

**SPATIAL EXTENSION OF THE THEORETICAL
FRAMEWORK OF THE ADAPTIVE
VIRULENCE EVOLUTION HYPOTHESIS**

by

© *Abdou Moutalab Fofana*

A thesis submitted to the School of Graduate Studies
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Department of *Biology*
Memorial University of Newfoundland

March 2020

St. John's

Newfoundland

Abstract

The adaptive virulence hypothesis states that parasites cause death to their hosts because virulence is beneficial for the transmission and spread of parasites. A growing body of empirical evidence supports the adaptive virulence hypothesis but more examples are needed for its empirical validation. The classic mathematical framework of the adaptive virulence hypothesis does not account for host population structure which can have important implications for virulence evolution. The goal of this thesis is to address the broad applicability problems and extend spatially the mathematical framework of the adaptive virulence hypothesis by accounting for host movement in the model. My thesis examines whether virulence is adaptive at the species level by investigating the relationship between virulence and parasite fitness using simulation data. I find that virulence and parasite fitness, measured as the basic reproduction number R_0 , are correlated at the between-species level and the exact form of the relationship depends on the selective pressures within each group of parasite species. Also, I break free from the classic framework of the adaptive virulence hypothesis to investigate virulence evolution when parasites reduce host movement. The results explain the transient coexistence of low- and high-virulence strains in Avian influenza viruses. I reviewed epidemic models with host movement to understand what aspects of disease spread are important to develop a spatially extended model for virulence evolution. I find that epidemic models with spatially heterogeneous epidemiological parameters, like disease transmissibility, are suitable for the spatial extension of the mathematical framework of the adaptive virulence hypothesis. Finally, I investigated the relationship between the temporal and the spatial spread of infectious diseases. I

show an inverse relationship between the initial epidemic growth rate and the spatial spread rate which may reflect a trade-off between parasites dispersal and transmission. Overall my thesis opens interesting research avenues for future works to formulate spatially explicit models for the evolution of virulence and makes significant contributions to the empirical investigation of the adaptive virulence hypothesis and the role of animal movement for disease spread and virulence evolution.

— MUN School of Graduate Studies

Acknowledgements

I thank my supervisor Amy Hurford for her guidance and precious support and making me a better research scientist. Amy Hurford was supported by an NSERC Discovery Grant (RGPIN 2014-05413). I thank my committee members, Eric Vander Wal and Sharene D. Bungay, for their helpful advice throughout this Ph.D. I thank Oliver Stueker for helping me to run the disease outbreak simulations of the chapter one on Beluga and Cedar clusters, which are Compute Canada computing resources. I thank S. Andrews, S.R. Chalise J. Ebel, F. Frazão, J. MacSween, J. Mariño, A. McLeod, E.J. Moran, M. Rittenhouse, M. Rizzuto, S. Yalcin for their helpful comments on the different manuscripts of this thesis. I thank Ted Miller and his wife Margarita for their warm and wide opened arms and for giving me a family home here in Newfoundland. I thank my wife G. Fafa for her patience and precious support during this wonderful Ph.D experience. Special thanks to my mum, Abdoulaye Issifou, Azia G. Ayeva, Amda M. Madougou and D. Ouro-Sama and his family. I also thank M. Gadegbeku, J. Baboulene and his wife C. Salmon for their support.

— MUN School of Graduate Studies

Contents

Abstract	ii
Acknowledgements	iv
1 Introduction and overview	2
Co-authorship Statement	8
2 Chapter one: Is virulence adaptive? A numerical investigation with cross-species disease outbreak data	
This chapter is a prepared manuscript for submission to the journal	
<i>PLOS Computational Biology</i>	9
2.1 Introduction	10
2.2 Methods	13
2.2.1 Phylogeny and species traits simulation	16
2.2.2 Disease outbreaks simulation	19
2.2.3 Parasite fitness and virulence measures	21
2.2.4 Statistical analysis	22
2.3 Results	22
2.3.1 Virulence is adaptive within most of the species	23
2.3.2 Cross-species data show that virulence is adaptive at least for low virulence levels	25
2.4 Discussion	28
2.4.1 Why is virulence beneficial for parasite transmission in some species?	28

2.4.2	Can we determine whether virulence is adaptive using empirical cross-species data ?	33
2.5	Conclusion	35
3	Chapter two: Parasite-induced shifts in host movement may explain the transient coexistence of high- and low-pathogenic disease strains	
	This chapter has been submitted to the <i>Journal of Evolutionary Biology</i> and was rejected with re-submission allowed	44
3.1	Introduction	46
3.2	Epidemiological model	49
3.3	Evolution model	53
3.4	Results	56
3.4.1	The evolutionarily stable within-host parasite net replication rate (ESS α^*)	57
3.4.2	Evolutionary bistability arises when hosts make contacts during lethargy	58
3.4.3	The effects of model parameters on the evolutionary dynamics	63
3.4.4	Evolutionary dynamics when parasite infection is non-lethal	64
3.5	Discussion	66
4	Chapter three: Mechanistic movement models to understand epidemic spread	
	This chapter has been published in the journal <i>Philosophical Transactions of the Royal Society</i>	87

4.1	Introduction	89
4.2	The mass-action law	92
4.3	Epidemics when host movement is random	93
4.4	Epidemics when host movement direction is biased or temporally au- tocorrelated	99
4.5	Epidemics when host individuals are discrete	101
4.5.1	Uncorrelated random walks	102
4.5.2	Levy walks	104
4.6	Case study: rabies	105
4.7	Concluding remarks and perspectives	108

5 Chapter four: A counter-intuitive relationship between the temporal and spatial spread of diseases

This chapter is a prepared manuscript for submission to the journal

	<i>Emerging Infectious Diseases</i>	118
5.1	Introduction	119
5.2	Method	124
5.2.1	Disease outbreak data	124
5.2.2	Estimation of disease spatial spread rate c	125
5.2.3	Estimation of disease epidemic growth rate r	128
5.2.4	Statistical analysis	130
5.3	Results	130
5.3.1	Epidemic growth rate r and spatial spread rate c are inversely related	131

5.3.2	Airborne infections grow slowly and spread faster compared to tick-borne infections	132
5.4	Discussion	137
5.4.1	The inverse relationship between c and r can reflect a link between the spatial dispersal and the transmission of parasites .	137
5.4.2	Disease spread rate c and epidemic growth rate r in tick-borne and airborne infections	139
5.5	Concluding remarks	141
6	Summary	152
	Appendix A	154
	Appendix B	160
	Appendix C	176

List of Tables

2.1	List of symbols used in the main text	16
2.2	Results summary of fitness-virulence correlations	23
3.1	List of notations and definitions	50
4.1	List of abbreviations and symbols used in the main text	91
4.2	Summary movement types and R_0	106
5.1	Model comparison	132
5.2	Model comparison	134
C.1	Model comparison r and years	180

List of Figures

2.1	Cross-species data simulation framework	14
2.2	Phylogenetic tree, strains diversification and trade-offs	15
2.3	Abouheif's phylogenetic correlation tests	20
2.4	Within-species fitness-virulence correlation	24
2.5	Cross-species fitness-virulence correlation	26
2.6	Within-selective regimes fitness-virulence correlation	27
2.7	R_0 , infectious period and transmissibility	30
3.1	Epidemiological model	49
3.2	The evolutionary trade-offs	58
3.3	PIP and dynamical simulations	60
3.4	R_0 in moving and resting states	61
3.5	Backward bifurcation	64
3.6	PIP when infection is non-lethal	66

4.1	URWs, BRWs and LWs	95
4.2	Mass-action law	97
4.3	PDE vs IBM	103
5.1	GAM and contour plots	127
5.2	Exponential function fit	129
5.3	Correlation c and r and years	133
5.4	Airborne vs tick-borne infections	135
5.5	Airborne and tick-borne infections	136
A.1	Epidemic Movie and Graph	155
A.2	Proportion of within-species data	156
A.3	within-taxa R-virulence	157
A.4	Concave-down fitness-virulence with proportions	158
A.5	Linear fitness-virulence with proportions	159
B.1	PIP Movie	172
B.2	Movie PIP	172
B.3	Movie PIP	174
C.1	Preliminary analysis c data	177
C.2	Preliminary analysis transmission route c	178
C.3	Preliminary analysis r	179

1. Introduction and overview

Virulence is a polysemic term which is often measured as host death due to parasite infection or parasite-induced host sterility by theoretical biologists, case fatality by epidemiologists and the degree of anemia, lethargy or weight loss due to infection by experimental biologists (Poulin and Combes 1999; Casadevall and Pirofski 1999; Mackinnon and Read 1999; Casadevall and Pirofski 2001; Thomas and Elkinton 2004; Nishiura 2010; Poulin 2011; Ejima et al. 2012; Casadevall 2017). The general definition of virulence is the harm parasites cause to their hosts, and the harm can be host mortality and sub-lethal measures like anemia, sterility and reduced host mobility (Ewald 1994).

Ebola virus, smallpox virus and Rhinovirus (which causes common cold in humans) can cause 80 %, 10 % and less than 0.0001 % death in infected humans respectively (Georges et al. 1999; Walther and Ewald 2004; Beeching et al. 2014). Because the harm parasites cause to their hosts can reduce parasite transmission and survival, it is paradoxical why some parasites kill, cause severe lethargy or castrate their hosts. The adaptive virulence hypothesis, also known as the trade-off hypothesis, suggests that virulence is an adaptation that is beneficial for parasite transmission and spread in host populations (Anderson and May 1982; Ewald 1983; Alizon and Michalakis 2015). The adaptive virulence hypothesis has received some empirical support during the past three decades (Alizon et al. 2009; Alizon and Michalakis 2015; Cressler et al. 2016) and the theoretical framework has been extended to include parasite transmission modes and routes (Ewald 1983; 1991; Berngruber et al. 2015), parasite competition within the host (Levin and Bull 1994; de Roode et al. 2005; Bell et al. 2006) and the spatial structure of host populations (Boots and Meador 2007; Kamo et al. 2007;

Messinger and Ostling 2013).

The extent of spatial spread of infectious diseases can be an indication of parasite fitness because a parasite can achieve higher fitness by spreading globally, and as such the spatial aspects of disease spread can affect virulence evolution. Also, parasite infection often affect host movement by inducing lethargy which can impact the contact rate between infected and susceptible hosts, epidemic spread and ultimately the evolution of virulence. As such, the explicit incorporation of host movement in the theoretical framework of the adaptive virulence evolution hypothesis is a promising research avenue for understanding why some parasites kill their hosts.

This thesis will investigate whether different types of host movement affect the transmission potential and the spatial propagation of parasites in host populations. The main goal of this thesis is to extend the adaptive virulence evolution hypothesis by accounting for host movement in the mathematical framework. The following three reasons explain why the incorporation of the spatial aspects of disease spread in the mathematical framework of the adaptive virulence hypothesis is important. Firstly, the spatial extension of the mathematical framework of the adaptive virulence hypothesis is necessary to investigate the often suspected trade-off between parasite transmission and host movement (e.g., transmission-host-dispersal distance trade-off). Secondly, a trade-off between parasite transmission and host movement can naturally emerge from the spatiotemporal dynamics of host-parasite interactions, and the spatial extension of the mathematical framework of the adaptive virulence hypothesis can help in opening this research avenue. Thirdly, the adaptive virulence hypothesis must be widely applicable to be used as a framework for virulence management. The spatial extension of the mathematical framework of the adaptive virulence

hypothesis can help in improving the predictive power of the current model and help in making effective disease control decisions.

This thesis is divided into 4 chapters corresponding to 4 different research papers. The first chapter addresses the issues of the broad applicability of the adaptive virulence hypothesis. I investigate whether we can determine if virulence is adaptive using simulated cross-species data. In the second chapter, I develop a mathematical model to investigate the evolution of virulence when parasites cause lethargy, which is decreased host movement due to infection. The third chapter reviews epidemic models with animal movement, and investigates whether different types of animal movement affects the risk of disease spread. The last chapter investigates the relationship between the temporal and the spatial spread of infectious diseases. Finally, I summarize the main results of the thesis and discuss the significance of my findings.

Literature cited

- S. Alizon and Y. Michalakis. Adaptive virulence evolution: the good old fitness-based approach. *Trends in Ecology & Evolution*, 30(5):248–254, 2015.
- S. Alizon, A. Hurford, N. Mideo, and M. Van Baalen. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *Journal of Evolutionary Biology*, 22(2):245–259, 2009.
- R.M. Anderson and R.M. May. Coevolution of hosts and parasites. *Parasitology*, 85(02):411–426, 1982.
- N.J. Beeching, M. Fenech, and C.F. Houlihan. Ebola virus disease. *BMJ*, 349:g7348, 2014.
- A.S. Bell, J.C. de Roode, D. Sim, and A.F. Read. Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success. *Evolution*, 60(7):

- 1358–1371, 2006.
- T.W. Berngruber, S. Lion, and S. Gandon. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS pathogens*, 11(4):e1004810, 2015.
- M. Boots and M. Meador. Local interactions select for lower pathogen infectivity. *Science*, 315(5816):1284–1286, 2007.
- A. Casadevall. The pathogenic potential of a microbe. *mSphere*, 2(1):e00015–17, 2017.
- A. Casadevall and L.A. Pirofski. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. *Infection and immunity*, 67(8):3703–3713, 1999.
- A. Casadevall and L.A. Pirofski. Host-pathogen interactions: the attributes of virulence. *Journal of infectious Diseases*, 184(3):337–344, 2001.
- C.E. Cressler, D.V. McLeod, C. Rozins, J. Van Den Hoogen, and T. Day. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology*, 143(7):915–930, 2016.
- J.C. de Roode, R. Pansini, S.J. Cheesman, M.E. Helinski, S. Huijben, A.R. Wargo, A.S. Bell, B.H. Chan, D. Walliker, and A.F. Read. Virulence and competitive ability in genetically diverse malaria infections. *Proceedings of the National Academy of Sciences*, 102(21):7624–7628, 2005.
- K. Ejima, R. Omori, B.J. Cowling, K. Aihara, and H. Nishiura. The time required to estimate the case fatality ratio of influenza using only the tip of an iceberg: joint estimation of the virulence and the transmission potential. *Computational and mathematical methods in medicine*, 2012, 2012.
- P.W. Ewald. Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics*, pages 465–485, 1983.
- P.W. Ewald. Waterborne transmission and the evolution of virulence among gastrointestinal bacteria. *Epidemiology and infection*, 106(01):83–119, 1991.

- P.W. Ewald. *Evolution of infectious disease*. Oxford University Press on Demand, 1994.
- A.J. Georges, E.M. Leroy, A.A. Renaut, C.T. Benissan, R.J. Nabias, M.T. Ngoc, P. Obiang, J.P.M. Lepage, E.J. Bertherat, D.D. Bénoni, et al. Ebola hemorrhagic fever outbreaks in Gabon, 1994–1997: epidemiologic and health control issues. *The Journal of infectious diseases*, 179(Supplement_1):S65–S75, 1999.
- M. Kamo, A. Sasaki, and M. Boots. The role of trade-off shapes in the evolution of parasites in spatial host populations: an approximate analytical approach. *Journal of Theoretical Biology*, 244(4):588–596, 2007.
- B.R. Levin and J.J. Bull. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends in Microbiology*, 2(3):76–81, 1994.
- M.J. Mackinnon and A. Read. Selection for high and low virulence in the malaria parasite. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1420):741–748, 1999.
- S.M. Messinger and A. Ostling. The influence of host demography, pathogen virulence, and relationships with pathogen virulence on the evolution of pathogen transmission in a spatial context. *Evolutionary ecology*, 27(2):353–380, 2013.
- H. Nishiura. Case fatality ratio of pandemic influenza. *The Lancet infectious diseases*, 10(7):443–444, 2010.
- R. Poulin. *Strategies of host exploitation*, chapter 5, page 66. Evolutionary ecology of parasites. Princeton university press, 2011.
- R. Poulin and C. Combes. The concept of virulence: interpretations and implications. *Parasitology Today*, 15(12):474–475, 1999.
- S.R. Thomas and J.S. Elkinton. Pathogenicity and virulence. *Journal of invertebrate pathology*, 85(3):146–151, 2004.
- B.A. Walther and P.W. Ewald. Pathogen survival in the external environment and the

evolution of virulence. *Biological Reviews*, 79(4):849–869, 2004.

Co-authorship Statement

A.M Fofana is principal author, contributed to identify research proposal, performed mathematical and statistical analysis and prepared the manuscripts for publication for all the chapters.

A. Hurford supervised the research, provided financial assistance, is co-author, contributed to identify research proposal and provided guidance for analysis and to prepare all the manuscripts for publication.

S. Alizon is co-author, contributed to identify research proposal and provided guidance for analysis and to prepare the manuscript of the chapter I (Is virulence adaptive? A numerical investigation with cross-species disease outbreak data) for publication.

— MUN School of Graduate Studies

2. Chapter one: Is virulence adaptive? A numerical investigation with cross-species disease outbreak data

This chapter is a prepared manuscript for submission to the journal *PLOS Computational Biology*

Is virulence adaptive? A numerical investigation with cross-species disease outbreak data

Abdou M. Fofana, Amy Hurford, Samuel Alizon

Abstract

The adaptive virulence hypothesis states that virulence can persist because it is linked with increased parasite spread. This hypothesis is supported by empirical data but validation is lacking at the cross-species level because of empirical and statistical challenges. To investigate if virulence is adaptive, we simulated the evolution of 50 species with 30 strains within each species and assumed that life history trade-offs are different within species, but qualitatively similar across species. We then simulated outbreak data for the 1500 phylogenetically related parasites and used the phylogenetic comparative approach to investigate whether higher virulence parasites have higher fitness, virulence being defined as the inverse of the time to host death due to infection and fitness being measured as the basic or the effective reproduction number (R_0 or R). We find that virulence and R_0 are correlated at the cross-species level, and that the exact form of the relationship depends on the selective pressures within each group of parasite species, allowing us to distinguish extreme parasite lifestyles (obligate killers and sub-lethal infections). Finally, we discuss open methodological challenges for testing the adaptive virulence hypothesis at the cross-species level.

Keywords: Basic reproduction number, trade-off, virulence, phylogenetic comparative method, Ornstein-Uhlenbeck.

2.1. Introduction

1 The conventional wisdom states that parasites and their hosts will always coevolve
2 towards low virulence because a highly virulent parasite that kills its host reduces its

3 own transmission, reproduction and survival (Smith 1904; Ball 1943; Burnet and
4 White 1972; Méthot 2012). This view has been challenged by the adaptive virulence
5 hypothesis, commonly known as the trade-off hypothesis, which states that if viru-
6 lence is beneficial for parasite transmission and spread in host populations then high
7 virulence will be maintained (Ewald 1983; May and Anderson 1983). Within-species
8 data support the adaptive virulence hypothesis for some parasites species, but cross-
9 species empirical tests are lacking mostly due to methodological challenges (Alizon
10 and Michalakis 2015; Cressler et al. 2016). This study addresses these challenges using
11 simulation data.

12 At the within-species level, HIV-1 in humans (Fraser et al. 2007), the protozoan
13 parasite *Ophryocystis elektroscirrha* in monarch butterflies (de Roode et al. 2008) and
14 the malaria parasite *Plasmodium falciparum* in humans (Mackinnon and Read 2004),
15 Cauliflower mosaic virus in turnips (Doumayrou et al. 2013) are some of the host-
16 parasite interactions that exhibit the strongest evidence that intermediate virulence
17 maximizes parasite fitness. More recently, evidence from dengue virus infections has
18 been put forward suggesting that intermediate viral loads in human could maximize
19 the transmission potential of the virus to mosquito vectors (Ben-Shachar and Koelle
20 2018). In this latter example, the trade-off is governed by the host recovery rate rather
21 than the virulence. Other examples have been reported and appropriately reviewed
22 elsewhere, and more examples are needed because the validity of the adaptive virulence
23 hypothesis depends on its wide applicability (Alizon et al. 2009; Froissart et al. 2010;
24 Bolker et al. 2010; Chapuis et al. 2012; Råberg 2012; Williams et al. 2014; Alizon and
25 Michalakis 2015; Cressler et al. 2016).

26 A number of studies have also compared parasite virulence at the cross-species

27 level for different transmission routes. One important prediction of the adaptive viru-
28 lence hypothesis is that high virulence can be maintained when parasites rely less on
29 the host for their transmission (Ewald 1983; Frank 1996). As such, parasites that sur-
30 vive longer in the environment, waterborne and vector-borne parasites are expected
31 to evolve higher virulence compared to directly transmitted parasites, a hypothesis
32 known as the ‘Curse of the pharaoh’ (Bonhoeffer et al. 1996; Gandon 1998; Boldin
33 and Kisdi 2012). Empirical evidence supporting this prediction has been found in
34 human gastrointestinal bacteria, where a positive relationship between virulence and
35 the proportion of disease outbreaks caused by waterborne bacteria has been reported
36 (Ewald 1991a;b). Similarly, Walther and Ewald (2004) compared virulence of infec-
37 tions caused by human respiratory viruses and found that the ones that survive longer
38 in the environment cause higher mortality rate in humans. A more recent study has
39 investigated the relationship between within-host parasite growth rate and virulence
40 and parasite transmission using between-parasite-species data (Leggett et al. 2017),
41 but no previous study has directly tested whether higher virulence parasites have
42 higher fitness at the cross-species level.

43 Cross-species data can be used to test the adaptive virulence hypothesis because
44 variations in virulence across parasite species can be large, and cross-species compar-
45 ative approaches can help in understanding why some parasite species cause higher
46 virulence than others (Frank and Schmid-Hempel 2008). However, cross-species data
47 is phylogenetically structured such that closely related species are more similar than
48 distant species, and as such, virulence measured from different parasite species can-
49 not be considered as statistically independent observations (Rohle 2006; Pavoine et al.
50 2008; Jombart and Dray 2010). Therefore, statistical tools for investigating whether

51 virulence is adaptive using cross-species data must correct for the effect of phylogeny
52 (Felsenstein 1985; Grafen 1989; Pagel and Harvey 1988; Harvey and Pagel 1991; Pagel
53 and Harvey 1992; Hadfield and Nakagawa 2010). The problem is that reconstructing
54 the phylogeny is often unfeasible. The empirical data that is often available is infec-
55 tion cases caused by parasites from a wide range of taxa, including viruses, bacteria,
56 fungi, protozoa and worms and constructing a phylogenetic tree for these infectious
57 organisms can be complicated (but see Iyer et al. 2001; Kühnert et al. 2011; Koonin
58 et al. 2015; Leggett et al. 2017).

59 To investigate whether we can determine if virulence is adaptive, we simulated
60 within- and cross-species data and analyzed the relationship between virulence, de-
61 fined as the inverse of the time to host death due to infection, and parasite fitness,
62 defined as the basic reproduction number R_0 and the effective reproductive number
63 R . Then we performed a phylogenetic comparative analysis to determine the correla-
64 tion between virulence and parasite fitness. We find that virulence and parasite fitness
65 are correlated at the cross-species level, and the exact form of the fitness-virulence
66 correlation depends on the selective pressures within each biological and ecological
67 group of species.

68 *2.2. Methods*

69 We investigated whether virulence is adaptive by simulating cross-species data and
70 by analyzing the relationship between parasite fitness and virulence. To test whether
71 parasite species that are less virulent have lower fitness, we estimated virulence and
72 parasite fitness from simulated disease outbreak data, and we did phylogenetic com-
73 parative analysis to determine the correlation between virulence and parasite fitness.
74 In this section we present the simulated phylogeny, the parasite traits and the disease

75 outbreak data that we used to estimate parasite fitness and virulence (Figures 2.1 and
 76 2.2). The phylogenetic tree simulation was implemented in *phytools* (Revell 2012), the
 77 traits simulation in *mvMORPH* (Clavel et al. 2015), the phylogenetic analyses in *ade-*
 78 *phylo*, and the estimation of parasite fitness in *EpiEstim* R packages (Cori et al. 2013).
 79 All the codes that we used are available as electronic supplementary materials S2.1-
 80 S2.4, and are publicly available at Figshare doi:10.6084/m9.figshare.11387865.v2. The
 81 simulations were run on Compute Canada *cedar* and *beluga* clusters, and took a total
 82 of 5.7 core-years to complete.

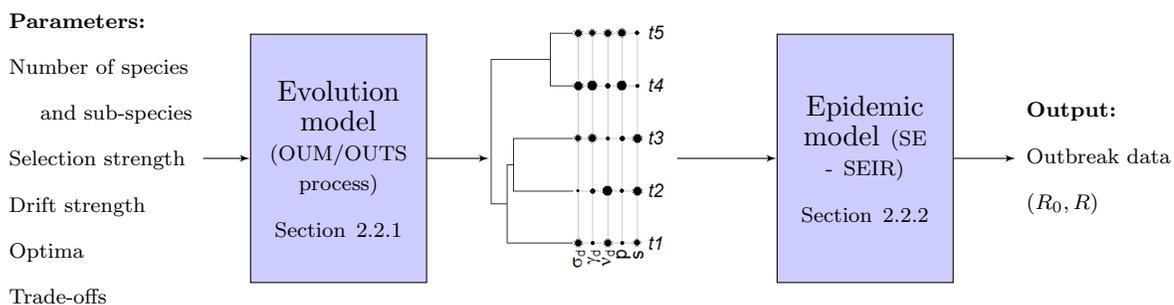
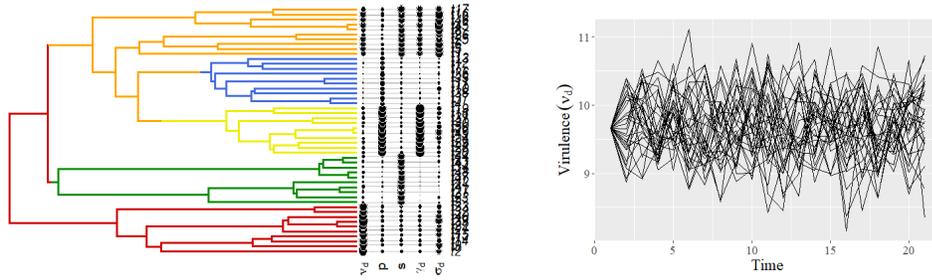
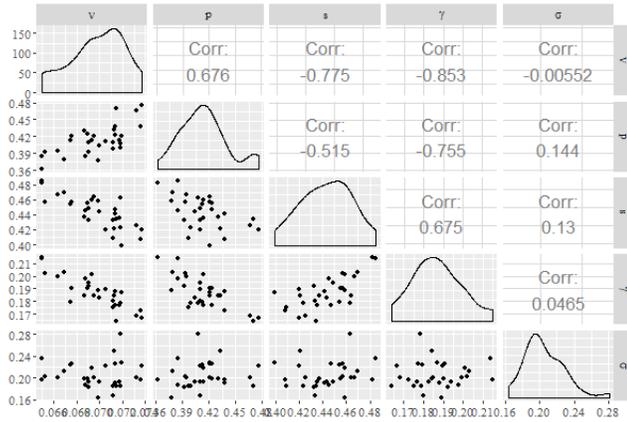


Figure 2.1: We developed a framework whereby simulated cross-species data are used to determine if virulence is adaptive by simulating the evolution of parasites, simulating the epidemiological dynamics arising from the evolved parasite trait values and calculating parasite fitness from these epidemics. The compartments are the simulation models, and the input and the output variables are at the left and the right of the compartments respectively. We simulated 1500 parasites (50 species with 30 strains within each species) that are phylogenetically correlated and have 5 parameters, the incubation period (σ_d), the time to host recovery (γ_d), the time to host death (ν_d), the probability of disease transmission given an infectious contact (p) and infected host movement step length reduction factor (s). To simulate the evolution of parasite species trait values we used the Ornstein-Uhlenbeck process with multiple optima (OUM process), the diversification of each species into strains was simulated using Time series Ornstein-Uhlenbeck process (OUTS) and we considered 5 within-species evolutionary trade-offs. We used the simulated 5 traits as input for a spatially explicit SEIR epidemic model (SE-SEIR) to generate disease outbreak data for each parasite and calculate fitness.



