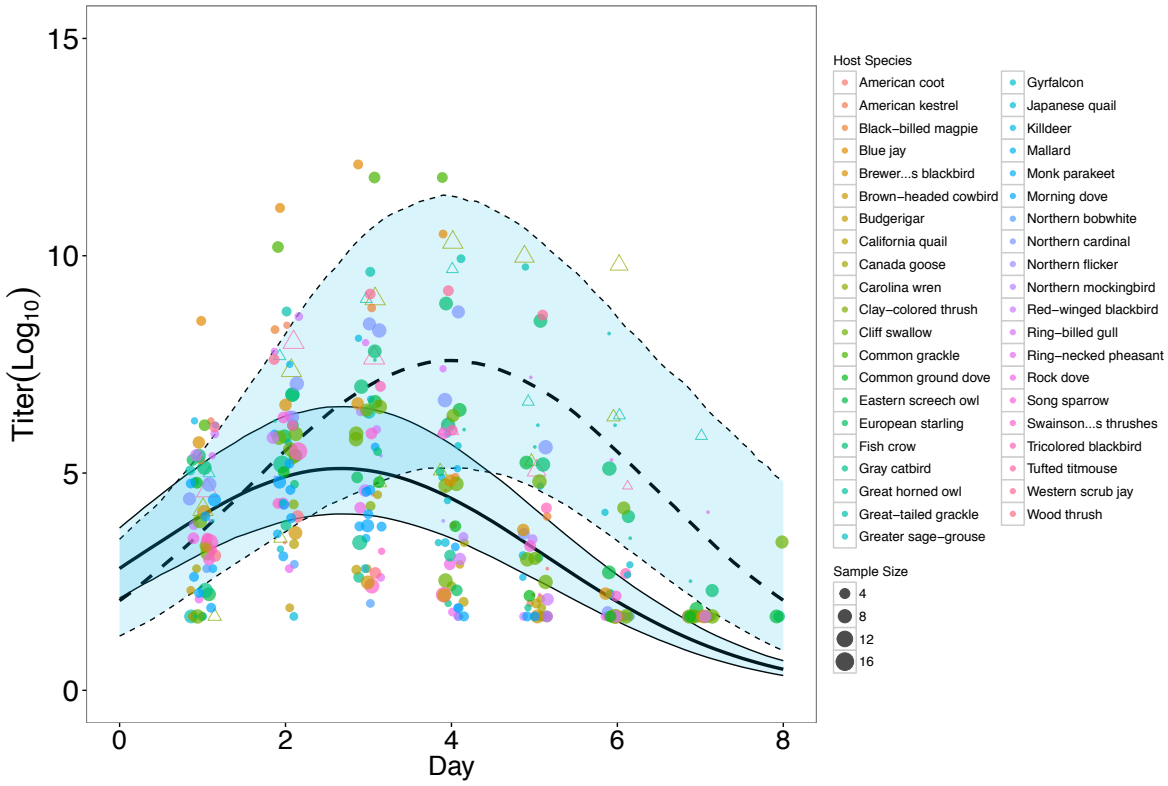

Figure S3.1: Titer Profiles for all other birds.

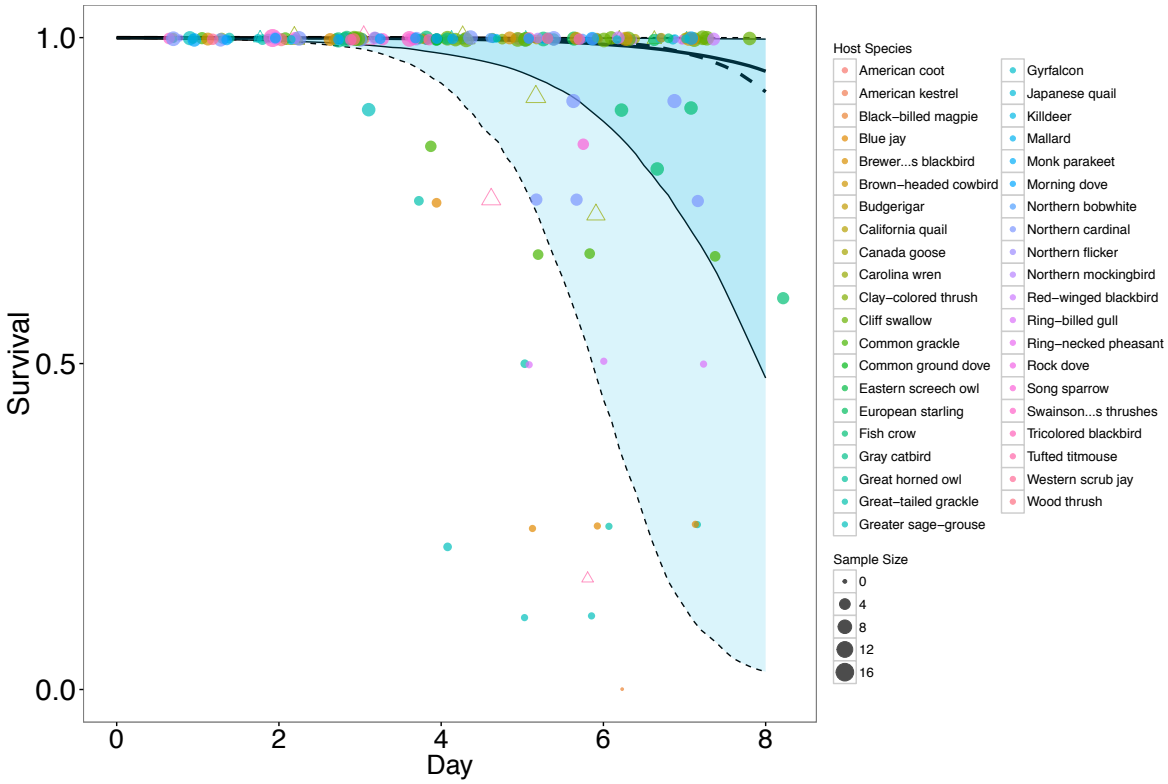

Figure S3.2: Survival for all other birds.

(4) Mosquito to bird transmission adjustments

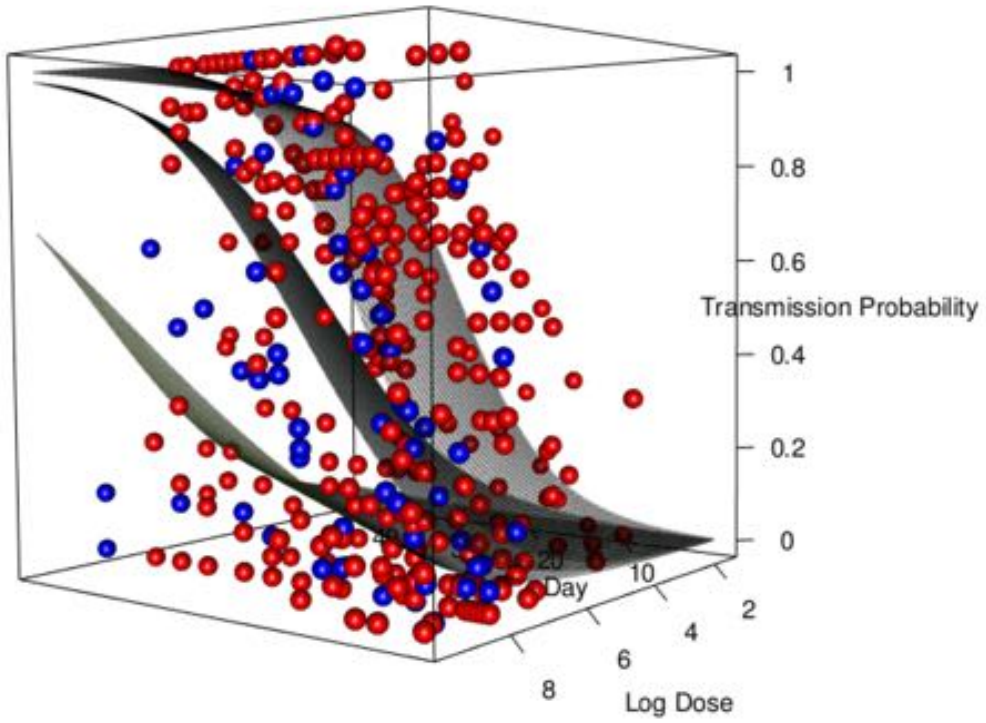

Figure S4.1: 3d figure of Mosquito to Bird model fit to raw data for *NY99 with JEV data*. Red points are NY99 data, blue points are WN02 data. Surfaces are predicted probabilities of transmission from an infected mosquito to a naive bird (Z-axis) *for NY99 with JEV data*. X-axis is days from 1-40, y-axis is Log Dose from 2 to 8. Light green surface is fitted surface at 16 degrees Celcius, darker green surface is 20 degrees Celcius, and black surface is 24 degrees Celcius.

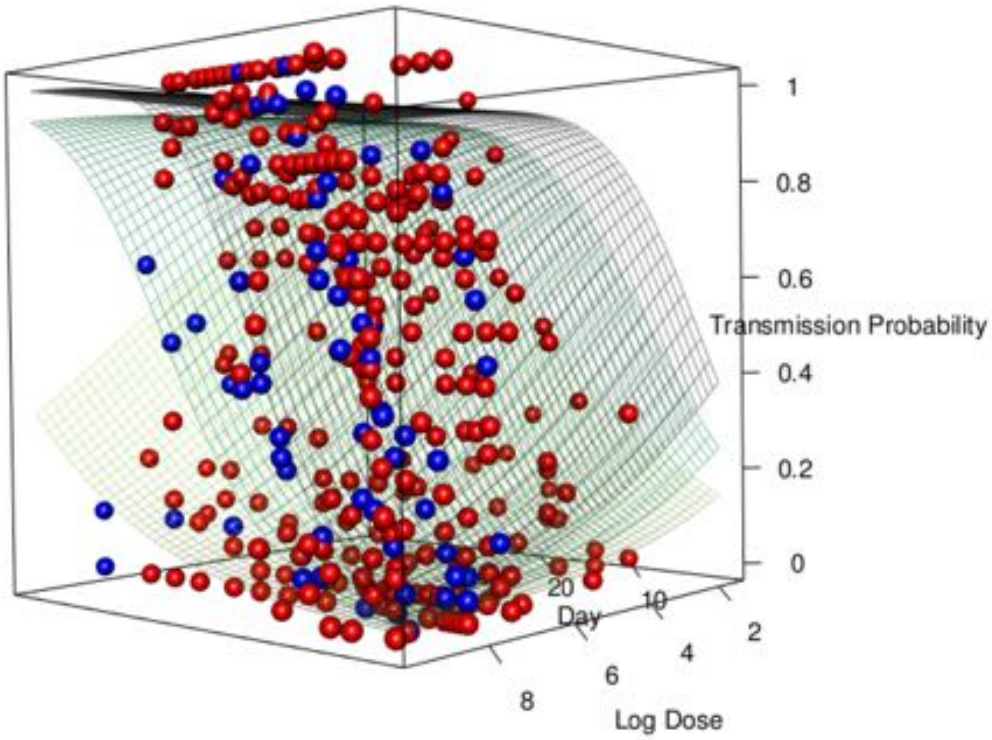

Figure S4.2: 3d figure of Mosquito to Bird model fit to raw data for *WN02 with JEV data*. Red points are NY99 data, blue points are WN02 data. Surfaces are predicted probabilities of transmission from an infected mosquito to a naive bird (Z-axis) *for WN02 with JEV data*. X-axis is days from 1-40, y-axis is Log Dose from 2 to 8. Light green surface is fitted surface at 16 degrees Celcius, darker green surface is 20 degrees Celcius, and black surface is 24 degrees Celcius.

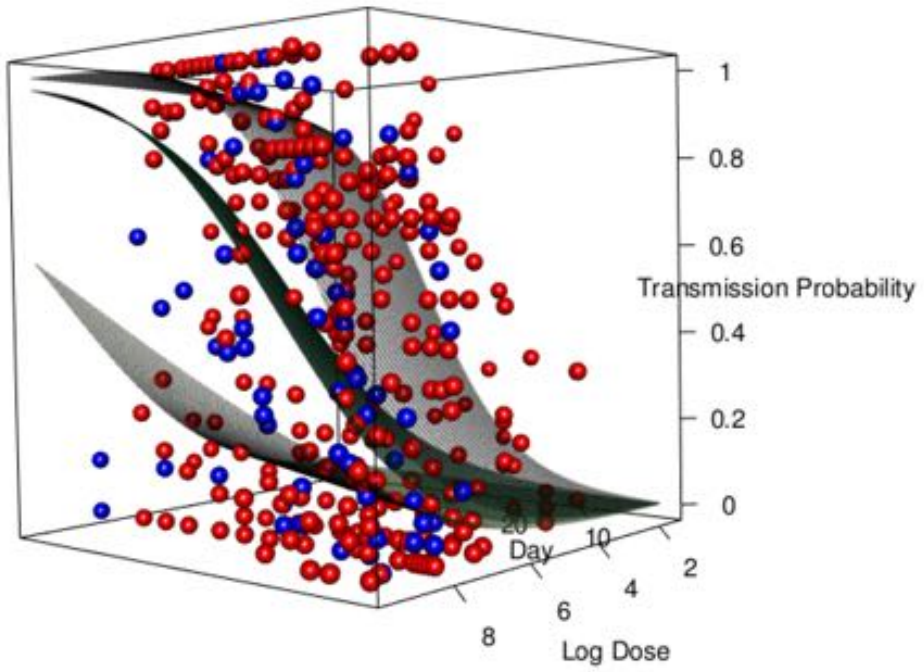

Figure S4.3: 3d figure of Mosquito to Bird model fit to raw data for *NY99 without JEV data*. Red points are NY99 data, blue points are WN02 data. Surfaces are predicted probabilities of transmission from an infected mosquito to a naive bird (Z-axis) *for NY99 without JEV data*. X-axis is days from 1-40, y-axis is Log Dose from 2 to 8. Light green surface is fitted surface at 16 degrees Celcius, darker green surface is 20 degrees Celcius, and black surface is 24 degrees Celcius.

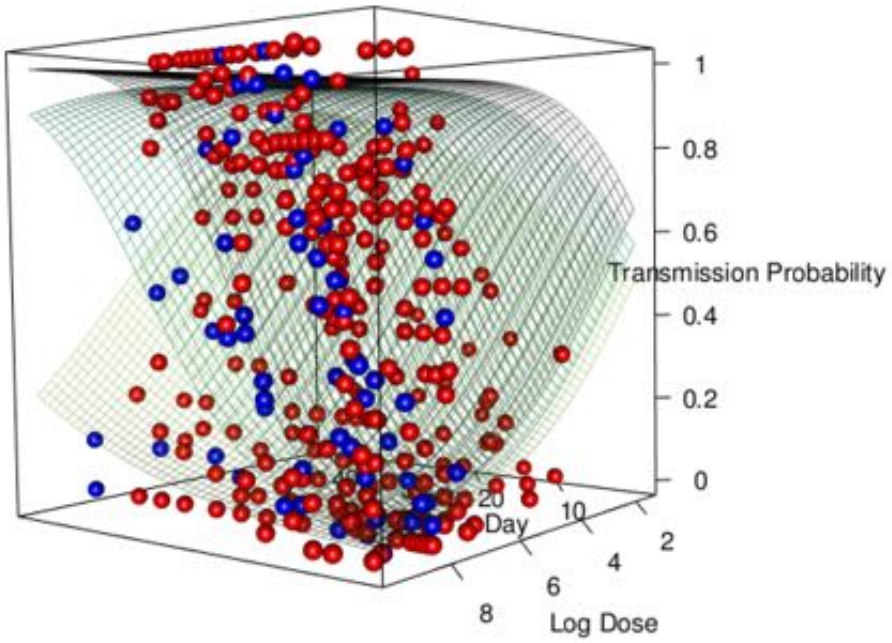

Figure S4.4: 3d figure of Mosquito to Bird model fit to raw data for *WN02 without JEV data*. Red points are NY99 data, blue points are WN02 data. Surfaces are predicted probabilities of transmission from an infected mosquito to a naive bird (Z-axis) *for WN02 without JEV data*. X-axis is days from 1-40, y-axis is Log Dose from 2 to 8. Light green surface is fitted surface at 16 degrees Celcius, darker green surface is 20 degrees Celcius, and black surface is 24 degrees Celcius.

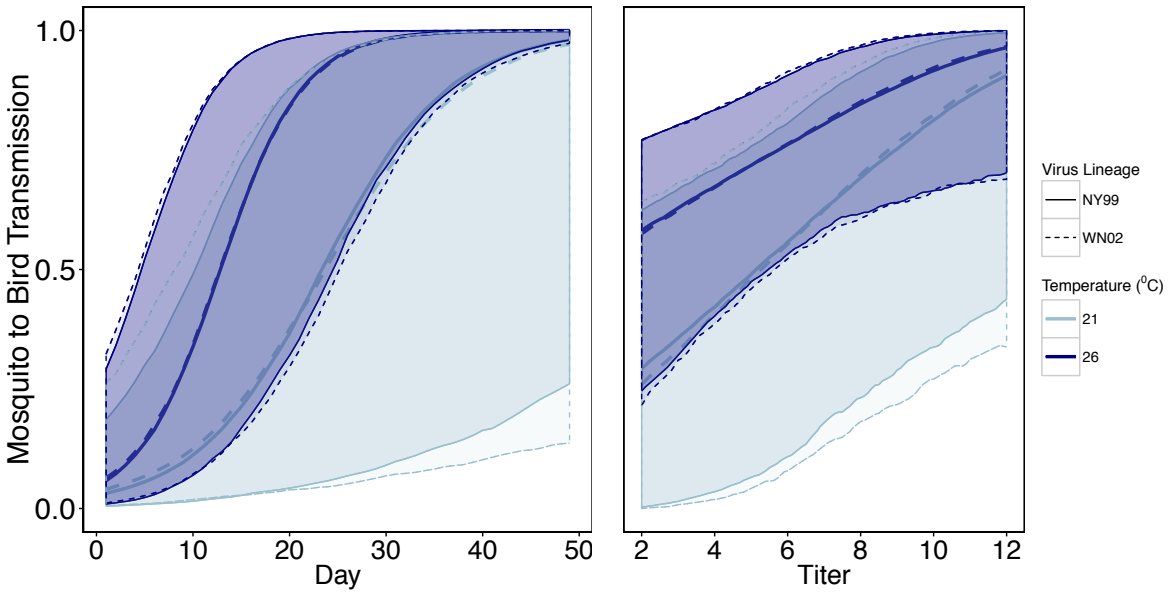

Figure S4.5: Figure 4a from the main text without encorporating mosquito survival

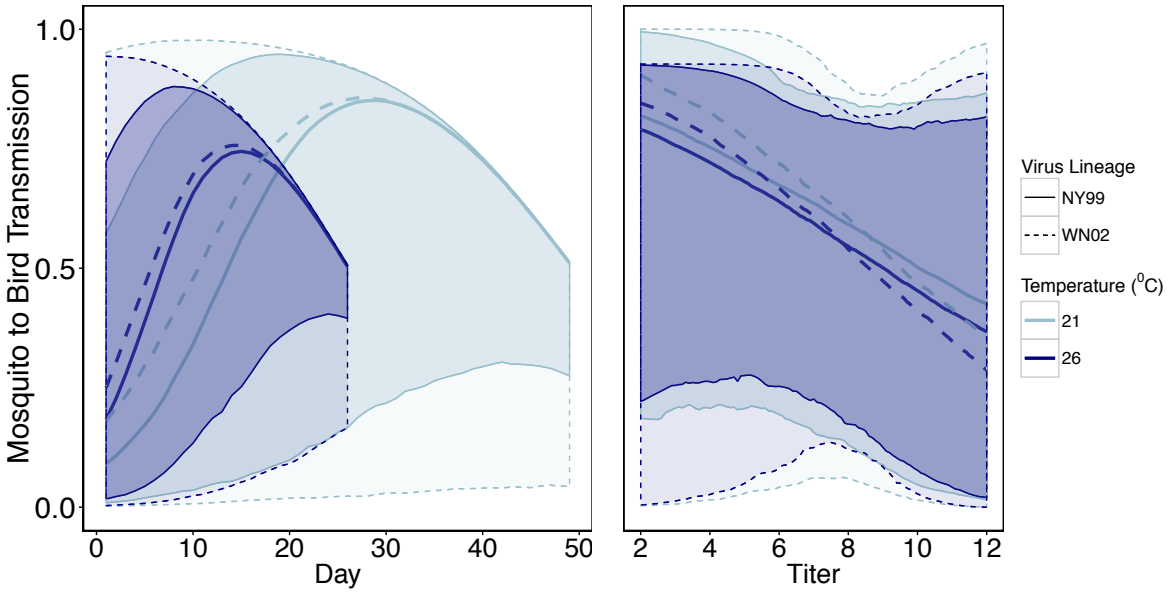

Figure S4.6: Figure 4a, b from the main text without JEV data. See Section 5 for coefficient plots for mosquito to bird transmission without JEV data.

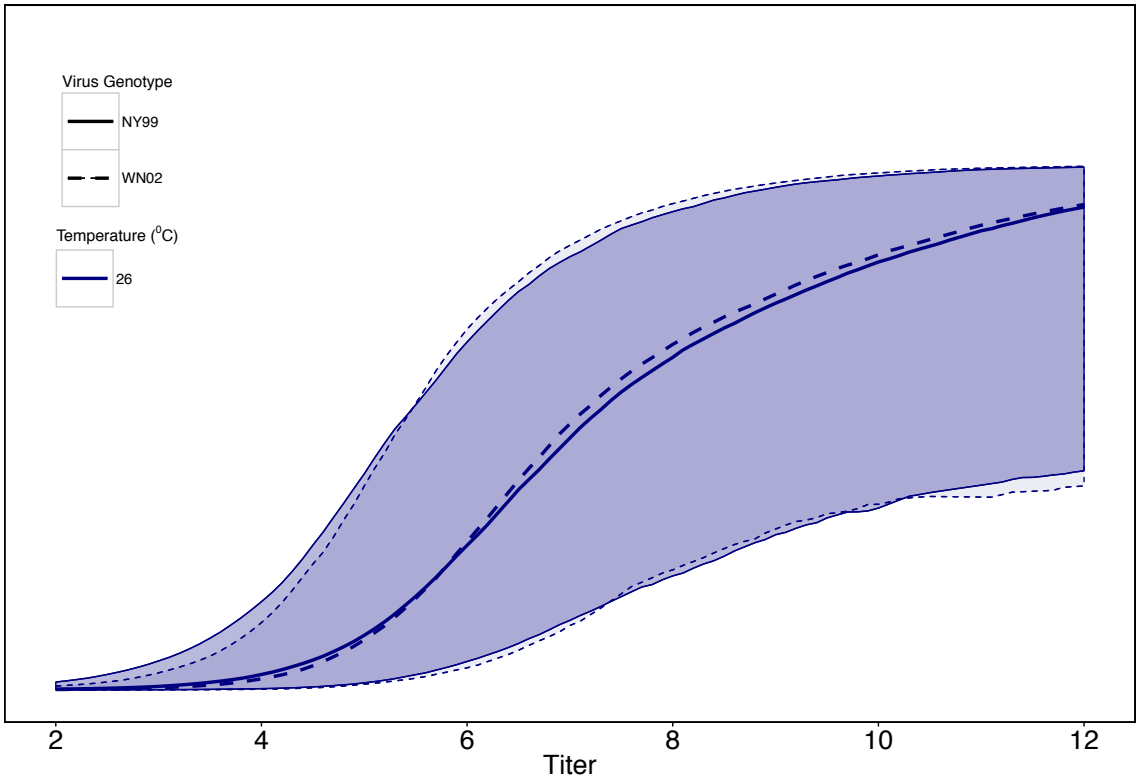

Figure S4.7: Vector-Competence for NY99 and WN02. Vector Competence at 26 Degrees Celcius, generated by combining Bird-to-Mosquito and Mosquito-to-Bird transmission (conditioning Mosquito-to-Bird transmission on all mosquitos that fed on an infected blood sample).

(5) R_0 Caculations without Japanese Encephalitis Virus (JEV) data

Here we present the analysis presented in the primary manuscript removing all "prior information" on mosquito transmission at lower titers using transmission of the closely realted JEV virus. Here we present parameter estimates for the Mosquito to Bird transmission model and R_0 estimates with and without the JEV transmission data.

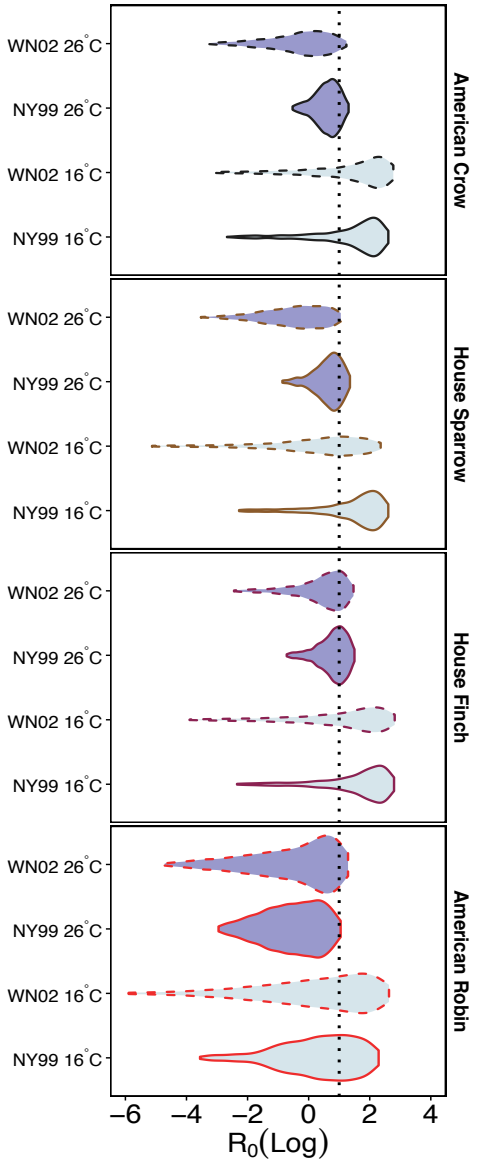

Figure S5: R_0*without* JEV data. Panels correspond to Figure 5 in the main text.

Community R_0s without JEV

In the Chicago, IL community *with* "other" birds median R_0 for NY99 was greater than WN02, but credible intervals overlap:

NY99 at 16^0C, Median: 0.83, CI: 0.02-3.06;

WN02 at 16^0C, Median: 1.54, CI: 0.01-4.93;

NY99 at 26^0C, Median: 0.30, CI: 0.07-0.88;

WN02 at 26^0C, Median: 0.55, CI: 0.06-1.33.

In the Chicago, IL community *without* "other" birds median R_0 WN02 was greater than NY99, but credible intervals also overlap:

NY99 at 16^0C, Median: 1.18, CI: 0.03-5.70;

WN02 at 16^0C, Median: 1.06, CI: 0.002-7.38;

NY99 at 26^0C, Median: 0.42, CI: 0.08-1.63;

WN02 at 26^0C, Median: 0.49, CI: 0.02-1.96.

(6) Coefficient plots for all models

In this section we include all coefficient plots for the fixed effects, random effects, and for the linear predictors from one of the random effects of our choice for each model.

Titer

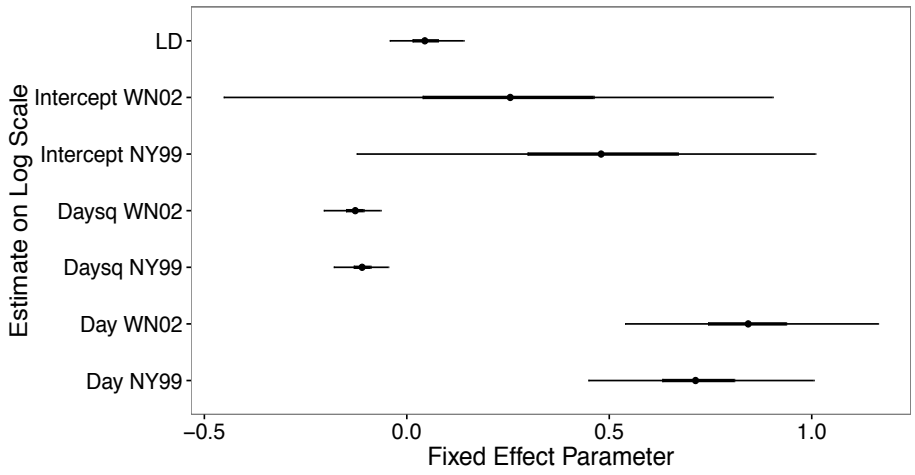

Figure S6.1: Fixed Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
LD	0.046	0.047	-0.042	0.014	0.044	0.077	0.140
Intercept WN02	0.249	0.336	-0.452	0.039	0.255	0.462	0.904
Intercept NY99	0.472	0.295	-0.124	0.298	0.480	0.670	1.009
Day WN02	0.843	0.156	0.539	0.744	0.843	0.937	1.164
Day NY99	0.719	0.139	0.448	0.631	0.713	0.809	1.005
Daysq WN02	-0.129	0.036	-0.205	-0.150	-0.127	-0.107	-0.064
Daysq NY99	-0.111	0.034	-0.180	-0.131	-0.110	-0.089	-0.046

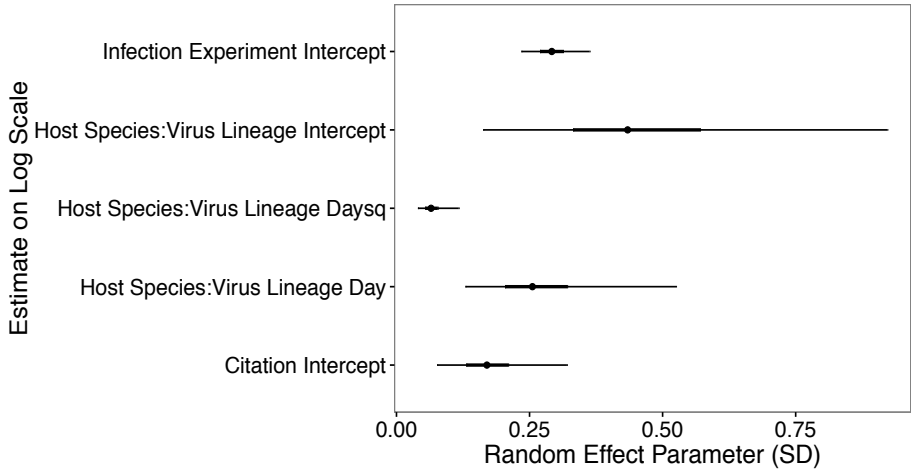

Figure S6.2: Random Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Infection Experiment Intercept	0.293	0.032	0.235	0.270	0.292	0.313	0.363
Host Species:Virus Lineage Daysq	0.068	0.021	0.040	0.054	0.065	0.077	0.117
Host Species:Virus Lineage Intercept	0.462	0.194	0.163	0.332	0.434	0.570	0.922
Host Species:Virus Lineage Day	0.272	0.101	0.130	0.204	0.255	0.321	0.525
Citation Intercept	0.176	0.062	0.076	0.131	0.170	0.210	0.320

Survival

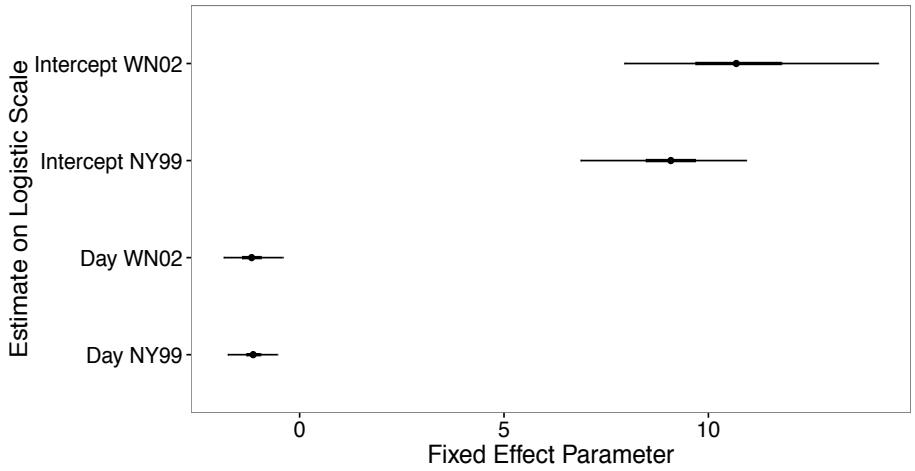

Figure S6.3: Fixed Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Intercept WN02	10.767	1.575	7.944	9.685	10.682	11.783	14.152
Intercept NY99	9.044	1.038	6.874	8.470	9.081	9.673	10.923
Day WN02	-1.163	0.365	-1.857	-1.406	-1.174	-0.949	-0.412
Day NY99	-1.135	0.286	-1.758	-1.299	-1.136	-0.960	-0.547

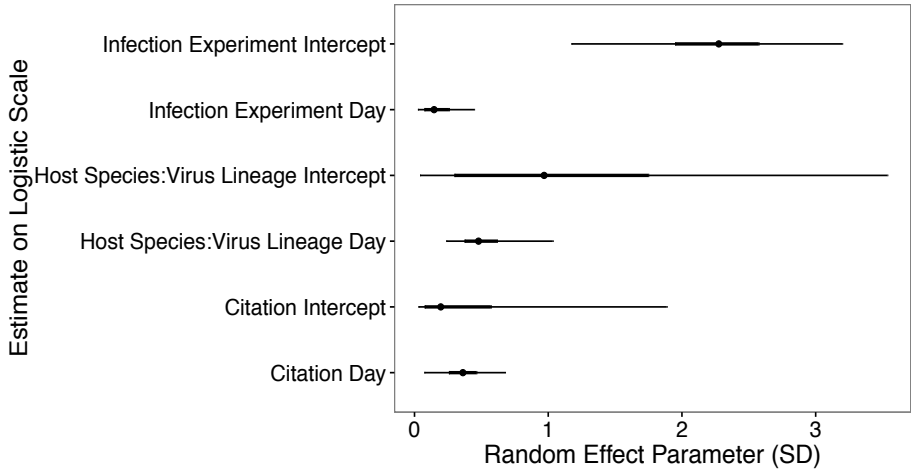

Figure S6.4: Random Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Infection Experiment Intercept	2.254	0.503	1.174	1.949	2.275	2.573	3.199
Infection Experiment Day	0.176	0.122	0.026	0.074	0.146	0.257	0.446
Host Species:Virus Lineage Intercept	1.181	1.068	0.042	0.299	0.969	1.748	3.536
Host Species:Virus Lineage Day	0.519	0.211	0.238	0.375	0.479	0.617	1.034
Citation Intercept	0.419	0.508	0.030	0.076	0.196	0.571	1.886
Citation Day	0.363	0.156	0.074	0.257	0.362	0.463	0.677

Bird to Mosquito Transmission

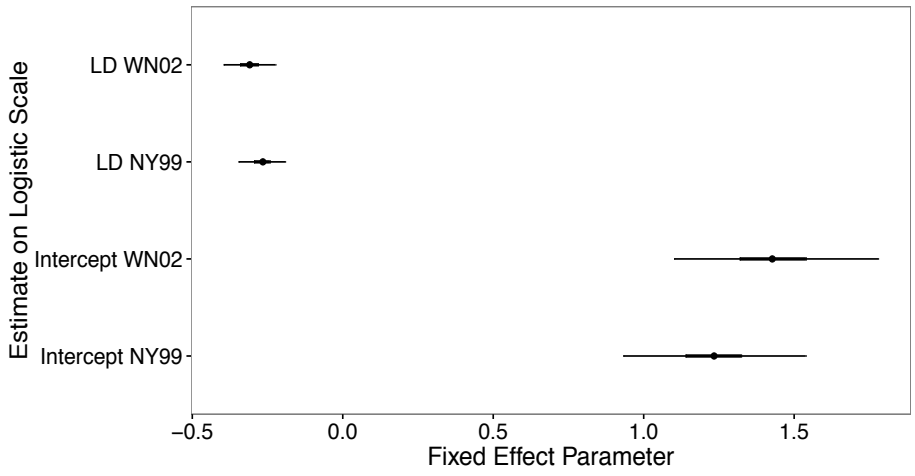

Figure S6.5: Fixed Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Intercept WN02	1.431	0.174	1.101	1.319	1.427	1.539	1.779
Intercept NY99	1.233	0.152	0.932	1.139	1.234	1.324	1.539
LD WN02	-0.310	0.044	-0.396	-0.341	-0.309	-0.282	-0.224
LD NY99	-0.267	0.039	-0.346	-0.295	-0.265	-0.242	-0.191

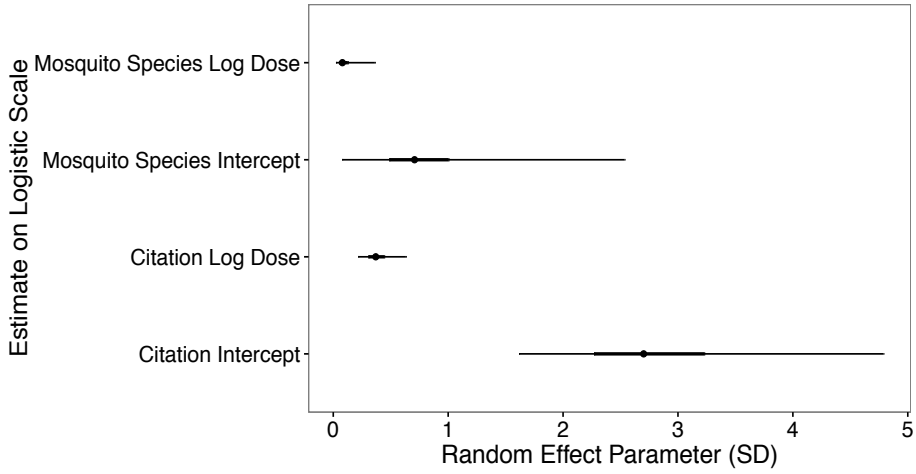

Figure S6.6: Random Effects (no intercept displayed)

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Mosquito Species Log Dose	0.106	0.097	0.024	0.049	0.080	0.126	0.362
Mosquito Species Intercept	0.839	0.620	0.078	0.488	0.707	1.003	2.535
Citation Log Dose	0.384	0.107	0.217	0.306	0.370	0.443	0.632
Citation Intercept	2.817	0.787	1.617	2.271	2.701	3.229	4.788

Mosquito to Bird Transmission: *With* JEV

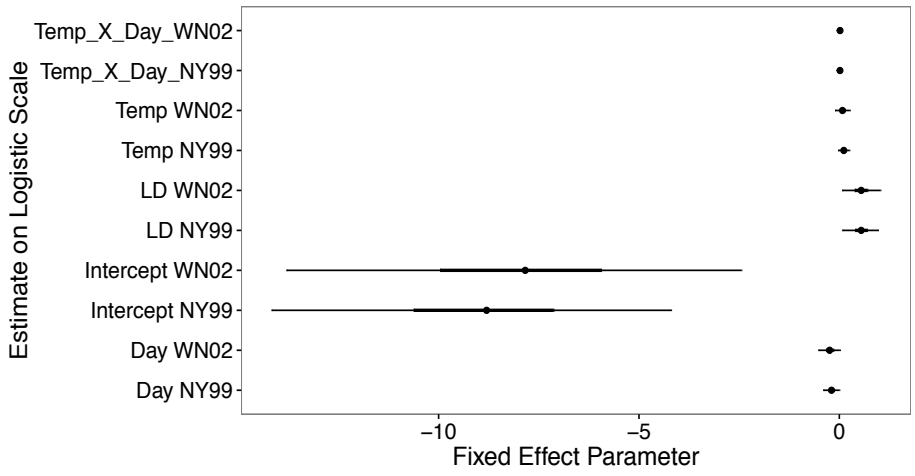

Figure S6.7: Fixed Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Temp_X_Day_WN02	0.018	0.005	0.009	0.015	0.018	0.021	0.028
Temp_X_Day_NY99	0.017	0.004	0.010	0.014	0.017	0.019	0.024
Temp WN02	0.078	0.092	-0.106	0.018	0.078	0.138	0.261
Temp NY99	0.112	0.072	-0.028	0.066	0.112	0.158	0.251
LD WN02	0.546	0.240	0.074	0.386	0.546	0.699	1.025
LD NY99	0.539	0.222	0.077	0.398	0.546	0.689	0.965
Intercept WN02	-7.957	2.954	-13.795	-9.961	-7.838	-5.947	-2.444
Intercept NY99	-8.919	2.518	-14.166	-10.620	-8.802	-7.134	-4.198
Day WN02	-0.242	0.135	-0.522	-0.328	-0.242	-0.149	0.014
Day NY99	-0.197	0.100	-0.399	-0.261	-0.196	-0.131	0.000

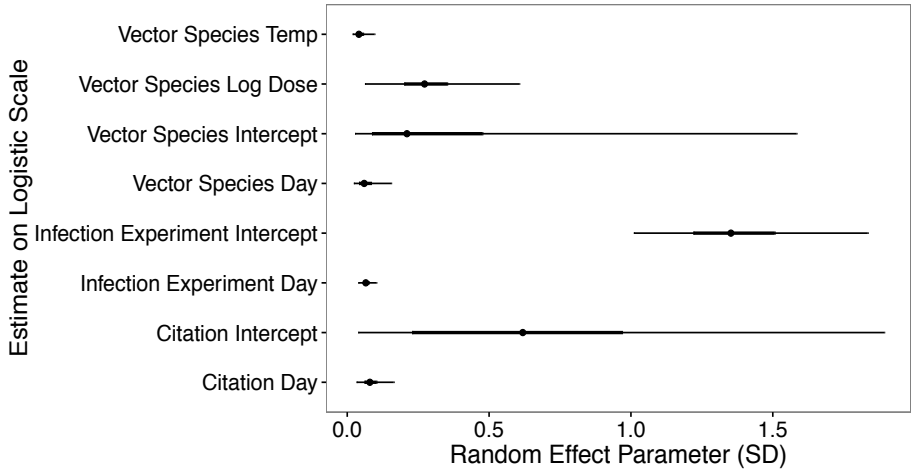

Figure S6.8: Random Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Vector Species Temp	0.046	0.020	0.019	0.031	0.041	0.056	0.096
Vector Species Log Dose	0.286	0.132	0.062	0.201	0.273	0.352	0.607
Vector Species Intercept	0.370	0.432	0.028	0.087	0.210	0.476	1.584
Vector Species Day	0.068	0.036	0.023	0.041	0.060	0.084	0.155
Infection Experiment Intercept	1.372	0.211	1.010	1.220	1.352	1.507	1.835
Infection Experiment Day	0.067	0.016	0.039	0.056	0.065	0.076	0.103
Citation Intercept	0.671	0.508	0.039	0.229	0.619	0.969	1.893
Citation Day	0.084	0.034	0.033	0.061	0.080	0.103	0.164

Mosquito to Bird Transmission: *Without* JEV

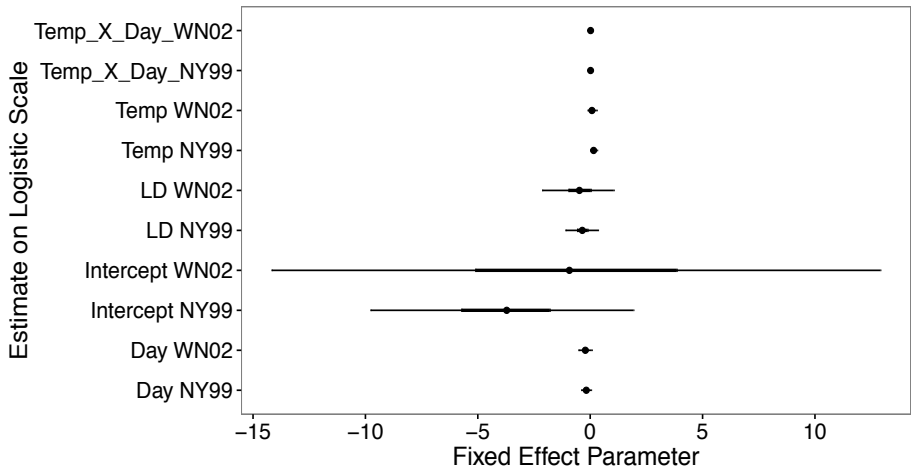

Figure S6.9: Fixed Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Temp_X_Day_WN02	0.018	0.006	0.007	0.015	0.018	0.022	0.029
Temp_X_Day_NY99	0.017	0.004	0.010	0.015	0.017	0.019	0.025
Temp WN02	0.079	0.104	-0.111	0.011	0.078	0.146	0.293
Temp NY99	0.153	0.078	0.004	0.103	0.152	0.204	0.313
LD WN02	-0.484	0.782	-2.134	-0.964	-0.480	0.027	1.050
LD NY99	-0.350	0.364	-1.097	-0.584	-0.352	-0.115	0.349
Intercept WN02	-0.644	6.983	-14.172	-5.112	-0.926	3.849	12.911
Intercept NY99	-3.752	2.972	-9.772	-5.732	-3.715	-1.800	1.925
Day WN02	-0.220	0.149	-0.520	-0.312	-0.216	-0.126	0.074
Day NY99	-0.174	0.111	-0.396	-0.246	-0.173	-0.106	0.041

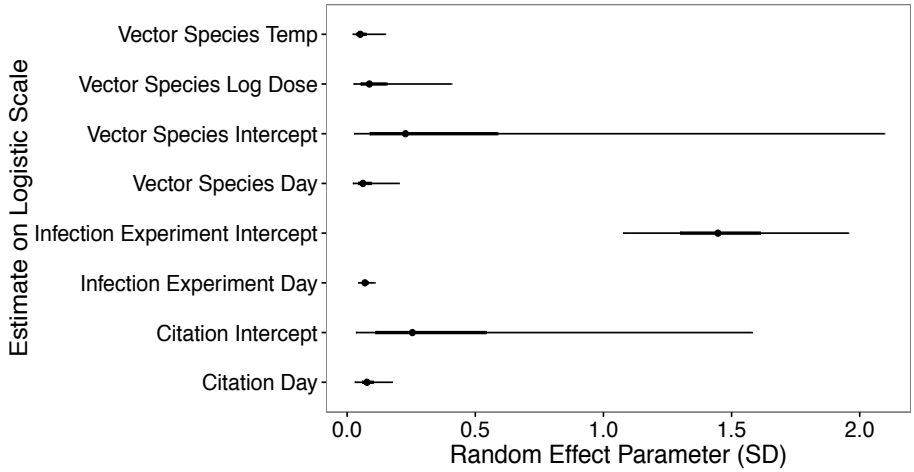

Figure S6.10: Random Effects

Parameter	Mean	SD	ci2.5	ci25	ci50	ci75	ci97.5
Vector Species Temp	0.060	0.036	0.022	0.036	0.051	0.072	0.148
Vector Species Log Dose	0.123	0.105	0.025	0.053	0.087	0.154	0.407
Vector Species Intercept	0.457	0.580	0.027	0.088	0.227	0.586	2.096
Vector Species Day	0.076	0.053	0.022	0.043	0.061	0.094	0.202
Infection Experiment Intercept	1.463	0.225	1.077	1.300	1.447	1.611	1.955
Infection Experiment Day	0.071	0.016	0.044	0.060	0.069	0.081	0.108
Citation Intercept	0.392	0.405	0.034	0.111	0.255	0.541	1.579
Citation Day	0.083	0.037	0.030	0.058	0.077	0.101	0.175

(7) Amplification Fraction Table

Species	Lower	Median	Upper
American Crow	0.000	0.000	0.003
House Sparrow	0.089	0.551	0.939
House Finch	0.004	0.031	0.222
American Robin	0.017	0.391	0.892

(8) Stan model notes

All stan models are available as .stan files in the online supplement and in the Github repository `https://github.com/morgankain/WNV_Synthesis.git`

For the titer profiles model, fixed effect parameters were given uninformative cauchy priors: intercepts were given cauchy(0, 10) priors and slopes were given cauchy(0, 2.5) priors. Variance parameters with positive constraints were given uninformative inverse gamma priors.

For the bird survival model, bird to mosquito transmission, and mosquito to bird transmission models parameters without constraints such as intercept or slope coefficients, were given normal(0.0, 1.0E3) priors. Variance parameters with positive constraints were given gamma(1.0E-3, 1.0E-3) priors.

References

Andreadis, S., O. Dimotsiou, and M. Savopoulou-Soultani 2014. Variation in adult longevity of *Culex pipiens f. pipiens*, vector of the West Nile Virus. *Parisitology Research 113*(11), 4315–4319.

Moudy, R. M., M. A. Meola, L.-L. L. Morin, G. D. Ebel, and L. D. Kramer 2007. A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by *Culex* mosquitoes. *The American Journal of Tropical Medicine and Hygiene 77*(2), 365–370.

Nash, J. C. 2012. nlmrt-vignette.

Appendix : Citation information for the data used in Chapter 2

Supplemental information on citations from which data was extracted, originally presented as a supplemental component of the published work presented as Chapter 2.

Citation Information for: Can existing data on WNV infection in birds and mosquitos explain strain replacement?