(a)

(b)



(c)

Figure 2.2: We simulated a phylogenetic tree with 50 species and the evolution of 5 parasite traits by Multiple optima Ornstein-Uhlenbeck process along the tree, where each optimum corresponds to a selective regime (a) the diversification of each species into 30 strains (b), and we assume that the traits are correlated within the species (c). In (a) the size of the symbol is relative trait value, and the five traits are the incubation period (σ_d), the time before the host recovers from an infection (γ_d), the time before the host dies due to infection (ν_d), the probability of disease transmission given a contact (p) and the infected host movement step length reduction factor (s). Each selective regime (R1 = red, R2 = green, R3 = yellow, R4 = blue and R5 = orange) represents biological and environmental conditions that drive the evolution of a trait towards an optimum trait value. In (b) the lines are the evolutionary trajectories of the 30 strains within the species t_2 for virulence trait and initial $\nu_d = 9.6$. We assume that virulence-recovery $\nu - \gamma$, virulence-host-movement $\nu - s$ and recovery-transmission $\gamma - p$ trade-offs are negative, virulence-transmission $\nu - p$ and recovery-host-movement $\gamma - s$ trade-offs are positive, and the correlation coefficients for the species t_2 are shown in the upper diagonal diagonal matrix in (c) as example. We transformed the length of times ν_d and γ_d into rates $\nu = 1/\nu_d$, $\gamma = 1/\gamma_d$ respectively to illustrate the trade-offs in a more intuitive way.

Table 2.1: List of symbols used in the main text

Symbols	Definitions	Units
σ_d	The incubation period	days
γ_d	Time before host recovers from infection	days
ν_d	Time before host dies due to infection	days
$\sigma = 1/\sigma_d$	Rate of exposed becoming infectious	day ⁻¹
$\nu = 1/\nu_d$	Virulence	day ⁻¹
$\gamma = 1/\gamma_d$	Host recovery rate	day ⁻¹
p	Probability of disease transmission given a contact	unitless (0-1)
s	Infected host movement step length reduction factor	unitless (0-1)
R_0	The expected number secondary cases by the primary case in a completely susceptible host population	unitless
R	The average number of infections during the infectious period	unitless
OUM	Multiple optima Ornstein-Uhlenbeck process	
OUTS	Time series Ornstein-Uhlenbeck process	

83 *2.2.1. Phylogeny and species traits simulation*

84 We generated numerically a total of 1500 parasites, 50 species and 30 strains within
85 each species, that are phylogenetically related and each parasite has five parasite traits
86 that we used as input for an epidemic model (Figure 2.1). We assumed a number of
87 evolutionary trade-offs within each species and these trade-offs can emerge from the
88 dynamics of parasite replication and the immune response within the host, as well as
89 ecological constraints (Ewald 1983; Antia et al. 1994; Day 2001; Gilchrist and Sasaki
90 2002; Alizon and van Baalen 2005; Alizon 2008).

91 First, we simulated a phylogenetic tree of 50 species by stochastic birth-death
92 process and along the tree we simulated the evolution of five parasite traits by Multiple

93 optima Ornstein-Uhlenbeck process (OUM). The five parasite traits are the incubation
94 period (σ_d), the time before the host recovers from an infection (γ_d), the time before
95 the host dies due to infection (ν_d), the probability of disease transmission given a
96 contact (p) and the infected host movement step length reduction factor (s), which
97 indicates the effect of infection on host movement (see Figure 2.2a and Table 2.1
98 for the definition of the abbreviations and symbols used in this paper). Second, we
99 simulated the diversification of each species into 30 strains by Time Series Ornstein-
100 Uhlenbeck process (OUTS) where the diversification of parasite traits is constrained
101 by evolutionary trade-offs within each species (Figures 2.2b and 2.2c).

102 The OUM process is a model of continuous trait evolution that simulates the
103 adaptive evolution of traits by selection and drift towards multiple optima that rep-
104 resent different evolutionarily stable strategies (Hansen 1997; Butler and King 2004;
105 Monteiro and Nogueira 2011; Beaulieu et al. 2012; Cressler et al. 2015; Citadini et al.
106 2018). For each trait, we defined five optima and each optimum corresponds to the ex-
107 pected evolutionarily stable strategy for one trait under a particular selective regime.
108 We sampled randomly the optima p and s in the range $[0.1, 0.9]$, γ_d and ν_d in the
109 range $[2, 14]$ days and σ_d in the range $[2, 6]$ days, and the values within the ranges
110 are assumed to be uniformly distributed. Here, a selective regime corresponds to any
111 biological and environmental conditions that drive the evolution of a trait towards
112 an adaptive optimum, and parasite species that have similar transmission route or
113 exploit similar host tissues can be in the same selective regime. For example, the
114 immune responses of the host can impose a selective pressure on the rate of parasite
115 replication within a host, and different sites of parasite growth within a host can con-
116 strain the evolution of parasite traits towards different adaptive optima (Engwerda

117 and Kaye 2000; Engwerda et al. 2004). The OUM process is appropriate for simulating
118 parasite trait values because trait variation between the evolved species is large, and
119 the distribution of the traits reflect evolution by drift and selection (Butler and King
120 2004).

121 We further simulated the diversification of each species into 30 strains that are
122 slightly different (Figure 2.2b) using Time Series Ornstein-Uhlenbeck processes (OUTS).
123 Based on a number of evolutionary trade-offs that have been previously reported in the
124 literature, we assumed that virulence-recovery $\nu - \gamma$, virulence-host-movement $\nu - s$
125 and recovery-transmission $\gamma - p$ trade-offs are negative, and virulence-transmission
126 $\nu - p$ and recovery-host-movement $\gamma - s$ trade-offs are positive (Figure 2.2c). These
127 evolutionary trade-offs often emerge from physiological constraints and the dynamics
128 of within-host immune response and parasite replication rate (Alizon and van Baalen
129 2005; Alizon 2008). Higher within-host parasite replication rate, which is positively
130 correlated with virulence, can result in higher parasite transmission because more in-
131 fectious stages are produced (Fraser et al. 2007; de Roode et al. 2008; 2009; de Roode
132 and Altizer 2010; Fraser et al. 2014) and lower host movement because the host is
133 severely lethargic (see Day 2001; Zitzow et al. 2002; Lion et al. 2006; Belser et al.
134 2013; Osnas et al. 2015; Finnerty et al. 2018; Fofana and Hurford 2019). The relation-
135 ship between within-host replication and host recovery rate is not well understood,
136 but some studies suggest that fast replicating parasites are not cleared rapidly by
137 the immune system of the host and induce lower recovery rate (Mackinnon and Read
138 2004; Metcalf et al. 2011; Klein et al. 2014; Greischar et al. 2019). The correlations
139 are qualitatively similar across parasite species, but each parasite species has its own
140 trade-off curves (its correlation coefficients) which means that the exact form of the

141 trade-offs depend on the biology and the ecology of the parasite and the host species
142 (Ewald 1983; Ewald and Giulio 2002; Alizon and van Baalen 2005; Alizon 2008; Alizon
143 et al. 2009).

144 The goal of the phylogeny and parasite traits simulation is to generate phylogenet-
145 ically correlated cross-species data with epidemiological parameter values that emerge
146 from an underlying evolution model and account for evolutionary trade-offs. As such,
147 the epidemiological parameter values that we used for disease outbreak simulation are
148 biologically meaningful and can be used to test evolutionary hypotheses at the cross-
149 species level. To verify whether the simulated traits are phylogenetically correlated,
150 we performed Abouheif’s test at the species level, and the results show that all the
151 simulated traits are phylogenetically correlated (Figure 2.3).

152 *2.2.2. Disease outbreaks simulation*

153 To generate disease outbreak data, we used a spatially explicit stochastic Susceptible-
154 Exposed-Infectious-Removed (SEIR) model with host movement to simulate the spread
155 of different parasite strains in host populations. We modelled host movement as ran-
156 dom walks in two-dimensional spatial domain with periodic boundaries. We simulated
157 the spread of the different strains in host populations of different size (250-4000 in-
158 dividuals corresponding to 62.5-1000 individuals per km²) to mimic empirical disease
159 outbreak data and inform best ways to correct for the effect of host population density
160 in future empirical data analysis.

161 We model the infection and disease progression as SEIR, such that at a given time
162 a host individual can be either Susceptible, Exposed (infected but not yet infectious),
163 Infectious or Recovered from the infection. When an infectious and a susceptible host
164 are spatially close to each other then a contact can occur. We defined 10 meters as

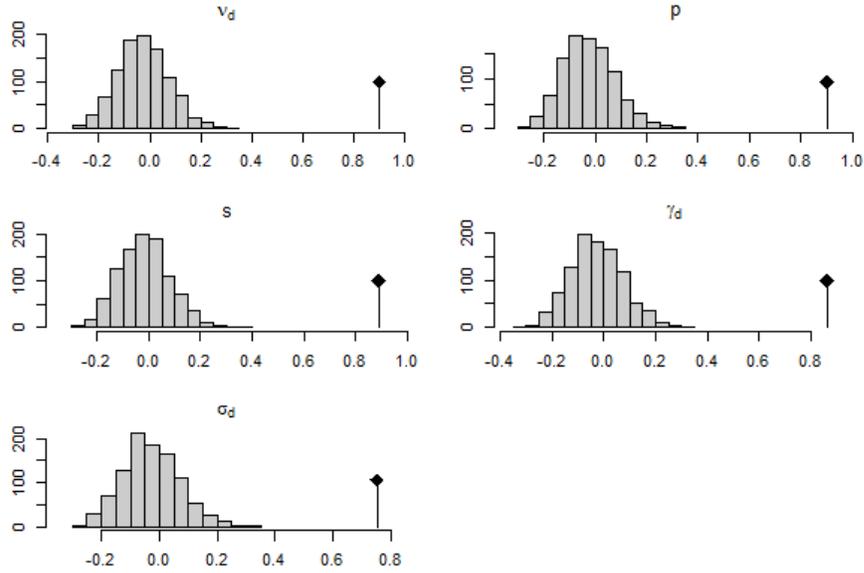


Figure 2.3: The graphs are the results of Abouheif's test based on Moran's I index to detect phylogenetic correlation in the simulated parasite traits at the species level. The results are permutation tests in which the data is resampled randomly (ignoring the phylogeny) to perform Abouheif's test and the correlation index is compared to the true phylogenetic correlation. The vertical black line is the true phylogenetic correlation, measured as Moran's Index, for each trait ($I_{\nu_d} = 0.90$, $I_p = 0.90$, $I_s = 0.89$, $I_{\gamma_d} = 0.86$, and $I_{\sigma_d} = 0.75$), and all the observed values tend to one which suggests that all the simulated traits are phylogenetically correlated. The histogram is the distribution of the phylogenetic correlations obtained from the permutations, and the indexes from random samples are different from the observed index for all the simulated traits (P-values < 0.0015).

165 the maximum spatial radius within which contact can occur and the probability of a
 166 contact is exponentially distributed. If an infectious contact occurs then an infection
 167 can occur with a probability p which is the probability of disease transmission given an
 168 infectious contact. An infected host is exposed upon infection, and becomes infectious
 169 after an average of σ_d days which is the mean incubation period of the parasite. An
 170 infected host can either recover from the disease after an average of γ_d days or die due
 171 to infection after an average of ν_d days, which are the average times to host recovery
 172 and host death due to infection respectively. The incubation period, the time to
 173 host recovery and the time to host death due to infection are gamma-distributed with

174 means of σ_d , γ_d and ν_d respectively (Brauer 2008; Bretscher et al. 2011). Our epidemic
175 model has no recruitment of new susceptible hosts through birth and recovery, and
176 as such the outbreaks will always end (see Figures A.1 and A.2 in appendix). Finally,
177 we used the infections, the serial interval and other counts data from the outbreak
178 simulation to estimate parasite fitness.

179 *2.2.3. Parasite fitness and virulence measures*

180 We measured parasite fitness as the basic reproduction number (R_0) and the ef-
181 fective reproduction number (R), and these two quantities measure the ability of a
182 parasite to invade and spread in a host population. We use R_0 as a measure of parasite
183 fitness because it is a threshold quantity that informs whether a parasite can generate
184 an outbreak in a completely susceptible host population (Lipsitch et al. 2003; Brauer
185 2008; Diekmann et al. 2012). The quantity R_0 measures the growth potential of the
186 infected class on generation basis and R_0 can be seen as absolute fitness (Alizon and
187 Michalakis 2015). To estimate R_0 we counted the total number of secondary cases
188 generated by the primary case in a completely susceptible host population in each
189 outbreak simulation. The effective reproductive number (R), is the average number
190 of infections per infected host when the parasite is established in the host population
191 (Lipsitch et al. 2003). In evolutionary epidemiology the quantity R indicates whether
192 a new mutant strain can replace an established one at the endemic equilibrium, and
193 R can be seen as the relative fitness (Alizon and Michalakis 2015). Previous studies
194 showed that evolution maximizes R at endemic equilibrium or during the late phase
195 of an epidemic (Osnas et al. 2015; Cressler et al. 2016). To estimate R , we used the
196 incidence and the serial interval data from the outbreak simulation to calculate the
197 average transmission potential of each parasite, and the estimation of R was imple-

198 mented in the *EpiEstim* R package. For both R_0 and R the results are averaged over
199 30 simulated outbreak runs.

200 2.2.4. *Statistical analysis*

201 To test whether parasite species that are less virulent have lower fitness we used
202 an autoregressive model, which is a regression model that accounts for the phylo-
203 genetic relationships and non-independence of observations (Cheverud et al. 1985;
204 Gittleman and Kot 1990). We formulated parasite fitness (R_0 and R) as a function of
205 phylogenetic and virulence components, and closely related species are assigned more
206 weights than distant species. Observations in cross-species data are not independent
207 and are phylogenetically correlated, and as such traditional statistical methods, such
208 as classic regression model, are not appropriate for the analysis of cross-species data
209 (Pagel and Harvey 1988; Harvey and Pagel 1991; Pagel and Harvey 1992; Martins
210 and Hansen 1997). We divided R_0 and R by host population density N to correct for
211 host population density, and to meet the assumptions of the linear model we square
212 root transformed R_0 and R .

213 2.3. *Results*

214 We measured parasite fitness as the basic reproduction number R_0 and the effec-
215 tive reproduction number R for 1500 parasites (50 species and 30 strains within each
216 species) that are phylogenetically related, virulence as the length of time before an
217 infected host dies due to infection (ν_d), thus shorter ν_d means higher virulence, and
218 we investigated the fitness-virulence correlation using phylogenetically corrected sta-
219 tistical model. We find that virulence and parasite fitness are correlated at the cross-
220 species level which suggests that virulence is adaptive, however the exact form of the

221 fitness-virulence correlation depends on the selective pressures within each biological
 222 and ecological group of species.

Table 2.2: Summary of the proportion ($\times 100$) of cross-species and within-species data for which the best model is positive, negative, concave-down or concave-up correlation between virulence and parasite fitness measures (R_0 or R). We fit linear and parabolic models to within- and cross-species data (sample sizes are 26-30 and 200 observations per within-species and cross-species data respectively) and we performed a likelihood ratio test to select the best fit model based on Log-likelihood. For each species data (50 species) or cross-species data (1000 random samples) the proportion ($\times 100$) of model fit that is significant (p-value < 0.05) is shown in parenthesis.

Data and fitness measure	% Concave-down	% Positive	% Concave-up	% Negative
Within-species R_0	2 (0)	70 (71)	2 (100)	26 (46)
Within-species R	6 (66)	47 (13)	4 (50)	43 (0)
Cross-species R_0	97.8 (100)	0	0	2.2 (100)
Cross-species R	99.5 (100)	0.5 (100)	0	0

223 *2.3.1. Virulence is adaptive within most of the species*

224 We compared linear and polynomial model fits to each species' data, we selected
 225 the best fit model by likelihood ratio test. We found that the correlation between
 226 virulence and R_0 is positive for 70 % of species, and concave-up for 2 % of the species
 227 (Figures 2.4a and 2.4b, and Table 2.2). As such, higher virulence strains have higher
 228 R_0 within most of our simulated parasite species. However the relationship between
 229 virulence and R_0 is negative in 26 % of the species, and this result suggests that there
 230 are conditions where virulence is not beneficial for parasite transmission (Figures 2.4c,
 231 and Table 2.2). For R fitness measures only 16 % of the best model fit models are
 232 significant, and the results are presented in A.3 and A.4 as electronic supplementary
 233 material.

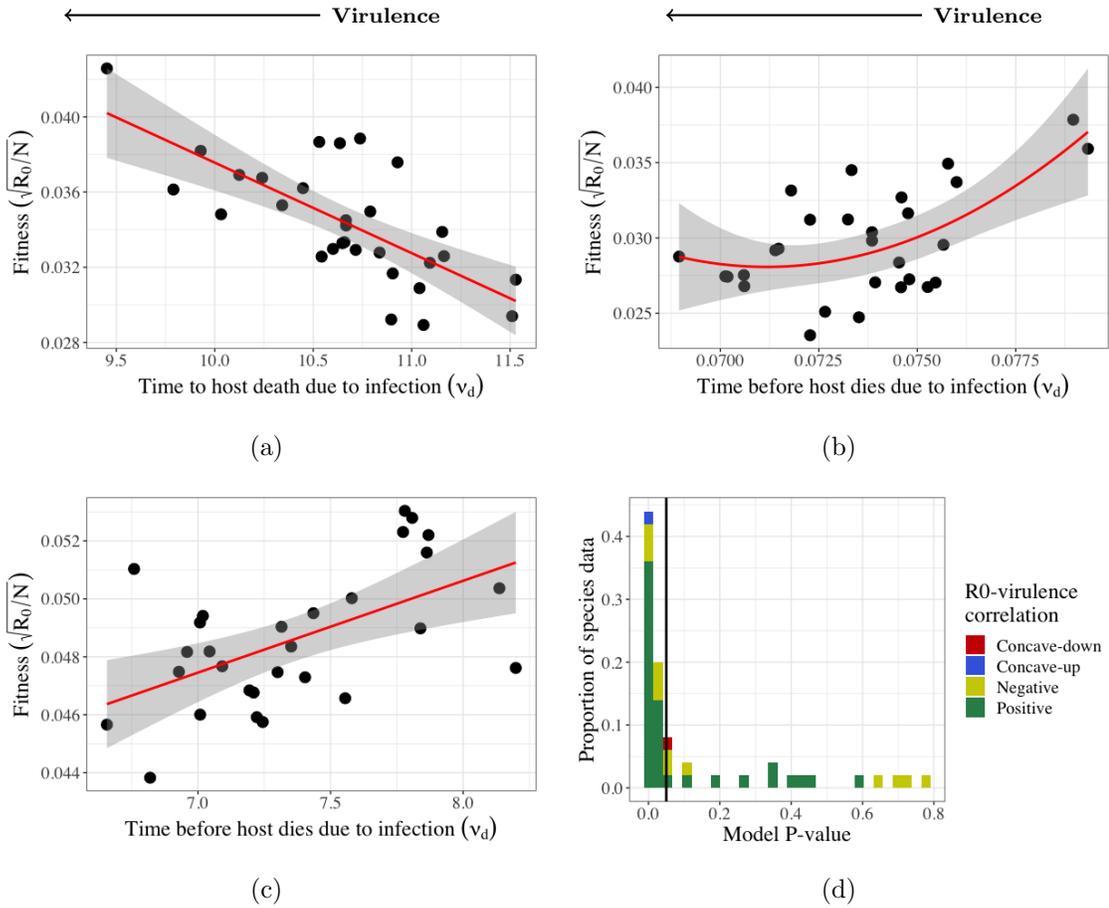
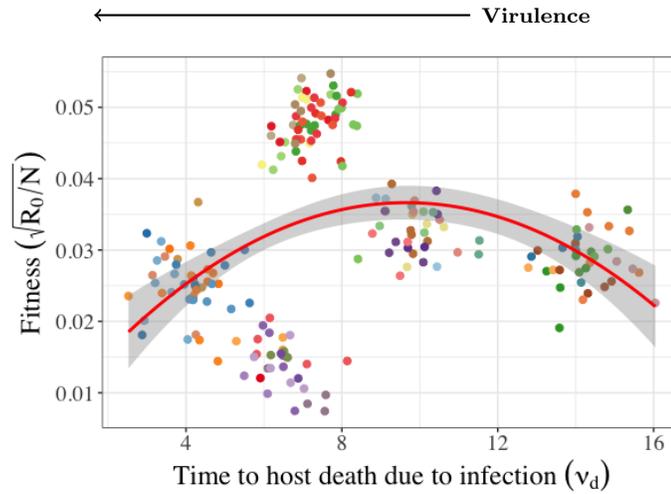


Figure 2.4: Most of the within-species data show that virulence is adaptive and the correlation between virulence and R_0 is positive (a), however, in a few species the correlation between virulence and R_0 is concave-up (b) and negative (c) (We graph only examples of significant model fits). The proportion ($\times 100$) of species data for which the relationship between R_0 and virulence is positive (green bar), negative (yellow bar) or concave-up (blue bar) and their corresponding model p-values are presented in (d), where the vertical black line is the 0.05 significance level. We fit linear and parabolic models to our within-species data (sample sizes are 26-30 observations per species data, with 50 species in total) and we performed a likelihood ratio test to select the best fit model. The dots are the estimated fitness measures for each parasite strain of the same species (averaged over 30 outbreaks simulation runs), the line through the data is the best fit model, and the grey area is the 95% confidence interval. Model p-values are less than 0.001 for all graphs and the adjusted $R^2 = 0.48, 0.28, 0.32$ for (a), (b) and (c) respectively. We divided R_0 by host population density (N) to correct for the effect of host population density and we square root transformed the response variables to meet the assumptions of the linear model.

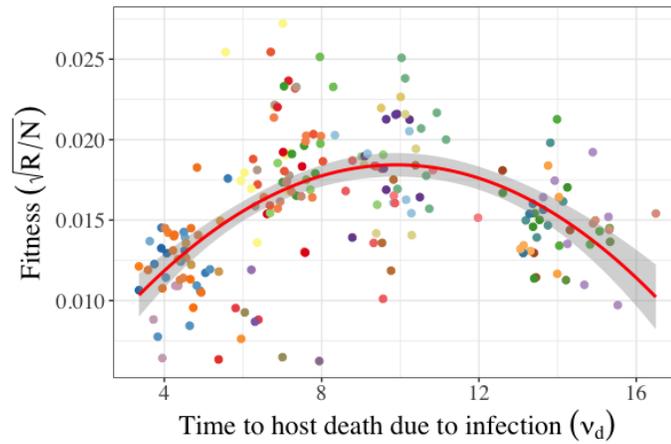
234 *2.3.2. Cross-species data show that virulence is adaptive at least for low virulence*
235 *levels*

236 Most of the random cross-species data samples show a concave-down relationship
237 between virulence and both R_0 and R measures (97.8 % for R_0 and 99.5 % for R)
238 and this result suggests that the relationship between virulence and parasite fitness
239 is positive for low virulence levels ($\nu_d > 10$) and negative for high virulence levels
240 (Figures 2.5a, 2.5b and Table 2.2). In very few cases, our results show that the fitness-
241 virulence relationship can be linear (2.2 % for R_0 and 0.5 % for R), but the model fits
242 are relatively poor and model diagnostics show that the relationship between virulence
243 and R_0 and R is probably not linear (See Figures A.5 in appendix).

244 We investigated whether the relationship between virulence and parasite fitness
245 depends on the selective regimes, and we fit samples of cross-species data to (1) a
246 phylogenetically corrected linear model without interaction and (2) a phylogenetically
247 corrected linear model with interaction between virulence and selective regime and
248 we performed a likelihood ratio test to select the best fit model. We found that the
249 model with interaction is significantly more likely for R_0 fitness measures, and this
250 result suggests that the relationship between virulence and R_0 is not the same within
251 the selective regimes (p-value < 0.0001 and $\Delta \text{Log likelihood} = 44$). The virulence-
252 R_0 relationship is positive for the selective regimes R1, R2 and R5, negative for R3
253 and concave-down for R4 (Figure 2.6a). However for R fitness measures the fitness-
254 virulence correlations within the different selective regimes are unclear (Figure 2.6b
255 and $\Delta \text{Log likelihood} = 2$ and p-value > 0.5 for the comparison of model with and
256 without interactions).

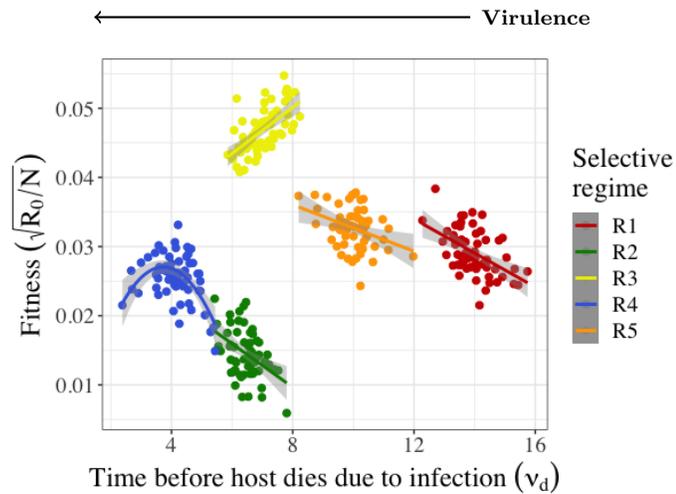


(a)

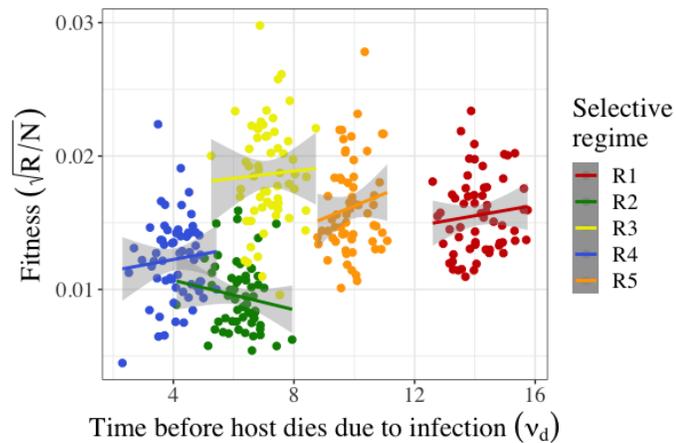


(b)

Figure 2.5: The correlation between virulence and the basic reproduction number R_0 in (a) and the effective reproduction number R in (b) is concave down. The dots are the estimated fitness measures of each parasite (averaged over 30 outbreaks simulation runs), the colour indicates parasite strains of the same species, the line through the data is the best fit model, and the grey area is the 95% confidence interval. For both graphs the best fit models are polynomials, model p-value < 0.0001 and adjusted $R^2 = 0.50$ and $R^2 = 0.38$ in (a) and (b) respectively. We divided R_0 and R by host population density (N) to correct for the effect of host population density, and we square root transformed the response variables to meet the assumptions of the linear model. We fit phylogenetically corrected linear and parabolic models to 1000 randomly sampled cross-species data (sample size is 200 observations per randomly sampled data) and we performed likelihood ratio tests to select the best fit model. The best model fit is the concave-down correlation between virulence and parasite fitness (99.9 % of samples for R_0 and 99.5 % for R).