Morgan P. Kain[1†], Benjamin M. Bolker[1,2]

[1]Department of Biology, McMaster University, 1280 Main St. West, Hamilton, ON, Canada, L8S 4K1

[2]Department of Mathematics and Statistics, McMaster University, 1280 Main St. West, Hamilton, ON, Canada, L8S 4L8

[†]Correspondence author.
LSB-215
1280 Main St. West
Hamilton, ON
L8S 4K1
E-mail: kainm@mcmaster.ca

Notes on methodological decisions and oddities found in data extraction and full citation list sorted by data type

Titer and Survival

(Brault et al., 2004)

* An error bar on titer when there appeared to be only one surviving bird

* Overlapping error bars due to a lack of jitter. Measured carefully to connect error to the appropriate means

* No dates given for host capture (or dates overlooked if given).

* After-hatch-year birds

* No sex given

(Brault et al., 2007)

* After-hatch-year birds

* No sex given

(Brault et al., 2011)

* An error bar on titer when there appeared to be only one surviving bird

* Overlapping error bars due to a lack of jitter. Measured carefully to connect error to the appropriate means

* No dates given for host capture (or dates overlooked if given).

* No age given

* No sex given

(Clark et al., 2006)

* Birds were determined to be hatch-year birds by weight and plumage

* No sex given

(Duggal et al., 2014)

* Infection profiles of individual birds received from Dr. Nisha Duggal

* SW03 genotypes treated as genotypes of WN02 and NY2001 treated as NY99 (see main text)

* No age given

* No sex given

(Fang and Reisen, 2006)

* An error bar on titer when there appeared to be only one surviving bird

* Range given for titer dose. Used center of range in analysis.

* Death of birds from days 6-7 given as a total. Assumed 1/2 died each day

* No age given

* No sex given

(Grubaugh et al., 2015)

* An error bar on titer when there appeared to be only one surviving bird

* No dates given for host capture (or dates overlooked if given).

* No age given

* No sex given

(Guerrero-Sánchez et al., 2011)

* Overlapping error bars due to a lack of jitter. Measured carefully to connect error to the appropriate means

* Some instances of data description in text not matching appropriately to data depicted in figure. Used data from figure

* No dates given for host capture (or dates overlooked if given).

* Adult birds

* No sex given

(Kilpatrick et al., 2010)

* Hatch-year birds, 3-5 weeks old

* No sex given

(Kilpatrick et al., 2013)

* An error bar on titer when there appeared to be only one surviving bird

* Hatch-year and after-hatch year birds

* No sex given

(Kinney et al., 2006)

* No dates given for host capture (or dates overlooked if given).

* Hatch-year birds

(Kipp et al., 2006)

* No dates given for host capture (or dates overlooked if given)

* No age given

* No sex given

(Komar et al., 2003)

* Some oddities in the calculation of ranges.

* No age given

* Mixture of Males and Females. Sex given for some individual-level data not used in this study (titer in organs)

(Komar et al., 2005)

* No age given

* No sex given

(Langevin et al., 2005)

* No dates given for host capture (or dates overlooked if given).

* No age given

* No sex given

(Langevin et al., 2014)

* An error bar on titer when there appeared to be only one surviving bird

* Overlapping error bars due to a lack of jitter. Measured carefully to connect error to the appropriate means

* After-hatch year birds

* No sex given

(Melian et al., 2014)

* No dates given for host capture (or dates overlooked if given).

* No age given

* No sex given

(Nemeth et al., 2006)

* No dates given for host capture (or dates overlooked if given).

* Range given for titer dose. Used center of range in analysis.

* Most birds needle injected while a single bird was orally injected. Removed orally injected bird because it was at odds with the rest of the experiment.

* Juvenile birds

* No sex given

(Nemeth et al., 2009)

* Error bar on titer when there appeared to be only one surviving bird

* Range given for titer dose. Used center of range in analysis.

* Death of birds from days 5-9 given as a total. Assumed even mortality

* Adult birds

* No sex given

(Nemeth et al., 2011)

* No age given, but at least > 6 months old based on elapsed time between capture and experimentation

* No sex given

(Oesterle et al., 2009)

* Nestlings (8-17 days post-hatch)

* No sex given

(Owen et al., 2006)

* No mention of mortality. Activity levels were listed as not being affected, and given other language assumed no birds died.

* Combined all data from migrant and control birds because of no direct manipulation by the

authors

* Hatching-year birds

* No sex given (Authors include a statement that the species used in this study are monomorphic)

(Owen et al., 2012)

* Adult birds

* 15 females and 20 males

(Reisen and Fang, 2007)

* No age given

* No sex given

(Reisen and Hahn, 2007)

* No dates given for host capture (or dates overlooked if given).

* No age given

* No sex given

(Reisen et al., 2005)

* Death of House Finches in Figure 3B given as a total over the whole study duration. Due to too large of a time window left these data out. For sample size weighting for titer model death assumed to take place in the last 3 days of data, where the lack of data past certain day taken as complete mortality

* Death of some birds from days 4-7 given as a total. Assumed even mortality

* < 0.3 log10 titer units given. Used 0.3

* No dates given for host capture (or dates overlooked if given).

* No age given (Mosquitos were allowed to feed on adult birds in one trial, but it is unclear if the experimental birds were also adults)

* No sex given

(VanDalen et al., 2013)

* Adult birds

* No sex given

(Worwa et al., 2015)

* Hatching-year birds

* No sex given

(Ziegler et al., 2013)

* No dates given for host capture (or dates overlooked if given).

* Adolescent birds (>6 months)

* No sex given

Bird to Mosquito and Mosquito to Bird Transmission

(Anderson et al., 2012)

* Range given for titer dose. Used center of range in analysis.

* Virus retrieved from mosquitos by allowing them to feed on suckling mice

(Bolling et al., 2012)

* control used from control and coinfected

(Ciota et al., 2013)

Ciota, Alexander T, Chin, Pamela A, & Kramer, Laura D. (2013). The effect of hybridization of Culex pipiens complex mosquitoes on transmission of West Nile virus. Parasit Vectors, 6, 305.

* sample size given as 65-75. Used 70

* data from hybrids given. Just used non-hybrids

(Danforth et al., 2015)

* Virus retrieved from mosquitos using capillary tube method (20 min of feeding)

(Dodson et al., 2011)

* Multiple studies averaged

* control used from control and nutritionally deprived

(Dodson et al., 2014)

* control used from control and coinfected

(Dohm et al., 2002)

* Titer converted to transmission probability using the fitted relationship using the data in Moudy et al. 2007

* Reported transmission given as dissemination with the note that at least 90% of mosquitos with disseminated virus are able to transmit (Turell et al. 2000, 2001).

(Ebel et al., 2005)

(Goddard et al., 2002)

(Goenaga et al., 2015)

* Virus retrieved by collecting saliva using capillary tube method

(Hanley et al., 2005)

(Johnson et al., 2003)

* Titer converted to transmission probability using the fitted relationship using the data in Moudy et al. 2007

(Kilpatrick et al., 2008)

* Range given for titer dose. Used center of range in analysis.

* Transmission converted to Transmission | Infection

* Virus retrieved by collecting saliva using capillary tube method

(Moudy et al., 2007)

* Transmission converted to Transmission | Infection

* Data from intrathoracic inoculation of Culex pipiens excluded

* Virus retrieved by collecting saliva using capillary tube method

(Moudy et al., 2009)

* Range given for titer dose. Used center of range in analysis.

* Virus retrieved by collecting saliva using capillary tube method

(Reisen et al., 2005)

* Range given for log10 dose. Used center of range

* Range given for sample size. Used center of range

(Reisen et al., 2006a)

* Range given for titer dose. “Fed on sparrow at peak viremia”. Taken as 6.5 (could be off and also more variable)

* Virus retrieved by collecting saliva using capillary tube method

(Reisen et al., 2006)

(Richards et al., 2007)

(Richards et al., 2014)

(Sardelis and Turell, 2001)

(Sardelis et al., 2001)

* Transmission converted to Transmission | Infection

(Tiawsirisup et al., 2005)

* Transmission converted to Transmission | Infection

(Turell et al., 2000)

(Turell et al., 2001)

(Vanlandingham et al., 2004)

(Vanlandingham et al., 2007)

(Vanlandingham et al., 2008)

(Worwa et al., 2015)

* Transmission converted to Transmission | Infection

JEV

(Gould et al., 1962)

(MACKENZIE-IMPOINVIL et al., 2015)

(Muangman et al., 1972)

(Van Den Hurk et al., 2003)

Case Study data for mosquito to bird ratio, bird community composition, mosquito bite preference

(Hamer et al., 2009)

* Odd confidence intervals given binomial error distribution

(Simpson et al., 2012)

* Some oddities with confidence intervals

(Loss et al., 2009)

(Ruiz et al., 2010)

(Newman et al., 2011)

Seroprevalence Data

Using the search algorithm <West Nile Virus Seroprevalence> in google scholar we located 12 studies within the first 80 hits that presented seroprevalence data for WNV. These studies included:

(Bell et al., 2006)

* North Dakota and Minnesota, 2003-2005

(Bernard et al., 2001)

* New York, 2000

(Beveroth et al., 2006)

* Illinois, 2002-2004

(Chaves et al., 2016)

* Mexico, 2012

(Dusek et al., 2009)

* Many locations along east coast and midwest, 2001-2003

(Komar, 2001)

* New York, 2000

(Komar et al., 2005)

* Louisiana, 2002

(Loss et al., 2009)

* Illinois, 2005-2006

(O'Brien et al., 2010)

* Nebraska, 2008

(Reisen et al., 2004)

* California, 2003

(Reisen et al., 2006b)

* California, 2004

(Ringia et al., 2004)

* Illinois, 2002

References

Anderson, J. F., A. J. Main, G. Cheng, F. J. Ferrandino, and E. Fikrig 2012. Horizontal and vertical transmission of West Nile virus genotype NY99 by *Culex salinarius* and genotypes NY99 and WN02 by *Culex tarsalis*. *The American Journal of Tropical Medicine and Hygiene 86*(1), 134–139.

Bell, J. A., C. M. Brewer, N. J. Mickelson, G. W. Garman, and J. A. Vaughan 2006. West nile virus epizootiology, central red river valley, north dakota and minnesota, 2002–2005. *Emerg Infect Dis 12*(8), 1245–1247.

Bernard, K. A., J. G. Maffei, S. A. Jones, E. B. Kauffman, G. Ebel, A. Dupuis 2nd, K. A. Ngo, D. C. Nicholas, D. M. Young, P.-Y. Shi, et al. 2001. West nile virus infection in birds and mosquitoes, new york state, 2000. *Emerging infectious diseases 7*(4), 679.

Beveroth, T. A., M. P. Ward, R. L. Lampman, A. M. Ringia, and R. J. Novak 2006. Changes in seroprevalence of west nile virus across illinois in free-ranging birds from 2001 through 2004. *The American journal of tropical medicine and hygiene 74*(1), 174–179.

Bolling, B. G., F. J. Olea-Popelka, L. Eisen, C. G. Moore, and C. D. Blair 2012. Transmission dynamics of an insect-specific flavivirus in a naturally infected *Culex pipiens* laboratory colony and effects of co-infection on vector competence for West Nile virus. *Virology 427*(2), 90–97.

Brault, A. C., C. Y. Huang, S. A. Langevin, R. M. Kinney, R. A. Bowen, W. N. Ramey, N. A. Panella, E. C. Holmes, A. M. Powers, and B. R. Miller 2007. A single positively selected West Nile viral mutation confers increased virogenesis in American crows. *Nature Genetics 39*(9), 1162–1166.

Brault, A. C., S. A. Langevin, R. A. Bowen, N. A. Panella, B. J. Biggerstaff, B. R. Miller, and N. Komar 2004. Differential virulence of West Nile strains for American crows. *Emerging Infectious Diseases 10*(12), 2161.

Brault, A. C., S. A. Langevin, W. N. Ramey, Y. Fang, D. W. Beasley, C. M. Barker, T. A. Sanders, W. K. Reisen, A. D. Barrett, and R. A. Bowen 2011. Reduced avian virulence and viremia of West Nile virus isolates from Mexico and Texas. *The American Journal of Tropical Medicine and Hygiene 85*(4), 758–767.

Chaves, A., J. Sotomayor-Bonilla, O. Monge, A. Ramírez, F. Galindo, R. E. Sarmiento-Silva, G. A. Gutiérrez-Espeleta, and G. Suzán 2016. West nile virus in resident birds from yucatan, mexico. *Journal of wildlife diseases 52*(1), 159–163.

Ciota, A. T., P. A. Chin, and L. D. Kramer 2013. The effect of hybridization of *Culex pipiens* complex mosquitoes on transmission of West Nile virus. *Parasites & Vectors 6*, 305.

Clark, L., J. Hall, R. McLean, M. Dunbar, K. Klenk, R. Bowen, and C. A. Smeraski 2006. Susceptibility of greater sage-grouse to experimental infection with West Nile virus. *Journal of Wildlife Diseases 42*(1), 14–22.

Danforth, M. E., W. K. Reisen, and C. M. Barker 2015. Extrinsic incubation rate is not accelerated in recent California strains of West Nile virus in *Culex tarsalis* (Diptera: Culicidae). *Journal of Medical Entomology 52*(5), 1083–1089.

Dodson, B. L., G. L. Hughes, O. Paul, A. C. Matacchiero, L. D. Kramer, and J. L. Rasgon 2014. Wolbachia enhances West Nile virus (wnv) infection in the mosquito *Culex tarsalis*. *PLoS Neglected Tropical Diseases 8*(7), e2965.

Dodson, B. L., L. D. Kramer, and J. L. Rasgon 2011. Larval nutritional stress does not affect vector competence for West Nile virus (WNV) in *Culex tarsalis*. *Vector-Borne and Zoonotic Diseases 11*(11), 1493–1497.

Dohm, D. J., M. L. O'Guinn, and M. J. Turell 2002. Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *Journal of Medical Entomology 39*(1), 221–225.

Duggal, N. K., A. Bosco-Lauth, R. A. Bowen, S. S. Wheeler, W. K. Reisen, T. A. Felix, B. R. Mann, H. Romo, D. M. Swetnam, and A. D. Barrett 2014. Evidence for co-evolution of West Nile Virus and house sparrows in North America. *PLoS Neglected Tropical Diseases 8*(10), e3262.

Dusek, R. J., R. G. McLean, L. D. Kramer, S. R. Ubico, A. P. Dupuis, G. D. Ebel, and S. C. Guptill 2009. Prevalence of west nile virus in migratory birds during spring and fall migration. *The American journal of tropical medicine and hygiene 81*(6), 1151–1158.

Ebel, G. D., I. Rochlin, J. Longacker, and L. D. Kramer 2005. *Culex restuans* (Diptera: Culicidae) relative abundance and vector competence for West Nile virus. *Journal of Medical Entomology 42*(5), 838–843.

Fang, Y. and W. K. Reisen 2006. Previous infection with West Nile or St. Louis encephalitis viruses provides cross protection during reinfection in house finches. *The American Journal of Tropical Medicine and Hygiene 75*(3), 480–485.

Goddard, L. B., A. E. Roth, W. K. Reisen, and T. W. Scott 2002. Vector competence of California mosquitoes for West Nile virus. *Emerging Infectious Diseases 8*(12), 1385–1391.

Goenaga, S., J. L. Kenney, N. K. Duggal, M. Delorey, G. D. Ebel, B. Zhang, S. C. Levis, D. A. Enria, and A. C. Brault 2015. Potential for co-infection of a mosquito-specific flavivirus, nhumirim virus, to block West Nile Virus transmission in mosquitoes. *Viruses 7*(11), 5801–5812.

Gould, D. J., H. C. Barnett, and W. Suyemoto 1962. Transmission of japanese encephalitis virus by culex gelidus theobald. *Transactions of the Royal Society of Tropical Medicine and Hygiene 56*(5), 429–435.

Grubaugh, N. D., D. R. Smith, D. E. Brackney, A. M. Bosco-Lauth, J. R. Fauver, C. L. Campbell, T. A. Felix, H. Romo, N. K. Duggal, and E. A. Dietrich 2015. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. *PLoS Pathogens 11*(5), e1004874.

Guerrero-Sánchez, S., S. Cuevas-Romero, N. M. Nemeth, M. Trujillo-Olivera, G. Worwa, A. Dupuis, A. C. Brault, L. D. Kramer, N. Komar, and J. G. Estrada-Franco 2011. West Nile virus infection of birds, Mexico. *Emerging Infectious Diseases 17*(12), 2245–2252.

Hamer, G. L., U. D. Kitron, T. L. Goldberg, J. D. Brawn, S. R. Loss, M. O. Ruiz, D. B. Hayes, and E. D. Walker 2009. Host selection by *Culex pipiens* mosquitoes and West Nile virus amplification. *The American Journal of Tropical Medicine and Hygiene 80*(2), 268–278.

Hanley, K. A., L. B. Goddard, L. E. Gilmore, T. W. Scott, J. Speicher, B. R. Murphy, and A. G. Pletnev 2005. Infectivity of West Nile/dengue chimeric viruses for West Nile and dengue mosquito vectors. *Vector-Borne and Zoonotic Diseases 5*(1), 1–10.

Johnson, B., T. Chambers, M. Crabtree, J. Arroyo, T. Monath, and B. Miller 2003. Growth characteristics of the veterinary vaccine candidate ChimeriVax™-West Nile (WN) virus in *Aedes* and *Culex* mosquitoes. *Medical and Veterinary Entomology 17*(3), 235–243.

Kilpatrick, A. M., A. P. Dupuis, G.-J. J. Chang, and L. D. Kramer 2010. Dna vaccination of American robins (*Turdus migratorius*) against West Nile virus. *Vector-Borne Zoonot 10*(4), 377–380.

Kilpatrick, A. M., M. A. Meola, R. M. Moudy, and L. D. Kramer 2008. Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathogens 4*(6), e1000092.

Kilpatrick, A. M., R. J. Peters, A. P. Dupuis, M. J. Jones, P. Daszak, P. P. Marra, and L. D. Kramer 2013. Predicted and observed mortality from vector-borne disease in wildlife: West Nile virus and small songbirds. *Biological Conservation 165*, 79–85.

Kinney, R. M., C. Y.-H. Huang, M. C. Whiteman, R. A. Bowen, S. A. Langevin, B. R. Miller, and A. C. Brault 2006. Avian virulence and thermostable replication of the North American strain of West Nile virus. *Journal of General Virology 87*(12), 3611–3622.

Kipp, A. M., J. A. Lehman, R. A. Bowen, P. E. Fox, M. R. Stephens, K. Klenk, N. Komar, and M. L. Bunning 2006. West Nile virus quantification in feces of experimentally infected American and fish crows. *The American Journal of Tropical Medicine and Hygiene 75*(4), 688–690.

Komar, N. 2001. West nile virus surveillance using sentinel birds. *Annals of the New York Academy of Sciences 951*(1), 58–73.

Komar, N., S. Langevin, S. Hinten, N. Nemeth, E. Edwards, D. Hettler, B. Davis, R. Bowen, and M. Bunning 2003. Experimental infection of North American birds with the New York 1999 strain of West Nile virus. *Emerging Infectious Diseases 9*(3), 311–322.

Komar, N., N. A. Panella, S. A. Langevin, A. C. Brault, M. Amador, E. Edwards, and J. C. Owen 2005. Avian hosts for West Nile virus in St. Tammany Parish, Louisiana, 2002. *The American Journal of Tropical Medicine and Hygiene 73*(6), 1031–1037.

Langevin, S. A., R. A. Bowen, W. K. Reisen, C. C. Andrade, W. N. Ramey, P. D. Maharaj, M. Anishchenko, J. L. Kenney, N. K. Duggal, and H. Romo 2014. Host competence and helicase activity differences exhibited by West Nile viral variants expressing NS3-249 amino acid polymorphisms. *PLoS ONE 9*(6), e100802.

Langevin, S. A., A. C. Brault, N. A. Panella, R. A. Bowen, and N. Komar 2005. Variation in virulence of West Nile virus strains for house sparrows (*Passer domesticus*). *The American Journal of Tropical Medicine and Hygiene 72*(1), 99–102.

Loss, S. R., G. L. Hamer, E. D. Walker, M. O. Ruiz, T. L. Goldberg, U. D. Kitron, and J. D. Brawn 2009. Avian host community structure and prevalence of West Nile virus in Chicago, Illinois. *Oecologica 159*(2), 415–424.

MACKENZIE-IMPOINVIL, L., D. Impoinvil, S. Galbraith, R. Dillon, H. Ranson, N. Johnson, A. Fooks, T. Solomon, and M. Baylis 2015. Evaluation of a temperate climate mosquito, *Ochlerotatus detritus* (= *Aedes detritus*), as a potential vector of Japanese encephalitis virus. *Medical and Veterinary Entomology 29*(1), 1–9.

Melian, E. B., S. Hall-Mendelin, F. Du, N. Owens, A. M. Bosco-Lauth, T. Nagasaki, S. Rudd, A. C. Brault, R. A. Bowen, and R. A. Hall 2014. Programmed ribosomal frameshift alters expression of West Nile Virus genes and facilitates virus replication in birds and mosquitoes. *PLoS Pathogens 10*(11), e1004447.

Moudy, R. M., M. A. Meola, L.-L. L. Morin, G. D. Ebel, and L. D. Kramer 2007. A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by *Culex* mosquitoes. *The American Journal of Tropical Medicine and Hygiene 77*(2), 365–370.

Moudy, R. M., B. Zhang, P.-Y. Shi, and L. D. Kramer 2009. West Nile virus envelope protein glycosylation is required for efficient viral transmission by *Culex* vectors. *Virology 387*(1), 222–228.

Muangman, D., R. Edelman, M. J. Sullivan, and D. J. Gould 1972. Experimental transmission of japanese encephalitis virus by culex fuscocephala. *The American Journal of Tropical Medicine and Hygiene 21*(4), 482–6.

Nemeth, N., B. Thomsen, T. Spraker, J. Benson, A. Bosco-Lauth, P. Oesterle, J. Bright, J. Muth, T. Campbell, and T. Gidlewski 2011. Clinical and pathologic responses of American crows (*Corvus brachyrhynchos*) and fish crows (*C. ossifragus*) to experimental West Nile virus infection. *Veterinary Pathology 48*(6), 1061–1074.

Nemeth, N. M., D. C. Hahn, D. H. Gould, and R. A. Bowen 2006. Experimental West Nile virus infection in eastern screech owls (*Megascops asio*). *Avian Diseases 50*(2), 252–258.

Nemeth, N. M., P. T. Oesterle, and R. A. Bowen 2009. Humoral immunity to West Nile virus is long-lasting and protective in the house sparrow (*Passer domesticus*). *The American Journal of Tropical Medicine and Hygiene 80*(5), 864–869.

Newman, C. M., F. Cerutti, T. K. Anderson, G. L. Hamer, E. D. Walker, U. D. Kitron, M. O. Ruiz, J. D. Brawn, and T. L. Goldberg 2011. *Culex* flavivirus and West Nile virus mosquito coinfection and positive ecological association in Chicago, United States. *Vector-Borne and Zoonotic Diseases 11*(8), 1099–1105.

O'Brien, V. A., C. U. Meteyer, W. K. Reisen, H. S. Ip, and C. R. Brown 2010. Prevalence and pathology of west nile virus in naturally infected house sparrows, western nebraska, 2008. *The American journal of tropical medicine and hygiene 82*(5), 937–944.

Oesterle, P. T., N. M. Nemeth, K. VanDalen, H. Sullivan, K. T. Bentler, G. R. Young, R. G. McLean, L. Clark, C. Smeraski, and J. S. Hall 2009. Experimental infection of cliff swallows (petrochelidon pyrrhonota) with varying doses of West Nile virus. *The American Journal of Tropical Medicine and Hygiene 81*(6), 1159–1164.

Owen, J., F. Moore, N. Panella, E. Edwards, R. Bru, M. Hughes, and N. Komar 2006. Migrating birds as dispersal vehicles for West Nile virus. *EcoHealth 3*(2), 79–85.

Owen, J. C., A. Nakamura, C. A. Coon, and L. B. Martin 2012. The effect of exogenous corticosterone on West Nile virus infection in northern cardinals (*Cardinalis cardinalis*). *Veterinary Research 43*(1), 34.

Reisen, W., Y. Fang, and V. Martinez 2005. Avian host and mosquito (Diptera: Culicidae) vector competence determine the efficiency of West Nile and St. Louis encephalitis virus transmission. *Journal of Medical Entomology 42*(3), 367–375.

Reisen, W., H. Lothrop, R. Chiles, M. Madon, C. Cossen, L. Woods, S. Husted, V. Kramer, and J. Edman 2004. West nile virus in california. *Emerging infectious diseases 10*, 1369–1378.

Reisen, W. K. and Y. Fang 2007. Does feeding on infected mosquitoes (Diptera: Culicidae) enhance the role of song sparrows in the transmission of arboviruses in California? *Journal of Medical Entomology 44*(2), 316–319.

Reisen, W. K., Y. Fang, H. D. Lothrop, V. M. Martinez, J. Wilson, P. O'Connor, R. Carney, B. Cahoon-Young, M. Shafii, and A. C. Brault 2006a. Overwintering of West Nile virus in southern California. *Journal of Medical Entomology 43*(2), 344–355.

Reisen, W. K., Y. Fang, H. D. Lothrop, V. M. Martinez, J. Wilson, P. O'Connor, R. Carney, B. Cahoon-Young, M. Shafii, and A. C. Brault 2006b. Overwintering of west nile virus in southern california. *Journal of Medical Entomology 43*(2), 344–355.

Reisen, W. K., Y. Fang, and V. M. Martinez 2006. Vector competence of *Culiseta incidens* and *Culex thriambus* for West Nile Virus 1. *Journal of the American Mosquito Control Association 22*(4), 662–665.

Reisen, W. K. and D. C. Hahn 2007. Comparison of immune responses of brown-headed cowbird and related blackbirds to West Nile and other mosquito-borne encephalitis viruses. *Journal of Wildlife Diseases 43*(3), 439–449.

Richards, S. L., S. L. Anderson, and C. C. Lord 2014. Vector competence of *Culex pipiens quinquefasciatus* (Diptera: Culicidae) for West Nile virus isolates from Florida. *Tropical Medicine and International Health 19*(5), 610–617.

Richards, S. L., C. N. Mores, C. C. Lord, and W. J. Tabachnick 2007. Impact of extrinsic incubation temperature and virus exposure on vector competence of *Culex pipiens quinquefasciatus* say (Diptera: Culicidae) for West Nile virus. *Vector-Borne and Zoonotic Diseases 7*(4), 629–636.

Ringia, A. M., B. J. Blitvich, H.-Y. Koo, M. Van de Wyngaerde, J. D. Brawn, and R. J. Novak 2004. Antibody prevalence of west nile virus in birds, illinois, 2002. *Emerg Infect Dis 10*(6), 1120–4.

Ruiz, M. O., L. F. Chaves, G. L. Hamer, T. Sun, W. M. Brown, E. D. Walker, L. Haramis, T. L. Goldberg, and U. D. Kitron 2010. Local impact of temperature and precipitation on West Nile virus infection in *Culex* species mosquitoes in northeast Illinois, USA. *Parasites & Vectors 3*(1), 1.

Sardelis, M. R. and M. J. Turell 2001. *Ochlerotatus j. japonicus* in Frederick County, Maryland: discovery, distribution, and vector competence for West Nile virus. *Journal of the American Mosquito Control Association 17*(2), 137–141.

Sardelis, M. R., M. J. Turell, D. J. Dohm, and M. L. O'Guinn 2001. Vector competence of selected North American *Culex* and *Coquillettidia* mosquitoes for West Nile virus. *Emerging Infectious Diseases 7*(6), 1018.

Simpson, J. E., P. J. Hurtado, J. Medlock, G. Molaei, T. G. Andreadis, A. P. Galvani, and M. A. Diuk-Wasser 2012. Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile virus as a model system. *Proceedings of the Royal Society of London B: Biological Sciences 279*(1730), 925–933.

Tiawsirisup, S., K. B. Platt, R. B. Evans, and W. A. Rowley 2005. A comparision of West Nile Virus transmission by *Ochlerotatus trivittatus* (coq.), *Culex pipiens* (L.), and *Aedes albopictus* (skuse). *Vector-Borne and Zoonotic Diseases 5*(1), 40–47.

Turell, M. J., M. O'Guinn, and J. Oliver 2000. Potential for New York mosquitoes to transmit West Nile virus. *The American Journal of Tropical Medicine and Hygiene 62*(3), 413–414.

Turell, M. J., M. L. O'Guinn, D. J. Dohm, and J. W. Jones 2001. Vector competence of North American mosquitoes (Diptera: Culicidae) for West Nile virus. *Journal of Medical Entomology 38*(2), 130–134.

Van Den Hurk, A., D. Nisbet, R. Hall, B. Kay, J. Mackenzie, and S. Ritchie 2003. Vector competence of australian mosquitoes (Diptera: Culicidae) for Japanese encephalitis virus. *Journal of Medical Entomology 40*(1), 82–90.

VanDalen, K. K., J. S. Hall, L. Clark, R. G. McLean, and C. Smeraski 2013. West Nile virus infection in American robins: new insights on dose response. *PLoS ONE 8*(7), e68537.

Vanlandingham, D. L., C. E. McGee, K. A. Klinger, N. Vessey, C. Fredregillo, and S. Higgs 2007. Relative susceptibilties of South Texas mosquitoes to infection with West Nile virus. *The American Journal of Tropical Medicine and Hygiene 77*(5), 925–928.

Vanlandingham, D. L., C. E. McGee, K. A. Klingler, S. E. Galbraith, A. D. Barrett, and S. Higgs 2008. Comparison of oral infectious dose of West Nile virus isolates representing three distinct genotypes in *Culex quinquefasciatus*. *The American Journal of Tropical Medicine and Hygiene 79*(6), 951–954.

Vanlandingham, D. L., B. S. Schneider, K. Klingler, J. Fair, D. Beasley, J. Huang, P. Hamilton, and S. Higgs 2004. Real-time reverse transcriptase–polymerase chain reaction quantification of West Nile virus transmitted by *Culex pipiens quinquefasciatus*. *The American Journal of Tropical Medicine and Hygiene 71*(1), 120–123.

Worwa, G., S. S. Wheeler, A. C. Brault, and W. K. Reisen 2015. Comparing competitive fitness of West Nile Virus strains in avian and mosquito hosts. *PLoS ONE 10*(5), e0125668.

Ziegler, U., J. Angenvoort, D. Fischer, C. Fast, M. Eiden, A. V. Rodriguez, S. Revilla-Fernández, N. Nowotny, J. G. de la Fuente, and M. Lierz 2013. Pathogenesis of West Nile virus lineage 1 and 2 in experimentally infected large falcons. *Veterinary Microbiology 161*(3), 263–273.

Appendix : Additional methods and results for Chapter 3

Supplemental methods and results, originally presented as a supplemental component of the work in revision in Parasites & Vectors, presented as Chapter 3.

Supplemental Material for: Predicting bird community reservoir competence for West Nile virus using phylogenetic mixed effects models and eBird citizen science data

Morgan P. Kain[1] **and Benjamin M. Bolker**[1,2]

[1]**Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1 Canada**
[2]**Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario**

Author's Email Addresses:
Corresponding Author: Morgan P. Kain (kainm@mcmaster.ca)
Benjamin M. Bolker (bolker@mcmaster.ca)

Methods

Community resampling

Here we present the results of our simulated resampling of the best sampled bird communities in Texas, USA in the eBird database (communities with greater than 1,300 complete lists). We resampled 5-120 complete lists from the 46 most sampled communities 100 times. From these results we determined, subjectively, that with fewer than 80 complete lists, the gain in total number of communities was not worth the increased error rate and loss of species representation; at greater than 80 complete lists the loss in communities was too large for the small decrease in error and species loss.

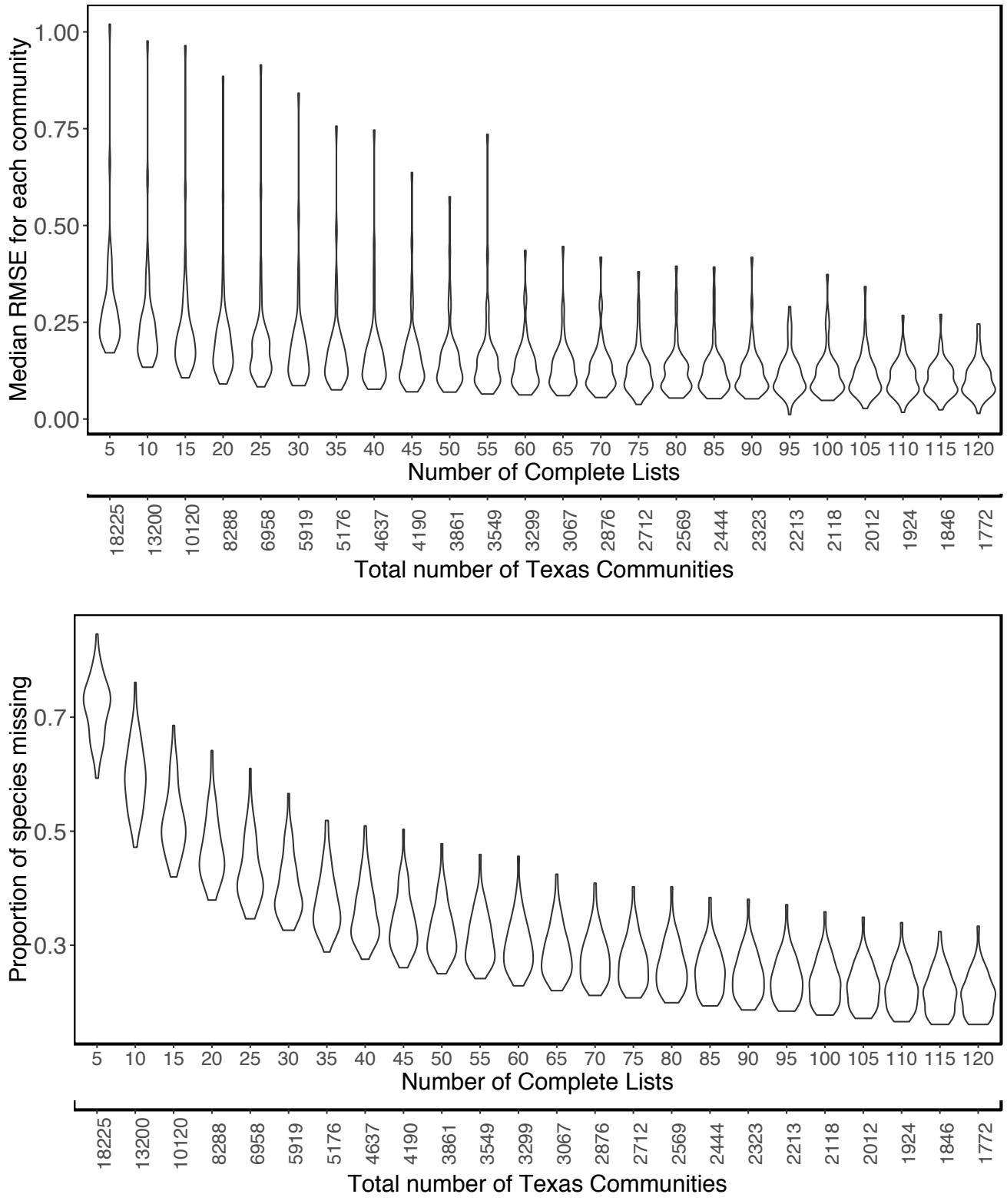

Figure S1. Result from 100 subsamples from each of the 46 communities with greater than 1,300 complete lists. Violin plots show distribution of medians across the 46 communities from the 100 resampling events. The top panel shows the RMSE between the relative proportions of all birds in the complete communities and subsampled communities as a function of the number of lists subsampled. The secondary axis shows the number of communities that would be left in the Texas dataset if the corresponding number of lists were used as a cutoff to define a well sampled community.

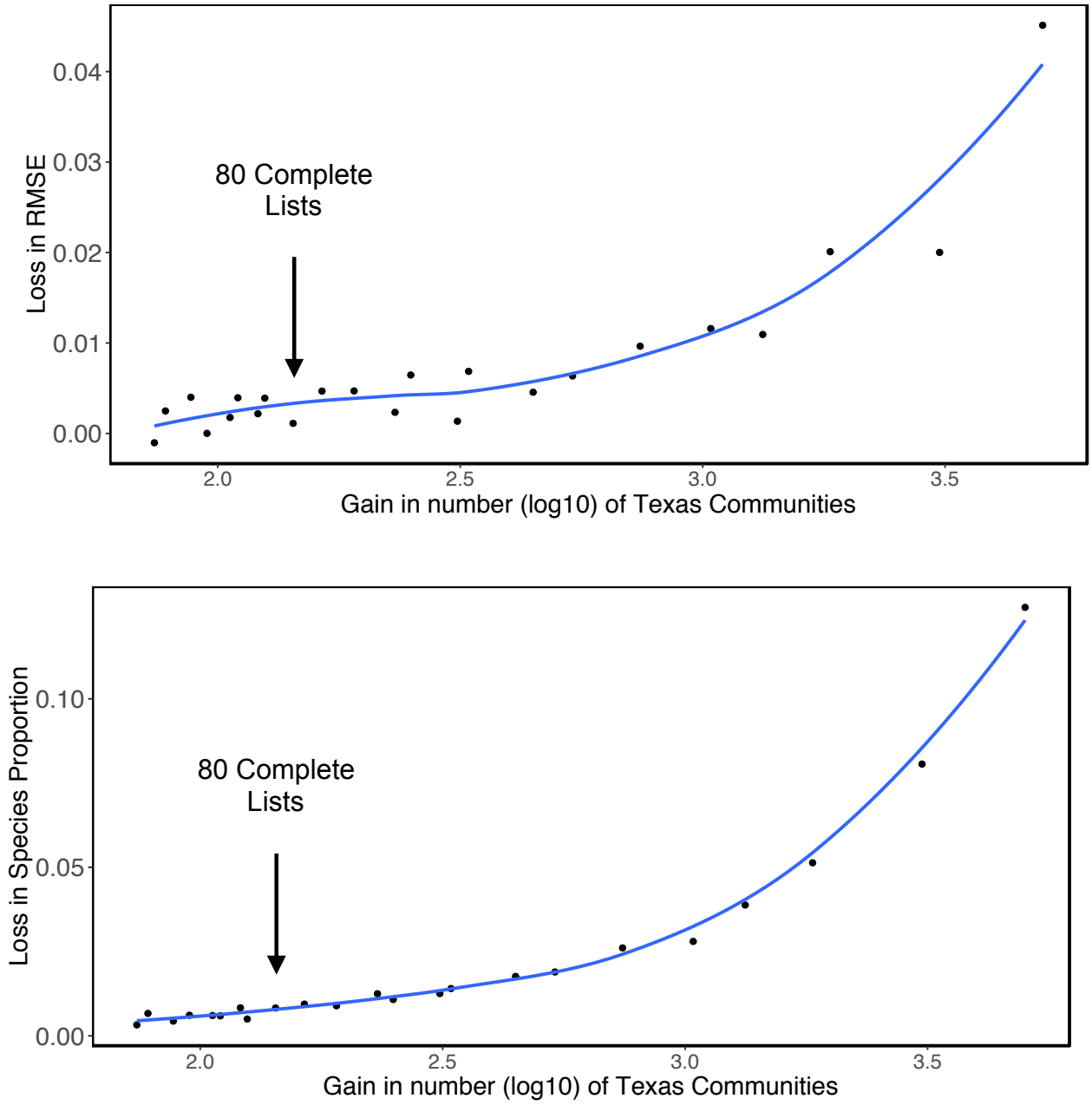

Figure S2. A second depiction of results from 100 subsamples from each of the 46 communities with greater than 1,300 complete lists. Top and bottom panels shows how much RMSE is gained, or the additional proportion of species that are lost, respectively, with an increase in retained Texas communities (which are gained when fewer complete lists are used as a cutoff, see Figure S1). Points in both panels show overall medians across all communities and 100 resampling events. Blue lines are smooths showing approximate slope. 80 complete lists resides at the beginning of the accelerating portion of this curve.

Sampling region for the mosquito biting preferences model

Figure S3 shows the eBird sampling region used to inform our prior in the Bayesian component of the mosquito biting preference model.

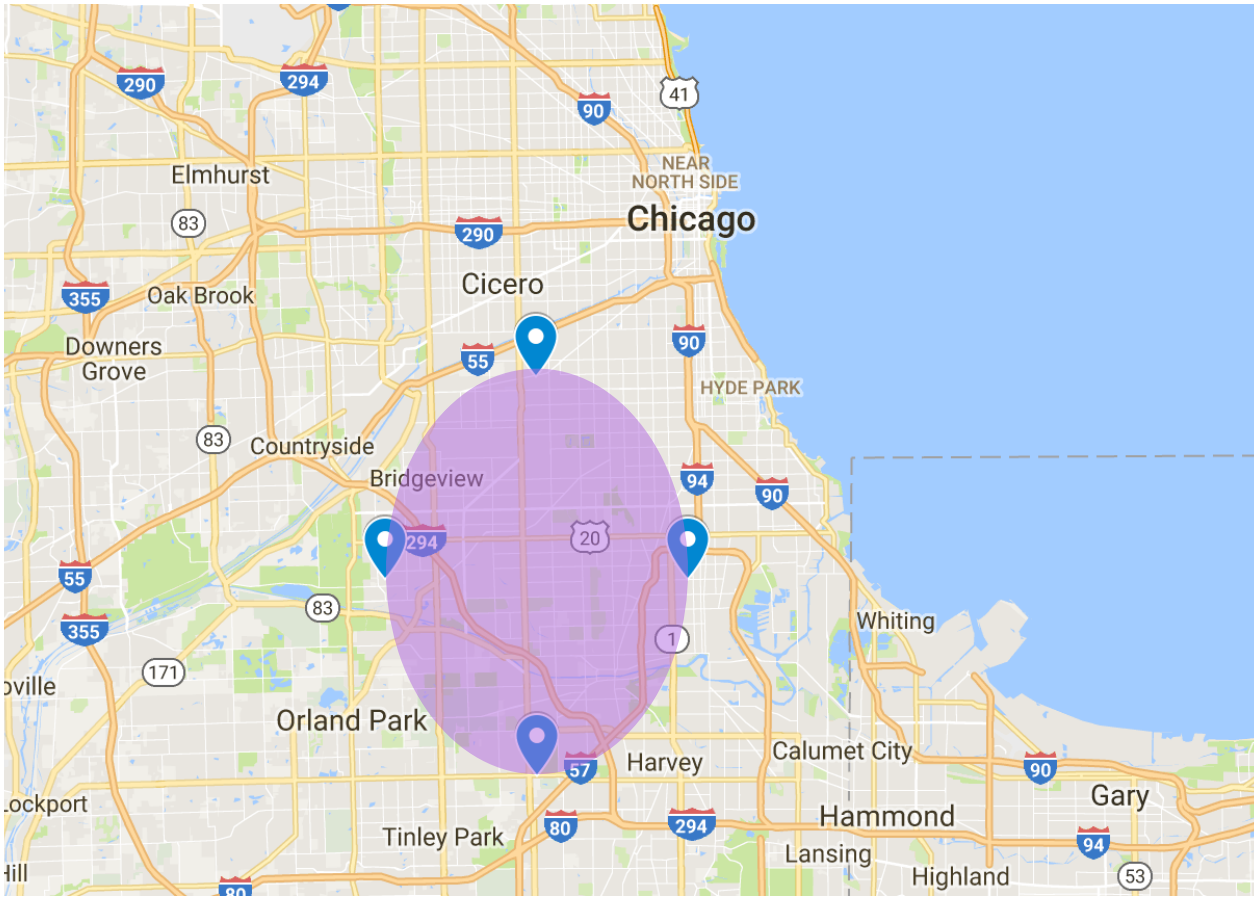

Figure S3. Delineation of the area where eBird lists were used as prior information for the bird community sampled in [1]. The center of this region was given as the location of the sampling conducted in [1]. The entire area is contained within Cook County, IL and is composed primarily of suburban Chicago, IL.

Phylogenetic mixed effects model validation

We present results for R^2 and blocked leave-one-out cross validation in Table S1. We calculated conditional R^2 (which measures the variance explained by both fixed and random factors; see [2, 3, 4]) for our phylogenetic mixed models using slightly modified code from `R` package `MuMIn` [5]). For blocked leave-one-out cross validation we left out all data for a single species at a time, refit our model with and without the species level phylogenetic random effect, and calculated goodness of fit of predicted values to the left out bird species' data for each model. For models with a normal or Poisson error distribution we use scaled root mean squared error (RMSE) between predicted values and the data with the following formula:

$$\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\bar{y}}. \tag{1}$$

For predicting responses using the titer model we use mean log dose across all infection experiments (which decreases fit to individual infection experiments, inflating RMSE slightly). For the binomial model (bird survival), we use area under the curve (AUC), calculated using the observed number of birds that died and did not die on each day and the estimated probabilities for each of these outcomes. Set up in this way, larger AUC values mean a better fit between the predicted values and the empirical data. AUC cannot be calculated for scenarios in which no birds died; in our calculation we ignore experiments in which no individual birds died. To calculate AUC we used the `R` package `MESS` [6]. Table S1 shows mean RMSE across all species for each method.

Despite an established method for calculating R^2 for mixed effects models [2, 3, 4], R^2 can be difficult to interpret for these models. Additionally, by convention, the biting preference model without the phylogenetic random effect (an intercept only model) has an $R^2 = 0$, which makes the ratio of R^2 values for this model difficult to interpret. The important point is that R^2 increases (in many cases greatly) for all four of our models when the species level phylogenetic random effect is included (Table S1).

The inclusion of a species level phylogenetic random effect increases the ability of three of four models to predict left-out species' responses (Table S1). For the survival model, AUC values are small overall because most bird species survive with high probability; it is difficult to predict the rare case of individual birds dying. AUC increases when a phylogenetic random effect is included in the survival model because this model is better able to predict the mortality of species that die with a higher frequency than the average species. For the biting preference model, the phylogenetic random effect marginally increases RMSE (Table S1); this may be due to minor overfitting.

Table S1. Phylogenetic mixed effects model validation. We use RMSE to quantify leave-one-out cross validation error for Titer, Biting preference, and detectability models, and AUC to quantify error for the survival model.

	Model							
	Titer		**Survival**		**Biting preference**		**Detectability**	
Phylogenetic random effect included?	No	Yes	No	Yes	No	Yes	No	Yes
Conditional R^2	0.65	0.93	0.805	0.898	0.00	0.97	0.51	0.92
Leave-one-out goodness of fit estimate	0.48	0.41	0.01	0.03	0.23	0.30	0.25	0.18

Ricker function

The Ricker function:

$$y = axe^{-bx} \tag{2}$$

is a hump-shaped curve that is commonly used for modeling right-skewed patterns in ecology [7] such as density dependence (e.g. [8]). When $a = 1$ this function is a Ricker function; when $a \neq 0$ this function is a Generalized Ricker. We use a Generalized Ricker function to model patterns in titer in birds, where x is given by day. Using log(titer) as our response variable, we estimate the a and b parameters in the Generalized Ricker function using a linear model with the form log(titer) $\sim$ day + log(day).

Results

Community $\mathcal{R}_0$

Complete eBird Dataset

Estimates for WNV $\mathcal{R}_0$ using the complete eBird dataset were more variable, but had similar monthly medians to the well sampled eBird dataset in most months (Figure S4, Figure S5).

The spatio-temporal GAM model explained 97% of the variation in WNV $\mathcal{R}_0$. WNV $\mathcal{R}_0$ was once again found to be driven mostly by temperature, and vary little across human population density and years. With the full dataset, bird communities most favorable for WNV transmission were once again estimated to be in the "Llano Uplift" ecoregion, while the least competent bird communities were estimated to be in the "High Plains" and "Oak Woods & Prairies" ecoregions. Overall $\mathcal{R}_0$ was estimated to be the highest in the "Piney Woods" and "Oak Woods & Prairies" (favorable temperature in the "Oak Woods & Prairies" has a much larger effect than the unfavorable bird communities in this region), and the lowest in the "High Plains".