(a)



(b)

Figure 2.6: The correlation between virulence and R_0 is positive for R1, R2, and R5, negative for R3 and concave-down for R3 (a), however, the correlation between virulence and R for R1-R5 is statistically unclear (b). For each selective regime, we fit phylogenetically corrected linear and parabolic models to random samples of cross-species data (60 observations per selective regime), and we performed a likelihood ratio test to select the best fit model. The dots are the estimated fitness measures for each parasite (averaged over 30 outbreaks simulation runs), the colours are the different selective regimes (R1 = red, R2 = green, R3 = yellow, R4 = blue and R5 = orange), the line through the data is the best fit model, and the grey area is the 95% confidence interval. In (a) model p-value is < 0.05 for all selective regimes, and in (b) model p-value is > 0.1 for all selective regimes. We divided R_0 and R by host population density (N) to correct for the effect of host population density and we square root transformed the response variables to meet the assumptions of the linear model.

257 *2.4. Discussion*

258 The adaptive virulence hypothesis suggests that virulence is maintained because
259 it is adaptive, and this hypothesis has been successful in theoretical studies but its
260 empirical validation is methodologically challenging, especially at the cross-species
261 level. In this paper we investigate whether we can determine if virulence is adaptive
262 using simulated cross-species data. We find that virulence and parasite fitness can
263 be correlated at the species level, the exact form of the fitness-virulence correlation
264 depends on the selective pressures within biological and ecological groups of species,
265 and this results suggests that ecological conditions determine whether virulence is
266 beneficial for parasite transmission.

267 *2.4.1. Why is virulence beneficial for parasite transmission in some species?*

268 Both within- and cross-species data show that the correlation between parasite
269 fitness and virulence is positive, which means that virulence is beneficial for parasite
270 transmission and disease spread in host populations (Figures 2.4a, 2.4b, 2.6a and Table
271 2.2). In the R1, R2, and R5 selective regimes, higher virulence parasites are more likely
272 to be transmitted and cause infections with longer infectious period during which new
273 infections can occur, and as such, higher virulence parasites generate more infections
274 and have higher fitness (Figures 2.7, see R1, R2 and R5). The positive relationship
275 between virulence and parasite transmissibility can emerge from the dynamics of
276 parasite replication rate and the immune response within the host, and previous
277 studies showed that parasites with higher within-host replication rate produce more
278 infectious stages and are not cleared rapidly by the immune system of the host, which
279 increase both the probability of host death due to infection and disease transmissibility
280 (Mackinnon and Read 2004; de Roode et al. 2008; Metcalf et al. 2011; Leggett et al.

281 2017; Acevedo et al. 2019; Greischar et al. 2019). This result is expected given the
282 linear virulence-transmission and virulence-recovery trade-offs that we assume within-
283 species.

284 The other reason why higher virulence parasites generate more infections in the R1,
285 R2 and R5 selective regimes is the relationship between virulence and the infectious
286 period. In the R1, R2 and R5 selective regimes, the parasites induce infections where
287 most of the hosts recover before the expected time to host death due to infection
288 because $\gamma_d < \nu_d$, and as such, the infectious period mainly depends on the time to
289 host recovery γ_d (Figure 2.7d, $\nu_d - \gamma_d > 0$). Higher virulence parasites (thus shorter
290 ν_d) have longer times to host recovery γ_d and induce longer infections, which is a
291 translation of the negative within-species virulence-recovery trade-offs at the species
292 level. Because higher virulence parasites are more likely to be transmitted given a
293 contact, these parasites generate more new infections over longer infectious period,
294 which explains why the relationship between virulence and parasite fitness is positive
295 at the cross-species level in the R1, R2 and R5 selective regimes.

296 Moreover in the R1, R2 and R5 selective regimes, parasite species cause few host
297 death ($\gamma_d < \nu_d$), most of the infected hosts recover from the disease sooner or later.
298 Lower virulence parasites are rapidly cleared by the immune system of the host, and
299 a host infected with a lower virulence parasite is more likely to recover earlier from
300 the infection. The species in the R1, R2 and R5 selective regimes can be seen as
301 sub-lethal parasites, and previous studies showed that host recovery rate ($1/\gamma_d$) is the
302 main factor that drives the evolution of virulence in sub-lethal parasites (Alizon 2008;
303 Bull and Lauring 2014). Our results are in accordance with previous studies, and we
304 show that within-species virulence-transmission and virulence-recovery trade-offs can

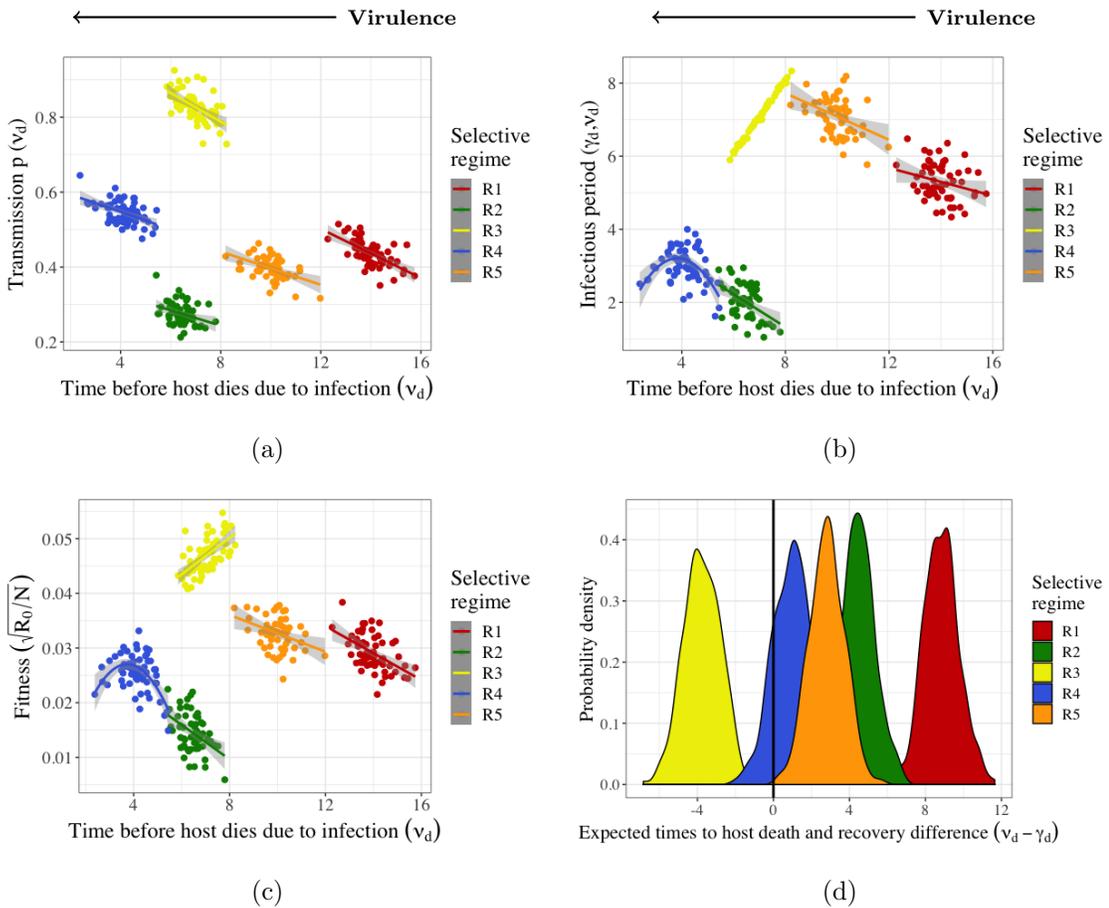


Figure 2.7: Virulence as a function of the probability of disease transmission given a contact p (a), the duration of the infectious period (b) and the basic reproduction number R_0 (c). In (d), we graph the distribution of the difference between the expected time to host death due to infection and the expected time to host recovery ($\nu_d - \gamma_d$) for all parasites and selective regimes. This graph illustrates how ν_d and γ_d can affect the duration of the infectious period, shown in (b), and when $\nu_d - \gamma_d > 0$ then γ_d is the main factor limiting the duration of the infectious period, whereas when $\nu_d - \gamma_d < 0$ then ν_d is the main limiting factor. Parasites within the same selective regime are biologically or ecologically similar and we have 5 selective regimes R1 = red, R2 = green, R3 = yellow, R4 = blue and R5 = orange. Parasite traits, ν_d , γ_d and p , were simulated to generate outbreaks and calculate R_0 , see section 2.2 for details.

305 explain why virulence is beneficial in sub-lethal parasites.

306 However, in the R3 selective regime species the correlation between parasite fit-
 307 ness and virulence is negative, and the explanation is that as virulence increases the
 308 probability of disease transmission given a contact, p , increases but the infectious pe-

309 riod becomes shorter. The species in the R3 selective regime induce death before the
310 expected time to host recovery ($\nu_d < \gamma_d$) and most of the infected hosts die from the
311 infections (Figures 2.7, $\nu_d - \gamma_d < 0$). In this situation, the duration of the infectious
312 period depends on the length of time to host death due to infection ν_d , which is shorter
313 for higher virulence parasites because the infected host is more likely to die earlier
314 from the infection. As such, higher virulence parasites have higher transmissibility but
315 generate overall less infections because the infected host is more likely to die earlier
316 from the infection.

317 The species in the R3 selective regime can be seen as obligate killers because in-
318 fected hosts will die sooner or later ($\nu_d < \gamma_d$). Previous studies show that for the
319 obligate killers, Nucleopolyhedrovirus in its insect host *Spodoptera exempta* and *Pas-*
320 *teuria ramosa* in its Crustacean host *Daphnia magna*, higher virulence strains produce
321 fewer transmission stages during their lifetime and have lower fitness (Redman et al.
322 2016; Ben-Ami 2017). Our results are in accordance with these previous studies, and
323 our simulation data suggest that the negative relationship between virulence and par-
324 asite fitness is valid at the cross-species level for obligate killers.

325 In the R4 selective regime the relationship between R_0 and virulence is concave-
326 down, which means that higher virulence parasites generate more infections until
327 a point, which is $\nu_d < \gamma_d$, where increased virulence is not beneficial for parasite
328 transmission and results in lower parasite transmission from one host to another
329 (Figure 2.7, see R4 selective regime). Due to the negative virulence-recovery trade-
330 off, a host infected with low virulence parasites is less likely to die due to infection
331 but more likely to recover earlier from the infection, and the likelihood of earlier
332 host recovery decreases as virulence increases because higher virulence parasites can

333 replicate faster within the host and avoid rapid clearance by the immune system of
334 the host. As such, higher virulence parasites cause longer infections during which
335 more new infections can occur until a point where the likelihood of host death due to
336 infection becomes very high, which happens when the expected time to host death is
337 less than the expected time to host recovery ($\nu_d < \gamma_d$). From this point, the expected
338 time to host recovery γ_d is irrelevant for the duration of the infectious period. A
339 host infected with higher virulence parasites is more likely to die earlier during the
340 infection which results in shorter infectious period and lower parasite transmission
341 before the host dies due to infection. Thus, intermediate virulence parasite species
342 with intermediate disease transmissibility generate the maximum number of infection
343 over longer infectious period in the R4 selective regime.

344 Concave-down relationship between virulence and parasite fitness often arises from
345 a saturating relationship between virulence and parasite transmission, where interme-
346 diate virulence maximizes the total number of new infections per infected host during
347 the infectious period (Fraser et al. 2007; de Roode et al. 2008; Chapuis et al. 2012;
348 Fraser et al. 2014). However, a concave-down relationship between virulence and par-
349 asite fitness can also arise due to a trade-off between parasite transmission and the
350 clearance rate of the parasite by the immune of the host (van Ballegooijen and Boer-
351 lijst 2004; Alizon 2008). Empirical evidence supporting the transmission-clearance
352 trade-off has been reported for the *Mycoplasma gallisepticum* bacterial infections in
353 the North American house finch *Haemorhous mexicanus* (Williams et al. 2014) and
354 Dengue virus infection in humans (Ben-Shachar and Koelle 2018). In complement to
355 previous results, our simulation data show that a concave-down relationship between
356 virulence and parasite fitness can arise at the cross-species level, and a non-linear

357 relationship between virulence and the infectious period can maintain intermediate
358 virulence.

359 *2.4.2. Can we determine whether virulence is adaptive using empirical cross-species*
360 *data ?*

361 Overall, our simulation data show that we can test the adaptive virulence hypothe-
362 sis at the cross-species level but the following three factors need careful consideration
363 when investigating the relationship between virulence and parasite fitness. Firstly,
364 the biological and ecological similarities between the species cannot be ignored in
365 the empirical investigation of the correlation between parasite fitness and virulence
366 at the species level. In our simulation the different selective regimes correspond to
367 groups of species that have evolved under similar selective pressures and have similar
368 epidemiological trait values. These similarities can be the transmission route of the
369 parasites or the type of host tissues in which the parasites replicate. The biological
370 and ecological distinctions between each group of species is important because the
371 contribution of the different parasite traits to parasite fitness can be different from
372 one group to another.

373 Second, phylogenetic correction is important in cross-species comparison because
374 it improves the quality of the statistical tests (Gittleman and Kot 1990; Garland et al.
375 2005). Testing the adaptive virulence hypothesis at the cross-species level is difficult
376 because reconstructing the phylogenetic relationship within and between viruses, bac-
377 teria and fungi is not straightforward (but see Iyer et al. 2001; James et al. 2006;
378 Kühnert et al. 2011; Koonin et al. 2015). However, investigating the fitness-virulence
379 correlation in parasites that are biologically or ecologically similar can be a step for-
380 ward because constructing a phylogenetic tree for species that share some biological

381 similarities may be less complicated. The difficulties in reconstructing the phyloge-
382 netic relationship between parasite species for phylogenetic correction is probably
383 one of the reasons why previous studies are limited to virulence comparison across
384 parasite species that exploit similar host tissues or have similar transmission routes
385 (Ewald 1983; 1991b; Walther and Ewald 2004). However, more recent works showed
386 that reconstructing the phylogenetic relationship between parasite species that have
387 diverged a long time ago is feasible using similarities between the structure of parasite
388 proteins (Leggett et al. 2017).

389 Thirdly, the host population density is another important factor that needs to be
390 considered when investigating the correlation between parasite fitness and virulence.
391 A preliminary investigation of our data shows that the most important contributor to
392 both R_0 and R is the host population density. As such, to remove the density effect
393 we divided R_0 and R by the density of the host population. The fitness measures,
394 R_0 and R estimated from disease outbreak data, often carry some information about
395 host population density and appropriate methods to remove or control for the effect
396 of host population density will provide more accurate results. Furthermore, R_0 is
397 more precise at capturing the relationship between virulence and parasite fitness,
398 whereas the relationship between R measures and virulence is unclear. One possible
399 explanation is that during the late phase of the outbreak, the availability of susceptible
400 hosts often limit the spread of infectious diseases and this phenomenon does not
401 depend on the initial host population density (Lipsitch et al. 1995; Dieckmann 2002;
402 Otto and Day 2007; Lion and Metz 2018). As such, dividing R by host population
403 density may be insufficient to correct for the effect of the initial host population
404 density.

405 *2.5. Conclusion*

406 The main challenges with testing the adaptive virulence hypothesis at the cross-
407 species level are methodological, but we use simulated cross-species data to test
408 whether higher virulence parasites have higher fitness. We conclude that the empirical
409 investigation of whether virulence is adaptive using cross-species data is feasible with
410 species that are biologically or ecologically similar for the following reasons: (1) the
411 phylogenetic correction for comparative analysis can be simpler, (2) the main para-
412 site traits that directly affect fitness can be similar and (3) the interpretation of the
413 fitness-virulence relationship may be more straightforward for species under similar
414 selective pressures.

Literature cited

- M.A. Acevedo, F.P. Dillemuth, A.J. Flick, M.J. Faldyn, and B.D. Elderd. Virulence-driven trade-offs in disease transmission: A meta-analysis. *Evolution*, 2019.
- S. Alizon and Y. Michalakis. Adaptive virulence evolution: the good old fitness-based approach. *Trends Ecol. Evol.*, 30(5):248–254, 2015.
- S. Alizon and M. van Baalen. Emergence of a convex trade-off between transmission and virulence. *Am. Nat.*, 165(6):E155–E167, 2005.
- S. Alizon, A. Hurford, N. Mideo, and M. van Baalen. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.*, 22(2):245–259, 2009.
- Samuel Alizon. Transmission-recovery trade-offs to study parasite evolution. *Am. Nat.*, 172(3):E113–E121, 2008.

- R. Antia, B.R. Levin, and R.M. May. Within-host population dynamics and the evolution and maintenance of microparasite virulence. *Am. Nat.*, pages 457–472, 1994.
- Gordon H Ball. Parasitism and evolution. *Am. Nat.*, 77(771):345–364, 1943.
- J.M. Beaulieu, D.C. Jhwueng, C. Boettiger, and B.C. O’Meara. Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution*, 66(8):2369–2383, 2012.
- J.A. Belser, K.M. Gustin, M.B. Pearce, T.R. Maines, H. Zeng, C. Pappas, X. Sun, P.J. Carney, J.M. Villanueva, and J. Stevens. Pathogenesis and transmission of avian influenza a (H7N9) virus in ferrets and mice. *Nature*, 501(7468):556–559, 2013.
- F. Ben-Ami. The virulence–transmission relationship in an obligate killer holds under diverse epidemiological and ecological conditions, but where is the tradeoff? *Ecol. Evol.*, 7(24):11157–11166, 2017.
- R. Ben-Shachar and K. Koelle. Transmission-clearance trade-offs indicate that dengue virulence evolution depends on epidemiological context. *Nat. Commun.*, 9(1):2355, 2018.
- B. Boldin and É. Kisdi. On the evolutionary dynamics of pathogens with direct and environmental transmission. *Evolution*, 66(8):2514–2527, 2012.
- B.M. Bolker, A. Nanda, and D. Shah. Transient virulence of emerging pathogens. *J. R. Soci. Interface*, 7(46):811–822, 2010.
- S. Bonhoeffer, R.E. Lenski, and D. Ebert. The curse of the pharaoh: the evolution of virulence in pathogens with long living propagules. *Proc. R. Soc. Lond., B, Biol. Sci.*, 263(1371):715–721, 1996.
- F. Brauer. Compartmental models in epidemiology. In *Mathematical epidemiology*, chapter 2, pages 19–79. Springer, 2008.
- M.T. Bretscher, N. Maire, N. Chitnis, I. Felger, S. Owusu-Agyei, and T. Smith. The distribution of Plasmodium falciparum infection durations. *Epidemics*, 3(2):109–118,

- 2011.
- J.J. Bull and A.S. Luring. Theory and empiricism in virulence evolution. *PLoS Pathog.*, 10(10):e1004387, 2014.
- M. Burnet and D.O. White. *Natural history of infectious disease*. CUP Archive, 1972.
- M.A. Butler and A.A. King. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.*, 164(6):683–695, 2004.
- E. Chapuis, A. Arnal, and J.B. Ferdy. Trade-offs shape the evolution of the vector-borne insect pathogen *Xenorhabdus nematophila*. *Proc. R. Soc. B*, 279(1738):2672–2680, 2012.
- J.M. Cheverud, M.M. Dow, and W. Leutenegger. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. *Evolution*, 39(6):1335–1351, 1985.
- J.M. Citadini, R. Brandt, C.R. Williams, and F.R. Gomes. Evolution of morphology and locomotor performance in anurans: relationships with microhabitat diversification. *J. Evol. Biol.*, 31(3):371–381, 2018.
- J. Clavel, G. Escarguel, and G. Merceron. mvmorph: an R package for fitting multivariate evolutionary models to morphometric data. *Methods Ecol. Evol.*, 6(11):1311–1319, 2015.
- A. Cori, N.M. Ferguson, C. Fraser, and S. Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.*, 178(9):1505–1512, 2013.
- C.E. Cressler, D.V. McLeod, C. Rozins, J. Van Den Hoogen, and D.T. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology*, 143(7):915–930, 2016.
- Clayton E Cressler, Marguerite A Butler, and Aaron A King. Detecting adaptive evolution in phylogenetic comparative analysis using the Ornstein–Uhlenbeck model. *Syst. Biol.*, 64(6):953–968, 2015.

- T. Day. Parasite transmission modes and the evolution of virulence. *Evolution*, 55(12): 2389–2400, 2001.
- J.C. de Roode and S. Altizer. Host–parasite genetic interactions and virulence–transmission relationships in natural populations of monarch butterflies. *Evolution*, 64(2):502–514, 2010.
- J.C. de Roode, A.J. Yates, and S. Altizer. Virulence–transmission trade–offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proc. Natl. Acad. Sci. U.S.A.*, 105(21):7489–7494, 2008.
- J.C. de Roode, J. Chi, R.M. Rarick, and S. Altizer. Strength in numbers: high parasite burdens increase transmission of a protozoan parasite of monarch butterflies (*Danaus plexippus*). *Oecologia*, 161(1):67–75, 2009.
- U. Dieckmann. *Adaptive Dynamics of pathogen-host interactions*, chapter 4, pages 39–59. *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*. Cambridge University Press, 2002.
- O. Diekmann, H. Heesterbeek, and T. Britton. The epidemic in a closed population. In *Mathematical tools for understanding infectious diseases dynamics*, chapter 1, pages 3–36. Princeton University Press, 2012.
- J. Doumayrou, A. Avellan, R. Froissart, and Y. Michalakis. An experimental test of the transmission-virulence trade-off hypothesis in a plant virus. *Evolution: International Journal of Organic Evolution*, 67(2):477–486, 2013.
- C.R. Engwerda and P.M. Kaye. Organ-specific immune responses associated with infectious disease. *Immunol. Today*, 21(2):73–78, 2000.
- C.R. Engwerda, M. Ato, and P.M. Kaye. Macrophages, pathology and parasite persistence in experimental visceral leishmaniasis. *Trends Parasitol.*, 20(11):524–530, 2004.
- P.W. Ewald. Host-parasite relations, vectors, and the evolution of disease severity. *Annu.*

- Rev. Ecol. Evol. Syst.*, pages 465–485, 1983.
- P.W. Ewald. Waterborne transmission and the evolution of virulence among gastrointestinal bacteria. *Epidemiol. Infect.*, 106(01):83–119, 1991a.
- P.W. Ewald. Transmission modes and the evolution of virulence. *Hum. Nat.*, 2(1):1–30, 1991b.
- P.W. Ewald and D.L. Giulio. *Alternative transmission modes and the evolution of virulence*, chapter 2, pages 10–25. *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*. Cambridge University Press, 2002.
- J. Felsenstein. Phylogenies and the comparative method. *Am. Nat.*, 125(1):1–15, 1985.
- P.B. Finnerty, R. Shine, and G.P. Brown. The costs of parasite infection: Effects of removing lungworms on performance, growth and survival of free-ranging cane toads. *Funct. Ecol.*, 32(2):402–415, 2018.
- A.M. Fofana and A. Hurford. Parasite-induced shifts in host movement may explain the transient coexistence of high-and low-pathogenic disease strains. *BioRxiv*, page 623660, 2019.
- S.A. Frank. Models of parasite virulence. *Q. Rev. Biol.*, pages 37–78, 1996.
- S.A. Frank and P. Schmid-Hempel. Mechanisms of pathogenesis and the evolution of parasite virulence. *J. Evol. Biol.*, 21(2):396–404, 2008.
- C. Fraser, T.D. Hollingsworth, R. Chapman, F. de Wolf, and W.P. Hanage. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. U.S.A.*, 104(44):17441–17446, 2007.
- C. Fraser, K. Lythgoe, G.E. Leventhal, G. Shirreff, T.D. Hollingsworth, S. Alizon, and S. Bonhoeffer. Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. *Science*, 343(6177):1243727, 2014.
- R. Froissart, J. Doumayrou, F. Vuillaume, S. Alizon, and Y. Michalakis. The virulence–

- transmission trade-off in vector-borne plant viruses: a review of (non-) existing studies. *Philos. Trans. R. Soc. B*, 365(1548):1907–1918, 2010.
- S. Gandon. The curse of the pharaoh hypothesis. *Proc. R. Soc. Lond., B, Biol. Sci.*, 265(1405):1545–1552, 1998.
- T. Garland, A.F. Bennett, and E.L. Rezende. Phylogenetic approaches in comparative physiology. *J. Exp. Biol.*, 208(16):3015–3035, 2005.
- Michael A Gilchrist and Akira Sasaki. Modeling host–parasite coevolution: a nested approach based on mechanistic models. *J. Theor. Biol.*, 218(3):289–308, 2002.
- J.L. Gittleman and M. Kot. Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst. Zool.*, 39(3):227–241, 1990.
- A. Grafen. The phylogenetic regression. *Philos. Trans. R. Soc. B*, 326(1233):119–157, 1989.
- M.A. Greischar, L.M. Beck-Johnson, and N. Mideo. Partitioning the influence of ecology across scales on parasite evolution. *Evolution*, 73(11):2175–2188, 2019.
- J.D. Hadfield and S. Nakagawa. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.*, 23(3):494–508, 2010.
- T.F. Hansen. Stabilizing selection and the comparative analysis of adaptation. *Evolution*, 51(5):1341–1351, 1997.
- P.H. Harvey and M.D. Pagel. The comparative method for studying adaptation. In *The comparative method in evolutionary biology*, chapter 1, pages 1–34. Oxford University Press Oxford, 1991.
- L.M. Iyer, L. Aravind, and E.V. Koonin. Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.*, 75(23):11720–11734, 2001.
- T.Y. James, F. Kauff, C.L. Schoch, P.B. Matheny, V. Hofstetter, C.J. Cox, G. Celio, C. Guerdan, E. Fraker, J. Miadlikowska, et al. Reconstructing the early evolution of fungi using

- a six-gene phylogeny. *Nature*, 443(7113):818–822, 2006.
- T. Jombart and S. Dray. adephylo: exploratory analyses for the phylogenetic comparative method. *Bioinformatics*, 26(15):1–21, 2010.
- E.Y. Klein, A.L. Graham, M. Llinás, and S. Levin. Cross-reactive immune responses as primary drivers of malaria chronicity. *Infect. Immun.*, 82(1):140–151, 2014.
- E.V. Koonin, V.V. Dolja, and M. Krupovic. Origins and evolution of viruses of eukaryotes: the ultimate modularity. *Virology*, 479:2–25, 2015.
- D. Kühnert, C. Wu, and A.J. Drummond. Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. *Infect. Genet. Evol.*, 11(8):1825–1841, 2011.
- H.C. Leggett, C.K. Cornwallis, A. Buckling, and S.A. West. Growth rate, transmission mode and virulence in human pathogens. *Phil. Trans. R. Soc. B*, 372(1719):20160094, 2017.
- S. Lion and J.A.J Metz. Beyond R0 maximisation: on pathogen evolution and environmental dimensions. *Trends Ecol. Evol.*, 33(6):458–473, 2018.
- S. Lion, M. van Baalen, and W.G. Wilson. The evolution of parasite manipulation of host dispersal. *Proc. R. Soc. B.*, 273(1590):1063–1071, 2006.
- M. Lipsitch, E.A. Herre, and M.A. Nowak. Host population structure and the evolution of virulence: a “law of diminishing returns”. *Evolution*, 49(4):743–748, 1995.
- M. Lipsitch, T. Cohen, B. Cooper, J.M. Robins, S. Ma, L. James, G. Gopalakrishna, S.K. Chew, C.C. Tan, M.H. Samore, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970, 2003.
- M.J. Mackinnon and A.F. Read. Virulence in malaria: an evolutionary viewpoint. *Philos. Trans. R. Soc. B*, 359(1446):965–986, 2004.
- E.P. Martins and T.F. Hansen. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data.