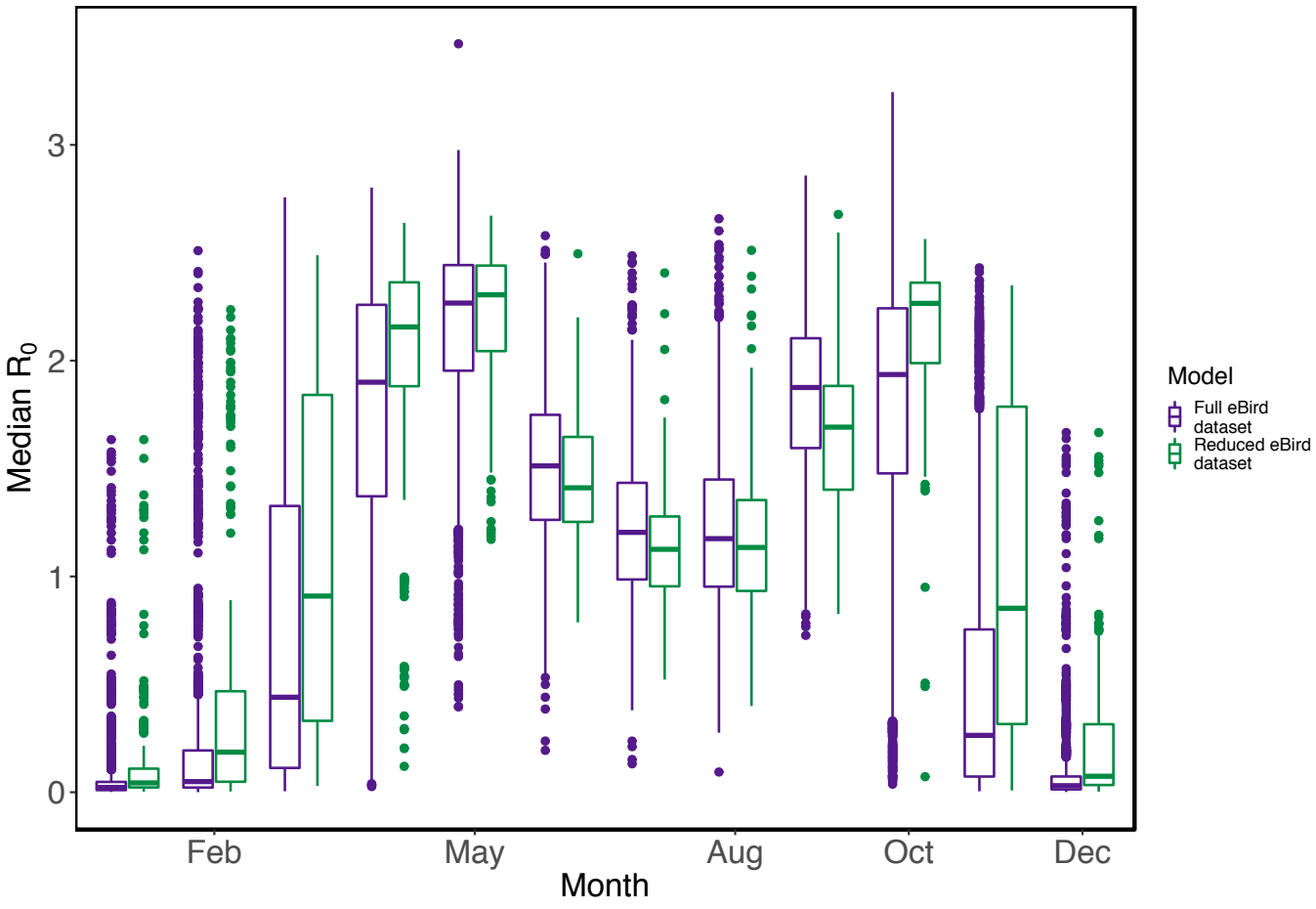

Figure S4. A comparison of $\mathcal{R}_0$ estimates. Purple and green boxplots show $\mathcal{R}_0$ estimates from models fit to the full and reduced eBird datasets respectively.

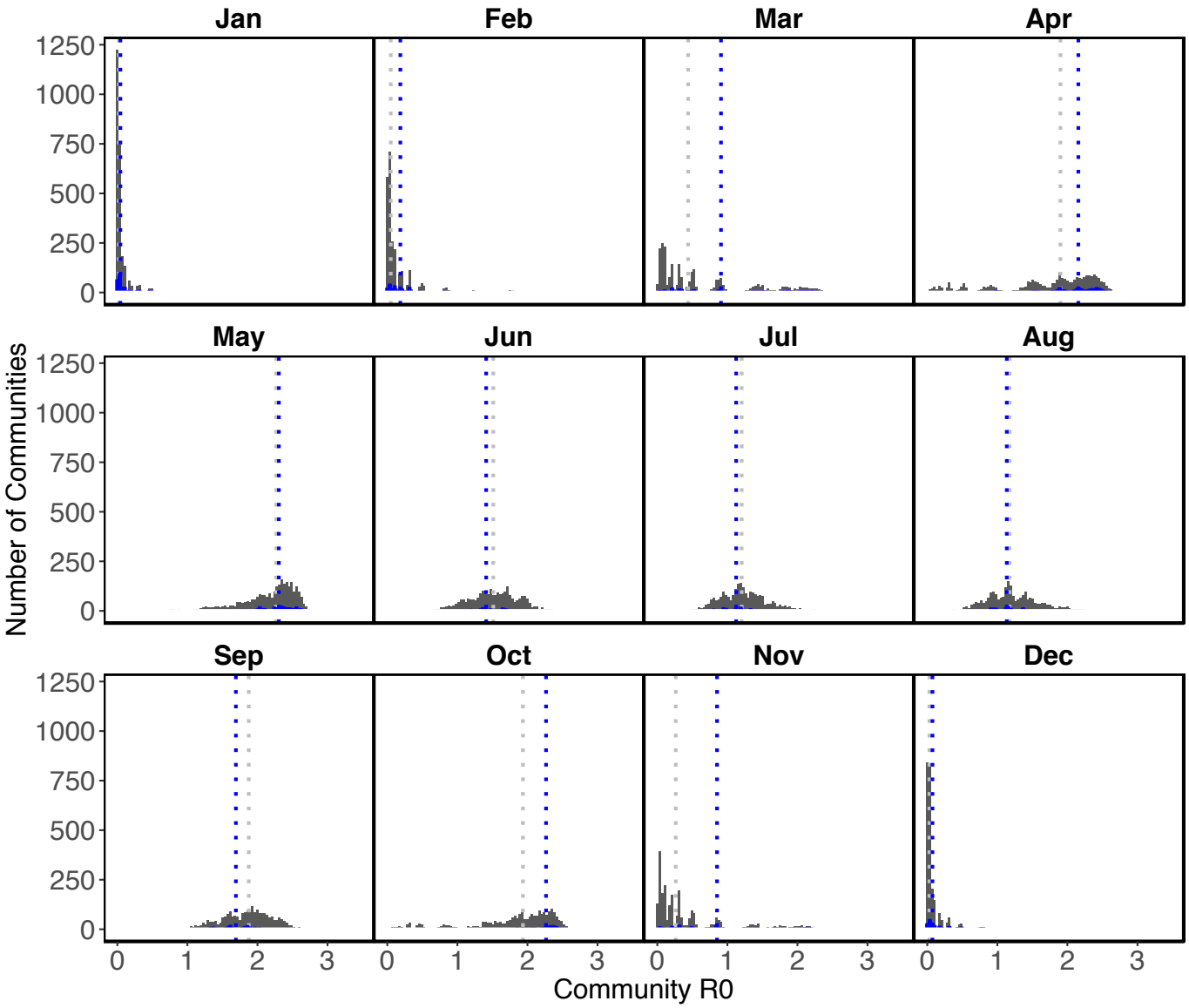

Figure S5. A comparison of $\mathcal{R}_0$ estimates. The gray histograms shows estimates for the full ebird data set and the blue histograms shows estimates for the reduced eBird data set by month. The vertical dotted gray and blue lines show the medians of the distributions.

No uncertainty propagated: Reduced eBird Dataset

Estimates for WNV $\mathcal{R}_0$ when propagating no uncertainty were more variable than when all uncertainty was propagated (Figure S6 vs main text Figure 1). No uncertainty propagation increased the impact of variation in the bird community on variation in $\mathcal{R}_0$ estimates within months (Figure S6).

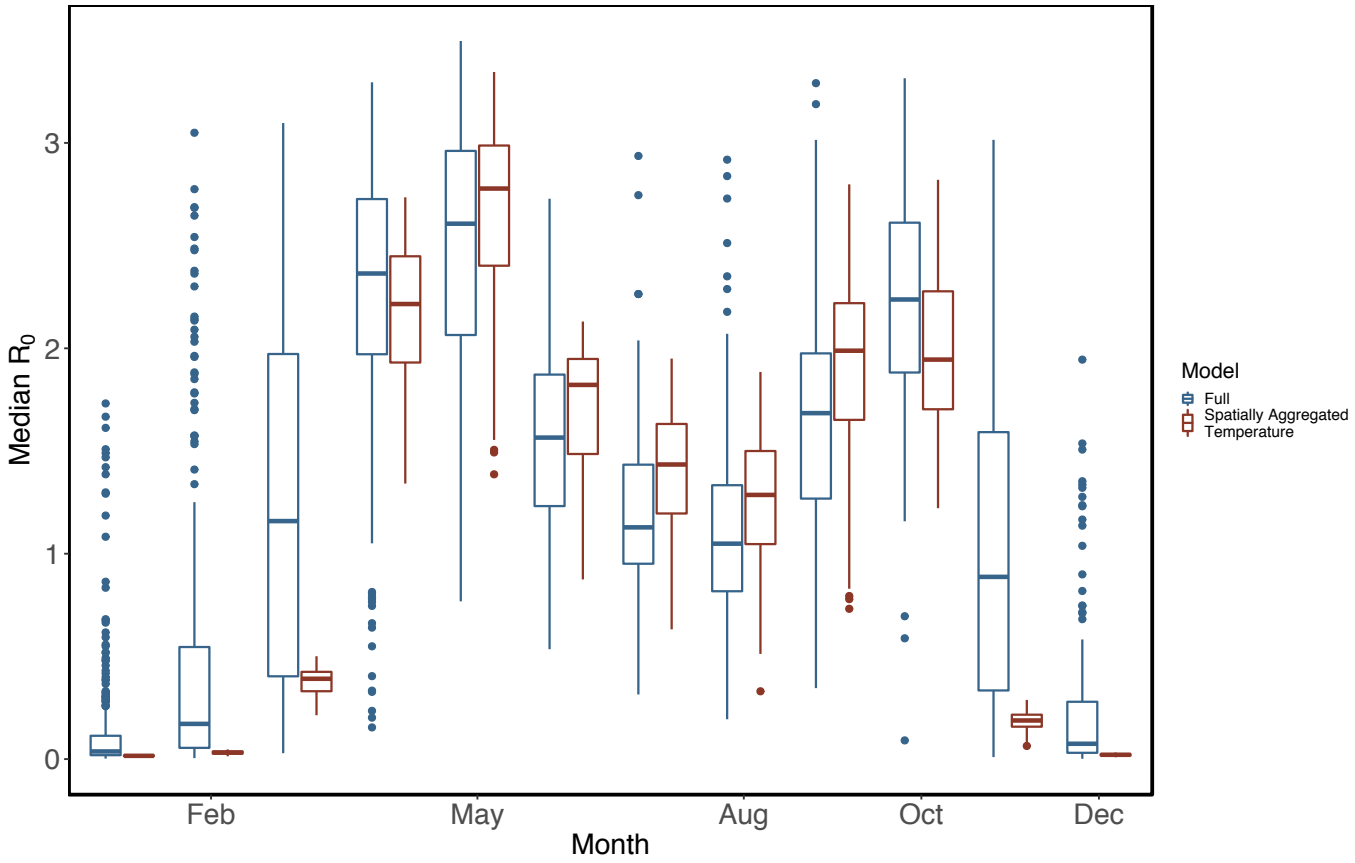

Figure S6. WNV $\mathcal{R}_0$ estimates between months and among Texas counties when no uncertainty is propagated. Blue boxplots show $\mathcal{R}_0$ estimates across Texas counties within months for a "Full" model, which used the eBird community and NOAA temperature data for each community. Red boxplots show $\mathcal{R}_0$ estimates from a model where each community retained their specific eBird community, but whose temperature was replaced with average temperature over all of Texas for that month. With no uncertainty propagated, variation in $\mathcal{R}_0$ within months attributable to variation in the bird communities (red boxplots) is much larger than when all uncertainty is propagated (see main text Figure 1). The increase in variation explained by the bird community is due to the translation of titer into bird-to-mosquito transmission probability, which has a maximum of one (see main text: Discussion for a full description of this phenomenon). Increases or decreases in medians between the models within months is due to the effects of averaging temperature prior to predicting $\mathcal{R}_0$ using the nonlinear functions for mosquito-to-bird transmission and mosquito survival across temperature.

Single sources of uncertainty: Reduced eBird Dataset

For each model and for most sources of uncertainty, using point estimates for a single parameter decreases the CV of $\mathcal{R}_0$ and increases the amount of variance explained in the spatio-temporal model. However, for the titer model, ignoring uncertainty in fixed effects increases the CV because of a smaller decrease in standard deviation than in the mean. In the titer model, ignoring uncertainty in the amount of evolutionary change along the most recent branch in time leading to species i results in the largest change in the CV of $\mathcal{R}_0$ for any source of uncertainty for any model, regardless of direction: a median of a 1.3 fold increase in the CV of $\mathcal{R}_0$.

Species-specific contributions to $\mathcal{R}_0$

Complete eBird Dataset

Here we present which species have the largest impact on $\mathcal{R}_0$ when they are removed from each of the 30,188 communities that they occupy, which gives a distribution of effects across communities. Northern cardinals were estimated to have a median effect on $\mathcal{R}_0$ similar in magnitude to their effect in the reduced dataset, but with a larger range of effects (recorded in 68% of the bird communities in the full dataset; median: 0.91 fold decrease in $\mathcal{R}_0$, 0.76-0.99 in 95% of communities). Jay species appear as the most impactful amplifiers in the full dataset as well. As in the reduced data set, Mourning doves (*Zenaida macroura*, recorded in 78% of the bird communities in the full dataset) accounted for the largest median dilution effect, but with an increased range of effect sizes (median: 1.05 fold increase in $\mathcal{R}_0$, 1.01-1.30 in 95% of communities). Again, only 2 species (Mourning doves: *Zenaida macroura* and White-winged doves: *Zenaida asiatica*) had a median dilution effect greater than a 1.01 fold increase in $\mathcal{R}_0$.

These analyses also provide a clear example of the type of problems that arise when using undersampled bird communities. In November of 2004 in Lamb county, a single complete list was submitted with two species, Sharp-tailed sandpiper (*Calidris acuminata*) and Ring-necked pheasant (*Phasianus colchicus*). Sharp-tailed sandpipers are estimated to be moderately competent hosts, while Ring-necked pheasants are estimated to be very poor hosts. This is the only "community" in which Sharp-tailed sandpipers were recorded in all of Texas. Sharp-tailed sandpipers were estimated to have a dilution effect of 0.29, which is implausible given the magnitude of the effects we estimate for the most important amplifier and diluter species.

Data citations

Laboratory infections of birds: [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38]

Laboratory infections of mosquitoes: bird-to-mosquito and mosquito-to-bird transmission: [35, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69]

Bird detectability: [70, 71, 72, 73], and [74, 75, 76, 77, 78, 79, 80] (cited in [81])

References

[1] Hamer GL, Kitron UD, Goldberg TL, Brawn JD, Loss SR, Ruiz MO, et al. Host selection by *Culex pipiens* mosquitoes and West Nile virus amplification. The American Journal of Tropical Medicine and Hygiene. 2009;80(2):268–278.

[2] Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution. 2013;4(2):133–142.

[3] Johnson PC. Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. Methods in Ecology and Evolution. 2014;5(9):944–946.

[4] Nakagawa S, Johnson PC, Schielzeth H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface. 2017;14(134):20170213.

[5] Barton K, Barton MK. Package 'MuMIn'. 2018;.

[6] Ekstrøm CT. MESS: Miscellaneous Esoteric Statistical Scripts; 2019. R package version 0.5.5. Available from: https://CRAN.R-project.org/package=MESS.

[7] Bolker BM. Ecological models and data in R. Princeton, NJ: Princeton University Press; 2008.

[8] Elliott J. Mechanisms responsible for population regulation in young migratory trout, *Salmo trutta*. The critical time for survival. The Journal of Animal Ecology. 1989;p. 987–1001.

[9] Brault AC, Langevin SA, Bowen RA, Panella NA, Biggerstaff BJ, Miller BR, et al. Differential virulence of West Nile strains for American crows. Emerging Infectious Diseases. 2004;10(12):2161.

[10] Brault AC, Huang CY, Langevin SA, Kinney RM, Bowen RA, Ramey WN, et al. A single positively selected West Nile viral mutation confers increased virogenesis in American crows. Nature Genetics. 2007;39(9):1162–1166.

[11] Brault AC, Langevin SA, Ramey WN, Fang Y, Beasley DW, Barker CM, et al. Reduced avian virulence and viremia of West Nile virus isolates from Mexico and Texas. The American Journal of Tropical Medicine and Hygiene. 2011;85(4):758–767.

[12] Clark L, Hall J, McLean R, Dunbar M, Klenk K, Bowen R, et al. Susceptibility of greater sage-grouse to experimental infection with West Nile virus. Journal of Wildlife Diseases. 2006;42(1):14–22.

[13] Duggal NK, Bosco-Lauth A, Bowen RA, Wheeler SS, Reisen WK, Felix TA, et al. Evidence for co-evolution of West Nile Virus and house sparrows in North America. PLoS Neglected Tropical Diseases. 2014;8(10):e3262.

[14] Fang Y, Reisen WK. Previous infection with West Nile or St. Louis encephalitis viruses provides cross protection during reinfection in house finches. The American Journal of Tropical Medicine and Hygiene. 2006;75(3):480–485.

[15] Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, et al. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. PLoS Pathogens. 2015;11(5):e1004874.

[16] Guerrero-Sánchez S, Cuevas-Romero S, Nemeth NM, Trujillo-Olivera M, Worwa G, Dupuis A, et al. West Nile virus infection of birds, Mexico. Emerging Infectious Diseases. 2011;17(12):2245–2252.

[17] Hofmeister E, Porter RE, Franson JC. Experimental susceptibility of wood ducks (Aix sponsa) for West Nile virus. Journal of wildlife diseases. 2015;51(2):411–418.

[18] Hofmeister EK, Lund M, Shearn-Bochsler V, Balakrishnan CN. Susceptibility and antibody response of the laboratory model zebra finch (Taeniopygia guttata) to West Nile virus. PloS one. 2017;12(1):e0167876.

[19] Kilpatrick AM, Dupuis AP, Chang GJJ, Kramer LD. DNA vaccination of American robins (*Turdus migratorius*) against West Nile virus. Vector-Borne Zoonot. 2010;10(4):377–380.

[20] Kilpatrick AM, Peters RJ, Dupuis AP, Jones MJ, Daszak P, Marra PP, et al. Predicted and observed mortality from vector-borne disease in wildlife: West Nile virus and small songbirds. Biological Conservation. 2013;165:79–85.

[21] Kinney RM, Huang CYH, Whiteman MC, Bowen RA, Langevin SA, Miller BR, et al. Avian virulence and thermostable replication of the North American strain of West Nile virus. Journal of General Virology. 2006;87(12):3611–3622.

[22] Kipp AM, Lehman JA, Bowen RA, Fox PE, Stephens MR, Klenk K, et al. West Nile virus quantification in feces of experimentally infected American and fish crows. The American Journal of Tropical Medicine and Hygiene. 2006;75(4):688–690.

[23] Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, Hettler D, et al. Experimental infection of North American birds with the New York 1999 strain of West Nile virus. Emerging Infectious Diseases. 2003;9(3):311–322.

[24] Komar N, Panella NA, Langevin SA, Brault AC, Amador M, Edwards E, et al. Avian hosts for West Nile virus in St. Tammany Parish, Louisiana, 2002. The American Journal of Tropical Medicine and Hygiene. 2005;73(6):1031–1037.

[25] Langevin SA, Brault AC, Panella NA, Bowen RA, Komar N. Variation in virulence of West Nile virus strains for house sparrows (*Passer domesticus*). The American Journal of Tropical Medicine and Hygiene. 2005;72(1):99–102.

[26] Langevin SA, Bowen RA, Reisen WK, Andrade CC, Ramey WN, Maharaj PD, et al. Host competence and helicase activity differences exhibited by West Nile viral variants expressing NS3-249 amino acid polymorphisms. PLoS ONE. 2014;9(6):e100802.

[27] Melian EB, Hall-Mendelin S, Du F, Owens N, Bosco-Lauth AM, Nagasaki T, et al. Programmed ribosomal frameshift alters expression of West Nile Virus genes and facilitates virus replication in birds and mosquitoes. PLoS Pathogens. 2014;10(11):e1004447.

[28] Nemeth NM, Hahn DC, Gould DH, Bowen RA. Experimental West Nile virus infection in eastern screech owls (*Megascops asio*). Avian Diseases. 2006;50(2):252–258.

[29] Nemeth NM, Oesterle PT, Bowen RA. Humoral immunity to West Nile virus is long-lasting and protective in the house sparrow (*Passer domesticus*). The American Journal of Tropical Medicine and Hygiene. 2009;80(5):864–869.

[30] Nemeth N, Thomsen B, Spraker T, Benson J, Bosco-Lauth A, Oesterle P, et al. Clinical and pathologic responses of American crows (*Corvus brachyrhynchos*) and fish crows (*C. ossifragus*) to experimental West Nile virus infection. Veterinary Pathology. 2011;48(6):1061–1074.

[31] Owen J, Moore F, Panella N, Edwards E, Bru R, Hughes M, et al. Migrating birds as dispersal vehicles for West Nile virus. EcoHealth. 2006;3(2):79–85.

[32] Owen JC, Nakamura A, Coon CA, Martin LB. The effect of exogenous corticosterone on West Nile virus infection in Northern Cardinals (*Cardinalis cardinalis*). Veterinary Research. 2012;43(1):34.

[33] Reisen WK, Fang Y. Does feeding on infected mosquitoes (Diptera: Culicidae) enhance the role of song sparrows in the transmission of arboviruses in California? Journal of Medical Entomology. 2007;44(2):316–319.

[34] Reisen WK, Hahn DC. Comparison of immune responses of brown-headed cowbird and related blackbirds to West Nile and other mosquito-borne encephalitis viruses. Journal of Wildlife Diseases. 2007;43(3):439–449.

[35] Reisen W, Fang Y, Martinez V. Avian host and mosquito (Diptera: Culicidae) vector competence determine the efficiency of West Nile and St. Louis encephalitis virus transmission. Journal of Medical Entomology. 2005;42(3):367–375.

[36] VanDalen KK, Hall JS, Clark L, McLean RG, Smeraski C. West Nile virus infection in American robins: new insights on dose response. PLoS ONE. 2013;8(7):e68537.

[37] Worwa G, Wheeler SS, Brault AC, Reisen WK. Comparing Competitive Fitness of West Nile Virus Strains in Avian and Mosquito Hosts. PLoS ONE. 2015;10(5):e0125668.

[38] Ziegler U, Angenvoort J, Fischer D, Fast C, Eiden M, Rodriguez AV, et al. Pathogenesis of West Nile virus lineage 1 and 2 in experimentally infected large falcons. Veterinary Microbiology. 2013;161(3):263–273.

[39] Anderson JF, Main AJ, Cheng G, Ferrandino FJ, Fikrig E. Horizontal and vertical transmission of West Nile virus genotype NY99 by *Culex salinarius* and genotypes NY99 and WN02 by *Culex tarsalis*. The American Journal of Tropical Medicine and Hygiene. 2012;86(1):134–139.

[40] Bolling BG, Olea-Popelka FJ, Eisen L, Moore CG, Blair CD. Transmission dynamics of an insect-specific flavivirus in a naturally infected *Culex pipiens* laboratory colony and effects of co-infection on vector competence for West Nile virus. Virology. 2012;427(2):90–97.

[41] Ciota AT, Chin PA, Kramer LD. The effect of hybridization of *Culex pipiens* complex mosquitoes on transmission of West Nile virus. Parasites & Vectors. 2013;6:305.

[42] Danforth ME, Reisen WK, Barker CM. Extrinsic incubation rate is not accelerated in recent California strains of West Nile virus in *Culex tarsalis* (Diptera: Culicidae). Journal of Medical Entomology. 2015;52(5):1083–1089.

[43] Dodson BL, Kramer LD, Rasgon JL. Larval nutritional stress does not affect vector competence for West Nile virus (WNV) in *Culex tarsalis*. Vector-Borne and Zoonotic Diseases. 2011;11(11):1493–1497.

[44] Dodson BL, Hughes GL, Paul O, Matacchiero AC, Kramer LD, Rasgon JL. Wolbachia enhances West Nile virus (WNV) infection in the mosquito *Culex tarsalis*. PLoS Neglected Tropical Diseases. 2014;8(7):e2965.

[45] Dohm DJ, O'Guinn ML, Turell MJ. Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. Journal of Medical Entomology. 2002;39(1):221–225.

[46] Ebel GD, Rochlin I, Longacker J, Kramer LD. *Culex restuans* (Diptera: Culicidae) relative abundance and vector competence for West Nile virus. Journal of Medical Entomology. 2005;42(5):838–843.

[47] Goddard LB, Roth AE, Reisen WK, Scott TW. Vector competence of California mosquitoes for West Nile virus. Emerging Infectious Diseases. 2002;8(12):1385–1391.

[48] Goenaga S, Kenney JL, Duggal NK, Delorey M, Ebel GD, Zhang B, et al. Potential for Co-Infection of a Mosquito-Specific Flavivirus, Nhumirim Virus, to Block West Nile Virus Transmission in Mosquitoes. Viruses. 2015;7(11):5801–5812.

[49] Hanley KA, Goddard LB, Gilmore LE, Scott TW, Speicher J, Murphy BR, et al. Infectivity of West Nile/dengue chimeric viruses for West Nile and dengue mosquito vectors. Vector-Borne and Zoonotic Diseases. 2005;5(1):1–10.

[50] Johnson B, Chambers T, Crabtree M, Arroyo J, Monath T, Miller B. Growth characteristics of the veterinary vaccine candidate ChimeriVax^{TM}-West Nile (WN) virus in *Aedes* and *Culex* mosquitoes. Medical and Veterinary Entomology. 2003;17(3):235–243.

[51] Kilpatrick AM, Meola MA, Moudy RM, Kramer LD. Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. PLoS Pathogens. 2008;4(6):e1000092.

[52] Moudy RM, Meola MA, Morin LLL, Ebel GD, Kramer LD. A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by *Culex* mosquitoes. The American Journal of Tropical Medicine and Hygiene. 2007;77(2):365–370.

[53] Moudy RM, Zhang B, Shi PY, Kramer LD. West Nile virus envelope protein glycosylation is required for efficient viral transmission by *Culex* vectors. Virology. 2009;387(1):222–228.

[54] Reisen WK, Fang Y, Lothrop HD, Martinez VM, Wilson J, O'Connor P, et al. Overwintering of West Nile virus in southern California. Journal of Medical Entomology. 2006;43(2):344–355.

[55] Reisen WK, Fang Y, Martinez VM. Vector competence of *Culiseta incidens* and *Culex thriambus* for West Nile Virus 1. Journal of the American Mosquito Control Association. 2006;22(4):662–665.

[56] Richards SL, Mores CN, Lord CC, Tabachnick WJ. Impact of extrinsic incubation temperature and virus exposure on vector competence of *Culex pipiens quinquefasciatus* Say (Diptera: Culicidae) for West Nile virus. Vector-Borne and Zoonotic Diseases. 2007;7(4):629–636.

[57] Richards SL, Anderson SL, Lord CC. Vector competence of *Culex pipiens quinquefasciatus* (Diptera: Culicidae) for West Nile virus isolates from Florida. Tropical Medicine and International Health. 2014;19(5):610–617.

[58] Sardelis MR, Turell MJ. *Ochlerotatus j. japonicus* in Frederick County, Maryland: discovery, distribution, and vector competence for West Nile virus. Journal of the American Mosquito Control Association. 2001;17(2):137–141.

[59] Sardelis MR, Turell MJ, Dohm DJ, O'Guinn ML. Vector competence of selected North American *Culex* and *Coquillettidia* mosquitoes for West Nile virus. Emerging Infectious Diseases. 2001;7(6):1018.

[60] Tiawsirisup S, Platt KB, Evans RB, Rowley WA. A comparision of West Nile Virus transmission by *Ochlerotatus trivittatus* (COQ.), *Culex pipiens* (L.), and *Aedes albopictus* (Skuse). Vector-Borne and Zoonotic Diseases. 2005;5(1):40–47.

[61] Turell MJ, O'Guinn M, Oliver J. Potential for New York mosquitoes to transmit West Nile virus. The American Journal of Tropical Medicine and Hygiene. 2000;62(3):413–414.

[62] Turell MJ, O'Guinn ML, Dohm DJ, Jones JW. Vector competence of North American mosquitoes (Diptera: Culicidae) for West Nile virus. Journal of Medical Entomology. 2001;38(2):130–134.

[63] Vanlandingham DL, Schneider BS, Klingler K, Fair J, Beasley D, Huang J, et al. Real-Time reverse transcriptase–polymerase chain reaction quantification of West Nile virus transmitted by *Culex pipiens quinquefasciatus*. The American Journal of Tropical Medicine and Hygiene. 2004;71(1):120–123.

[64] Vanlandingham DL, McGee CE, Klinger KA, Vessey N, Fredregillo C, Higgs S. Relative susceptibilties of South Texas mosquitoes to infection with West Nile virus. The American Journal of Tropical Medicine and Hygiene. 2007;77(5):925–928.

[65] Vanlandingham DL, McGee CE, Klingler KA, Galbraith SE, Barrett AD, Higgs S. Comparison of oral infectious dose of West Nile virus isolates representing three distinct genotypes in *Culex quinquefasciatus*. The American Journal of Tropical Medicine and Hygiene. 2008;79(6):951–954.

[66] Gould DJ, Barnett HC, Suyemoto W. Transmission of Japanese encephalitis virus by Culex gelidus Theobald. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1962;56(5):429–435.

[67] MACKENZIE-IMPOINVIL L, Impoinvil D, Galbraith S, Dillon R, Ranson H, Johnson N, et al. Evaluation of a temperate climate mosquito, *Ochlerotatus detritus* (= *Aedes detritus*), as a potential vector of Japanese encephalitis virus. Medical and Veterinary Entomology. 2015;29(1):1–9.

[68] Muangman D, Edelman R, Sullivan MJ, Gould DJ. Experimental transmission of Japanese encephalitis virus by Culex fuscocephala. The American Journal of Tropical Medicine and Hygiene. 1972;21(4):482–6.

[69] Van Den Hurk A, Nisbet D, Hall R, Kay B, Mackenzie J, Ritchie S. Vector competence of Australian mosquitoes (Diptera: Culicidae) for Japanese encephalitis virus. Journal of Medical Entomology. 2003;40(1):82–90.

[70] Dixon T. The distance at which sitting birds can be seen at sea. Ibis. 1977;119(3):372–375.

[71] Gibbs JP, Melvin SM. Call-response surveys for monitoring breeding waterbirds. The Journal of Wildlife Management. 1993;p. 27–34.

[72] Ronconi RA, Burger AE. Estimating seabird densities from vessel transects: distance sampling and implications for strip transects. Aquatic Biology. 2009;4(3):297–309.

[73] Pagano AM, Arnold TW. Detection probabilities for ground-based breeding waterfowl surveys. The Journal of Wildlife Management. 2009;73(3):392–398.

[74] Smith PM. Yuma Clapper Rail Study Mohave County, Arizona, 1973. State of California, The Resources Agency, Department of Fish and Game; 1974.

[75] Bart J, Stehn RA, Herrick JA, Heaslip NA, Bookhout TA, Stenzel JR. Survey methods for breeding Yellow Rails. The Journal of wildlife management. 1984;48(4):1382–1386.

[76] Piest L, Campoy J. Report of Yuma Clapper Rail Surveys at Cienega de Santa Clara, Sonora. Unpublished Report. 1998;.

[77] Kerlinger P, Sutton C. Black rail in New Jersey. Records of New Jersey Birds. 1989;15(2):22–26.

[78] Legare ML, Eddleman WR, Buckley P, Kelly C. The effectiveness of tape playback in estimating Black Rail density. The Journal of wildlife management. 1999;p. 116–125.

[79] Tecklin J. Distribution and abundance of the California black rail (*Laterallus jamaicensis coturniculus*) in the Sacramento valley region with accounts of ecology and call behavior of the subspecies. Contract Nos FG6154WM and FG6154-1WM. 1999;.

[80] Spear LB, Terrill SB, Lenihan C, Delevoryas P. Effects of Temporal and Environmental Factors on the Probability of Detecting California Black Rails. Journal of Field Ornithology. 1999;p. 465–480.

[81] Conway CJ, Gibbs J. Factors influencing detection probability and the benefits of call broadcast surveys for monitoring marsh birds. Final Report, USGS Patuxent Wildlife Research Center, Laurel, MD. 2001;.

Appendix : Details on code used in Chapter 3

Supplemental code information, originally presented as a supplemental component of the work in revision in Parasites & Vectors, presented as Chapter 3.

Supplementary readme for code used in: Predicting bird community reservoir competence for West Nile virus using phylogenetic mixed effects models and eBird citizen science data

Morgan P. Kain[1] and Benjamin M. Bolker[1,2]
[1]Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1 Canada
[2]Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario

Author's Email Addresses:
Corresponding Author: Morgan P. Kain (kainm@mcmaster.ca)
Benjamin M. Bolker (bolker@mcmaster.ca)

Supplementary code is available at: https://github.com/morgankain/WNV_Mechanistic_Model
---

Code can be run using two methods:

1) Clean eBird .zip file and run R code "manually" in two separate steps

A) In the command prompt run "sh ebird_data_clean.sh" once all eBird .zip files have been placed in the folder titled "ebird_zip_fresh". Note: Requires "Command Line Tools" on Mac.

B) Open "top_level_script.R" (Details about each R script can be found in "top_level_script.R")

i) Read the instructions at the top of this script

ii) Adjust parameters and options as desired

iii) Run the script as desired

2) Clean eBird .zip file and run R code in a single step

A) Run sh ebird_bash_run.sh

i) This requires some additional setup on the part of the user (see Step 4 in "ebird_bash_run.sh")

ii) This method cannot be used with manual matching of scientific names if scientific names are not found (see "saved_matching" below)

The code .zip contains a number of folders, many of which are empty. The folders are one of three types:

1) Folders that need to have data placed in them prior to running the code.

A) trees -- Contains phylogenetic tree data

B) ebird_zip_fresh -- Contains the .zip ebird file

C) data -- Contains all other data (bird responses, county data etc.)

2) Folders that contain model components

A) stan -- Contains stan model definitions

3) Folders that start empty and get filled when output is saved to disk while running code

A) ebird_data_for_R -- Will contain all of the eBird data once it is extracted

B) ebird_zip_dump -- Will contain .zip files after they get extracted

C) ebird_pieces -- One of two intermediate folders to momentarily house extracted eBird pieces

D) ebird_unzip -- Two of two intermediate folders to momentarily house extracted eBird pieces

E) saved_fits -- Houses intermediate fitted model results to expedite code in future runs

F) saved_matching -- Houses matched scientific names and bird body sizes

G) saved_output -- Houses all other fits including final "product"

Further detail for the purpose of each R script can be found in "top_level_script.R"

Appendix : A general model exploration for Chapter 4

Supplemental figures, originally presented as a supplemental component of the published work presented as Chapter 4.

The evolutionary response of virulence to host heterogeneity: a general model with application to myxomatosis in rabbits co-infected with intestinal helminths

Morgan P. Kain[1], Isabella M. Cattadori[2,3], and Benjamin M. Bolker[1,4]

[1]**Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1 Canada**
[2]**Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, 16802 PA, United States of America**

[3]**Department of Biology, The Pennsylvania State University, University Park, 16802 PA, United States of America**

[4]**Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L8 Canada**

Correspondence: kainm@mcmaster.ca

APPENDIX

GENERAL MODEL EXPLORATION

To explore the results presented here visually use the parameter values listed in Table S1 with the `Shiny R` applications at: https://morganpkain.shinyapps.io/myxo_power_hmf/ & https://morganpkain.shinyapps.io/myxo_piecewise_hmf/. In general, because of the focus on the piecewise heterogeneity map and its effects on MYXV ESS virulence, here we primarily focus on detailed analysis of qualitative results using the saturating exponential heterogeneity map. The complexity of the piecewise heterogeneity map, which has 4 parameters, makes it difficult to present all qualitative results across a range of parameter values. In general, qualitative patterns will differ greatly depending on the location of the median of the Gamma distribution of observed heterogeneity relative to the minimum of the quadratic portion of the piecewise function (h). Here we show one scenario where the median is less than h. Interested readers are encouraged to examine other parameter values using the online `Shiny R` application.

Relative virulence: What are the effects of a change in *median trait value* on pathogen virulence at the population level? (Figure S3 and Figure S4, panels A1 and A2)

With the saturating exponential heterogeneity map (Figure S3), an increase in mean trait value, such as a secondary infection with helminth macro-parasites, increases relative virulence monotonically. Specifically, the fold change in pathogen virulence that optimizes pathogen fitness in a homogeneous population with a non-zero worm burden follows the inverse of the heterogeneity map. An increase in median/mean trait value in a heterogeneous population, regardless of the shape of the heterogeneity distribution, will therefore decrease ESS relative virulence for both a power-law or sigmoidal tradeoff curve. Alternatively, with the non-monotonic (piecewise) heterogeneity map, an increase in the homogeneous population trait value can increase ESS virulence, as is shown in (Figure S4). The full inverse piecewise heterogeneity map is shown as a small insert in Figure S4, panel B2). At a high trait value ESS virulence decreases; these regions of the piecewise heterogeneity map are not captured in panels A1 and A2 of Figure S4) because the heterogeneity distribution used here.

Relative virulence: What are the effects of *host heterogeneity* on pathogen virulence at the population level? (Figure S3 and Figure S4, panels B1 and B2)

The effects of a change in heterogeneity on ESS virulence can be separated from the effects of a change in median (mean) trait value by comparing a heterogeneous population to a homogeneous population that share the same median (mean). The *combined effects* of a change in median and a change in heterogeneity is captured in panels B1 and B2 of Figure S3 and Figure S4. With the saturating exponential heterogeneity map variation tends to have have a small effect of further decreasing ESS virulence, but can result in a small increase in ESS virulence (e.g. blue curve, Figure S3 panel B2). With the piecewise heterogeneity map heterogeneity has a much larger overall effect of increasing or decreasing ESS virulence Figure S4. The fold change in ESS virulence attributable to a change in host heterogeneity also depends on the functional form (power-law or sigmoidal) of the tradeoff curve. For example, for either the saturating exponential or piecewise heterogeneity map, for a given host heterogeneity distribution, an increase in the slope (curvature: γ) of the power-law tradeoff curve decreases ESS relative virulence, but an increase in the slope of the sigmoidal tradeoff curve at its inflection point tends to increases ESS relative virulence (Figure S3).

Median exploitation: How does host heterogeneity affect pathogen severity in a given host type? (Figure S3 and Figure S4, panels C1 and C2)

Power-law and sigmoidal tradeoff curves for the relationship between transmission and virulence lead to qualitatively different patterns for median exploitation, though the difference is more pronounced with a saturating exponential heterogeneity map. Focusing on a saturating exponential heterogeneity map here, the median-host tends to be, but is not always, under-exploited. However, peak over-exploitation of the median host occurs at a larger mean trait value (larger relative virulence) for a power-law tradeoff curve than for a sigmoidal tradeoff curve. Peak over-exploitation of the median host occurs at small slopes of the power law tradeoff curve, but at large slopes of the sigmoidal tradeoff curve. Additionally, at very large mean trait value the majority of hosts have relative virulence approaching the saturating point of the het-

erogeneity map, lowering over-exploitation with either tradeoff curve. That is, as heterogeneity decreases, even in the presence of high helminth burdens, the median host experiences lower virulence because of a reduction in over-exploitation.

Efficiency: How much does host heterogeneity affect the ability of the pathogen to spread? (Figure S3 and Figure S4, panels D1 and D2)

A change in the heterogeneity distribution or slope of the tradeoff curve has similar effects for both power-law and sigmoidal tradeoff curves (Figure S3). For example, efficiency decreases with an increase in the slope of the tradeoff curve or an increase in the coefficient of variation (captured here by an increase in the median of the distribution with a changing scale parameter and constant shape parameter) of the heterogeneity distribution. Because of our assumptions of homogeneous mixing and an implastic pathogen, efficiency is constrained to be < 1. Within this constraint, an increase in individual variation increases the total amount of over- or under-exploitation by the implastic pathogen strain, decreasing pathogen efficiency. For example, in Figure S3 panel D2 efficiency is the lowest at the point that there is the highest over- or under-exploitation of the median host (Figure S3, panel C2). Though efficiency is the integral of over- and under-exploitation over all hosts in the population, the median-exploitation provides a clear illustration of the idea.

Generally, efficiency is lower when the tradeoff curve is sigmoidal (C and D—note vertical axis scale) because the $\mathcal{R}_0$ curve for the sigmoidal tradeoff curve is steeper than the $\mathcal{R}_0$ curve for the power-law tradeoff curve (Figure 1). That is, a given fold change away from α^* results in a larger reduction in $\mathcal{R}_0$ with a sigmoidal tradeoff curve than in a power-law tradeoff curve.

Sensitivity

Relative virulence, median-exploitation and efficiency are the most sensitive to changes in parameter values of each model component when the relative virulence distribution (Figure 2 panel C) is broad. The relative virulence distribution depends both on the distribution of host heterogeneity and on the heterogeneity map. When the host heterogeneity distribution has a mode equal to 0, which is assumed in our model exploration and is the case for MYXV in both

Scotland and Australia (e.g. a Gamma distribution with a shape parameter less than 1), the relative virulence distribution can become bimodal (as shown in Figure 2), which maximizes its variance. Bimodality arises when the heterogeneity map approaches its asymptote near the mean of the host heterogeneity distribution (Figure 2). A bimodal relative virulence distribution has two effects. First, any change in either the measured host heterogeneity distribution or the heterogeneity map causes a shift in which mode of the bimodal relative virulence distribution has higher density. For example, in the extreme of a relative virulence distribution with half low-virulence and half high-virulence (for example, in a population with 50% helminth-free rabbits and 50% rabbits with heavy helminth burden), a change to a distribution of either 51%:49% or 49%:51% causes a shift in the preferred host-type and a large difference in model outcomes.

Second, the sensitivity of model outcomes to changes in the functional form (power-law or sigmoidal) and slope of the tradeoff curve are similarly maximized when the relative virulence distribution is bimodal—small changes in the tradeoff curve have a large impact on which hosts contribute the most (are weighted the highest) in determining optimal virulence. However, the exact parameter space where results are most sensitive to changes in the parameter values of the tradeoff curve differ by the form of the tradeoff curve (power-law or sigmoidal) and among measured outcomes, because of differences in the asymmetry of the relationship between virulence and $\mathcal{R}_0$ for each tradeoff curve (Figure 1). For example, in these simulations the slope of the power-law tradeoff curve has the largest impact on ESS relative virulence when the mean of the observed heterogeneity distribution has relative virulence ≈ 4.5, which is $3/4$ of the saturation point of the saturating exponential heterogeneity map used here. Alternatively, the slope of the sigmoidal tradeoff curve has the largest impact on ESS relative virulence when the mean of the observed heterogeneity distribution has relative virulence ≈ 3, which is the half saturation point of the heterogeneity map (Figure S3). Additionally, the asymmetry of the $\mathcal{R}_0$ curve for a power-law tradeoff curve favors the over-exploitation of hosts at small tradeoff curve slopes, while a sigmoidal tradeoff curve favors the over-exploitation of hosts at large slopes. For example, an increase in the slope of the power-law tradeoff curve decreases relative virulence while a *decrease* in the slope of the sigmoidal tradeoff curve decreases relative virulence

(Figure S3, panel A1 and A2). Furthermore, the greater asymmetry of the sigmoidal tradeoff curve (Figure 1) causes all model components to be more sensitive to changes in the parameter values of the sigmoidal tradeoff curve.

With a heterogeneity distribution with mode greater than 0 (results not pictured) (e.g. a Gamma distribution with a shape parameter greater than 1), the sensitivity of model outcomes to changes in parameter values is also largest when the relative virulence distribution is variable. However, a heterogeneity distribution with mode greater than 0 does not allow for a bimodal relative virulence distribution (to examine parameter values see the online `Shiny R` applications). Instead, a symmetric, unimodal distribution maximizes variation in the relative virulence distribution. This distribution has a smaller maximum variance than a bimodal distribution, which greatly reduces the overall sensitivity of the model results to changes in any model component.

Table S1. Simulation parameter values

Model Component	**Tradeoff Curve**						**Gamma Distribution**		**Heterogeneity Map**						$\alpha \sim$ **titer**	**Background Mortality**
Functional Form	Power-Law		Sigmoidal						Saturating		piecewise					
	γ	c	a	b	c	n	shape	scale	z	r	h	k	z	r	b	μ
Figure																
3	–	–	2.04×10^6	**59.77**	**320.22**	**5.96**	**0.45**	**100.1**	–	–	10-400	0.33-0.99	3	0.020	0.1	0.006
4	–	–	2.04×10^6	**59.77**	**320.22**	**5.96**	**0.45**	**717.6, 276.1**	–	–	10-400	0.33-0.99	2	0.0125	0.1	0.006
s3	1.05-2	2	1	1	5	1.1-5	0.5	1-400	3	0.020	–	–	–	–	0.1	0.006
S4	1.05-2	2	1	1	5	1.1-5	0.5	1-400	–	–	100	0.5	3	0.020	0.1	0.006
S5	–	–	2.04×10^6	**59.77**	**320.22**	**5.96**	**0.45**	**717.6, 276.1**	–	–	10–400	0.33–0.99	10	0.025	0.1	0.006
S6	–	–	2.04×10^6	**59.77**	**320.22**	**5.96**	**0.45**	**717.6, 100.1**	–	–	10-325	0.33-0.99	2	0.0125	0.1	0.006

Parameter values obtained from fits to empirical data are listed in bold.

APPENDIX FIGURES

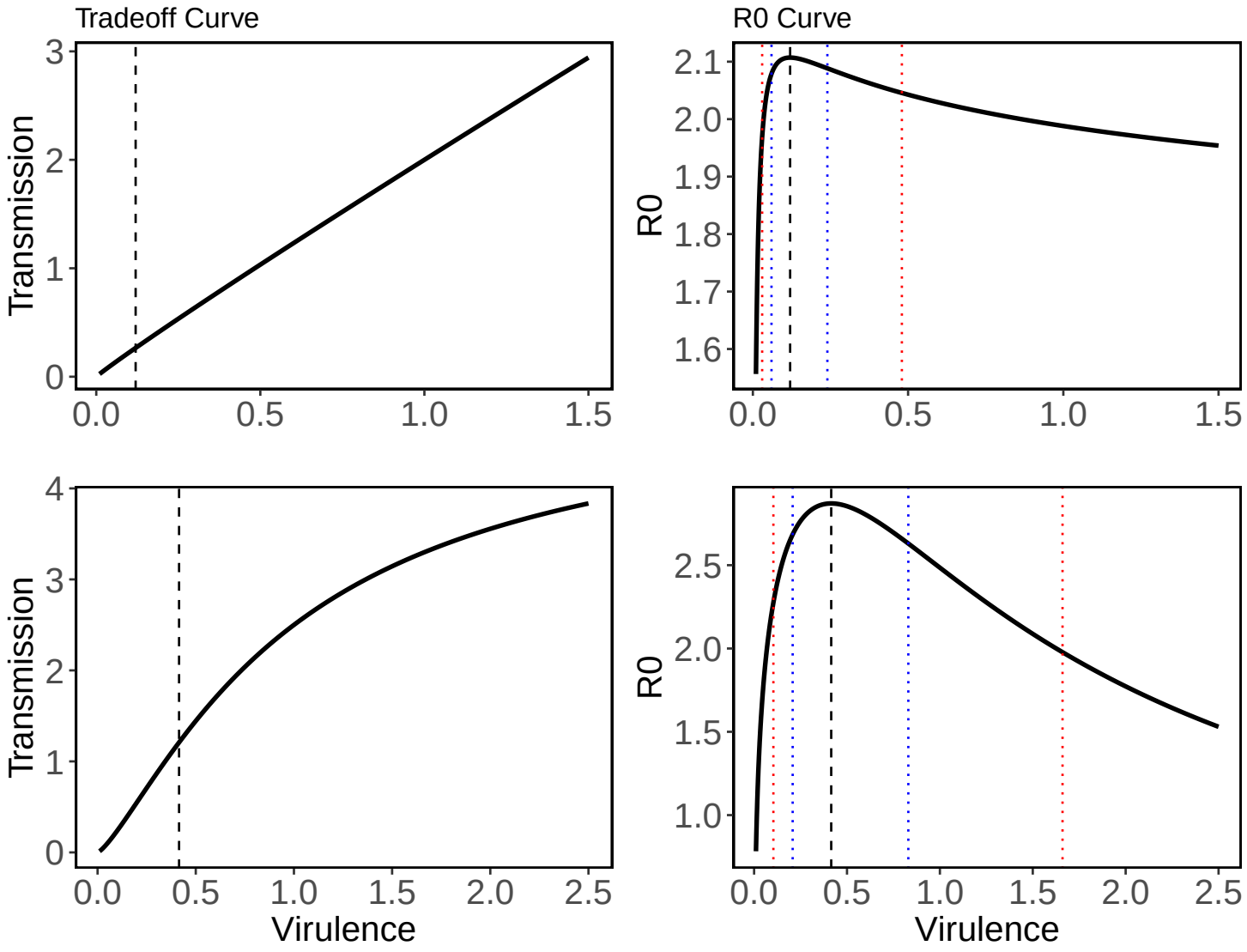

Figure S1. The top (bottom) row shows an example of the power-law (sigmoidal) tradeoff and $\mathcal{R}_0$ curves using the same parameters as the simulation used to generated Figures 2-4 (with $\gamma = 1.05$ for the power-law relationship and $c = 1.3$ for the sigmoidal relationship). The vertical dashed black line for each curve shows the optimum α. The vertical dotted blue (red) lines show the $\mathcal{R}_0$ for a 2 (4) $\times$ higher or lower α than optimum).