- Am. Nat.*, 149(2):646–667, 1997.
- R.M. May and R.M. Anderson. Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. B.*, 219(1216):281–313, 1983.
- C.J.E. Metcalf, A.L. Graham, S. Huijben, V.C Barclay, G.H. Long, B.T. Grenfell, A.F. Read, and O.N. Bjørnstad. Partitioning regulatory mechanisms of within-host malaria dynamics using the effective propagation number. *Science*, 333(6045):984–988, 2011.
- P.O. Méthot. Why do parasites harm their host? on the origin and legacy of theobald smith’s “law of declining virulence”–1900-1980. *Hist. Philos. Life Sci.*, pages 561–601, 2012.
- L.R. Monteiro and M.R. Nogueira. Evolutionary patterns and processes in the radiation of phyllostomid bats. *BMC Evol. Biol.*, 11(1):137, 2011.
- E.E. Osnas, P.J. Hurtado, and A.P. Dobson. Evolution of pathogen virulence across space during an epidemic. *Am. Nat.*, 185(3):332–342, 2015.
- S.P. Otto and T. Day. *Evolutionary invasion analysis*, chapter 12, pages 454–566. A biologist’s guide to mathematical modeling in ecology and evolution. Princeton University Press, 2007.
- M.D. Pagel and P.H. Harvey. Recent development in the analysis of comparative data. *Q. Rev. Biol.*, 63(4):413–440, 1988.
- M.D. Pagel and P.H. Harvey. A method for the analysis of comparative data. *J. Theor. Biol.*, 156(4):431–442, 1992.
- S. Pavoine, S. Ollier, D. Pontier, and D. Chessel. Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theor. Popul. Biol.*, 73(1):79–91, 2008.
- L. Råberg. Infection intensity and infectivity of the tick-borne pathogen *Borrelia afzelii*. *J. Evol. Biol.*, 25(7):1448–1453, 2012.

- E.M. Redman, K. Wilson, and J.S. Cory. Trade-offs and mixed infections in an obligate-killing insect pathogen. *J. Anim. Ecol.*, 85(5):1200–1209, 2016.
- L.J. Revell. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.*, 3(2):217–223, 2012.
- F.J. Rohle. A comment on phylogenetic correction. *Evolution*, 60(7):1509–1515, 2006.
- T. Smith. Some problems in the life history of pathogenic microorganisms. *Science*, 20(520):817–832, 1904.
- W.M. van Ballegooijen and M.C. Boerlijst. Emergent trade-offs and selection for outbreak frequency in spatial epidemics. *Proc. Natl. Acad. Sci. U.S.A.*, 101(52):18246–18250, 2004.
- B.A. Walther and P.W. Ewald. Pathogen survival in the external environment and the evolution of virulence. *Biol. Rev.*, 79(4):849–869, 2004.
- P.D. Williams, A.P. Dobson, K.V. Dhondt, D.M. Hawley, and A.A. Dhondt. Evidence of trade-offs shaping virulence evolution in an emerging wildlife pathogen. *J. Evol. Biol.*, 27(6):1271–1278, 2014.
- L.A. Zitzow, T. Rowe, T. Morken, W.J. Shieh, S. Zaki, and J.M. Katz. Pathogenesis of avian influenza a (h5n1) viruses in ferrets. *J. Virol.*, 76(9):4420–4429, 2002.

3. **Chapter two: Parasite-induced shifts in host movement may explain the transient coexistence of high- and low-pathogenic disease strains**

This chapter has been submitted to the *Journal of Evolutionary Biology* and was rejected with re-submission allowed

Parasite-induced shifts in host movement may explain the transient coexistence of high- and low-pathogenic disease strains

Abdou M. Fofana, Amy Hurford

Abstract

Many parasites induce decreased host movement, known as lethargy, which can impact disease spread and the evolution of virulence. Mathematical models have investigated virulence evolution when parasites cause host death, but disease-induced decreased host movement has received relatively less attention. Here, we consider a model where, due to the within-host parasite replication rate, an infected host can become lethargic and shift from a moving to a resting state, where it can die. We find that when the lethargy and disease-induced mortality costs to the parasites are not high, then evolutionary bistability can arise, and either moderate or high virulence can evolve depending on the initial virulence and the magnitude of mutation. These results suggest, firstly, the transient coexistence of strains with different virulence, which may explain the coexistence of low- and high-pathogenic strains of avian influenza and human immunodeficiency viruses, and secondly, that medical interventions to treat the symptoms of lethargy or prevent disease-induced host deaths can result in a large jump in virulence and the rapid evolution of high virulence. In complement to existing results that show bistability when hosts are heterogeneous at the population-level, we show that evolutionary bistability may arise due to transmission heterogeneity at the individual host-level.

Keywords: Bistability, Evolutionarily stable strategy, lethargy, movement, trade-off, virulence.

3.1. Introduction

Reduced host movement due to infection, known as lethargy, is a commonly observed disease manifestation, which can affect the parasite transmission rate and disease spread (Eames et al. 2010; Perkins et al. 2016). Many parasites, including those responsible for common illnesses in humans such as measles and the flu, can alter host movement behaviour and induce lethargy, which can prevent infected individuals from socializing and going to work and school (Hart 1988; Holmstad et al. 2006; Eames et al. 2010; Van Kerckhove et al. 2013). Like parasite-induced host mortality, parasite-induced host lethargy can be a direct or an indirect consequence of the rate a parasite produces infectious stages using host resources and/or the clearance rate of the parasite by the immune system of the host (Zitzow et al. 2002; Belser et al. 2013). The severity of lethargy can affect the transmission of a parasite from one host to another because a lethargic host may be less likely to make a direct contact with a susceptible host than a moving host (Ewald 1983; 1994; Day 2001). Thus a trade-off can emerge between the rate of host lethargy and the rate a parasite produces infectious stages within a host.

Animal movement is frequently modelled as a Markov process with probabilistic transitions between discrete movement states, which are defined based on distributions of step lengths and turning angles recovered from animal movement data (Morales et al. 2004; Patterson et al. 2008; Gurarie et al. 2009; Moorter et al. 2010; McKellar et al. 2014; Edelhoff et al. 2016; Teimouri et al. 2018). These discrete state movement models inspire our model formulation, as our epidemiological model considers two infective classes: moving and resting (or lethargic), which have distinct epidemiological characteristics due to distinct movement behaviours. A number of previous studies

have proposed similar epidemic models with coupled behaviour-disease classes and transitions from one class to another (Perra et al. 2011; Fenichel et al. 2011; Wang et al. 2015; Verelst et al. 2016), and recent works highlight the need to combine modelling frameworks from the epidemiological and animal movement literatures (Fofana and Hurford 2017; Dougherty et al. 2017).

We formulate a behaviour-disease model to investigate the role of host movement as an underlying process for an evolutionary trade-off between the rate of parasite transmission and the production of parasite transmission stages within a host, which determines the level of virulence a parasite causes in its host. During the past three decades the trade-off theory has emerged as an accepted explanation for different levels of virulence (Read 1994; Bull 1994; Ebert and Herre 1996; Frank 1996; Lipsitch and Moxon 1997; Alizon et al. 2009; Alizon and Michalakis 2015; Cressler et al. 2016). This theory assumes that high virulence or slow recovery rates are the consequence of the parasite producing transmission stages at a high rate within a host (Anderson and May 1982; Antia et al. 1994; Gilchrist and Sasaki 2002; Alizon and van Baalen 2005). For example, when the transmission-virulence trade-off has a saturating form then parasites will evolve towards an intermediate level of virulence (Anderson and May 1982; Ebert and Herre 1996; Frank 1996).

The trade-off theory has received some empirical support (Paul et al. 2004; Fraser et al. 2007; de Roode et al. 2008; Doumayrou et al. 2013; Fraser et al. 2014; Williams et al. 2014; Blanquart et al. 2016), but has been criticized for its restrictive definition of the term virulence (Alizon et al. 2009). Theoretical analyses of the evolution of virulence frequently define virulence as parasite-induced host mortality and ignore non-lethal effects due to parasite infection (Anderson and May 1982; Frank 1996;

Alizon et al. 2009). Notable mathematical formulations that have investigated non-lethal parasite virulence have considered parasite-induced host sterility (O’Keefe and Antonovics 2002; Bonds 2006; Lively 2006; Abbate et al. 2015; Best et al. 2017) and parasite-induced reduced host growth (Schjørring and Koella 2003), but reduced host movement due to infection has received relatively less attention (but see Ewald 1983; Day 2001).

The aim of this paper is to explicitly represent parasite-induced effects on host movement as a process underlying the transmission-virulence trade-off. Notably, we consider that infected hosts can shift between two discrete movement states: moving and resting and we justify this formulation based on studies from the animal movement literature (Edelhoff et al. 2016; Teimouri et al. 2018). We investigate the evolution of the rate of parasite replication within a host when the infection is potentially lethal and when the infection is non-lethal. We find that the main drivers of the evolutionary dynamics are lethargy and disease-induced mortality costs to the parasite, and when the disease-induced mortality or the lethargy cost is high, then evolution converges towards a parasite strain that induces moderate virulence. For a range of parameter values, where the lethargy and the disease-induced mortality costs are not high, a bistable evolutionary equilibrium occurs. As such, depending on the initial virulence and the magnitude of the effect of mutation, either a parasite strain that induces moderate virulence or a parasite strain that induces high virulence in the host population can evolve. Finally, we discuss how our results can aid in understanding the transient coexistence of parasite strains with different virulence in avian influenza and human immunodeficiency viruses.

3.2. Epidemiological model

To formalize the epidemic model, we couple two discrete movement states (moving and resting) with a Susceptible-Infected-Susceptible (SIS) model. Figure 3.1 describes the epidemic model, and definitions for all the parameters and notations used in this paper are provided in Table 3.1.

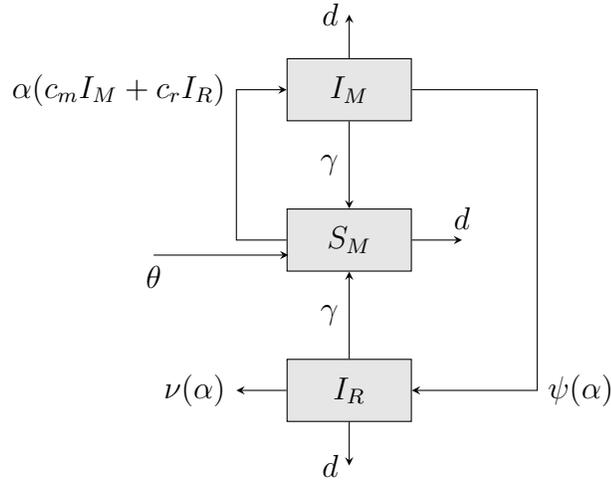


Figure 3.1: An epidemiological model where the compartments represent combinations of host epidemiological statuses and movement states. The symbols S and I indicate susceptible and infected, and the subscripts M and R indicate the moving and the resting states. The arrows indicate the flow of individuals between each compartment with the corresponding rates. Susceptible hosts, S_M , are recruited through immigration at the rate θ and become infected at a per capita rate $\alpha(c_m + c_r)$. Following infection the infected host enters the moving state (I_M). The infected host can become lethargic and enter the resting state (I_R) at the rate $\psi(\alpha)$ or recover from the disease before lethargy and become susceptible again at the rate γ . When the infected host becomes lethargic it can die from the disease at the rate $\nu(\alpha)$ or recover from the disease and become susceptible again at the rate γ . Finally, we assume that a host can die naturally at the rate d independently of the movement state and epidemiological status.

We assume that susceptible hosts are always in the moving state, and infected hosts are in the moving state before lethargy and in the resting state during lethargy. Let S_M , I_M and I_R denote the numbers of susceptible hosts in the moving state, infected hosts in the moving state, and infected hosts in the resting state respectively.

Table 3.1: List of notations and definitions

Symbols	Definitions
(Epidemic)	
α	Within-host parasite net replication rate.
$\psi(\alpha)$	Parasite-induced host lethargy rate.
$\nu(\alpha)$	Parasite-induced host mortality rate.
c_m	The per capita host-host contact rate in the moving state.
c_r	The per capita host-host contact rate in the resting state.
γ	Host recovery rate.
d	Host natural mortality rate.
θ	Host immigration rate.
R_0	The expected number of secondary cases by a primary case in a susceptible population.
ρ	The fraction of infected hosts that experience lethargy.
σ	Case fatality ratio given lethargy.
χ	Case fatality ratio ($\rho \times \sigma$).
Symbols	Definitions
(Evolution)	
α_1	Within-host net replication rate of the resident strain.
α_2	Within-host net replication rate of the mutant strain.
α^*	Evolutionarily stable or convergence stable net replication rates (ESS or CSS).
$\psi(\alpha_1)$	Parasite-induced host lethargy rate of the resident strain.
$\psi(\alpha_2)$	Parasite-induced host lethargy rate of the mutant strain.
$\nu(\alpha_1)$	Parasite-induced host mortality rate of the resident strain.
$\nu(\alpha_2)$	Parasite-induced host mortality rate of the mutant strain.

The epidemiological dynamics of the host population are described by,

$$\frac{dS_M}{dt} = \theta + \gamma (I_M + I_R) - S_M (\Lambda + d) \quad (3.1)$$

$$\frac{dI_M}{dt} = \Lambda S_M - I_M [d + \gamma + \psi(\alpha)] \quad (3.2)$$

$$\frac{dI_R}{dt} = \psi(\alpha) I_M - I_R [d + \gamma + \nu(\alpha)], \quad (3.3)$$

where $\Lambda = \alpha (c_m I_M + c_r I_R)$ is the force of infection.

We assume that a susceptible host becomes infected by making a direct contact

with an infected host that is either in the moving or the resting state. We formulate these two infection events separately because a lethargic host in the resting state is less likely to make a direct contact than a non-lethargic host in the moving state. In order to capture this idea, we decompose the transmission coefficient frequently denoted β , into two components: the rate of direct contact and the probability of disease transmission given a contact between a susceptible host and an infected host (Day 2001). The first component depends on the movement state of the host, and we assume that an infected host is less likely to make a direct contact in the resting state compared to the moving state ($c_r < c_m$) (Ewald 1983; 1994; Day 2001; Lloyd-Smith et al. 2004). The second component depends on parasite properties only, and we assume that the probability of disease transmission given an infectious contact is proportional to the net replication rate of a parasite within a host which we denote α (Brauer 2008; Diekmann et al. 2012). The within-host parasite net replication rate (α) is the difference between parasite replication rate and the parasite clearance rate by the immune system of the host (Lipsitch and Moxon 1997). To formulate the infection process we apply the mass-action law, thus the number of new infections per unit time due to one infected host in the moving state is $\alpha c_m S_M$ and in the resting state is $\alpha c_r S_M$. We assume that the parasite has a short incubation period, meaning that an infected host is immediately infectious.

An infected host in the moving state can become lethargic and enter the resting state at the rate $\psi(\alpha)$, which is the parasite-induced host lethargy rate, and the infected host can die from the disease in the resting state at the rate $\nu(\alpha)$, which is the parasite-induced host mortality rate. Both the rate of lethargy and the rate of host death due to infection depend on the within-host parasite net replication rate,

and we ignore the details of the dynamics between the parasite replication rate and the immune system within the host for simplicity. We assume that an infected host can recover either in the moving or the resting state at a constant rate γ and become susceptible again. A host can be reinfected multiple times during the course of its life, thus this type of model is appropriate for infectious diseases that confer no immunity such as rhinoviruses responsible for the common cold in humans (May 1986; Brauer 2008). Finally, we assume that susceptible and infected hosts can die naturally at a constant rate d , and new susceptible hosts are recruited through immigration at the rate θ .

The system of equations (3.1-3.3) exhibits two equilibria: one disease-free and one endemic equilibrium. We use the next-generation matrix approach (see van den Driessche and Watmough 2002) to derive the basic reproduction number which is given by,

$$R_0 = \left[\frac{\alpha c_m}{d + \gamma + \psi(\alpha)} + \frac{\alpha c_r}{d + \gamma + \nu(\alpha)} \times \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)} \right] S_M^*, \quad (3.4)$$

where $S_M^* = \theta/d$ represents the number of susceptible hosts in the absence of the disease (see Appendix B of supporting information for the derivation of R_0). Equation (3.4) is the expected number of secondary cases generated by a primary case in a completely susceptible host population, and it informs the outcome of the disease when rare in the host population (Diekmann et al. 2012). If equation (3.4) is less than one then no outbreak occurs, and if equation (3.4) is greater than one then an epidemic occurs and the system reaches a stable endemic equilibrium as long as the input of susceptible hosts through immigration and recovery is permanent (Brauer 2008). Equation (3.4) is the sum of the expected number of new infections generated

by an infected host in the moving state,

$$\frac{\alpha c_m S_M^*}{d + \gamma + \psi(\alpha)},$$

and the resting state multiplied by the probability of entering the resting state,

$$\frac{\alpha c_r S_M^*}{d + \gamma + \nu(\alpha)} \times \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)}.$$

To characterize the degree of non-lethal and lethal virulence associated with the net replication rate of a parasite within a host (α) we define:

$$\rho = \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)}, \quad (3.5)$$

$$\sigma = \frac{\nu(\alpha)}{d + \gamma + \nu(\alpha)}, \quad (3.6)$$

and

$$\chi = \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)} \times \frac{\nu(\alpha)}{d + \gamma + \nu(\alpha)}, \quad (3.7)$$

where equations (3.5), (3.6) and (3.7) are the fraction of hosts that become lethargic, the case fatality ratio given lethargy and the case fatality ratio, respectively. We consider equation (3.5) as a measure of non-lethal virulence and equation (3.7) as a measure of lethal virulence a parasite causes to the host.

3.3. Evolution model

To investigate the evolution of the within-host parasite net replication rate, we assume that a resident parasite strain with a net replication rate α_1 is present in the host population at a locally stable endemic equilibrium and a rare mutant strain with

a net replication rate α_2 arises in the population. Assuming that only one strain can infect one host at the same time, the evolutionary dynamics are described by the following system of differential equations:

$$\frac{dS_M}{dt} = \theta + \gamma (I_{M1} + I_{M2} + I_{R1} + I_{R2}) - S_M (\Lambda_1 + \Lambda_2 + d) \quad (3.8)$$

$$\frac{dI_{M1}}{dt} = \Lambda_1 S_M - I_{M1} [d + \gamma + \psi(\alpha_1)] \quad (3.9)$$

$$\frac{dI_{M2}}{dt} = \Lambda_2 S_M - I_{M2} [d + \gamma + \psi(\alpha_2)] \quad (3.10)$$

$$\frac{dI_{R1}}{dt} = \psi(\alpha_1) I_{M1} - I_{R1} [d + \gamma + \nu(\alpha_1)] \quad (3.11)$$

$$\frac{dI_{R2}}{dt} = \psi(\alpha_2) I_{M2} - I_{R2} [d + \gamma + \nu(\alpha_2)], \quad (3.12)$$

where $\Lambda_1 = \alpha_1 (c_m I_{M1} + c_r I_{R1})$ and $\Lambda_2 = \alpha_2 (c_m I_{M2} + c_r I_{R2})$ are the force of infections associated with the resident and the mutant strains respectively. Let I_{M1} and I_{R1} denote the number of infected hosts in the moving and the resting states respectively infected with the resident strain, and I_{M2} and I_{R2} denote the number of infected hosts in the moving and the resting states respectively infected with the mutant strain.

To investigate the evolutionary dynamics, we analyze the stability of the mutant-free equilibrium (the endemic equilibrium of the system (3.1-3.3)) using the next-generation matrix approach for evolutionary invasion analysis (see, Hurford et al. 2010). We derive the expression for the invasion fitness, $R(\alpha_2, \alpha_1)$, which is the expected lifetime infection success of a rare mutant strain, α_2 , in a host population where the resident strain, α_1 , is at endemic equilibrium, and it gives the conditions for α_2 to replace α_1 (see, Otto and Day 2007; Dieckmann 2002). The stability analysis of the mutant-free equilibrium reveals that α_2 replaces α_1 at the endemic equilibrium

if,

$$R(\alpha_2, \alpha_1) = \frac{\alpha_2 \left(c_m [d + \gamma + \nu(\alpha_2)] + c_r \psi(\alpha_2) \right)}{\left[d + \gamma + \nu(\alpha_2) \right] \left[d + \gamma + \psi(\alpha_2) \right]} \times \frac{\left[d + \gamma + \nu(\alpha_1) \right] \left[d + \gamma + \psi(\alpha_1) \right]}{\alpha_1 \left(c_m [d + \gamma + \nu(\alpha_1)] + c_r \psi(\alpha_1) \right)} - 1 > 0. \quad (3.13)$$

Details of the derivation of $R(\alpha_2, \alpha_1)$ are provided in Appendix B as supporting information. Equation (3.13) suggests that if,

$$\frac{\alpha_2 c_m}{d + \gamma + \psi(\alpha_2)} + \frac{\alpha_2 c_r}{d + \gamma + \nu(\alpha_2)} \times \frac{\psi(\alpha_2)}{d + \gamma + \psi(\alpha_2)} > \frac{\alpha_1 c_m}{d + \gamma + \psi(\alpha_1)} + \frac{\alpha_1 c_r}{d + \gamma + \nu(\alpha_1)} \times \frac{\psi(\alpha_1)}{d + \gamma + \psi(\alpha_1)}, \quad (3.14)$$

then the mutant strain (α_2) replaces the resident strain (α_1). Therefore, a resident strain that maximizes

$$R^*(\alpha_1) = \frac{\alpha_1 c_m}{d + \gamma + \psi(\alpha_1)} + \frac{\alpha_1 c_r}{d + \gamma + \nu(\alpha_1)} \times \frac{\psi(\alpha_1)}{d + \gamma + \psi(\alpha_1)}, \quad (3.15)$$

can not be invaded, and as such its net replication rate is evolutionarily stable (ESS α^*). This result is valid provided that there is always susceptible hosts to infect (Dieckmann 2002). The expression $R^*(\alpha)$ is the expected secondary infections by a single infected host per susceptible host in the moving and the resting states.

To investigate the evolutionary dynamics we need to determine the general form of the functions $\psi(\alpha)$ and $\nu(\alpha)$ which are parasite-induced host lethargy and parasite-induced host mortality rates respectively. We assume that both $\psi(\alpha)$ and the $\nu(\alpha)$ are determined by the rate a parasite replicates within a host (α), meaning that a trade-off exists between α and $\psi(\alpha)$ on one hand and $\nu(\alpha)$ on the other hand. We consider parasite-induced host lethargy rate ($\psi(\alpha)$) to reflect a form of non-lethal virulence because the lethargic state is harmful to the host as the host is less able to engage in activities essential to survival (i.e., foraging, provisioning for offspring,

evading predators), and parasite-induced host mortality rate ($\nu(\alpha)$) reflects a form of lethal virulence because the host dies due to infection. A number of studies support the existence of a positive correlation between parasite load (a measure of the net replication rate of a parasite within a host, α) and host survival (Timms et al. 2001; Paul et al. 2004; Brunner et al. 2005; Bell et al. 2006; de Roode et al. 2006; 2008; 2009; de Roode and Altizer 2010). Finnerty et al. (2018) found that both the total running time and the running distance of infected cane toads decreases as the number of lungworms increases. Therefore, we assume that both $\psi(\alpha)$ and $\nu(\alpha)$ are increasing functions.

To investigate the within-host parasite net replication rate that is evolutionarily stable (ESS α^*) and to determine the conditions for the ESS α^* to be convergence stable (CSS α^*), we perform an evolutionary invasion analysis (Dieckmann 2002; Otto and Day 2007). When a parasite strain with the α value that is evolutionarily stable is dominant in the host population then no parasite strain with a different α value can replace it. An evolutionarily stable within-host net replication rate (ESS α^*) that is also convergence stable (CSS α^*) is an evolutionary attracting equilibrium, in other words parasites evolve towards α^* by a succession of small mutations and selection (Eshel 1983; Dieckmann 2002; Dieckmann 2004; Otto and Day 2007). To illustrate our analytical results we use Pairwise Invasibility Plot (PIP), which is a graphical representation used for evolutionary invasion analysis, and numerical simulation (Geritz et al. 1998; Dieckmann 2002; Dieckmann 2004).

3.4. Results

We derive the within-host parasite net replication rate that is evolutionarily stable, and the conditions for an ESS α^* to be convergence stable. Also, we investigate the

effects of some important parameter values on the ESS α^* and the corresponding virulence (equations 3.5 and 3.7).

3.4.1. The evolutionarily stable within-host parasite net replication rate (ESS α^)*

At the within-host parasite net replication rate that is evolutionarily stable, the expected number of new infections generated by an infected host in the moving and the resting state is maximal (equation 3.15). To determine the ESS α^* , we evaluate the first derivative of the invasion fitness ($R(\alpha_2, \alpha_1)$, equation 3.13) equal to zero at $\alpha_1 = \alpha_2 = \alpha^*$, and we solve for α^* . To verify under what conditions α^* is a maximum, we require the second derivative of $R(\alpha_2, \alpha_1)$ at $\alpha_1 = \alpha_2 = \alpha^*$ to be less than zero. We find that when

$$\alpha^* = \frac{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \psi][d + \gamma + \nu]}{[(c_m - c_r)(d + \gamma) + c_m\nu][d + \gamma + \nu]\psi' + [d + \gamma + \psi]c_r\psi\nu'}, \quad (3.16)$$

equation (3.15) is maximal, where ψ' and ν' respectively, are the first derivatives of $\psi(\alpha)$ and $\nu(\alpha)$ with respect to α_2 evaluated at α^* . When $\psi(\alpha)$ is a concave up trade-off and $\nu(\alpha)$ has a concave up or a linear form, then equation (3.16) is a maximum and a biologically feasible evolutionarily stable within-host parasite net replication rate, ESS α^* (Equation 3.16 is an implicit expression of the ESS and see Appendix B for details). Where PIP and dynamical simulation are presented, we model the parasite-induced host lethargy rate as $\psi(\alpha) = \alpha^a$ ($a > 1$ is the exponent parameter), and the parasite-induced host mortality rate proportional to the parasite-induced host lethargy rate, $\nu(\alpha) = b\psi(\alpha)$ (where b is the ratio of the parasite-induced mortality rate to the lethargy rate). The lower the parameter b the lower the fraction of lethargic hosts that die in the resting state, and so decreasing b may represent host adaptations (host resistance) or medical interventions that prevent disease-induced host death.

Figures 3.2a and 3.2b illustrate the concave-up trade-offs ($\psi(\alpha)$ and $\nu(\alpha)$), and the non-lethal and lethal virulence (equations 3.5 and 3.7) a parasite with a given net replication rate causes to its host. Supporting details for the evolutionary invasion analysis and description of the simulation are provided in Appendix B. The code used for PIP, simulation and Movie is available as electronic supplementary materials S3.1-S3.3, and is publicly available on Figshare doi:10.6084/m9.figshare.8059781.v2.