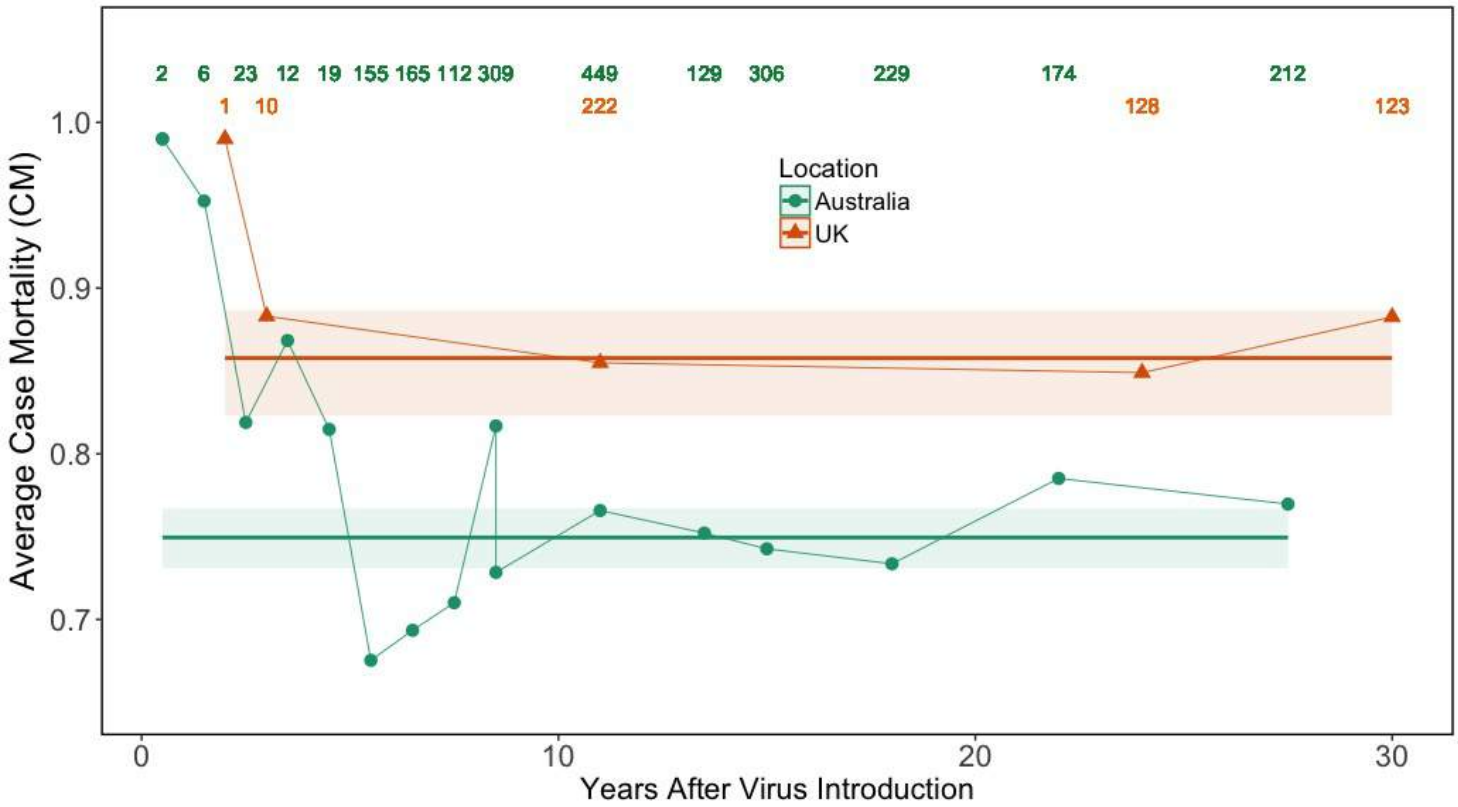

Figure S2. Virulence of MYXV strains in Australia and the UK following the introduction of the virus. Points are the average virulence of circulating MYXV strains in the two locations. Lines and ribbons are estimates $\pm$ 2 SE from the binomial glm model fit to average MYXV virulence in years 4-30 post introduction (described in the main text). Numbers at the top of the plot are the total number of strains sampled in a given location in a given year. The data presented here is not the result of a rigorous meta-analysis, but simply the result of a search through the literature; data was taken from 7 total publications (raw data is provided in a csv file in the online supplemental material). Other caveats include: data being grouped across many years (Fenner and Marshall, 1957) and the transition from using 5 virulence classes (Fenner and Marshall, 1957), to using 6 virulence classes (Fenner and Chapple, 1965).

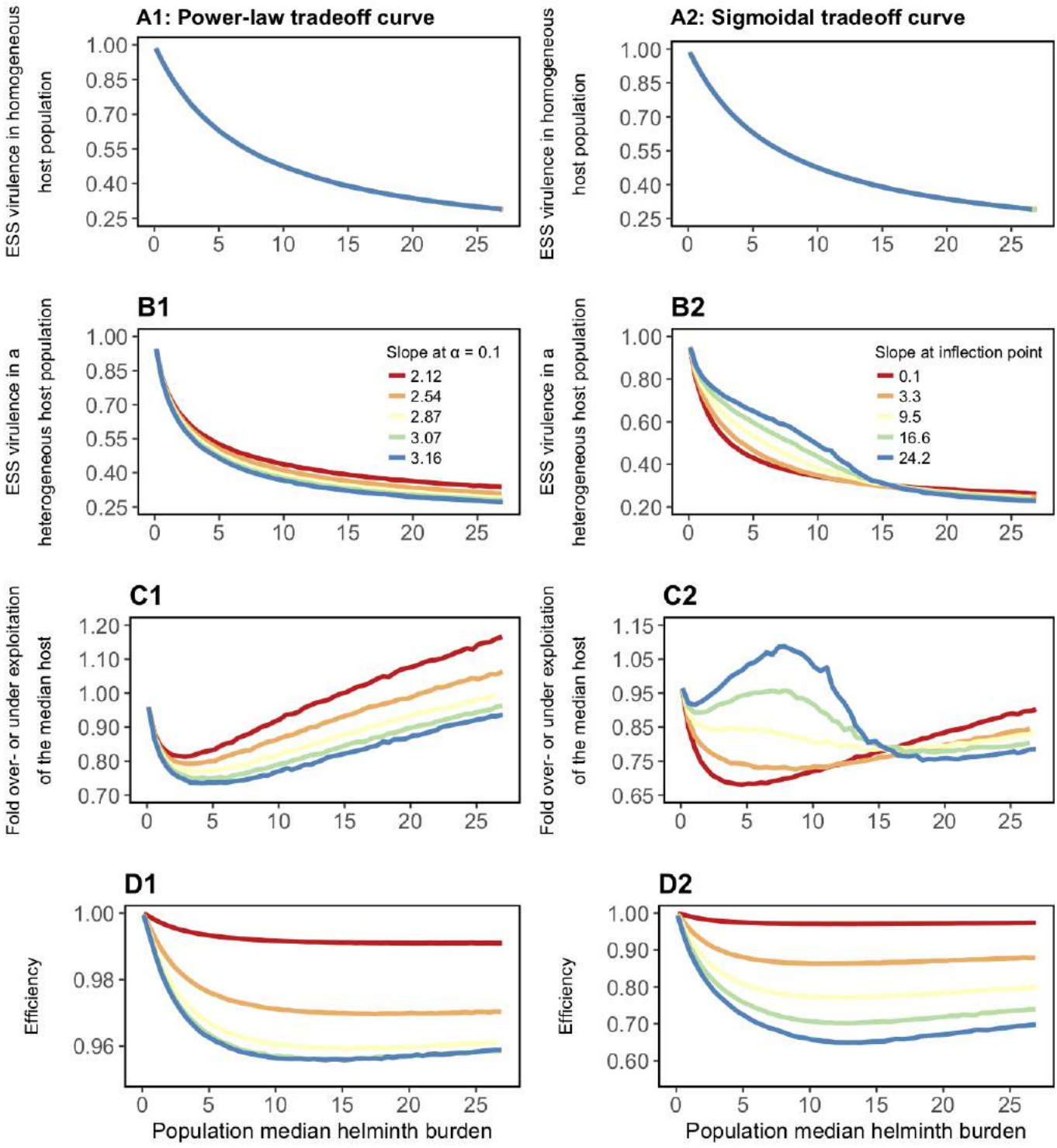

Figure S3. Caption on next page

Caption for Figure S3: General model exploration using a saturating exponential heterogeneity map. Panels A1, B1, C1, and D1 show results using a power-law tradeoff curve and panels A2, B2, C2, and D2 show results using a sigmoidal tradeoff curve. Panels A1 and A2 show the fold change in virulence in a homogeneous population with a helminth burden given by the x-axis relative to a homogeneous population with zero helminths. The curve shown in panels A1 and A2 is the inverse of the saturating exponential heterogeneity map over the range of helminth burdens shown. The complete saturating heterogeneity map is shown in Figure 2 and in the `Shiny R` application. Panels A1 and A2 show the fold change in pathogen virulence attributable to a change in helminth burden. Panels B1 and B2 show the fold change in virulence in a heterogeneous population with the same median helminth burdens as the populations pictured in panels A1 and A2 relative to a homogeneous population with zero helminths. Comparing panels A1 and B1 or A2 and B2 show the effects of changing heterogeneity (e.g. variance) without a change in median. Panels C1 and C2 show median-host exploitation, which is the fold over- or under- exploitation of the host with a median helminth burden when it is infected with the ESS MYXV strain in the heterogeneous population. Specifically, each of the curves in panels C1 and C2 are calculated by dividing the curve in B1 or B2 (which is the ESS virulence of the strain in the heterogeneous population) by the curve is panels A1 or A2 (which is the ESS of a strain optimized to the median of the heterogeneous population). For example, in panel C1 at a median helminth burden of 5, the median host is under-exploited when it is infected with the MYXV strain that is optimized to the entire heterogeneous population; this under-exploitation occurs for all parameters tradeoff curves. Panels D1 and D2 show the efficiency of the ESS pathogen strain in the heterogeneous population, which is defined as the ratio of $\mathcal{R}_0$ at α_f^* in the heterogeneous population and $\mathcal{R}_0$ at α_h^* in a homogeneous population. Efficiency decreases as over- and under-exploitation of hosts increases, such as over- or under-exploitation of the median host which is shown in panels C1 and C2. For parameter values see Table S1.

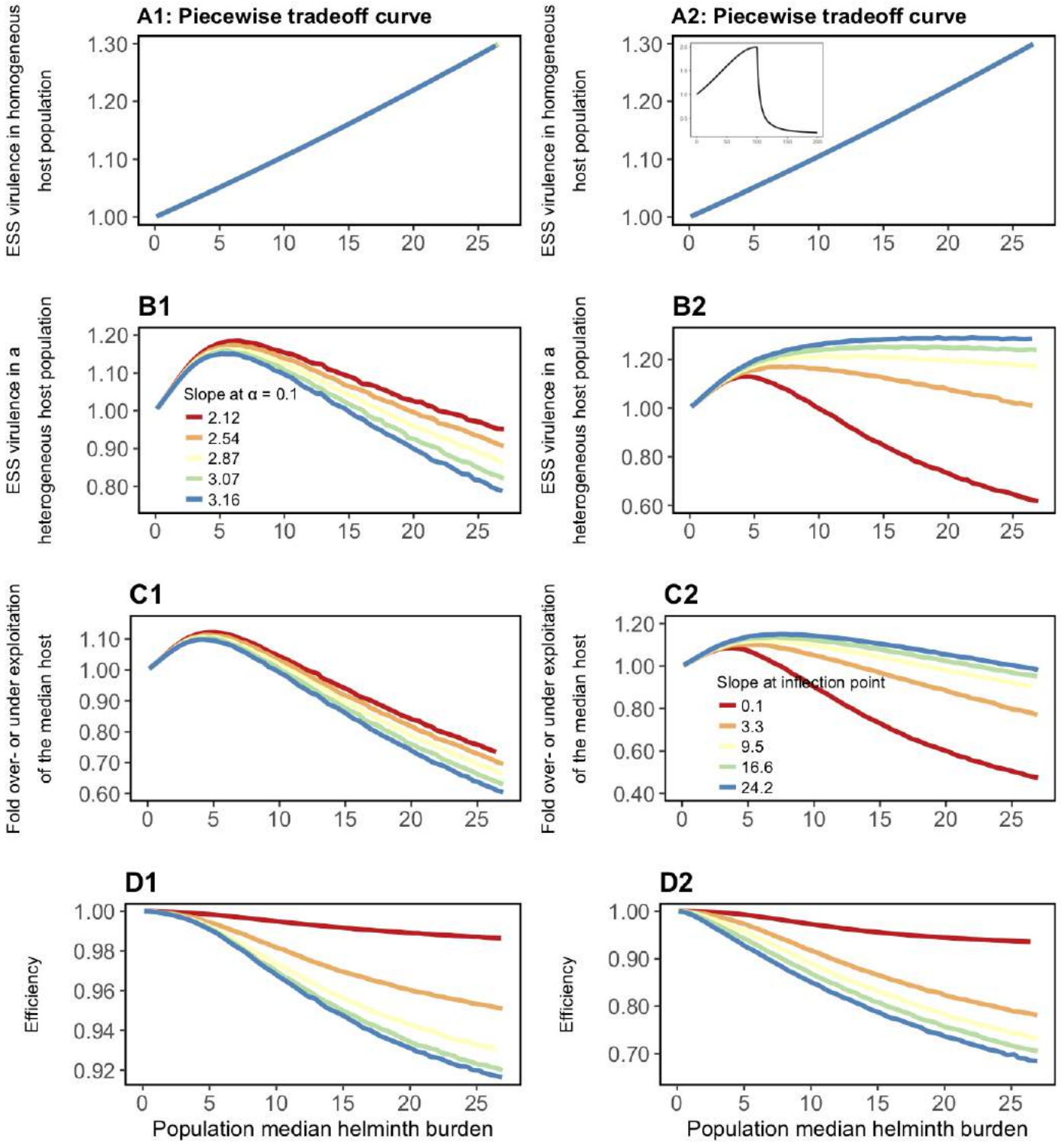

Figure S4. Caption on next page

Caption for Figure S4: General model exploration using a piecewise heterogeneity map. Panels A1, B1, C1, and D1 show results using a power-law tradeoff curve and panels A2, B2, C2, and D2 show results using a sigmoidal tradeoff curve. Panels A1 and A2 show the fold change in virulence in a homogeneous population with a helminth burden given by the x-axis relative to a homogeneous population with zero helminths. The curve shown in panels A1 and A2 is the inverse of the piecewise heterogeneity map over the range of helminth burdens shown. The complete piecewise heterogeneity map is shown in the `Shiny R` application. Panels A1 and A2 show the fold change in pathogen virulence attributable to a change in helminth burden. Panels B1 and B2 show the fold change in virulence in a heterogeneous population with the same median helminth burdens as the populations pictured in panels A1 and A2 relative to a homogeneous population with zero helminths. Comparing panels A1 and B1 or A2 and B2 show the effects of changing heterogeneity (e.g. variance) without a change in median. Panels C1 and C2 show median-host exploitation, which is the fold over- or under-exploitation of the host with a median helminth burden when it is infected with the ESS MYXV strain in the heterogeneous population. Specifically, each of the curves in panels C1 and C2 are calculated by dividing the curve in B1 or B2 (which is the ESS virulence of the strain in the heterogeneous population) by the curve is panels A1 or A2 (which is the ESS of a strain optimized to the median of the heterogeneous population). For example, in panel C1 at a median helminth burden of 5, the median host is under-exploited when it is infected with the MYXV strain that is optimized to the entire heterogeneous population; this under-exploitation occurs for all parameters tradeoff curves. Panels D1 and D2 show the efficiency of the ESS pathogen strain in the heterogeneous population, which is defined as the ratio of $\mathcal{R}_0$ at α_f^* in the heterogeneous population and $\mathcal{R}_0$ at α_h^* in a homogeneous population. Efficiency decreases as over- and under-exploitation of hosts increases, such as over- or under-exploitation of the median host which is shown in panels C1 and C2. For parameter values see Table S1.

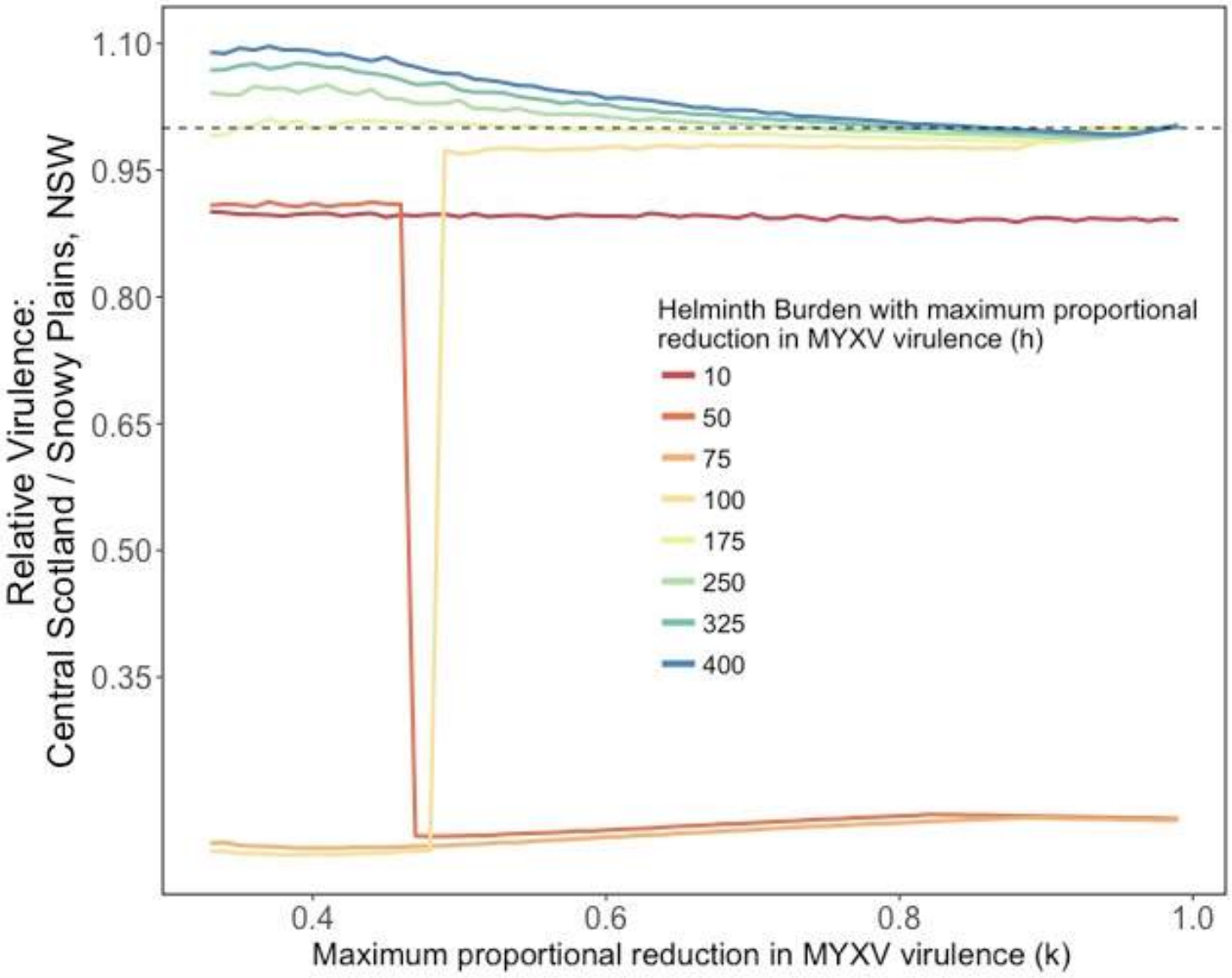

Figure S5. Ratio of ESS relative virulence in Snowy Plains-NSW to ESS virulence in Scotland. Regions above and below the gray reference plane show parameter combinations that lead to higher or lower virulence in Snowy Plains-NSW respectively. Results depicted here use different parameter values for the piecewise heterogeneity map than the main text Figure 4 (see Table S1 for parameter values). Here, with a lower rate of approach to the asymptote r, a larger maximum reduction in realized MYXV virulence k is required for higher ESS virulence in Australia than in Scotland. Regardless, the parameter space leading to higher virulence in Scotland remains small. The rapid change in the ratio of ESS virulence is also observed between Scotland and Australia between a piecewise minimum h of 75 and 100. This precipitous change occurs between 75 and 100 for this parameter set because of a high-variance realized virulence distribution in the Snowy Plains-NSW population.

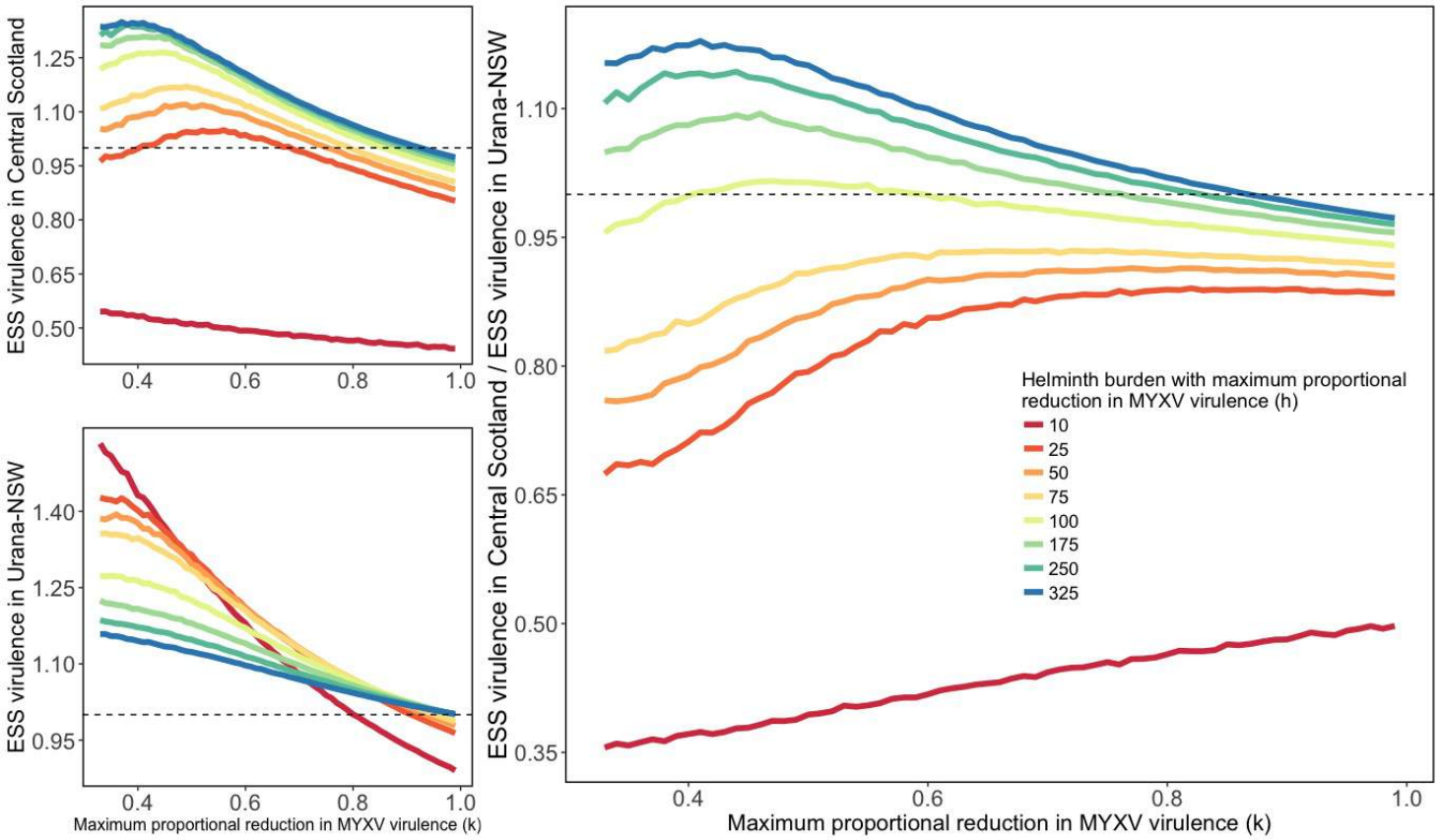

Figure S6. Panels A and B show the fold change in MYXV virulence in Central Scotland and Urana-NSW relative to MYXV virulence in a homogeneous helminth-free rabbit population. In panels A and B regions above and below the dashed reference line show parameter combinations that lead to a higher or lower ESS MYXV virulence in the wild rabbit populations, respectively. Panel C shows the ratio of ESS virulence in Scotland relative to ESS virulence in Snowy Plains-NSW. Here, regions above and below the dashed reference line show parameter combinations that lead to higher or lower virulence in Scotland respectively. For parameter values see Table S1.