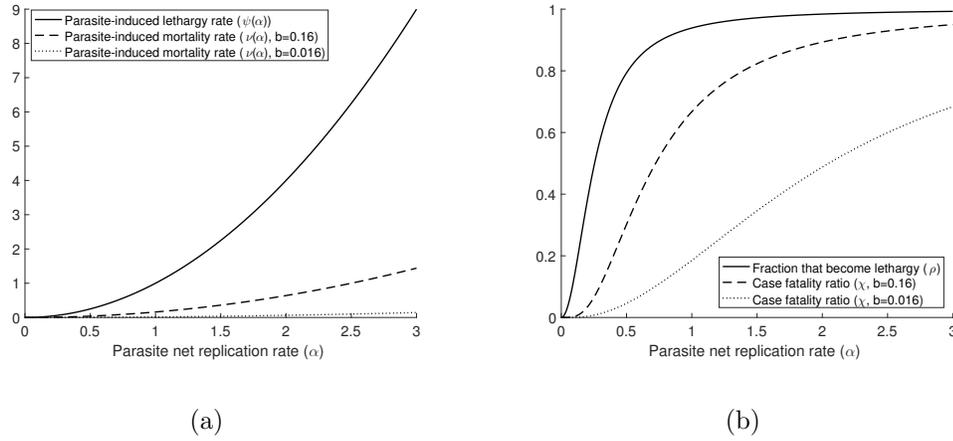


Figure 3.2: The concave up trade-offs in (a), where solid line is $\psi(\alpha)$, dashed and dotted lines are $\nu(\alpha)$ for different b , and the corresponding virulence in (b), where solid line is the fraction of hosts that experience lethargy (equation 3.5), dashed and dotted lines are the case fatality ratios (equation 3.7) for different b . The parameter b is the ratio of the parasite-induced mortality rate to the lethargy rate, and b can be seen as medical interventions that prevent disease-induced host deaths. For example when the net replication rate of the parasite strain that is present in the host population is $\alpha = 3$, then $\approx 100\%$ of infected hosts will experience lethargy and the case fatality ratio is $\approx 95\%$ and $\approx 68\%$ for $b = 0.16$ and $b = 0.016$ respectively. We model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = b\alpha^2$. We set $d = 0.0001$ and $\gamma = 0.065$.

3.4.2. Evolutionary bistability arises when hosts make contacts during lethargy

Hosts make no contacts in the resting state ($c_r = 0$)

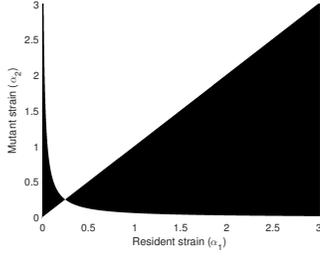
We found that if hosts make no contact in the resting state ($c_r = 0$) then the evolutionarily stable within-host net replication rate, ESS α^* , is also convergence stable,

suggesting that parasites will evolve towards α^* by a succession of small mutations and selection. The numerical results confirm that when hosts make no contacts in the resting state ($c_r = 0$) then parasites evolve towards an intermediate α^* which corresponds to moderate non-lethal virulence (moderate fraction of infected hosts becoming lethargic, equation 3.5), and moderate lethal virulence (moderate case fatality ratio, equation 3.7). The PIP shows that a resident strain with a within-host net replication rate corresponding to α^* can not be replaced by any rare mutant strain, and both the PIP and the dynamical simulation show that no matter the initial α evolution converges towards the intermediate ESS α^* (Figures 3.3a and 3.3b).

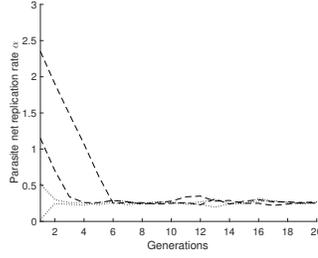
When $c_r = 0$, only the moving state contributes to the total number of secondary infections and an intermediate value of α_1^* maximizes equation (3.15), which is the expected secondary infections by a single infected host per susceptible host, $R^*(\alpha_1)$ (Figure 3.4a). An intermediate value of α_1 is optimal because for low values of α the probability of disease transmission given a contact is too low, and for high α values the duration of the moving state, which is the only state where parasites can be transmitted, is too short due to the concave up trade-off, $\psi(\alpha)$. Therefore, decreased $R^*(\alpha_1)$ due to lethargy occurring earlier in the infection, which we term the lethargy cost, is the main factor that maintains intermediate α^* and prevents evolution towards higher virulence.

Hosts make contacts in the resting state ($c_r > 0$)

We found that when hosts make contacts in the resting state ($c_r > 0$) evolutionary bistability, with a lower and an upper ESS, is possible for a set of parameter values (Figure 3.3c). As such, the evolutionary trajectory can depend on the initial value of α . For all initial within-host parasite net replication rates (α) below a critical level,

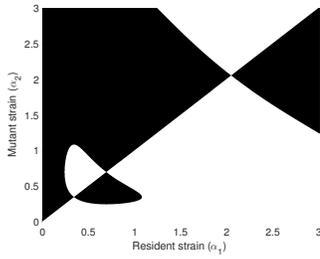


(a)

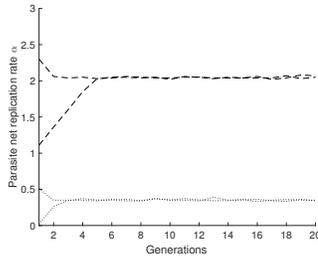


(b)

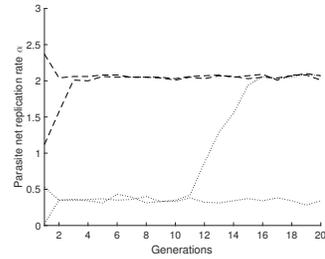
Hosts make no contacts in the resting state ($c_r = 0$)



(c)



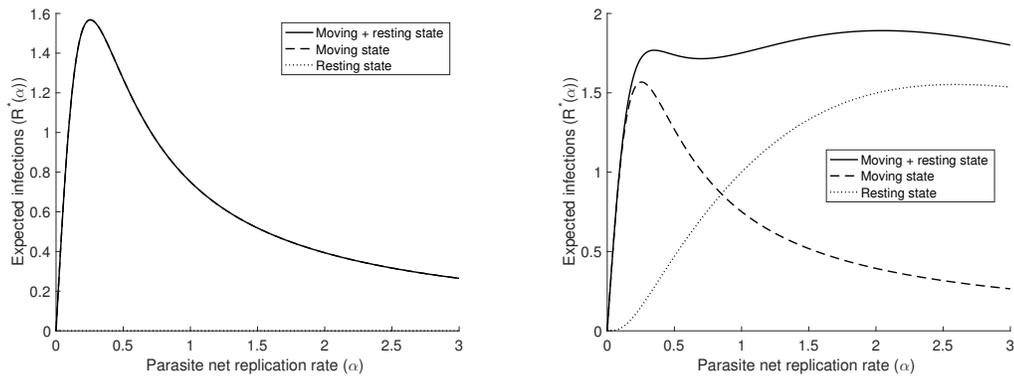
(d)



(e)

Hosts make contacts in the resting state ($c_r > 0$)

Figure 3.3: Pairwise Invasibility Plots (PIP) and dynamical simulations illustrating the evolutionary dynamics when hosts make no contacts in the resting state (top row) and when hosts make contacts in the resting state (bottom row). Panels (a) and (c) are PIPs, and for a given combination (α_1, α_2) , white indicates that the rare mutant goes extinct (equation 3.13 is negative), and black indicates that the rare mutant replaces the resident (equation 3.13 is positive). Panels (b), (d) and (e) are dynamical simulations of the evolution of parasite net replication rate (α) for different initial α . Dotted lines are evolutionary trajectories for initial α values below the *invasible repeller* ($\alpha \approx 0.7$) and dashed lines are evolutionary trajectories for initial α values above the *invasible repeller*. In (a) the unique intersection ($\alpha_1^* \approx 0.25$) is an ESS and in (b) evolution converges towards this ESS for all initial α . In (c) from low to high α , the first intersection ($\alpha^* \approx 0.33$) is an ESS (termed the lower ESS), the second ($\alpha \approx 0.7$) is an *invasible repeller*, and the third ($\alpha^* \approx 2$) is an ESS (termed the upper ESS). In (d) only small-effect mutations occur and in (e) large-effect mutations can occur. For all figures, we model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = 0.01\alpha^2$, and we set $c_m = 0.8$, $d = 0.0001$, $\gamma = 0.065$, and $c_r = 0.08$, except the top row figures where $c_r = 0$.



(a) Lethargic hosts make no contacts ($c_r = 0$) (b) Lethargic hosts make contacts ($c_r > 0$)

Figure 3.4: The expected secondary infections by a single infected host per susceptible host (equation 3.15) in the moving state (dashed black line), the resting states (dotted black line) and during the entire infectious period (solid black lines) as a function of within-host net replication rate α . In (a) we set $c_r = 0$, thus infection is possible only in the moving state. Equation (3.15) is maximized at $\alpha^* \approx 0.25$, and maximizing the number of secondary infections per susceptible hosts in the moving state also maximizes this quantity for the entire infectious period. In (b) we set $c_r = 0.08$, thus both moving and lethargic hosts contribute to the total infections and equation (3.15) has two local maxima: $\alpha^* \approx 0.33$ and $\alpha^* \approx 2$, corresponding to the lower and the upper ESS respectively. A parasite strain at the lower ESS is mainly transmitted in the moving state, whereas a parasite strain at the upper ESS is mainly transmitted in the resting state. For all graphs we model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$, $\nu(\alpha) = 0.01\psi(\alpha)$, and we set $d = 0.0001$ and $\gamma = 0.065$.

parasites evolve towards the lower ESS α^* by a succession of small mutations and selection. This critical level corresponds to an *invasible repeller* which is an invulnerable and non-convergent evolutionary equilibrium (Evans et al. 2010; Otto and Day 2007; Dieckmann 2004; Dieckmann 2002). In contrast, for initial α values above the *invasible repeller* parasites evolve towards the upper ESS α^* by a succession of small mutations and selection (Figures 3.3c and 3.3d).

When hosts make contacts in the resting state ($c_r > 0$) both moving and resting states contribute to the total number of secondary infections, and the expected secondary infections by one infected host per susceptible host ($R^*(\alpha_1)$) can have more than one maxima (i.e., local and global maxima). The parasite strain with the lower

ESS α^* is mostly transmitted in the moving state (Figure 3.4b, dashed line), whereas the parasite strain with the upper ESS α^* induces lethargy very early in the infection and is mostly transmitted in the resting state (Figure 3.4b, dotted line). For any α above the upper ESS the duration of the entire infectious period is too short due to the concave up trade-off $\nu(\alpha)$, and the overall infection success of a parasite, $R^*(\alpha_1)$, decreases. Therefore, decreased $R^*(\alpha_1)$ due to shorter infectious period, which we term the disease-induced mortality cost, limits evolution towards much higher α and maintains the upper ESS α^* (Figure 3.4b). In absence of a concave-up $\nu(\alpha)$ trade-off, ever increasing values of α will evolve.

In addition to the initial within-host parasite net replication rate (Figure 3.3d), its variability within the parasite population can play an important role in the evolutionary outcome. For example, when large-effect mutations can occur and a rare mutant strain can be very different from the resident strain, then parasites can evolve towards the upper ESS even if α is initially below the *invasible repellor* (Figure 3.3e). However when α is less variable within the parasite population, because only small-effect mutations can occur, the evolutionary outcome depends on the initial α value. The bistability suggests that a transient coexistence of two strains with different virulence is possible in the host population. For example, when a strain that induces moderate virulence corresponding to the lower ESS α^* is present in the host population and when large-effect mutations can occur, then any mutant strain with a net replication rate higher than the *invasible repellor* can emerge and produce an outbreak. As such, evolution can maintain two strains with low and high virulence in a transient coexistence, before eventually the elimination of one strain by competitive exclusion.

3.4.3. The effects of model parameters on the evolutionary dynamics

To gain a better understanding of how the parameter values affect the evolutionary dynamics, we graph the case fatality ratio (equation 3.7) corresponding to the evolutionary singular points as a function of the host contact rate during lethargy (c_r) and the constant b (the ratio of disease-induced host mortality rate to disease-induced host lethargy rate). We found that reduced c_r selects for parasite strains that induce lower virulence (Figure 3.5a, see also Movie in Figure B.1 in Appendix B). One way that c_r could be reduced is through interventions to reduce infectious contacts (e.g., isolation of infectious people), and our results suggest that these interventions would select for parasite strains that induce lower virulence. In contrast, medical interventions that treat the symptoms of lethargy, but do not prevent parasite transmission (e.g., painkillers), might increase c_r and select for parasite strains that induce higher virulence.

Similarly, as b decreases parasite strains that induce higher virulence are selected. One way that b might decrease is through medical interventions that reduce disease-induced host death rate, and our results suggest that these interventions are more likely to induce higher virulence (Figure 3.5b, see also Movie in Figure B.2 in Appendix B). Examples of these medical interventions are imperfect vaccines that decrease the probability that the host dies due to infection, but do not prevent the transmission of infectious stages (Gandon et al. 2001; 2003; Read et al. 2015). Moreover, when c_r as well as b increases then virulence increases slowly except for a range of c_r and b values where a backward bifurcation occurs with an evolutionary bistable equilibrium (Figures 3.5a and 3.5b). As such, a small increase in c_r or a small decrease in b within this range of values can result in a large increase in the evolutionary equilibrium and

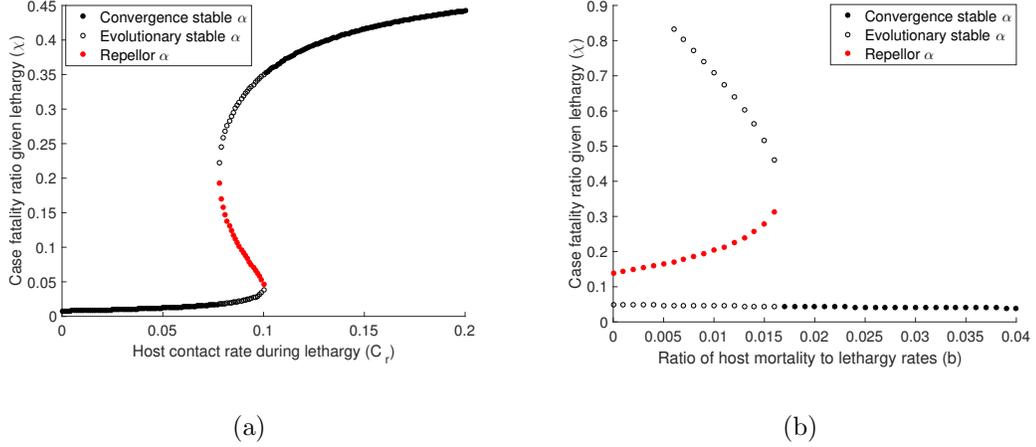


Figure 3.5: Increasing contact rate in the resting state (c_r) and increasing the ratio of host mortality to lethargy rates (b) induce a backward bifurcation in the evolutionary dynamics. In (a) we set $c_m = 0.8$, $b = 0.016$, $d = 0.0001$, $\gamma = 0.065$ and we graph the case fatality ratio (equation 3.7) corresponding to evolutionary equilibria for c_r values from 0 to 0.2. For c_r values between ≈ 0.07 and 0.1 there are two ESS (black open circles) separated by an *invasible repellor* (red filled circles), but outside this range there is only one ESS which is also a CSS (black filled circles). In (b) we set $c_m = 0.8$, $c_r = 0.08$, $d = 0.0001$, $\gamma = 0.065$ and we graph the case fatality ratio corresponding to the evolutionary equilibria for b values from 0 to 0.04. For b values between ≈ 0.006 and 0.016 there are two ESS separated by an *invasible repellor*, but outside this range there is only one ESS which is also a CSS. We choose to plot only the corresponding lethal virulence (equation 3.7) in function c_r and b , but the result is the same for non-lethal virulence (equation 3.5). For all graphs we model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = b\psi(\alpha)$.

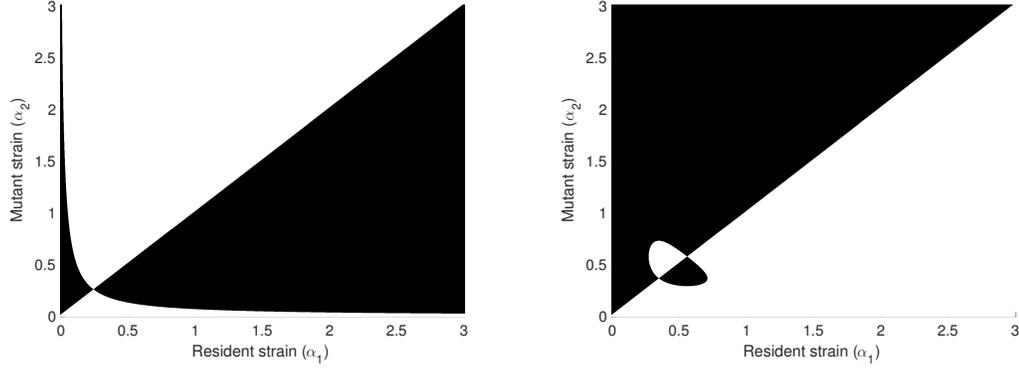
the corresponding virulence. When all other parameters are kept fixed, a 0.01 increase in c_r can select for a strain that is ≈ 12 -fold more virulent, and a 0.01 decrease in b can select for a strain that is ≈ 15 -fold more virulent.

3.4.4. Evolutionary dynamics when parasite infection is non-lethal

We investigate the evolution of the within-host parasite net replication rate (α) when no infected host dies from the disease ($b = 0$), and we derive the corresponding non-lethal virulence (the fraction of infected hosts that become lethargic, equation 3.5). Many human parasites such as rhinoviruses and chickenpox enter this category because they do cause lethargy, but negligible or no host mortality (Walther and

Ewald 2004). Also, for many human parasites a large proportion of infected individuals eventually recover from the disease after they receive appropriate medical treatment, and only a small proportion die from the disease. When no infected host dies from the disease then the cost of lethargy is the main factor that governs the evolutionary dynamics, and this cost is higher when $c_r = 0$ or c_r is small. Evolution converges towards a parasite strain that is mainly transmitted in the moving state resulting in a high fraction of hosts that avoid lethargy when $c_r = 0$ (Figure 3.6a, and details of the model and the evolutionary dynamics are provided in Appendix B as supporting information). In contrast, when the transmission rate in the resting state increases due to increased c_r , the incentive to avoid lethargy is lessened and without a disease-induced mortality cost ($b = 0$), the parasite can evolve ever increasing within host net replication rate with all infected hosts experiencing lethargy (Figure 3.6b). Finally, when there is no disease-induced mortality the evolutionary bifurcation diagram as a function of host contact rate in the resting state (c_r) is similar to Figure 3.5a but without the upper ESS.

Throughout this paper, we assumed that the probability of disease transmission given an infectious contact, which is proportional to the within-host parasite net replication rate (α), is the same in the moving and the resting states, but the probability of disease transmission given an infectious contact may be higher in the resting state because of a higher parasite load. We investigated the case where the probability of disease transmission given an infectious contact (proportional to α) is higher in the resting state than the moving state ($\alpha_m > \alpha_r$, where α_m and α_r are the within-host parasite net replication rates in the moving and the resting states respectively). We found that the results are qualitatively similar to the case where the probability of dis-



(a) Lethargic hosts make no contacts ($c_r = 0$) (b) Lethargic hosts make contacts ($c_r > 0$)

Figure 3.6: Pairwise Invasibility Plots (PIP) illustrating the evolutionary dynamics when parasite infection is non-lethal ($b = 0$). In (a) infected hosts make no contacts in the resting state ($c_r = 0$) and in (b) infected hosts make contacts in the resting state ($c_r > 0$). See the caption of Figure 3.3 for how to read a PIP. The unique equilibrium in (a) is an ESS ($\alpha^* \approx 0.25$) because it is non-invasible by any rare mutant strain, and a CSS because parasites evolve towards this evolutionary equilibrium by a succession of small mutations and selection independently of the initial α value. In (b) from low to high α , the first equilibrium ($\alpha^* \approx 0.35$) is an ESS and the second equilibrium ($\alpha \approx 0.58$) is an *invasible repellor*. For all figures we model the concave-up trade-off using a power function $\psi(\alpha) = \alpha^2$, and we set $d = 0.0001$, $\gamma = 0.065$, $c_m = 0.8$ and $c_r = 0.08$ except in (a) where $c_r = 0$.

ease transmission transmission given an infectious contact is the same in the moving and the resting states (Figure B.3 in Appendix B).

3.5. Discussion

Disease-induced mortality as an unavoidable consequence of increasing parasite transmission is the most frequently evoked explanation for the evolution and the maintenance of virulence. While parasites rarely induce death in their hosts, it is common that parasites cause reduced movement (lethargy), which can result in a behavioural shift from a moving to a resting state. As such, our epidemiological model considers discrete movement states, moving and resting, with a transition rate between the states, to understand how non-lethal in combination with lethal parasite-induced harm influences the evolution of the parasite net replication rate and the corresponding

virulence.

We found that when infected hosts make no contacts in the resting state, $c_r = 0$, or when the ratio of the disease-induced mortality rate to the lethargy rate (b) is high, then a parasite strain that is mainly transmitted in the moving state and induces moderate virulence (non-lethal and lethal virulence) will evolve (Figures 3.3a, 3.3b and 6a). In contrast, when $c_r > 0$ and the ratio b is low then high virulence can evolve, and a bistable evolutionary equilibrium is possible for a range of parameters values (Figures 3.3c, and 3.6b). As such, either a parasite strain that is mainly transmitted in the moving state and induces moderate virulence (lower ESS) or a parasite strain that is mainly transmitted in the resting state and induces high virulence (upper ESS) can evolve, depending on the initial virulence and the magnitude of the effect of mutation (Figures 3.3d and 3.3e). Furthermore, we show that medical interventions to treat the symptoms of lethargy (increased c_r) or reduce disease-induced host death (decreased b) can select for high virulence, and a small change in c_r and b can result in a large shift in the evolutionary dynamics due to the evolutionary bistability (Figures 3.5a and 3.5b).

Classic models of virulence evolution which ignore disease-induced lethargy and restrict virulence to parasite-induced host death suggest that the disease-induced mortality cost is the main factor that maintains intermediate virulence (Anderson and May 1982; Frank 1996). However, our results suggest that lethargy cost can also maintain an intermediate virulence whether parasite infection is lethal or non-lethal (see also Day 2001). It has been challenging to validate the tradeoff theory in the context of lethal virulence (Alizon et al. 2009; Alizon and Michalakis 2015; Cressler et al. 2016), but formulating the trade-off as a lethargy cost may facilitate experimental validation

of the trade-off theory.

Previous studies have demonstrated that evolutionary bistable virulence can emerge from a variety of ecological factors. Gandon et al. (2003) showed that imperfect vaccines that do not prevent infection, but limit parasite growth in infected hosts, can select for either low or highly virulent strains depending on the initial parasite virulence for intermediate vaccination coverage. Bistability occurs because intermediate vaccination coverage creates a heterogeneous host population, with vaccinated and unvaccinated hosts, and the anti-growth component of the vaccine can maintain high virulence whereas the anti-infection component can maintain low virulence. Similar conclusions are reached in the case where the vaccine increases the efficacy of host immunity, and the functional relationship between virulence and transmission emerges from within-host dynamics (André and Gandon 2006). Boots et al. (2004) found that for infectious diseases that confer long-lived immunity, when some of the infections occur globally, whereas others occur locally, then either an avirulent or a highly virulent strain can evolve depending on the initial parasite virulence, and several other examples of evolutionary bistability are given in van Baalen (1998), Boldin and Kisdi (2012) and Fleming-Davies et al. (2015). In our work, evolutionary bistability arises due to the two movement states with distinct epidemiological characteristics that create temporal heterogeneity in disease manifestation at individual host-level. As such, disease transmission in the moving state maintains parasite strains with moderate virulence, whereas disease transmission in the resting state maintains parasite strains that induce high virulence.

In the formulation of our model, we made several assumptions that require further discussion. We assumed that the parasite affects host movement via a trade-off

between the parasite net replication rate and the parasite-induced host lethargy rate ($\psi(\alpha)$). Finnerty et al. (2018) demonstrates this relationship, and a number of studies have reported that human and non-human parasites frequently induce lethargy in their hosts (Hart 1988; Holmstad et al. 2006; Ghai et al. 2015). This trade-off could be assessed experimentally by measuring the relationship between parasite load or within host parasite growth rate and the fraction of infected hosts that become lethargic using a scoring system based on the activity level of infected hosts (Reuman et al. 1989; Zitzow et al. 2002). We assumed that infected hosts shift from a moving to a resting state, where the host-host contact rate decreases and an infected host can die from the disease. The clinical manifestation of many infectious diseases that induce lethargy prior to host death can justify this assumption, and public health initiatives such as encouraging sick people to stay home from workplace and social distancing policies can also result in two infective classes with distinct epidemiological characteristics and a behavioural shift from moving to resting state (Hart 1988; Halloran et al. 2008; Fenichel et al. 2011; Ghai et al. 2015). We focus on parasite-induced reduced movement rates, while there are other examples of parasites (e.g., the so-called furious strain of rabies virus) that can cause increased movement in infected hosts (Bacon 1985; Hemachudha et al. 2002; Susilawathi et al. 2012). Evolutionary bistability may not arise in the case where the parasite increases host movement ($c_r > c_m$) because the lethargy cost is no more present, and the higher the disease-induced mortality cost the lower the ESS α^* that is favoured by natural selection. Our model formulation is not specific to parasites that cause lethargy, but is applicable to any host-parasite system with two infective classes with distinct epidemiological characteristics such as Ebola or human immunodeficiency viruses, which have asymptomatic and symptomatic disease

stages.

In the result section, we show that for a range of parameter values a bistable evolutionary equilibrium is possible, and as such, transient coexistence of low and high virulence is possible. The coexistence of two strains with different levels of virulence is not uncommon in nature, and we provide two examples where host movement and/or medical interventions can explain the coexistence of a low and a high virulent strains, and rapid emergence of high virulence.

Example 1: The emergence of highly pathogenic avian influenza (HPAI) viruses

Avian influenza is caused by a type A influenza virus which infects domestic poultry (e.g., chickens and turkeys), free-living and wild bird populations (e.g., ducks, gulls and terns) (Stallknecht 2003; Causey and Edwards 2008; Yoon et al. 2014). Our model assumptions are valid for the avian influenza virus because it is mainly transmitted through direct contact with infected hosts or their secretions and infection does not confer long-lasting immunity (Stallknecht et al. 1990; Alexander 2000; 2007). The avian influenza virus induces symptoms such as lethargy, depression and anorexia prior to death in infected hosts, and the different virus strains are often classified as low pathogenic (LPAI) and highly pathogenic (HPAI) strains based on the severity of lethargy and the case fatality rate/ratio they cause in birds (Perkins and Swayne 2001; Mutinelli et al. 2003; Bertran et al. 2011; Belser et al. 2013; Wu et al. 2017). Infected chickens may have more contacts before lethargy because they are more active in the chicken pen or more likely to be transported between locations. As symptoms of lethargy appear, infected chickens may experience a decrease in their

contact rate because they are less active in the chicken pen or less likely to enter the global poultry market.

Our results suggest an alternative to the current best explanations for the emergence of HPAI in domestic poultry: 1) that HPAI strains result from infection spillover from strains endemic to wild bird populations; and 2) that HPAI can arise in poultry as a consequence of genetic mechanisms such as mutation, insertion, substitution and reassortment from an already circulating LPAI strain (Perdue et al. 1997; Alexander 2000; Banks et al. 2000; 2001; Sims et al. 2005; Taubenberger and Kash 2010; Nao et al. 2017; Qi et al. 2018). We show that when lethargic infected chickens can transmit the disease ($c_r > 0$), then a HPAI strain can emerge rapidly even when a LPAI strain reached a local ESS. In addition, our results suggest that a HPAI strain will not evolve if chickens make no contacts during lethargy ($c_r = 0$) or if a high fraction of lethargic chickens die (b is high). Therefore, our results suggest a dual benefit of quarantining or culling lethargic chickens, in that not only is infection transmission prevented, but the evolution of highly pathogenic strains becomes less likely.

Example 2: Human Immunodeficiency Viruses 1 and 2 (HIV-1/HIV-2)

HIV is a human lentivirus that is transmitted through sexual contact, from mother-to-child, through transfusion and needle sharing (Jaffar et al. 2004; Shaw and Hunter 2012; Patel et al. 2014). HIV disease is characterized by an acute, an asymptomatic stage followed by a symptomatic stage with acquired immunodeficiency syndrome (AIDS), and HIV can be transmitted during all stages with variable probability (Moylett and Shearer 2002; Pinkerton 2008; Levy 2009; Maartens et al. 2014). Two

HIV types are known: HIV-1 which may originate from a virus that infects chimpanzees in central Africa (*Pan troglodytes*), and HIV-2 which has been traced back to a virus found in Sooty mangabey (*Cercocebus atys*) in west Africa (Sharp and Hahn 2010; Ndung'u and Weiss 2012).

The HIV symptomatic stage can be viewed as the resting state in our model because individuals with AIDS symptoms may experience a decrease in sexual contacts during the symptomatic stage. To apply our model to HIV, we set the recovery rate equal zero ($\gamma = 0$) because HIV infection is invariably lethal. For HIV, virulence is often measured as the rate of progression to AIDS in the absence of treatment, whereas in treated individuals plasma viral load, set-point viral load and CD4 T-cells decline rate are frequently used as proxies for virulence (Cheng-Mayer et al. 1988; Carré et al. 1997; Pantazis et al. 2014; Roberts et al. 2015). To be consistent with our model formulation, we measure virulence as the fraction of asymptomatic hosts that progress to AIDS ($\psi(\alpha)/[d + \psi(\alpha)]$) corresponding to a within-host parasite net replication rate (α).

Our results suggest that when symptomatic AIDS individuals do not transmit HIV ($c_r = 0$) then a strain with low replication rate and slow progression to AIDS will evolve. As such, reduced needle sharing and protected sex can reduce HIV transmission and prevent the evolution of HIV strains with a high replication rate and rapid progression to AIDS. In contrast, when symptomatic AIDS individuals can transmit HIV then one strain with a long asymptomatic stage and a second strain with a short asymptomatic stage can coexist. The strain with the long asymptomatic stage has a lower within-host replication rate and induces slower progression to AIDS, whereas the strain with the short asymptomatic stage has a higher within-host replication rate

and induces rapid progression to AIDS. These relationships are consistent with a number of studies showing that plasma viral load is ≈ 30 times lower in HIV-2-infected individuals than HIV-1-infected individuals, and this lower plasma viral load explains the observed faster progression to AIDS in HIV-1-infected individuals (Berry et al. 1998; Popper et al. 1999; Andersson et al. 2000; MacNeil et al. 2007; Drylewicz et al. 2008; Tchounga et al. 2016). Our findings suggest that the rapid progression to AIDS in HIV-1-infected individuals is due to a higher within-host replication rate, and most of the secondary infections are generated from an HIV-1-infected individual during the symptomatic stage. Moreover, our findings suggest that medical interventions that improve the health of HIV-infected individuals (e.g., antiretroviral treatments (ART)) can select for strains with higher replication rate and faster progression to AIDS. This result is in accordance with a number of studies that have shown that when ART is initiated early after infection at high coverage then HIV strains (whether HIV-1 or HIV-2) with higher virulence are favoured (Herbeck et al. 2016; Porco et al. 2005; Herbeck et al. 2012; Pantazis et al. 2014).

Inspired by Markov models used to describe animal movement, we considered an epidemic model with two movement states and a parasite-induced shift from a moving to a resting state. Previous studies have illustrated that evolutionary bistability can arise due to host population heterogeneity (Gandon et al. 2003) and transmission mode heterogeneity (Boldin and Kisdi 2012). We find that a parasite-induced shift from a moving to a resting state can also result in evolutionary bistability, and for our model the bistability arises due to heterogeneity at the individual host-level rather than at the host population-level or beyond.

Literature cited

- J.L. Abbate, S. Kada, and S. Lion. Beyond mortality: Sterility as a neglected component of parasite virulence. *PLoS Pathog*, 11(12):e1005229, 2015.
- D.J. Alexander. A review of avian influenza in different bird species. *Vet. Microbiol.*, 74(1): 3–13, 2000.
- D.J. Alexander. An overview of the epidemiology of avian influenza. *Vaccine*, 25(30): 5637–5644, 2007.
- S. Alizon and Y. Michalakis. Adaptive virulence evolution: the good old fitness-based approach. *Trends Ecol. Evol.*, 30(5):248–254, 2015.
- S. Alizon and M. van Baalen. Emergence of a convex trade-off between transmission and virulence. *Am. Nat.*, 165(6):E155–E167, 2005.
- S. Alizon, A. Hurford, N. Mideo, and M. van Baalen. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.*, 22(2):245–259, 2009.
- R.M. Anderson and R.M. May. Coevolution of hosts and parasites. *Parasitology*, 85(02): 411–426, 1982.
- S. Andersson, H. Norrgren, Z. da Silva, A. Biague, S. Bamba, S. Kwok, C. Christopherson, G. Biberfeld, and J. Albert. Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. *Arch. Intern. Med.*, 160(21):3286–3293, 2000.
- J.B. André and S. Gandon. Vaccination, within-host dynamics, and virulence evolution. *Evolution*, 60(1):13–23, 2006.
- R. Antia, B.R. Levin, and R.M. May. Within-host population dynamics and the evolution and maintenance of microparasite virulence. *Am. Nat.*, 144(3):457–472, 1994.

- P.J. Bacon. *Population dynamics of rabies in wildlife*. Academic Press Inc.(London) Ltd., 1985.
- J. Banks, E.C. Speidel, J.W. McCauley, and D.J. Alexander. Phylogenetic analysis of h7 haemagglutinin subtype influenza a viruses. *Arch. Virol.*, 145(5):1047–1058, 2000.
- J. Banks, E.S. Speidel, E. Moore, L. Plowright, A. Piccirillo, I. Capua, P. Cordioli, A. Fioretti, and D.J. Alexander. Changes in the haemagglutinin and the neuraminidase genes prior to the emergence of highly pathogenic H7N1 avian influenza viruses in italy. *Arch. Virol.*, 146(5):963–973, 2001.
- A.S. Bell, J.C. de Roode, D. Sim, and A.F. Read. Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success. *Evolution*, 60(7):1358–1371, 2006.
- J.A. Belser, K.M. Gustin, M.B. Pearce, T.R. Maines, H. Zeng, C. Pappas, X. Sun, P.J. Carney, J.M. Villanueva, and J. Stevens. Pathogenesis and transmission of avian influenza A (H7N9) virus in ferrets and mice. *Nature*, 501(7468):556–559, 2013.
- N. Berry, K. Ariyoshi, S. Jaffar, S. Sabally, T. Corrah, R. Tedder, and H. Whittle. Low peripheral blood viral HIV-2 RNA in individuals with high CD4 percentage differentiates HIV-2 from HIV-1 infection. *J. Hum. Virol.*, 1(7):457–468, 1998.
- K. Bertran, E. Pérez-Ramírez, N. Busquets, R. Dolz, A. Ramis, A. Darji, F.X. Abad, R. Valle, A. Chaves, J. Vergara-Alert, M. Barral, U Höfle, and N. Majó. Pathogenesis and transmissibility of highly (H7N1) and low (H7N9) pathogenic avian influenza virus infection in red-legged partridge (*alectoris rufa*). *Veterinary Research*, 42(1):24, 2011.
- A. Best, A. White, and M. Boots. The evolution of host defence when parasites impact reproduction. *Evol. Ecol. Res.*, 18(4):393–409, 2017.
- F. Blanquart, M.K. Grabowski, J. Herbeck, F. Nalugoda, D. Serwadda, M.A. Eller, M.L. Robb, R. Gray, G. Kigozi, O. Laeyendecker, K.A. Lythgoe, G. Nakigozi, T.C. Quinn, S.J.

- Reynolds, M.J. Wawer, and C. Fraser. A transmission-virulence evolutionary trade-off explains attenuation of HIV-1 in Uganda. *eLife*, 5:e20492, 2016.
- B. Boldin and É. Kisdi. On the evolutionary dynamics of pathogens with direct and environmental transmission. *Evolution*, 66(8):2514–2527, 2012.
- M.H. Bonds. Host life-history strategy explains pathogen-induced sterility. *Am. Nat.*, 168(3):281–293, 2006.
- M. Boots, P.J. Hudson, and A. Sasaki. Large shifts in pathogen virulence relate to host population structure. *Science*, 303(5659):842–844, 2004.
- Fred Brauer. Compartmental models in epidemiology. In *Mathematical epidemiology*, chapter 2, pages 19–79. Springer, 2008.
- J.L. Brunner, K. Richards, and J.P. Collins. Dose and host characteristics influence virulence of ranavirus infections. *Oecologia*, 144(3):399–406, 2005.
- J.J. Bull. Perspective: virulence. *Evolution*, 48(5):1423–1437, 1994.
- N. Carré, M. Prins, L. Meyer, R.P. Brettler, J.R. Robertson, H. McArdle, D.J. Goldberg, R. Zangerle, R.A. Coutinho, and A. van den Hoek. Has the rate of progression to AIDS changed in recent years? *AIDS*, 11(13):1611–1618, 1997.
- Douglas Causey and Scott V Edwards. Ecology of avian influenza virus in birds. *J. Infect. Dis.*, 197(Supplement_1):S29–S33, 2008.
- C. Cheng-Mayer, D. Seto, M. Tateno, and J.A. Levy. Biologic features of HIV-1 that correlate with virulence in the host. *Science*, 240(4848):80, 1988.
- C.E. Cressler, D.V. McLEOD, C. Rozins, J. van Den Hoogen, and T. Day. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology*, 143(7):915–930, 2016.
- T. Day. Parasite transmission modes and the evolution of virulence. *Evolution*, 55(12):2389–2400, 2001.

- J.C. de Roode and S. Altizer. Host-parasite genetic interactions and virulence-transmission relationships in natural populations of monarch butterflies. *Evolution*, 64(2):502–514, 2010.
- J.C. de Roode, L.R. Gold, and S. Altizer. Virulence determinants in a natural butterfly-parasite system. *Parasitology*, 134(5):657–668, 2006.
- J.C. de Roode, A.J. Yates, and S. Altizer. Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proc. Natl. Acad. Sci. U. S. A.*, 105(21):7489–7494, 2008.
- J.C. de Roode, J. Chi, R.M. Rarick, and S. Altizer. Strength in numbers: high parasite burdens increase transmission of a protozoan parasite of monarch butterflies (*Danaus plexippus*). *Oecologia*, 161(1):67–75, 2009.
- U. Dieckmann. Adaptive dynamics of pathogen-host interactions. In *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management*, chapter 4, pages 39–59. Cambridge University Press, 2002.
- O. Diekmann. A beginner’s guide to adaptive dynamics. *Banach Center Publications*, 63: 47–86, 2004.
- O. Diekmann, H. Heesterbeek, and T. Britton. The epidemic in a closed population. In *Mathematical tools for understanding infectious diseases dynamics*, chapter 1, pages 3–36. Princeton University Press, 2012.
- E.R. Dougherty, D.P. Seidel, C.J. Carlson, O. Spiegel, and W.M. Getz. Going through the motions: incorporating movement analyses into disease research. *Ecol. Lett.*, 21:588–604, 2017.
- J. Doumayrou, A. Avellan, R. Froissart, and Y. Michalakis. An experimental test of the transmission-virulence trade-off hypothesis in a plant virus. *Evolution*, 67(2):477–486, 2013.

- J. Drylewicz, S. Matheron, E. Lazaro, F. Damond, F. Bonnet, F. Simon, F. Dabis, F. Brun-Vezinet, G. Chêne, and R. Thiébaud. Comparison of viro-immunological marker changes between HIV-1 and HIV-2-infected patients in france. *AIDS*, 22(4):457, 2008.
- K.T. Eames, N.L. Tilston, P.J. White, E. Adams, and W.J. Edmunds. The impact of illness and the impact of school closure on social contact patterns. *Health Technol Assess*, 14(34):267–312, 2010.
- D. Ebert and E.A. Herre. The evolution of parasitic diseases. *Parasitol. Today*, 12(3):96–101, 1996.
- H. Edelhoff, J. Signer, and N. Balkenhol. Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement Ecology*, 4(1):21, 2016.
- Ilan Eshel. Evolutionary and continuous stability. *J. Theor. Biol.*, 103(1):99–111, 1983.
- T. Evans, R.G. Bowers, and M. Mortimer. Adaptive dynamics of temperate phages. *Evol. Ecol. Res.*, 12(4):413–434, 2010.
- P.W. Ewald. Host-parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Syst.*, 14(1):465–485, 1983.
- P.W. Ewald. Vectors, vertical transmission, and the evolution of virulence. In *Evolution of infectious disease*, chapter 3, pages 75–114. Oxford University Press, 1994.
- E.P. Fenichel, C. Castillo-Chavez, M.G. Ccedia, G. Chowell, P.A.G. Parra, G.J. Hickling, G. Holloway, R. Horan, B. Morin, C. Perrings, M. Springborn, L. Velazquez, and C. Villalobos. Adaptive human behavior in epidemiological models. *Proc. Natl. Acad. Sci. U. S. A*, 108(15):6306–6311, 2011.
- P.B. Finnerty, R. Shine, and G.P. Brown. The costs of parasite infection: Effects of removing lungworms on performance, growth and survival of free-ranging cane toads. *Funct. Ecol.*, 32(2):402–415, 2018.

- A.E. Fleming-Davies, V. Dukic, V. Andreasen, and G. Dwyer. Effects of host heterogeneity on pathogen diversity and evolution. *Ecol. Lett.*, 18(11):1252–1261, 2015.
- A.M. Fofana and A. Hurford. Mechanistic movement models to understand epidemic spread. *Phil. Trans. R. Soc. B*, 372(1719):20160086, 2017.
- S.A. Frank. Models of parasite virulence. *Q. Rev. Biol.*, 71(1):37–78, 1996.
- C. Fraser, T.D. Hollingsworth, R. Chapman, F. de Wolf, and W.P. Hanage. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. U. S. A.*, 104(44):17441–17446, 2007.
- C. Fraser, K. Lythgoe, G.E. Leventhal, G. Shirreff, T.D. Hollingsworth, S. Alizon, and S. Bonhoeffer. Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. *Science*, 343(6177):1243727, 2014.
- S. Gandon, M.J. Mackinnon, S. Nee, and A.F. Read. Imperfect vaccines and the evolution of pathogen virulence. *Nature*, 414(6865):751–756, 2001.
- S. Gandon, M. Mackinnon, S. Nee, and A. Read. Imperfect vaccination: some epidemiological and evolutionary consequences. *Proc. R. Soc. Lond. B Biol. Sci.*, 270(1520):1129–1136, 2003.
- S.A. Geritz, G. Mesze, and J.A.J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.*, 12(1):35–57, 1998.
- R.R. Ghai, V. Fugere, C.A. Chapman, T.L. Goldberg, and T.J. Davies. Sickness behaviour associated with non-lethal infections in wild primates. *Proc. R. Soc. Lond. B Biol. Sci.*, 282(1814):20151436, 2015.
- M.A. Gilchrist and A. Sasaki. Modeling host-parasite coevolution: a nested approach based on mechanistic models. *J. Theor. Biol.*, 218(3):289–308, 2002.
- E. Gurarie, R.D. Andrews, and K.L. Laidre. A novel method for identifying behavioural changes in animal movement data. *Ecol. Lett.*, 12(5):395–408, 2009.

- M.E. Halloran, N.M. Ferguson, S. Eubank, I.M. Longini, D.A. Cummings, B. Lewis, S. Xu, C. Fraser, A. Vullikanti, T.C. Germann, D. Wagener, R. Beckman, K. Kadau, C. Barrett, C.A. Macken, D.S. Burke, and P. Cooley. Modeling targeted layered containment of an influenza pandemic in the United States. *Proc. Natl. Acad. Sci. U. S. A.*, 105(12): 4639–4644, 2008.
- B.L. Hart. Biological basis of the behavior of sick animals. *Neurosci. Biobehav. Rev.*, 12(2): 123–137, 1988.
- T. Hemachudha, J. Laothamatas, and C.E. Rupprecht. Human rabies: a disease of complex neuropathogenetic mechanisms and diagnostic challenges. *Lancet Neurol.*, 1(2):101–109, 2002.
- J.T. Herbeck, V. Müller, B.S. Maust, B. Ledergerber, C. Torti, S. Di Giambenedetto, L. Gras, H.F. Günthard, L.P. Jacobson, J.I. Mullins, and G.S. Gottlieb. Is the virulence of HIV changing? a meta-analysis of trends in prognostic markers of HIV disease progression and transmission. *AIDS*, 26(2):193, 2012.
- J.T. Herbeck, J.E. Mittler, G.S. Gottlieb, S.M. Goodreau, J.T. Murphy, A. Cori, M. Pickles, and C. Fraser. Evolution of HIV virulence in response to widespread scale up of antiretroviral therapy: a modeling study. *Virus Evolution*, 2(2):vew028, 2016.
- P.R. Holmstad, K.H. Jensen, and A. Skorpung. Vector-borne parasites decrease host mobility: a field test of freeze or flee behaviour of willow ptarmigan. *Int. J. Parasitol.*, 36(7): 735–740, 2006.
- A. Hurford, D. Cownden, and T. Day. Next-generation tools for evolutionary invasion analyses. *J. Royal Soc. Interface*, 7(45):561–571, 2010.
- S. Jaffar, A.D. Grant, J. Whitworth, P.G. Smith, and H. Whittle. The natural history of HIV-1 and HIV-2 infections in adults in Africa: a literature review. *Bull. W.H.O.*, 82(6): 462–469, 2004.

- J.A. Levy. HIV pathogenesis: 25 years of progress and persistent challenges. *AIDS*, 23(2): 147–160, 2009.
- M. Lipsitch and E.R. Moxon. Virulence and transmissibility of pathogens: what is the relationship? *Trends Microbiol.*, 5(1):31–37, 1997.
- C.M. Lively. The ecology of virulence. *Ecol. Lett.*, 9(10):1089–1095, 2006.
- J.O. Lloyd-Smith, W.M. Getz, and H.V. Westerhoff. Frequency-dependent incidence in models of sexually transmitted diseases: portrayal of pair-based transmission and effects of illness on contact behaviour. *Proc. R. Soc. Lond. B Biol. Sci.*, 271(1539):625–634, 2004.
- G. Maartens, C. Celum, and S.R. Lewin. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *The Lancet*, 384(9939):258–271, 2014.
- A. MacNeil, A.D. Sarr, J.L. Sankalé, S.T. Meloni, S. Mboup, and P. Kanki. Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection. *J. Virol.*, 81(10):5325–5330, 2007.
- R.M. May. Population biology of microparasitic infections. In *Mathematical ecology: An Introduction*, pages 405–442. Springer, 1986.
- A.E. McKellar, R. Langrock, J.R. Walters, and D.C. Kesler. Using mixed hidden markov models to examine behavioral states in a cooperatively breeding bird. *Behav. Ecol.*, 26(1):148–157, 2014.
- B. Moorter, D.R. Visscher, C.L. Jerde, J.L. Frair, and E.H. Merrill. Identifying movement states from location data using cluster analysis. *J. Wildl. Manag.*, 74(3):588–594, 2010.
- J.M. Morales, D.T. Haydon, J. Frair, K.E. Holsinger, and J.M. Fryxell. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, 85(9):2436–2445, 2004.
- E.H. Moylett and W.T. Shearer. HIV: clinical manifestations. *J. Allergy Clin. Immunol.*,

- 110(1):3–16, 2002.
- F. Mutinelli, I. Capua, C. Terregino, and G. Cattoli. Clinical, gross, and microscopic findings in different avian species naturally infected during the H7N1 low-and high-pathogenicity avian influenza epidemics in Italy during 1999 and 2000. *Avian Dis.*, 47(s3):844–848, 2003.
- N. Nao, J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, et al. Genetic predisposition to acquire a poly-basic cleavage site for highly pathogenic avian influenza virus hemagglutinin. *MBio*, 8(1):e02298–16, 2017.
- T. Ndung’u and R.A. Weiss. On HIV diversity. *Aids*, 26(10):1255–1260, 2012.
- K.J. O’Keefe and J. Antonovics. Playing by different rules: the evolution of virulence in sterilizing pathogens. *Am. Nat.*, 159(6):597–605, 2002.
- S.P. Otto and T. Day. Evolutionary invasion analysis. In *A biologist’s guide to mathematical modeling in ecology and evolution*, chapter 12, pages 454–566. Princeton University Press, 2007.
- N. Pantazis, K. Porter, D. Costagliola, A. De Luca, J. Ghosn, M. Guiguet, A.M. Johnson, A.D. Kelleher, C. Morrison, R. Thiebaut, L. Wittkop, and G. Touloumi. Temporal trends in prognostic markers of HIV-1 virulence and transmissibility: an observational cohort study. *Lancet HIV*, 1(3):e119–e126, 2014.
- P. Patel, C.B. Borkowf, J.T. Brooks, A. Lasry, A. Lansky, and J. Mermin. Estimating per-act HIV transmission risk: a systematic review. *Aids*, 28(10):1509–1519, 2014.
- T.A. Patterson, L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. State-space models of individual animal movement. *Trends Ecol. Evol.*, 23(2):87–94, 2008.
- R.E.L. Paul, T. Lafond, C.D.M. Müller-Graf, S. Nithiuthai, P.T. Brey, and J.C. Koella. Experimental evaluation of the relationship between lethal or non-lethal virulence and

- transmission success in malaria parasite infections. *BMC Evol. Biol.*, 4(1):30, 2004.
- M.L. Perdue, M. Garci'a, D. Senne, and M. Fraire. Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. *Virus Res.*, 49(2):173–186, 1997.
- L.E.L. Perkins and D.E. Swayne. Pathobiology of a/chicken/Hong Kong/220/97 (H5N1) avian influenza virus in seven gallinaceous species. *Veterinary Pathology*, 38(2):149–164, 2001.
- T.A. Perkins, V.A. Paz-Soldan, S.T. Stoddard, A.C. Morrison, B.M. Forshey, K.C. Long, E.S. Halsey, T.J. Kochel, J.P. Elder, U. Kitron, T.W. Scott, and G.M. Vazquez-Prokopec. Calling in sick: impacts of fever on intra-urban human mobility. *Proc. R. Soc. Lond. B Biol. Sci.*, 283(1834):20160390, 2016.
- N. Perra, D. Balcan, B. Gonçalves, and A. Vespignani. Towards a characterization of behavior-disease models. *PloS one*, 6(8):e23084, 2011.
- S.D. Pinkerton. Probability of HIV transmission during acute infection in Rakai, Uganda. *AIDS Behav*, 12(5):677–684, 2008.
- S.J. Popper, A.D. Sarr, K.U. Travers, A. Guèye-Ndiaye, S. Mboup, M.E. Essex, and P.J. Kanki. Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *J. Infect. Dis.*, 180(4):1116–1121, 1999.
- T.C. Porco, J.O. Lloyd-Smith, K.L. Gross, and A.P. Galvani. The effect of treatment on pathogen virulence. *J. Theor. Biol.*, 233(1):91–102, 2005.
- W. Qi, W. Jia, D. Liu, J. Li, Y. Bi, S. Xie, B. Li, T. Hu, Y. Du, L. Xing, et al. Emergence and adaptation of a novel highly pathogenic H7N9 influenza virus in birds and humans from a 2013 human-infecting low-pathogenic ancestor. *J. Virol.*, 92(2):e00921–17, 2018.
- A.F. Read. The evolution of virulence. *Trends Microbiol.*, 2(3):73–76, 1994.
- A.F. Read, S.J. Baigent, C. Powers, L.B. Kgosana, L. Blackwell, L.P. Smith, D.A. Kennedy,

- S.W. Walkden-Brown, and V.K. Nair. Imperfect vaccination can enhance the transmission of highly virulent pathogens. *PLoS Biol.*, 13(7):e1002198, 2015.
- P.D. Reuman, S. Keely, and G.M. Schiff. Assessment of signs of influenza illness in the ferret model. *J. Virol. Methods*, 24(1):27–34, 1989.
- H.E. Roberts, P.J.R. Goulder, and A.R. McLean. The impact of antiretroviral therapy on population-level virulence evolution of HIV-1. *J. Royal Soc. Interface*, 12(113):20150888, 2015.
- S. Schjørring and J.C. Koella. Sub-lethal effects of pathogens can lead to the evolution of lower virulence in multiple infections. *Proc. R. Soc. Lond. B Biol. Sci.*, 270(1511):189–193, 2003.
- P.M. Sharp and B.H. Hahn. The evolution of HIV-1 and the origin of AIDS. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 365(1552):2487–2494, 2010.
- G.M. Shaw and E. Hunter. HIV transmission. *Cold Spring Harbor perspectives in medicine*, 2(11):a006965, 2012.
- L.D. Sims, J. Domenech, C. Benigno, S. Kahn, A. Kamata, J. Lubroth, V. Martin, and P. Roeder. Origin and evolution of highly pathogenic H5N1 avian influenza in asia. *Veterinary Record*, 157(6):159, 2005.
- D.E. Stallknecht. Ecology and epidemiology of avian influenza viruses in wild bird populations: waterfowl, shorebirds, pelicans, cormorants, etc. *Avian Dis.*, 47:61–69, 2003.
- D.E. Stallknecht, S.M. Shane, M.T. Kearney, and P.J. Zwank. Persistence of avian influenza viruses in water. *Avian Dis.*, 34:406–411, 1990.
- N.M. Susilawathi, A.E. Darwinata, I.B. Dwija, N.S. Budayanti, G.A. Wirasandhi, K. Subrata, N.K. Susilarini, R.A. Sudewi, Frank S. Wignall, and G.N. Mahardika. Epidemiological and clinical features of human rabies cases in bali 2008-2010. *BMC Infect. Dis.*, 12(1):81, 2012.

- J.K. Taubenberger and J.C. Kash. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe*, 7(6):440–451, 2010.
- B. Tchounga, D.K. Ekouevi, and F. Balestre, E.and Dabis. Mortality and survival patterns of people living with HIV-2. *Current Opinion in HIV and AIDS*, 11(5):537–544, 2016.
- M. Teimouri, U.G. Indahl, H. Sickel, and H. Tveite. Deriving animal movement behaviors using movement parameters extracted from location data. *ISPRS Int. J. Geo-Inf.*, 7(2):78, 2018.
- R. Timms, N. Colegrave, B.H.K. Chan, and A.F. Read. The effect of parasite dose on disease severity in the rodent malaria *Plasmodium chabaudi*. *Parasitology*, 123(1):1–11, 2001.
- M. van Baalen. Coevolution of recovery ability and virulence. *Proc. R. Soc. Lond. B Biol. Sci.*, 265(1393):317–325, 1998.
- P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.*, 180(1):29–48, 2002.
- K. Van Kerckhove, N. Hens, W.J. Edmunds, and K.T. Eames. The impact of illness on social networks: implications for transmission and control of influenza. *Am. J. Epidemiol.*, 178(11):1655–1662, 2013.
- F. Verelst, L. Willem, and P. Beutels. Behavioural change models for infectious disease transmission: a systematic review (2010–2015). *J. Royal Soc. Interface*, 13(125):20160820, 2016.
- B.A. Walther and P.W. Ewald. Pathogen survival in the external environment and the evolution of virulence. *Biol. Rev.*, 79(4):849–869, 2004.
- Z. Wang, M.A. Andrews, Z.X. Wu, L. Wang, and C.T. Bauch. Coupled disease-behavior dynamics on complex networks: A review. *Phys. Life Rev.*, 15:1–29, 2015.