Appendix : Additional results for Chapter 5

Supplemental figures, planned as a supplemental component to the work presented as Chapter 5.

Supplemental results for: The evolution of parasites constrained by a virulence-transmission tradeoff and compensatory virulence factor "tuning"

Morgan P Kain[1]* and Benjamin M Bolker[1,2]

*Corresponding author: kainm@mcaster.ca
[1]Department of Biology, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada
[2]Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, L8S 4K1 Hamilton, ON, Canada

Figures

Parasite evolution from high virulence low efficiency

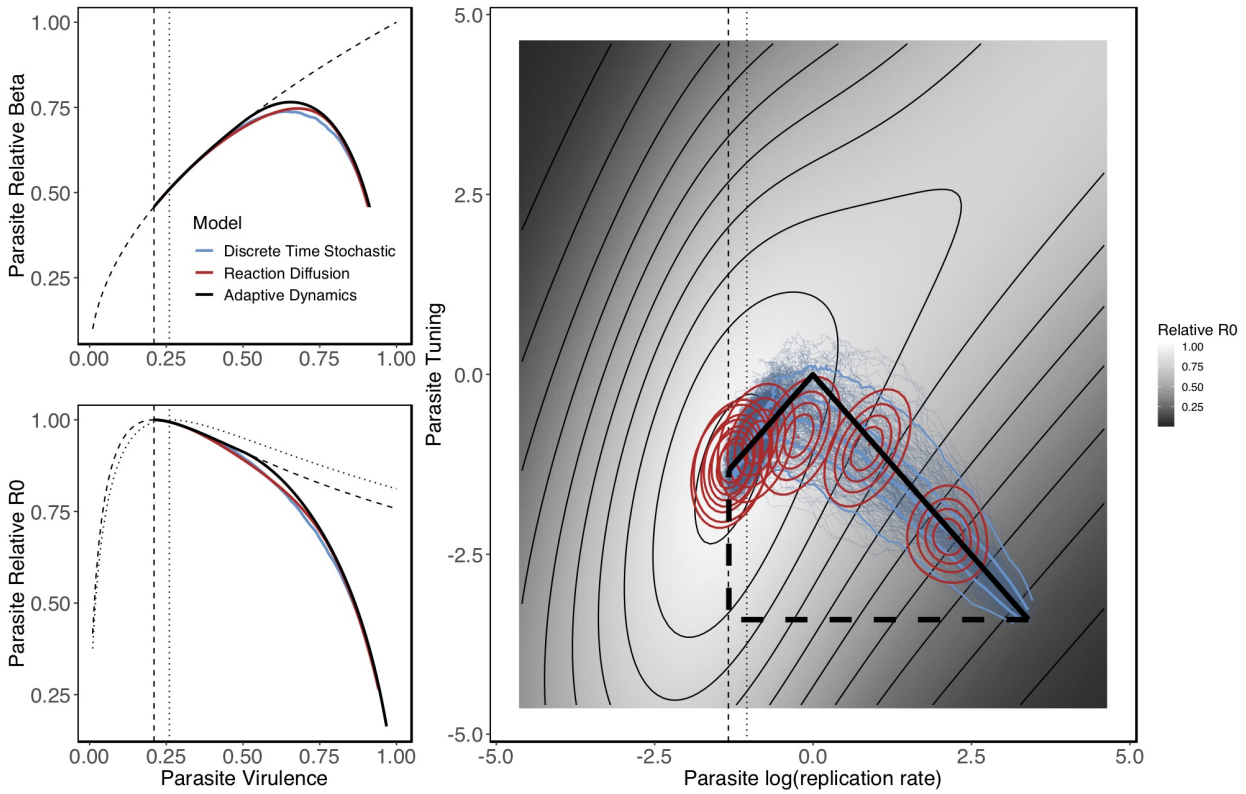

Figure S1: The large right panel shows the evolutionary trajectory of a parasite on its fitness landscape in the DTS model (thin blue lines show 250 stochastic simulations, while the solid blue lines show the median, 50%, and 95% quantities of these simulations), RD model (red ellipses show 95% of the distribution of parasite strains), and AD model (bold black solid line shows simultaneous evolution in two traits; dashed lines show evolution in a single trait per time step), using the same parameter values as main text Figure 1, except here parasites evolve from high replication rate and low tuning. The surface pictured is calculated using the $\mathcal{R}_0$ equation for rates (Eq.1), and has been scaled so that the adaptive peak has a value of 1 (shown in light grey). Due to the slight difference in $\mathcal{R}_0$ for the RD and DTS models (Eq.2 vs Eq.3), the DTS surface is marginally different (see Figure S3 for a comparison of these surfaces). The thin dashed vertical line shows the location of the adaptive peak for this surface; the thin dotted vertical line shows the location of the peak for the DTS model. The top left and bottom left panels show a parasite evolving in log(replication rate) and tuning, according to the right panel, translated into virulence and transmission following a power-law tradeoff (top panel) and virulence and $\mathcal{R}_0$ (bottom panel). The dashed curve in the bottom left panel shows parasite $\mathcal{R}_0$ for the RD and AD models; the dotted curves shows parasite $\mathcal{R}_0$ for the DTS model. Maximum $\mathcal{R}_0$ is designated by the vertical dashed and dotted lines respectively. All models have the same power-law tradeoff curve (top right panel).

Parasite evolution from low virulence and high efficiency

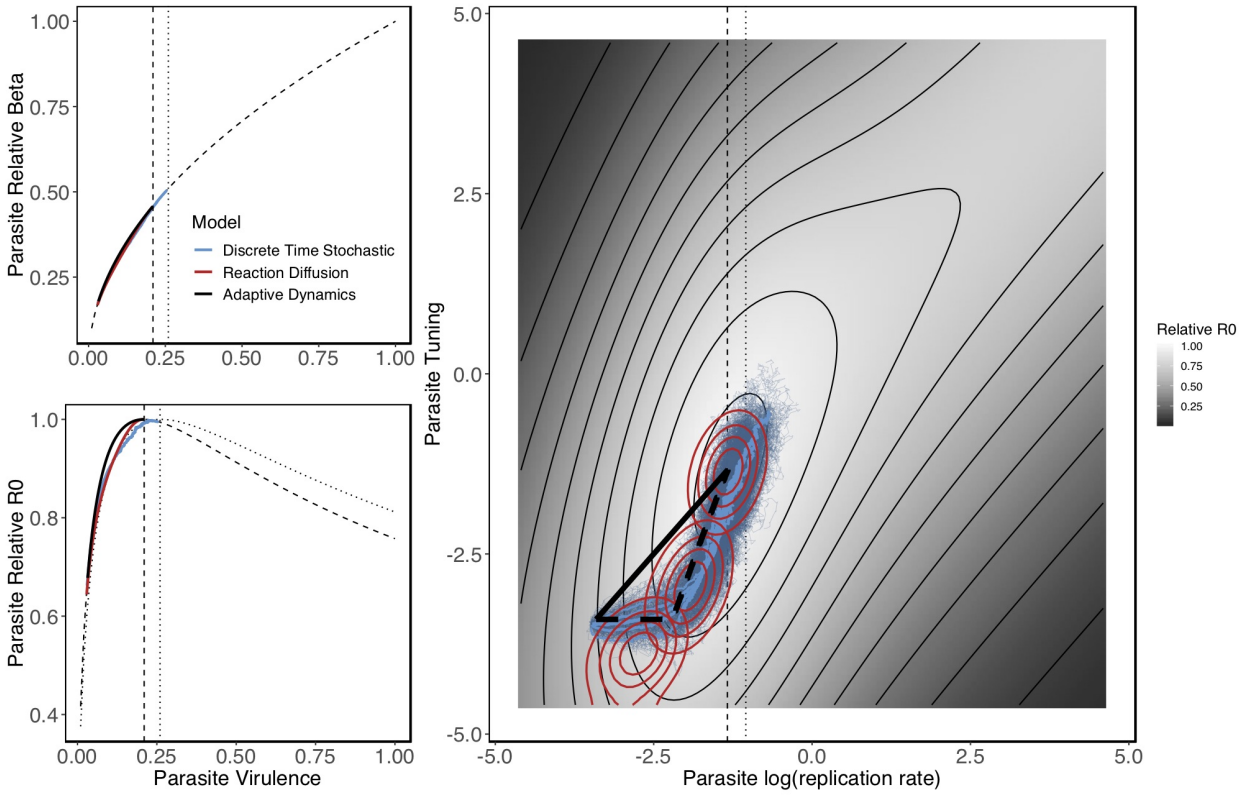

Figure S2: For a complete figure caption see Figure S1. Panels show the evolutionary trajectory of a parasite in the DTS (blue lines), RD (red ellipses), and AD (black solid and dashed lines) models, using the same parameter values as main text Figure 1, except here parasites evolve from low replication rate and low tuning.

Comparison of fitness landscapes

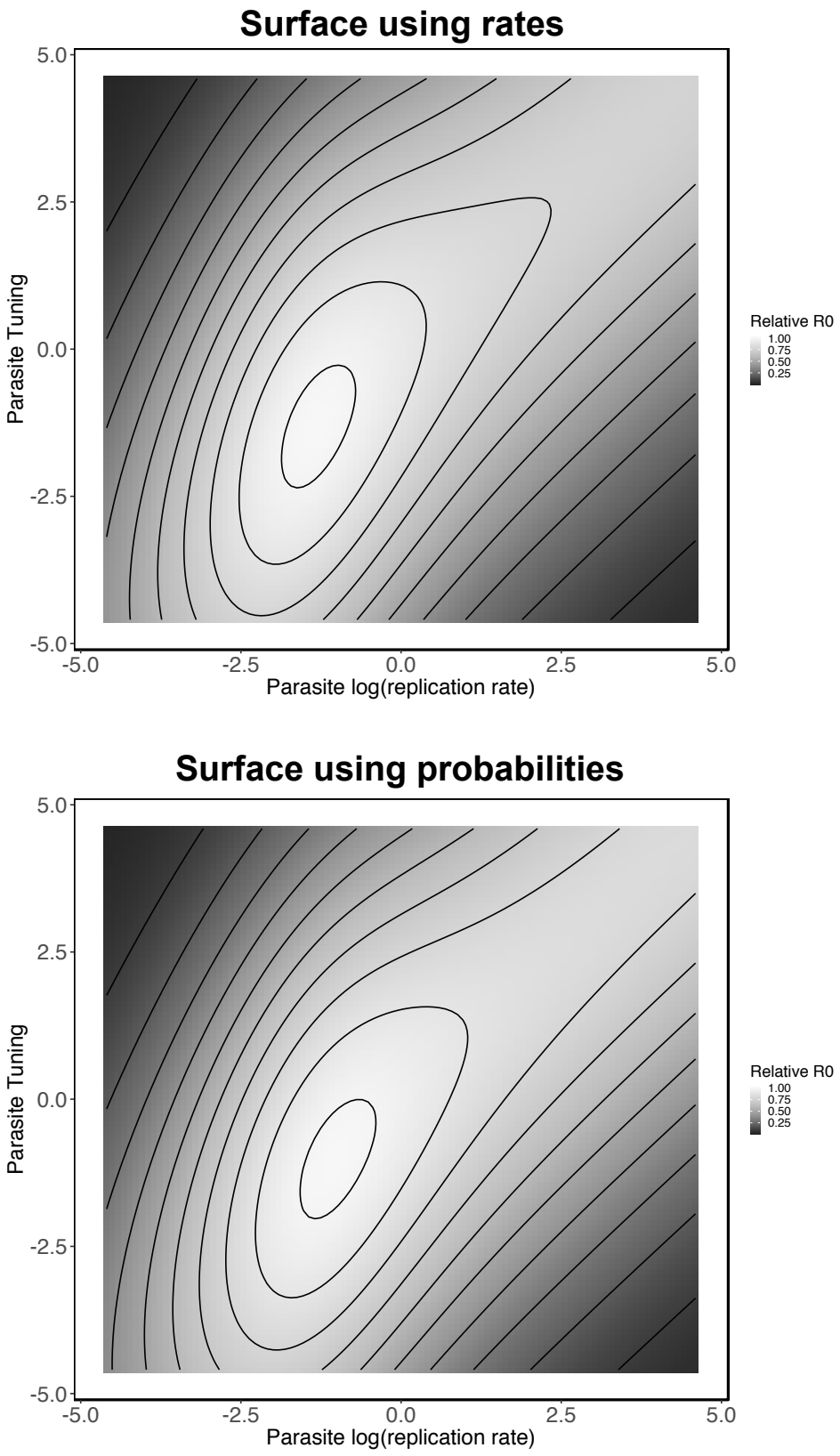

Figure S3: Top and bottom panels show a fitness landscape with rates (Main text Eq.2) and probabilities (Main text Eq.3) respectively using recovery (γ) of 0.2 and a background death (μ) of 0.01. Parasite optimal virulence in the top panel is 0.21 and 0.26 in the bottom panel.

DTS dynamics when a parasite evolves from high virulence and low efficiency

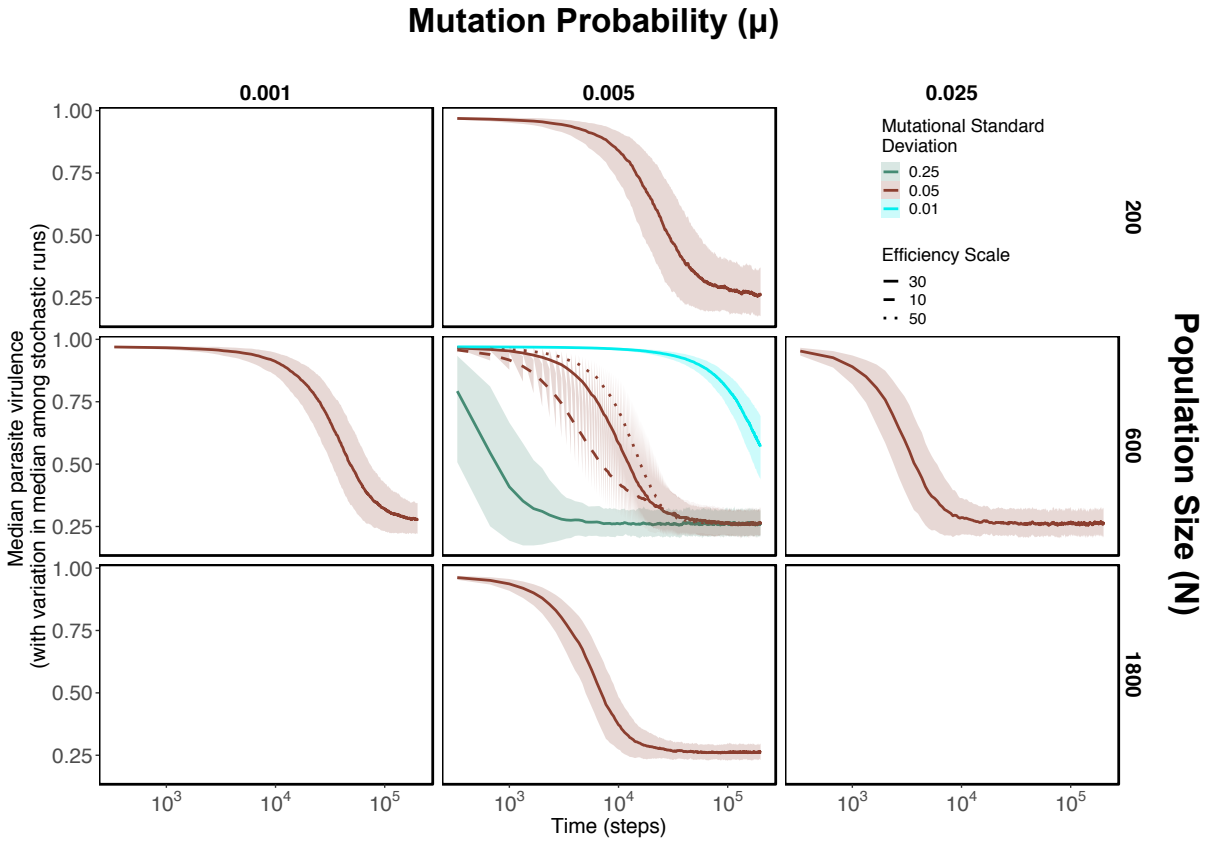

Figure S4: Parasite virulence for all parameter combinations presented in Table 2 for a parasite evolving from high virulence and low efficiency. Bold lines represent medians across 250 stochastic simulations; shaded envelopes show the 95% quantiles. The correspondence between each "Model Outcome" presented in Table 2, from left to right, are as follows: 1) Median number of circulating strains: not pictured; 2) Time to optimum virulence: Time at which median virulence first reaches 0.26; 3) Maximum transient virulence: not relevant for these starting values; 4) Time with greater than optimum virulence: same as time to optimum virulence for these starting values; 5) Departure from optimum virulence: width of the shaded envelope once a virulence of 0.26 is reached.)

Parasite evolution with high mutational standard deviation

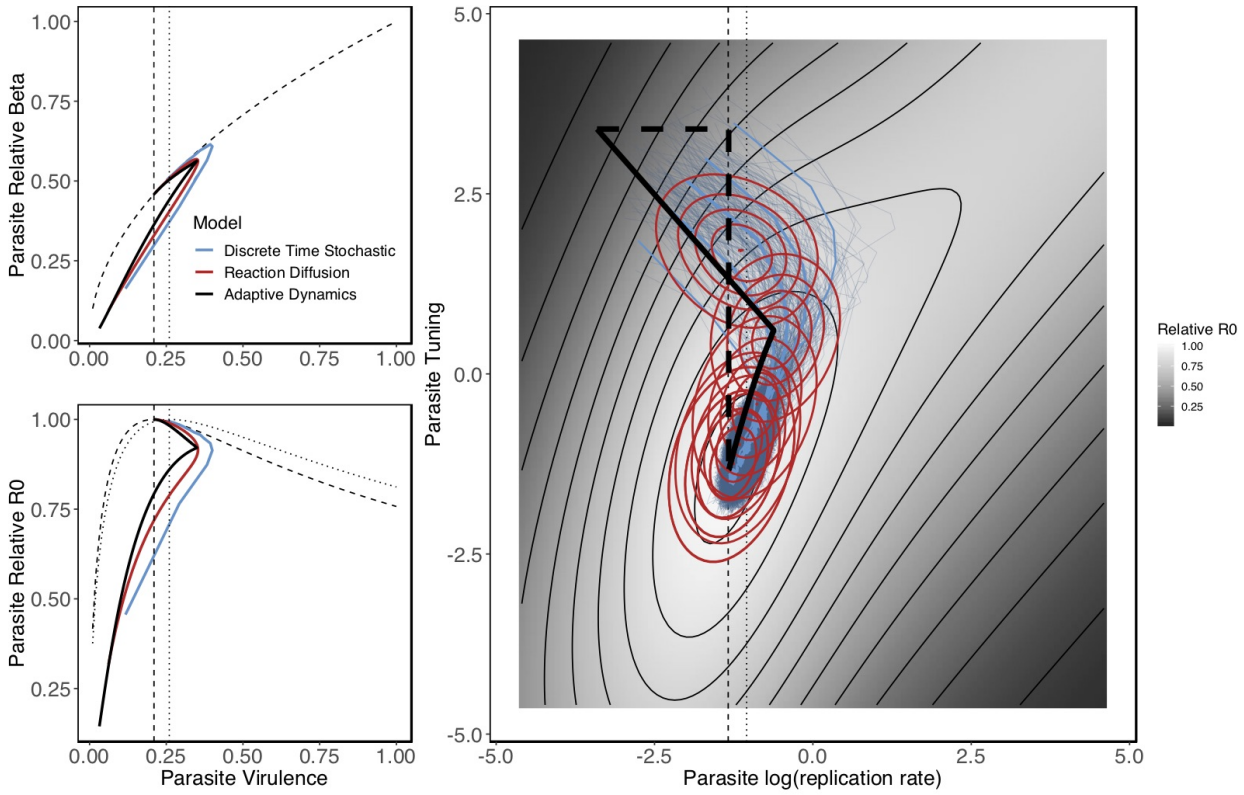

Figure S5: For a complete figure caption see Figure S1. Panels show the evolutionary trajectory of a parasite in the DTS (blue lines), RD (red ellipses), and AD (black solid and dashed lines) models, using the same parameter values as main text Figure 1, except with: DTS: $\sigma = 0.25$; RD: $\sigma = 0.39$. A numeric summary of the DTS model for these parameter values is available in main text Table 2 row 5.

Parasite evolution with a small population size

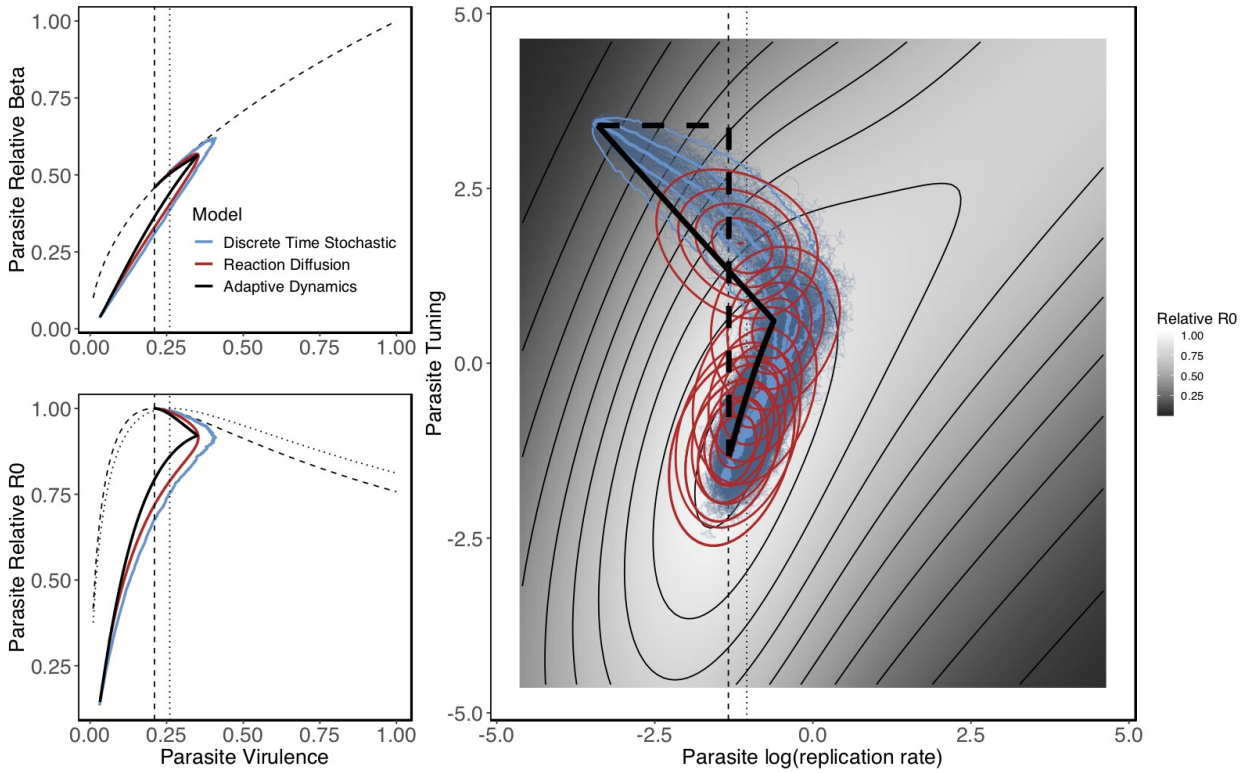

Figure S6: For a complete figure caption see Figure S1. Panels show the evolutionary trajectory of a parasite in the DTS (blue lines), RD (red ellipses), and AD (black solid and dashed lines) models, using the same parameter values as main text Figure 1, except with N = 200. A numeric summary of the DTS model for these parameter values is available in Table 2 row 7.