- P.D. Williams, A.P. Dobson, K.V. Dhondt, D.M. Hawley, and A.A. Dhondt. Evidence of trade-offs shaping virulence evolution in an emerging wildlife pathogen. *J. Evol. Biol.*, 27(6):1271–1278, 2014.
- Z.Q. Wu, Y. Zhang, N. Zhao, Z. Yu, H. Pan, T.C. Chan, Z.R. Zhang, and S.L. Liu. Comparative epidemiology of human fatal infections with novel, high (H5N6 and H5N1) and low (H7N9 and H9N2) pathogenicity avian influenza A viruses. *Int. J. Environ. Res. Public Health*, 14(3):263, 2017.
- S.W. Yoon, R.J. Webby, and R.G. Webster. Evolution and ecology of influenza a viruses. In *Influenza Pathogenesis and Control-Volume I*, pages 359–375. Springer, 2014.
- L.A. Zitzow, T. Rowe, T. Morken, W.J. Shieh, S. Zaki, and J.M. Katz. Pathogenesis of avian influenza A (H5N1) viruses in ferrets. *J. Virol.*, 76(9):4420–4429, 2002.

4. Chapter three: Mechanistic movement models to understand epidemic spread

This chapter has been published in the journal *Philosophical Transactions of the Royal Society*

Mechanistic movement models to understand epidemic spread

Abdou M. Fofana, Amy Hurford

Abstract

An overlooked aspect of disease ecology is considering how and why animals come into contact with one and other resulting in disease transmission. Mathematical models of disease spread frequently assume mass-action transmission, justified by stating that susceptible and infectious hosts mix readily, and foregoing any detailed description of host movement. Numerous recent studies have recorded, analyzed and modelled animal movement. These movement models describe how animals move with respect to resources, conspecifics, and previous movement directions and have been used to understand the conditions for the occurrence and the spread of infectious diseases when hosts perform a type of movement. Here, we summarize the effect of the different types of movement on the threshold conditions for disease spread. We identify gaps in the literature and suggest several promising directions for future research. The mechanistic inclusion of movement in epidemic models may be beneficial for the following two reasons. Firstly, the estimation of the transmission coefficient in an epidemic model is possible because animal movement data can be used to estimate the rate of contacts between conspecifics. Secondly, unsuccessful potential transmission events, where a susceptible host contacts an infectious host but does not become infected can be quantified. Following an outbreak, this enables disease ecologists to identify ‘near misses’ and to explore possible alternative epidemic outcomes given shifts in ecological or immunological parameters.

Keywords: Animal movement, random walks, levy walks, contact process, epidemic threshold, disease spread.

4.1. Introduction

Animal movement is essential for many ecological processes such as foraging, escaping from predators and finding a mate or new habitats. Movement determines the spatiotemporal distribution of populations, plays a major role in encounters between individuals (Preston et al. 2015; Ims 1995; Turchin 1991; Swingland and Greenwood 1983; Cronin 2003; Barry et al. 2016) and in turn affects the magnitude of ecological processes and the dynamics of interacting populations (Morales et al. 2010; Turchin 1998). In disease ecology, the transmission of many infectious diseases requires ‘contact’ between a susceptible and an infectious host. This contact process is traditionally modelled in a phenomenological fashion with few details on how and why individuals come into contact with one another (Kermack and McKendrick 1927; Hethcote 2000; Diekmann et al. 2012; Allen et al. 2008; Anderson et al. 1992). These traditional approaches assume homogeneous mixing of susceptible and infectious hosts and the spatial proximity between individuals is not explicitly acknowledged in disease transmission process. Although these traditional models have significantly contributed to understanding the conditions for epidemic occurrence (Diekmann et al. 1995), their spatial extension is necessary for capturing both the spatial and the temporal dynamics of infectious diseases (Cliff 1996; Durrett 1995).

During the past five decades, recording individual animal movement has been facilitated by Global Positioning Systems (GPS) and telemetry technology (Cagnacci et al. 2010; Hebblewhite and Haydon 2010). The *a posteriori* description obtained from successive positions data provides information about animal movement patterns, but contains limited information on why animals move as they do (Turchin 1998). During

the same period, many mathematical models have been developed with details on how individual animals move towards resources (for example, food, habitat and mates). In these models, individual movement follows specific rules describing movement direction, turning frequency and velocity, reflecting the resource distribution and how informed the mover is about resource locations (Berg 1983; Lewis et al. 2013; Okubo and Levin 2001). This detailed individual-based behaviour can be translated into a partial differential equation (PDE), describing the spatiotemporal distribution of the population (Codling et al. 2008, see Table 4.1 for the definition of the abbreviations and symbols used in this paper). Some models conserve the individual description (Lagrangian approach) whereas others focus on population-level consequences of these movement rules (Eulerian approach). These two approaches have been reviewed in detail in Smouse et al. (2010).

Different types of animal movement are uncorrelated, correlated, biased random walks (URW, CRW, BRW) and Levy walks (LWs). These models have been applied to ecological problems such as predator-prey dynamics (Merrill et al. 2010; Tiutiunov et al. 2013; McKenzie et al. 2012), biological invasions (Shigesada et al. 2015; Shaw et al. 2006) and have long attracted the interest of disease ecologists. The growing interest for these models in disease ecology is due to the following reasons. Firstly, in contrast to traditional epidemic models (see Kermack and McKendrick 1927; Hethcote 2000; Diekmann et al. 2012; Allen et al. 2008), the spatiotemporal distribution of the host population and the pattern of contacts between individuals emerges from individual movement rules rather than being simply homogeneous. For this reason, epidemic models with explicit individual movement are termed mechanistic, in contrast to traditional epidemic models which are phenomenological. Secondly, epidemic thresholds,

Table 4.1: List of abbreviations and symbols used in the main text

Abbreviations/ Symbols	Definitions
Movement models	
BCRW	Biased correlated random walk.
BRW	Biased random walk.
CRW	Correlated random walk.
LW	Levy walk.
URW	Uncorrected random walk.
Epidemic models	
IBM	Individual-based model.
PDE	Partial differential equation.
SI	Compartmental Susceptible-Infected model.
SIR	Compartmental Susceptible-Infected-Removed model.
SIS	Compartmental Susceptible-Infected-Susceptible model.
Movement parameters	
C_I	Advection rate of infectious hosts.
C_S	Advection rate of susceptible hosts.
D_I	The diffusion coefficient of infectious hosts.
D_S	The diffusion coefficient of susceptible hosts.
l	The degree of host movement between habitats.
Epidemic parameters	
β	Disease transmission coefficient.
β_i	The transmission coefficient for a defined habitat ($i = 1, 2, 3..$).
β_h^*	Spatially homogeneous vector-host transmission rate.
β_w^*	Spatially homogeneous host-vector transmission rate.
c	The speed of disease spread
c_0	The critical speed for disease propagation.
d	Natural host mortality rate.
d_w	Natural vector mortality rate for the host-vector model in the Table 4.2.
d_h	Natural host mortality rate for the host-vector model in the Table 4.2.
γ	Host recovery rate.
γ_i	Host recovery rate for a defined habitat ($i = 1, 2, 3..$).
I_0	The initial density/number of infectious hosts.
K_t	The critical carrying capacity of host population for epidemic occurrence.
λ	The probability of infection given a contact.
μ	Disease-induced host mortality rate.
μ_w	Disease-induced vector mortality rate for the host-vector model in the Table 4.2.
μ_h	Disease-induced host mortality rate for the host-vector model in the Table 4.2.
N	Total host population size.
p_j	Probability that a host performs a long ‘distance jump’ into a random location.
R_0	The expected number of secondary cases generated by a primary case in a completely susceptible host population (an epidemic occurs if $R_0 > 1$).
S_0	The initial density/number of susceptible hosts or the critical host population density for epidemic occurrence.
W_1^*	Density of susceptible vector population.
w_w	The incubation period of the parasite within vector individuals.
w_h	The incubation period of the parasite within host individuals.

in particular the basic reproduction number (R_0), which is the expected number of secondary cases generated by a primary case in a completely susceptible population

(Allen et al. 2008), depend on the density of the host population and the transmission rate including the host-host contact rate. Therefore, spatiotemporal distributions and contact patterns resulting from different types of animal movement might affect the spread of infectious diseases. In this paper, we review theoretical studies that account for mechanistic animal movement in disease ecology. Our objective is to summarize the effect of different types of animal movement on threshold conditions for disease spread.

4.2. *The mass-action law*

In disease ecology, any parasite transmission opportunity is considered a contact. Examples of contacts are a sexual contact between two partners for sexually transmitted diseases, a vector biting a host for vector-borne diseases and touching and exposure to aerosols emitted by another individual for directly transmitted diseases. The relationship between animal movement rules (how and why animals move) and the contact process is poorly understood (but see Rhodes and Anderson 2008). The formulation of the contact process for traditional models and directly transmitted diseases is generally based on two main assumptions. First, at every point in time, it is assumed that each individual has the same chance of making a contact with any other individual in the population. This is the so-called *homogeneous mixing* assumption which is a simplification aiming to keep the analysis of the mathematical equations tractable. Second, at any point in time, a fraction of these contacts are assumed to lead to the transmission of the disease. This is the so-called mass-action law which means that the total number of infectious contacts per unit of time increases with the densities of susceptible and infectious individuals (McCallum et al. 2001; Begon et al. 2002). In the next sections we ask, when a type of movement is explicitly considered

in the epidemic model, does the mass-action law hold?

4.3. Epidemics when host movement is random

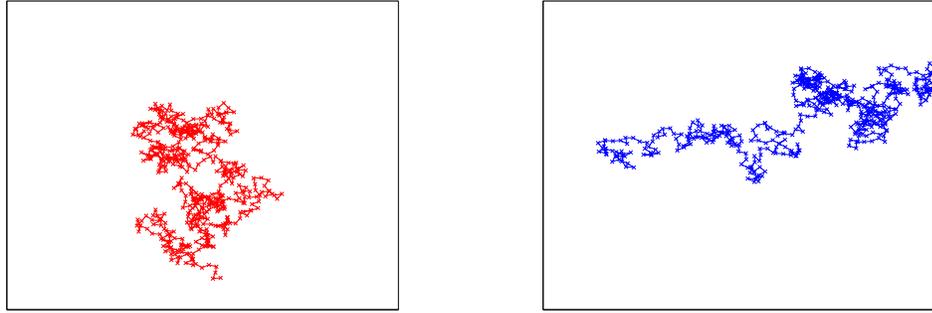
The *uncorrelated random walk* (URW) is considered as the starting point for animal movement models in ecology. It describes the non-persistent animal movement in a homogeneous environment (for example, a homogeneous food distribution). When performing the URW, an individual executes independent successive steps at a constant speed and turns in each direction with the same probability because it has no *a priori* information about the location of food. A sample movement path for an animal performing an URW is shown in Figure 4.1a. Over large spatial scales, a population of non-interacting individuals exhibiting such movement rules diffuses with time (Skellam 1951; Spitzer 1976). Using mark-recapture data from field studies, it has been shown that the foraging movement in some insect species reflect the URW when the food is homogeneously distributed (Kareiva 1983; 1982; Marchant et al. 2015). Accounting for the URW of host individuals in a Kermack-McKendrick epidemic model gives a system of partial differential equations (PDEs) of the following form:

$$\begin{aligned} \frac{\partial S}{\partial t} &= D_S \frac{\partial^2 S}{\partial x^2} - \beta SI \\ \frac{\partial I}{\partial t} &= D_I \frac{\partial^2 I}{\partial x^2} + \beta SI - \mu I, \end{aligned} \quad x \in \mathbb{R} \quad \text{and} \quad t \in \mathbb{R}^+, \quad (4.1)$$

where $I(x, t)$ and $S(x, t)$ are the densities of infectious and susceptible individuals respectively at location x at time t . The diffusion terms ($D_S \partial^2 S / \partial x^2$ and $D_I \partial^2 I / \partial x^2$) represent the URW of susceptible and infectious individuals and the remaining terms (called the reaction terms) are infection and disease-induced host mortality at each location. The parameters D_I and D_S represent the diffusion coefficients of infectious

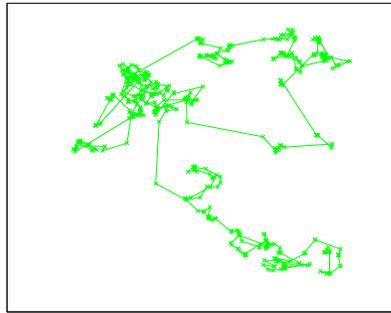
and susceptible individuals and β and μ represent the transmission coefficient and disease-induced host mortality rate respectively. The diffusion coefficient is a measure of how far a moving individual travels on average from its initial location during a fixed period of time (for details on how the diffusion coefficients can be estimated, see Murray et al. 1986). Assuming that the initial density of susceptible individuals is the same everywhere, $S(x, 0) = S_0$, and the disease is locally introduced, $I(x, 0) = I_0(x)$, Hosono and Ilyas (1995) showed that if $S_0 < \mu/\beta$, then the disease dies out. In contrast, if $S_0 > \mu/\beta$, then the disease spreads outward from the point of introduction as a travelling wave with a speed of propagation, c , satisfying $c \geq c_0 = 2\sqrt{\beta S_0 D_I (1 - \mu/\beta S_0)}$.

First, it can be noticed that the epidemic threshold given by the system (4.1) is independent of the movement parameters and is exactly the basic reproduction number given by traditional epidemic models (see for example Kermack and McKendrick 1927; Diekmann et al. 1995). This result suggests that the occurrence of an epidemic might be independent of the URW of host individuals. Second, the pattern of spatial spread of the disease exhibited by the system (4.1) is not captured by traditional models and it can be noticed that the critical speed for disease propagation, c_0 , increases with D_I . This suggests that the spatial spread of a disease, when it occurs, depends on the URWs of infectious individuals. Similar results were found for the spatial spread of rabies in the red fox (*Vulpes vulpes*), where only infectious individuals are assumed to be moving (Källén 1984; Källén et al. 1985; Murray et al. 1986). In recent studies, the system (4.1) has been modified by considering an incubation period (Li and Li 2015; Bai and Zhang 2015) and non-local (Wang and Wu 2010), non-linear (Li and Li 2015; Bai and Zhang 2015) and frequency-dependent infections (Wang et al. 2012).



(a) URWs

(b) BRWs



(c) LWs

Figure 4.1: Examples illustrating simulated URWs, BRWs and LWs of a single individual in 2 spatial dimensions. For the URWs, the individual chooses its movement direction and angle from a uniform distribution and moves a constant step at each time ($p_r = p_l$ and $1 - p_r - p_l$ is the probability of waiting). For the BRWs, we set the probability distribution of the movement directions such that the individual is more likely to move left ($p_l > p_r$ and $1 - p_r - p_l$ is the probability of waiting). For the LWs, the individual chooses its movement direction and the angle from a uniform distribution but the step length is chosen from a heavy-tailed distribution (Pareto distribution with infinite variance). For each simulated type of movement, the mean step length is equal.

The inclusion of the above factors did not change the main conclusion, which is that the threshold condition for the occurrence of an epidemic is independent of the URW of host individuals. Given that traditional epidemic models assume the mass-action law and that the basic reproduction number is the same for models assuming a URW, at least from the perspective of the basic reproduction number, a URW may be con-

sistent with the mass-action assumption. We simulated a URW and compared the rate of new infections with the rate assumed by the mass-action law. The simulation results suggested that the rate of new infections for a population performing a URW is consistent with the mass-action law model formulation (Figure 4.2). All the codes used for the simulation is available as electronic supplementary materials S4.1-S4.3 and is publicly available at Figshare doi:10.6084/m9.figshare.11389623.v2.

For all of the studies where the basic reproduction number was found to be independent of the diffusion coefficient, the epidemiological parameters (especially β , μ and γ , where γ is the recovery rate) as well as the movement parameters (D_I and D_S) were assumed to be spatially homogeneous. Wang and Zhao (2011) proposed a reaction-diffusion model for a dengue fever epidemic with spatially-dependent transmission rates (modelled using a periodic function) and non-local and delayed transmission (i.e, infections at a given location at a given time result from contacts at different locations at an earlier time). The results of this study showed that the occurrence of a dengue fever epidemic is independent of the URW of the host (human) and the vector (mosquito) only when the transmission rates are spatially homogeneous. In the case where the transmission rates are spatially heterogeneous the derivation of an analytical expression for R_0 is more complex, but using numerical methods Wang and Zhao (2011) showed that R_0 decreases with increasing values of the diffusion coefficients for host and vector. This result suggests that the less distance the vector and the host travel on average (when exhibiting URWs) the higher the risk of occurrence of a dengue fever epidemic. Moreover, other studies investigated the epidemiological dynamics of Susceptible-Infected-Susceptible (SIS) reaction-diffusion models with spatially heterogeneous transmission and recovery rates (Allen et al. 2007; Peng 2009;

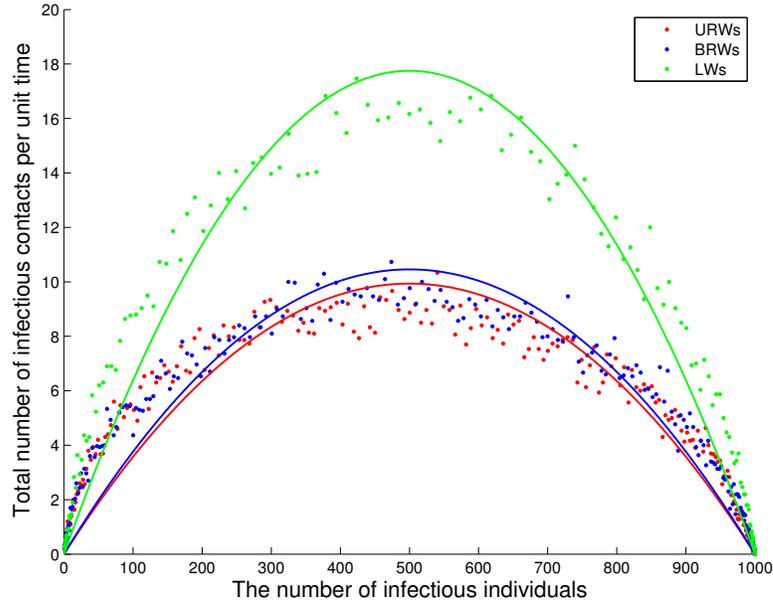


Figure 4.2: The number of infectious contacts per unit of time as a function of the number of infectious individuals when hosts perform URWs (red), BRWs (blue) and LWs (green curve). The circles represent the simulated epidemic data and the curves represent the fit of the mass action law to the simulation data. For the simulations, we use an SI model with no recovery and no disease-induced mortality. We assumed that an infectious contact occurs when the distance between a susceptible and an infectious individual is less than the interaction radius, $r = 1$, and the probability of disease transmission given a contact is 1. Thus, the total number of infections per unit of time is exactly the total number of infectious contacts per unit of time. We initially set the number of infectious individuals to 0.1% of the total host population which is $S + I = N = 1000$ and results are averaged over 30 runs for each simulated model. Under the mass-action law, the number of infections per unit of time is given by $\beta I(N - I)$ which is a quadratic function with one unknown parameter, β . To estimate β , we used the non-linear least squares method. The estimated values of β (with 95 % confidence intervals in the parentheses) are $3.9\text{e-}5$ ($[3.971\text{e-}5, 3.977\text{e-}5]$) for URW), $4.1\text{e-}5$ ($[4.180\text{e-}5, 4.186\text{e-}5]$) for BRW) and $7.1\text{e-}5$ ($[7.096\text{e-}5, 7.104\text{e-}5]$) for LW). For all model fits $R^2 = 0.999$.

Peng and Liu 2009). In these studies, a location x is defined as high-risk when the transmission rate is greater than the recovery rate ($\beta(x) > \gamma(x)$), otherwise it is a low-risk location. If the sum over the spatial domain of local transmission rates is less or equal to the sum of local recovery rates then it is a low-risk domain, otherwise it is a high-risk domain (Allen et al. 2007; Peng and Liu 2009). For a special case

($D_S = 0$), Allen et al. (2007) derive an analytical expression for R_0 for two adjacent habitats and showed that the occurrence of an epidemic depends on the epidemiological characteristics of the domain and the diffusion coefficient of infectious individuals D_I . Allen et al. (2007) showed that in a high-risk domain an epidemic occurs ($R_0 > 1$) no matter the value of D_I , in contrast, in a low-risk domain an epidemic occurs only if D_I is lower than a threshold diffusivity denoted D^* . This result suggests a relationship between the occurrence of the epidemic and the diffusive movement of host individuals.

Moreover, Peng (2009) and Peng and Liu (2009) investigated a reaction-diffusion SIS models with spatially heterogeneous transmission and recovery rates where susceptible hosts move more or less rapidly (D_S tends to 0 or ∞). These studies found that the extinction or the persistence of the epidemic depends on the epidemiological characteristics of the domain and the diffusion coefficients. Furthermore, epidemiological parameters (transmission and recovery rates) may vary not only spatially but also temporally due to seasonality. To fill this gap, Peng and Zhao (2012) incorporated spatially heterogeneous and temporally periodic epidemiological parameters to the reaction-diffusion epidemic model proposed by Allen et al. (2007). Their results show that, if the domain is high-risk or there is at least a high-risk location in the domain and if the diffusion coefficient of infectious individuals (D_I) tends to zero then an epidemic occurs. In contrast, if D_I is very high and if the domain is low-risk then the disease dies out.

In summary, it appears that when epidemiological parameters (transmission and recovery rates) are the same everywhere, the diffusion coefficient of infectious individuals (D_I) affects the speed of disease propagation once it occurs but not the occurrence

of the disease itself (Table 4.1). In contrast, when epidemiological parameters are spatially heterogeneous the URW of host individuals can affect the occurrence of an epidemic (R_0) via diffusion coefficients.

4.4. Epidemics when host movement direction is biased or temporally autocorrelated

The URW model assumes that successive steps moved by an individual are temporally independent. Including correlation between the direction of successive steps allows movement in a same direction relative to the previous one. This type of movement is termed the *correlated random walk* (CRW) and illustrates that the mover is informed about the location of food, prey or mate (Goldstein 1951; Okubo and Günbaum 2001). Empirical support for CRWs have been found in the oviposition movement of butterflies (Kareiva and Shigesada 1983), the foraging movement in bees (Marchand et al. 2015) and relatively short time scale movement of caribou (Bergman et al. 2000) and pea aphids (Nilsen et al. 2013). For both URWs and CRWs, the movement direction is chosen from a uniform distribution. When the URW or the CRW is more likely in a given direction (the movement direction is chosen from a non-uniform distribution), the resulting movement is a *biased random walk* (BRW) or a *biased correlated random walk* (BCRW) (Codling et al. 2008). Biased walks reflect a directed movement towards a specific point such as a foraging place or home and a sample movement path for an individual performing a BRW is shown in Figure 4.1b. Moreover, other models such as CRWs with heterogeneous distribution of resources and interactions between conspecifics have been developed and are appropriately reviewed in Okubo and Günbaum (2001) and Codling et al. (2008).

Including BRWs of hosts into a Kermack-McKendrick epidemic model gives a

system of PDEs of the following form:

$$\begin{aligned}\frac{\partial S}{\partial t} &= D_S \frac{\partial^2 S}{\partial x^2} - C_S \frac{\partial S}{\partial x} - \beta SI \\ \frac{\partial I}{\partial t} &= D_I \frac{\partial^2 I}{\partial x^2} - C_I \frac{\partial I}{\partial x} + \beta SI - \mu I,\end{aligned}\quad x \in \mathbb{R} \quad \text{and} \quad t \in \mathbb{R}^+, \quad (4.2)$$

where C_S and C_I represent the advection rates of susceptible and infectious individuals respectively and describe the speed of directed movement towards the focal point. The system (4.2) involves two components of individual movement. The random movement of susceptible and infected individuals represented by the diffusion terms ($D_S \partial^2 S / \partial x^2$ and $D_I \partial^2 I / \partial x^2$) and the directed movement of susceptible and infected individuals toward the focal point represented by the advection terms ($C_S \partial S / \partial x$ and $C_I \partial I / \partial x$). The remaining terms (called the reaction terms) are infection and disease-induced host mortality at each location. The focal point can be a fixed foraging location where food is more available, a den for animals such as foxes and badgers, or a workplace for humans. Beardmore and Beardmore (2003) investigated the system (4.2) on a bounded domain ($x \in [0, 5]$) and showed that $S_0 > \mu/\beta$ is a sufficient condition for the occurrence of an epidemic when host movement is biased. This result suggests that the occurrence of an epidemic might not depend on how host individuals move towards a preferred location. We performed numerical simulations and show that the rate that new infections occur for a population of individuals undergoing a BRW is consistent with the mass-action law assumed by traditional epidemic models (Figure 4.2).

In comparison to reaction-diffusion epidemic models, relatively few studies investigated the relationship between advection parameters and the pattern of spatial spread of infectious diseases (Gudelj and White 2004; Gudelj et al. 2004) and we are not aware of any studies that have determined if advection parameters affect R_0 for

spatially heterogeneous epidemiological parameters and environments. Finally, while analyses of movement data suggest CRWs as a possible model of animal movement, to date, no epidemiological model that consider host movement as a CRW have been investigated (for a mathematical formulation of the CRW model see Haderler 2015).

Our discussion in sections 4.3 and 4.4 has focused on PDE models, however, host movement may also be formulated mechanistically as an integro-differential equation. Under this formulation, movement from a location y to a location x is assumed to follow a probability density function specified by a kernel. This movement kernel might be skewed in a particular direction representing movement similar to a BRW. The theoretical framework as well as the epidemiological dynamics of integro-differential epidemic models are appropriately reviewed in Medlock and Kot (2003) and Ruan (2007). Similar to PDE-based epidemic models, there exists a critical velocity c_0 above which the disease spreads as a travelling wave from its introduction point. Medlock and Kot (2003) showed that the expression for c_0 depends on the choice of the kernel and c_0 is a function of the movement coefficients of host individuals. However, Medlock and Kot (2003) did not report any relationship between the disease outbreak itself ($R_0 > 1$) and the movement of host individuals or the choice of the kernel.

4.5. Epidemics when host individuals are discrete

In contrast to PDE-based models, individual-based models (IBMs) focus on a Lagrangian description of animal movement. For IBMs, host individuals are represented as discrete entities (the size of the total host population is a whole number) and each host is associated with a specific location, whereas hosts are represented as densities in PDE-based models (Figure 4.3). In the IBM formulation, at each time the location of every individual is updated following a set of movement rules (Preisler et al. 2004;

DeAngelis and Gross 1992). Also, the infection process is described using a set of rules governing contacts between individuals and the transmission of the disease. An epidemiological status (for example, susceptible or infectious) is attributed to each individual at each point in time. During an increment of time, a susceptible individual can become infected when it interacts with an infectious individual at a spatial location. An interaction radius r is defined and determines the spatial proximity required for potential infections. Thus, for IBM models the total number of infections at a time t depends on the total number of nearby susceptible and infectious hosts, whereas for PDE-based models the total number of infections within a small vicinity of the space at a time t is function of the densities of susceptible and infectious individuals on the interval (Figure 4.3; for a detailed description of an IBM epidemic model see Frasca et al. 2006).

4.5.1. *Uncorrelated random walks*

Buscarino et al. (2008) considered an IBM epidemic model in two spatial dimensions, where host individuals exhibit URWs and can perform long distance jumps to a random location with probability p_j . In the special case where $p_j = 1$ (host individuals perform only long distance jumps) the population mixes at random and the contact process is homogeneous. For this limiting case, an explicit expression for the epidemic threshold can be obtained and is given by $\lambda/\gamma > \sigma_c$ where, $\sigma_c = 1/\pi r^2 S_0$ and thus, $S_0 > \gamma/\lambda\pi r^2$ (λ is the probability of becoming infected given a contact, S_0 is the initial density of susceptible hosts and γ is host recovery rate). This epidemic threshold is equivalent to the one obtained from traditional epidemic model where $\beta = \lambda\pi r^2$. Buscarino et al. (2008) then investigated the relationship between the movement rules (URWs with different p_j) and σ_c , and found that for similar S_0 , σ_c decreases with p_j .

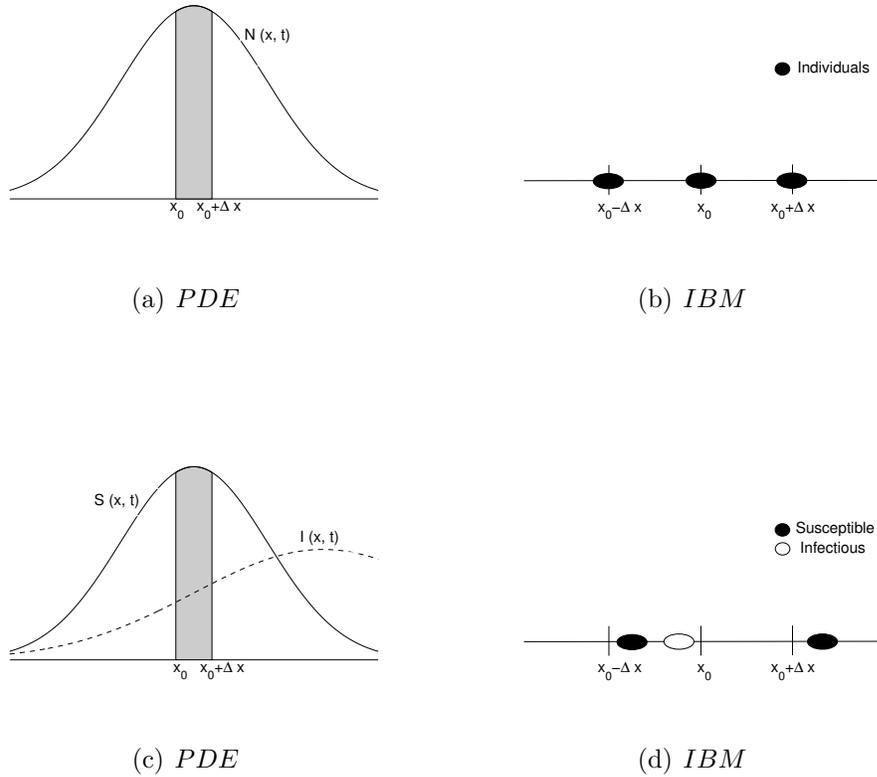


Figure 4.3: The population size and infectious contacts for PDE models (a and c) and IBMs (b and d). For PDEs, the number of individuals on an interval x_0 to $x_0 + \Delta x$ at a time t is given by the integral of the population density $N(x, t)$ over the interval $(\int_{x_0}^{x_0 + \Delta x} N(x, t) dx)$. The population size at each time is given by a probability distribution (a). In contrast, for IBMs individuals are discrete, the population size is represented by a whole number and each individual has a specific location at a given time (b). For PDEs, the number of infectious contacts on an interval x_0 to $x_0 + \Delta x$ at time t is given by the integral of the product of susceptible and infectious densities on the interval $(\int_{x_0}^{x_0 + \Delta x} S(x, t)I(x, t) dx)$ (c). In contrast, for IBMs, an interaction radius is defined because no two individuals will ever be located at exactly the same location at the same time. A contact occurs when two individuals fall in this interaction radius. The total number of infectious and susceptible individuals in spatial proximity determines the number of infectious contacts at a given time. As shown in (d), the interaction radius is Δx and a contact occurs on the interval $x_0 - \Delta x$ to x_0 .

This result suggests that an epidemic is less likely when individuals exhibit URWs ($p_j = 0$) compared to long distance jumps ($p_j = 1$). Long distance jumps may enhance the mixing process, and as such promote the occurrence of an epidemic. However, this

effect is less pronounced as S_0 becomes large and the epidemic threshold no longer depends on p_j .

4.5.2. *Levy walks*

Animal movement patterns can be described as clusters of short step lengths connected by persistent-like movement, reflecting a shift between intense and less intense search modes. This movement behaviour is termed a *Levy Walk* (LW) and is considered to be an efficient foraging strategy when food is rare and randomly distributed (Reynolds 2013). A sample movement path for an individual performing a LW is shown in Figure 4.1c. LWs have been reported in many species and ecological phenomenon including the foraging movement of spider monkeys (Ramos-Fernndez et al. 2004), the daily movement pattern of humans (Rhee et al. 2011) and the hunting-gathering movement of humans (Raichlen et al. 2014). For a complete review of LWs in movement ecology and its status as efficient foraging strategy see Reynolds (2015) and Pyke and Giuggioli (2015). Buscarino et al. (2010) modified the model proposed earlier in Buscarino et al. (2008) (see section 4.5.1) by considering a LW of host individuals and compared the risk of disease outbreak for URWs and LWs. They showed that for similar S_0 a disease outbreak may be more likely in a population of *Levy walkers* compared to a population of *uncorrected random walkers*. Few studies have investigated epidemics in populations where individuals perform a LW, however, the numerical simulations in Figure 4.2 illustrate that the mass-action assumption is consistent with the contact rate arising from LW movement.

In summary, the above IBM-based epidemic models (sections 4.5.1 and 4.5.2) suggest that the type of movement performed by host individuals may affect the critical quantity σ_c at least for relatively low population sizes (Frasca et al. 2006; Buscarino

et al. 2008; 2010). However, using an IBM framework, the analysis is often restricted to a quantitative description of the epidemiological dynamics. In particular, deriving an analytical expression for the epidemic threshold, R_0 , and a solution describing the spatial spread of the disease are mathematically challenging. The quantity σ_c derived in these studies does not involve movement parameters and it is difficult to conclude whether the type of movement performed by host individuals affects the epidemic threshold.

The effect of host movement on the spread of infectious diseases has also been studied using contact network models. In these models, a type of host movement is implicitly acknowledged and the contact structure of the population is explicitly modelled using networks. The nodes of the network represent either host individuals or neighbourhoods and the edges represent connections between individuals or neighbourhoods which is possible through movement. This class of model is appropriately reviewed in Keeling and Eames (2005) and Brauer (2008) and will not be discussed in the present paper.

4.6. Case study: rabies

Rabies is a viral infection which spreads mainly within wild carnivores including the red fox (*Vulpes vulpes*), the arctic fox (*Vulpes lagopus*), raccoons (*Procyon lotor*) and domestic carnivores such as dogs (*Canis familiaris*) and cats (*Felis catus*). The virus is present in the saliva of rabid hosts, is transmitted through direct contacts (especially bites), has a particularly long incubation period between 12 and 150 days and ultimately kills its host. Rabies causes a random-like movement when it affects the central nervous system of foxes (Baer 1991; Kaplan 1977; Bacon 1985). During the 1980s, particular attention was payed to the inclusion of animal movement (espe-

Table 4.2: Summary of different movement types, model formulations and the corresponding epidemic threshold, R_0 . Model abbreviations and parameters are defined in the Table 4.1.

Model formulation	Spatially heterogeneous β and γ	Movement type	R_0	References
Reaction-diffusion model. SI	No	URWs	$\frac{\beta S_0}{\mu}$, does not depend on movement parameters.	Hosono and Ilyas (1995); Källén (1984); Källén et al. (1985)
Reaction-diffusion model. SIR	No	URWs	$\frac{\beta S_0}{\mu + \gamma}$, does not depend on movement parameters.	Li and Li (2015); Wang and Wu (2010); Bai and Zhang (2015).
Host-vector epidemic model.	No/yes (numerically)	URWs	$\sqrt{\frac{e^{-d_w w w} W_1^* \beta_w^*}{d_h + \mu_h + \gamma} \times \frac{e^{-d_h w w} S^* \beta_h^*}{d_w + \mu_w}}$, depends on movement parameters for spatially heterogeneous β_w and β_h .	Wang and Zhao (2011).
Reaction-diffusion model. SIS	Yes	URWs	$\frac{\beta_2 \gamma_1 + \beta_1 \gamma_2 + D_I l (\beta_1 + \beta_2)}{2(\gamma_1 \gamma_2 + D_I l (\gamma_1 + \gamma_2))} + \sqrt{\frac{[\beta_2 \gamma_1 - \beta_1 \gamma_2 + D_I l (\beta_2 - \beta_1)]^2 + (2D_I l)^2 \beta_1 \beta_2}{2(\gamma_1 \gamma_2 + D_I l (\gamma_1 + \gamma_2))}}$	Allen et al. (2007); Peng (2009); Peng and Liu (2009).
Advection-diffusion model. SI	No	BRWs	$\frac{\beta S_0}{\mu}$, does not depend on movement parameters.	Beardmore and Beardmore (2003).
IBM SIR model.	No	URWs, LWs	$\frac{\lambda}{\gamma \sigma_c}$, σ_c is affected by movement parameters.	Frasca et al. (2006); Buscarino et al. (2008).

cially random movement of rabid foxes) in the mathematical models of rabies spread (Panjeti and Real 2011). Reaction-diffusion models have been mainly used to capture the spatial spread of rabies in the red fox in Western Europe as well as the arctic fox and raccoons in North America. In this section, we summarize some important results of these studies, their relationship with field data and some control measures implemented using these models.

Murray et al. (1986) proposed a reaction-diffusion model for the rabies epizootic that occurred in central Europe during the 1940s. The model assumes that rabies is transmitted among fox populations with density dependent growth. Susceptible foxes are considered territorial and are assumed to be homogeneously distributed. Rabid foxes move randomly, travel far away from their den, and may infect susceptible individuals they encounter during their wanderings. Murray et al. (1986) found that the occurrence of rabies epizootic depends on a critical carrying capacity of fox populations K_t which is analogous to the critical density S_0 for traditional models that utilize

different demographic assumptions. This critical carrying capacity is independent of the diffusion coefficient which is consistent with the finding that R_0 is independent of movement parameters reported in Section 4.3. If the carrying capacity of the fox population K is greater than the critical carrying capacity K_t , then the disease spreads outward from the endemic location to disease-free locations as a travelling wave at a critical speed of propagation $c_0 = \sqrt{D_I \beta K z}$ (where z is the unique root of a cubic function). From the expression for c_0 , it can be noticed that the speed of rabies propagation increases with the diffusion coefficient of rabid foxes. Moreover, it has been shown that the front of the wave (which is the first passage of rabies epizootic at a location) is followed by an oscillatory tail suggesting periodic outbreaks after the first outbreak. The front of the wave is characterized by a severe epizootic with a high number of foxes dying from rabies whereas each following outbreak is less severe than the previous one. A similar model assuming exponential growth for fox populations, exhibits the same qualitative behaviour, which has been shown to agree with field data (Källén et al. 1985). Furthermore, Murray et al. (1986) have estimated that D_I is between $50 \text{ km}^2 \text{ year}^{-1}$ and $330 \text{ km}^2 \text{ year}^{-1}$ using different data sources and methods. Varying D_I in this interval and keeping all the other parameters constant, the speed of the epidemic increases by a factor of 2.6. In addition, Murray et al. (1986) showed that for $D_I = 200 \text{ km}^2 \text{ year}^{-1}$ and fixing the fox population carrying capacity at 2 km^{-2} , rabies spreads at a velocity $c = 51 \text{ km year}^{-1}$.

The above reaction-diffusion framework has been used for the implementation and the evaluation of rabies control measures. Murray et al. (1986) suggested that the spatial propagation of rabies can be ‘broken’ by reducing the density of susceptible foxes below the persistence threshold K_t before the wave reaches a disease-free area.

However, the results of Källén et al. (1985) and Murray et al. (1986) do not account for environmental heterogeneity (resources and landscape) and spatially heterogeneous epidemiological parameters. In particular, landscape heterogeneity can play a major role in the spatial spread of rabies (Panjeti and Real 2011). For example, an immigration-based model of the spatial spread of rabies in raccoons across heterogeneous landscapes has revealed that large rivers can reduce the speed of propagation of rabies by 7 fold (Smith et al. 2002). As reported in Section 4.3, theoretical studies suggest that conclusions based on threshold quantities for a disease outbreak may be sensitive to assumptions of environmental homogeneity.

4.7. Concluding remarks and perspectives

Overall, including the URWs of host individuals in disease models reveals that the diffusion coefficients (D_S and D_I) affect the threshold condition for epidemic occurrence R_0 only when epidemiological parameters (the transmission and the recovery rates) are spatially heterogeneous (Table 4.1). An effect of host movement on R_0 was expected because how host individuals move affects the distribution of susceptible and infected individuals and the contact process which is represented by βSI in a mass-action model formulation. It is surprising however that spatially heterogeneous transmission and/or recovery rates are required for the epidemic occurrence (R_0) to be affected by diffusion coefficients. Frequently, when the law of mass action is assumed it is stated that this assumption implies homogeneous mixing, however the types of movement that are consistent with a mass-action model formulation may be much more general. We reviewed epidemiological studies that considered animals moving following URWs, BRWs and LWs and found limited evidence that the threshold for a disease outbreak was affected by the type of host movement (Table 4.1). In

addition, numerical simulations suggested that each of these three movement types (URWs, BRWs and LWs) produces an infection rate consistent with the infection rate assumed by the law of mass action (Figure 4.2). Despite the failure of animal movement models to affect the threshold condition for a disease outbreak there are several reasons why considering animal movement in epidemic models is useful. We suggest the following promising directions for future research:

1. The formulation of a mechanistic sub-model for the contact process in order to understand how different types of animal movement affect the mixing process for disease transmission.
2. The development of PDE-based epidemic models with underlying individual movement such as CRWs, BCRWs and LWs in order to investigate the effect of more realistic movement rules on disease spread. CRW, BCRWs and LW models are prevalent in the animal movement literature, but few epidemic models consider these types of movement.
3. The development of epidemic models that consider spatially dependent diffusion coefficients in order to investigate the spread of infectious diseases in non-homogeneous environments and landscapes.
4. Finally, the coupling of telemetry-derived and epidemiological data to parameterize and validate epidemiological models and the development of robust statistical tools to achieve this goal. In particular, if the contact rate could be estimated from GPS data then it is more likely that the probability of an infection given a contact could be estimated from epidemic data. This is valuable because it would help to estimate the prevalence of ‘near-misses’ occurring during an outbreak. Near-misses are contacts that did not result in infection and

it may be useful to explore alternative epidemic scenarios based on instances where near misses are instead realized.

Literature cited

- L.J. Allen, B.M. Bolker, Y. Lou, and A.L. Nevai. Asymptotic profiles of the steady states for an SIS epidemic patch model. *SIAM Journal on Applied Mathematics*, 67(5):1283–1309, 2007.
- L.J. Allen, F. Brauer, P. Van den Driessche, and J. Wu. *Mathematical epidemiology*. Springer, 2008.
- R.M. Anderson, R.M May, and B. Anderson. *Infectious diseases of humans: dynamics and control*, volume 28. Oxford University Press, 1992.
- P.J. Bacon. *Population dynamics of rabies in wildlife*. Academic Press Inc.(London) Ltd., 1985.
- G.M. Baer. *The natural history of rabies*. CRC press, 1991.
- Z. Bai and S. Zhang. Traveling waves of a diffusive SIR epidemic model with a class of nonlinear incidence rates and distributed delay. *Communications in Nonlinear Science and Numerical Simulation*, 22(1):1370–1381, 2015.
- J. Barry, M. Newton, J.A. Dodd, O.E. Hooker, P. Boylan, M.C. Lucas, and C.E. Adams. Foraging specialisms influence space use and movement patterns of the european eel *Anguilla anguilla*. *Hydrobiologia*, 766(1):333–348, 2016.
- I. Beardmore and R. Beardmore. The global structure of a spatial model of infectious disease. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 459:1427–1448, 2003.
- M. Begon, M. Bennett, R.G. Bowers, N.P. French, SM Hazel, and J. Turner. A clarifica-

- tion of transmission terms in host–microparasite models: numbers, densities and areas. *Epidemiology and infection*, 129(01):147–153, 2002.
- H.C. Berg. *Random Walks in Biology*. Princeton University Press, 1983.
- C.M. Bergman, J.A. Schaefer, and S.N. Luttich. Caribou movement as a correlated random walk. *Oecologia*, 123(3):364–374, 2000.
- F. Brauer. *An introduction to networks in epidemic modeling*, pages 133–146. Mathematical epidemiology. Springer, 2008.
- A. Buscarino, A. Di Stefano, L. Fortuna, M. Frasca, and V. Latora. Disease spreading in populations of moving agents. *EPL (Europhysics Letters)*, 82(3):38002, 2008.
- A. Buscarino, A. Di Stefano, L. Fortuna, M. Frasca, and V. Latora. Effects of motion on epidemic spreading. *International Journal of Bifurcation and Chaos*, 20(3):765–773, 2010.
- F. Cagnacci, L. Boitani, R.A. Powell, and M.S. Boyce. Animal ecology meets GPS–based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1550):2157–2162, 2010.
- A.D. Cliff. *Incorporating spatial components into models of epidemic spread*, pages 119–149. Epidemic Models: Their Structure and Relation to Data, D.Mollison, ed. Cambridge University Press, Cambridge, 1996.
- E.A. Codling, M.J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.
- J.T. Cronin. Movement and spatial population structure of a prairie planthopper. *Ecology*, 84(5):1179–1188, 2003.
- D.L. DeAngelis and L.J. Gross. *Individual-based models and approaches in ecology: populations, communities and ecosystems*. Chapman & Hall, 1992.
- O. Diekmann, J.A.P. Heesterbeek, and J.A.J. Metz. *The legacy of Kermack and McK-*

- endrick*, pages 95–115. *Epidemic Models: Their Structure and Relation to Data*, D.Mollison, ed. D. Mollison, ed, 1995.
- O. Diekmann, H. Heesterbeek, and T. Britton. *Mathematical tools for understanding infectious disease dynamics*. Princeton University Press, 2012.
- R. Durrett. Spatial epidemic models. *Epidemic Models: Their structure and relation to data*, pages 187–201, 1995.
- M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, and S. Boccaletti. Dynamical network model of infective mobile agents. *Physical Review E*, 74(3):036110, 2006.
- S. Goldstein. On diffusion by discontinuous movements, and on the telegraph equation. *The Quarterly Journal of Mechanics and Applied Mathematics*, 4(2):129–156, 1951.
- I. Gudelj and K.A.J White. Spatial heterogeneity, social structure and disease dynamics of animal populations. *Theoretical population biology*, 66(2):139–149, 2004.
- I. Gudelj, K.A.J White, and N.F. Britton. The effects of spatial movement and group interactions on disease dynamics of social animals. *Bulletin of mathematical biology*, 66(1):91–108, 2004.
- K.P. Hadeler. Stefan problem, traveling fronts, and epidemic spread. *Discrete and continuous dynamical systems.*, 21(2):417–436, 2015.
- M. Hebblewhite and D.T. Haydon. Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1550):2303–2312, 2010.
- H.W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.
- Y. Hosono and B. Ilyas. Traveling waves for a simple diffusive epidemic model. *Mathematical Models and Methods in Applied Sciences*, 5(07):935–966, 1995.
- R.A. Ims. *Movement patterns related to spatial structures*, pages 85–109. Mosaic landscapes and ecological processes. Springer, 1995.

- A. Källén. Thresholds and the natural history of rabies in an epidemic model for rabies. *Nonlinear Analysis: Theory, Methods & Applications*, 8(8):851–856, 1984.
- A. Källén, P. Arcuri, and J.D. Murray. A simple model for the spatial spread and control of rabies. *Journal of theoretical biology*, 116(3):377–393, 1985.
- C. Kaplan. *Rabies: the facts*. Oxford University Press, Walton Street, Oxford OX2 6DP, 1977.
- P.M. Kareiva. Experimental and mathematical analyses of herbivore movement: quantifying the influence of plant spacing and quality on foraging discrimination. *Ecological Monographs*, pages 261–282, 1982.
- P.M. Kareiva. Local movement in herbivorous insects: applying a passive diffusion model to mark-recapture field experiments. *Oecologia*, 57(3):322–327, 1983.
- P.M. Kareiva and N. Shigesada. Analyzing insect movement as a correlated random walk. *Oecologia*, 56(2-3):234–238, 1983.
- M.J. Keeling and K.T. Eames. Networks and epidemic models. *Journal of the Royal Society, Interface*, 2(4):295–307, Sep 22 2005.
- W.O. Kermack and A.G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, 115(772):700–721, 1927.
- M.A. Lewis, P.K. Maini, and S.V. Petrovskii. *Dispersal, Individual Movement and Spatial Ecology*. Springer Berlin Heidelberg, 2013.
- Y. Li and G. Li, W.T. and Lin. Traveling waves of a delayed diffusive SIR epidemic model. *Communications on pure and applied analysis*, 14(3):1001–1022, 2015.
- P. Marchand, A.N. Harmon-Threatt, and I. Chapela. Testing models of bee foraging behavior through the analysis of pollen loads and floral density data. *Ecological Modelling*, 313:41–49, 2015.

- N.C. Marchant, A. Purwanto, F.A. Harsanto, N.S. Boyd, M.E. Harrison, and P.R. Houlihan. ‘random-flight’ dispersal in tropical fruit-feeding butterflies? high mobility, long lifespans and no home ranges. *Ecological Entomology*, 40(6):696–706, 2015.
- H. McCallum, N. Barlow, and J. Hone. How should pathogen transmission be modelled? *Trends in ecology & evolution*, 16(6):295–300, 2001.
- H.W. McKenzie, E.H. Merrill, R.J. Spiteri, and M.A. Lewis. How linear features alter predator movement and the functional response. *Interface focus*, 2(2):205–216, 2012.
- J. Medlock and M. Kot. Spreading disease: integro–differential equations old and new. *Mathematical biosciences*, 184(2):201–222, 2003.
- E. Merrill, H. Sand, B. Zimmermann, H. McPhee, N. Webb, M. Hebblewhite, P. Wabakken, and J.L. Frair. Building a mechanistic understanding of predation with GPS–based movement data. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1550):2279–2288, 2010.
- J.M. Morales, P.R. Moorcroft, J. Matthiopoulos, J.L. Frair, J.G. Kie, R.A. Powell, E.H. Merrill, and D.T. Haydon. Building the bridge between animal movement and population dynamics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1550):2289–2301, 2010.
- J.D. Murray, E.A. Stanley, and D.L. Brown. On the spatial spread of rabies among foxes. *Proceedings of the Royal Society of London. Series B, Biological sciences*, 229(1255):111–150, 1986.
- C. Nilsen, J. Paige, O. Warner, B. Mayhew, R. Sutley, M. Lam, A.J. Bernoff, and C.M. Topaz. Social aggregation in pea aphids: experiment and random walk modeling. *PloS one*, 8(12):e83343–e83343, 2013.
- A. Okubo and D. Günbaum. *Mathematical treatment of biological diffusion*, pages 127–169. Diffusion and Ecological Problems: Modern Perspectives. Springer–Verlag, 2001.

- A. Okubo and S.A. Levin. *Diffusion and ecological problems : modern perspectives*. Springer-Verlag, 2001.
- V.G. Panjети and L.A. Real. Mathematical models for rabies. *Advances in Imaging and Electron Physics*, 79:377–395, 2011.
- R. Peng. Asymptotic profiles of the positive steady state for an SIS epidemic reaction–diffusion model. part i. *Journal of Differential Equations*, 247(4):1096–1119, 2009.
- R. Peng and S. Liu. Global stability of the steady states of an SIS epidemic reaction–diffusion model. *Nonlinear Analysis*, 71(1):239–247, 2009.
- R. Peng and X.Q. Zhao. A reaction–diffusion SIS epidemic model in a time–periodic environment. *Nonlinearity*, 25(5):1451, 2012.
- H.K. Preisler, A.A. Ager, B.K. Johnson, and J.G. Kie. Modeling animal movements using stochastic differential equations. *Environmetrics*, 15(7):643–657, 2004.
- M.D. Preston, M.L. Forister, J.W. Pitchford, and P.R. Armsworth. Impact of individual movement and changing resource availability on male–female encounter rates in an herbivorous insect. *Ecological Complexity*, 24:1–13, 2015.
- G.H. Pyke and L. Giuggioli. Understanding movements of organisms: it’s time to abandon the levy foraging hypothesis. *Methods in Ecology and Evolution*, 6(1):1–16, 2015.
- D.A. Raichlen, B.M. Wood, A.D. Gordon, A. Z. Mabulla, F.W. Marlowe, and H. Pontzer. Evidence of levy walk foraging patterns in human hunter–gatherers. *Proceedings of the National Academy of Sciences of the United States of America*, 111(2):728–733, 2014.
- G. Ramos-Fernndez, J.L. Mateos, O. Miramontes, G. Cocho, H. Larralde, and B. Ayala-Orozco. Levy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). *Behavioral Ecology and Sociobiology*, 55(3):223–230, 2004.
- A. Reynolds. *Beyond optimal searching: recent developments in the modelling of animal movement patterns as Levy walks*, pages 53–76. Dispersal, Individual Movement and

- Spatial Ecology. Springer, 2013.
- A. Reynolds. Liberating levy walk research from the shackles of optimal foraging. *Physics of life reviews*, 2015.
- I. Rhee, M. Shin, S. Hong, K. Lee, S.J. Kim, and S. Chong. On the levy-walk nature of human mobility. *IEEE/ACM transactions on networking (TON)*, 19(3):630–643, 2011.
- C.J. Rhodes and R.M. Anderson. Contact rate calculation for a basic epidemic model. *Mathematical biosciences*, 216(1):56–62, 2008.
- S. Ruan. *Spatial-temporal dynamics in nonlocal epidemiological models*, pages 97–122. Mathematics for life science and medicine. Springer, 2007.
- M.W. Shaw, T. D. Harwood, M.J. Wilkinson, and L. Elliott. Assembling spatially explicit landscape models of pollen and spore dispersal by wind for risk assessment. *Proceedings of the Royal Society B: Biological Sciences*, 273(1594):1705–1713, 2006.
- N. Shigesada, K. Kawasaki, and H.F. Weinberger. Spreading speeds of invasive species in a periodic patchy environment: effects of dispersal based on local information and gradient-based taxis. *Japan Journal of Industrial and Applied Mathematics*, 32(3):675–705, 2015.
- J. G. Skellam. Random dispersal in theoretical populations. *Biometrika*, 38(1–2):196, 1951.
- D.L. Smith, B. Lucey, L.A. Waller, J.E. Childs, and L.A. Real. Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6):3668–3672, 2002.
- P.E. Smouse, S. Focardi, P.R. Moorcroft, J.G. Kie, J.D. Forester, and J.M. Morales. Stochastic modelling of animal movement. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1550):2201–2211, 2010.
- F. Spitzer. *Principles of random walk*. Springer-Verlag, 1976.
- I.R. Swingland and P.J. Greenwood. *The Ecology of animal movement*. Clarendon Press, 1983.

- I. Tiutiunov, L.I. Titova, and S.V. Berdnikov. Mechanistic model for the allee effect and interference in predator population. *Biofizika*, 58(2):349–356, 2013.
- P. Turchin. Translating foraging movements in heterogeneous environments into the spatial distribution of foragers. *Ecology*, 72(4):1253–1266, 1991.
- P. Turchin. *Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants*. Sunderland: Sinauer Associates, 1998.
- W. Wang and X.Q. Zhao. A nonlocal and time–delayed reaction–diffusion model of dengue transmission. *SIAM Journal on Applied Mathematics*, 71(1):147–168, 2011.
- X.S. Wang, H. Wang, and J. Wu. Traveling waves of diffusive predator–prey systems: disease outbreak propagation. *Discrete and Continuous Dynamical Systems A*, 32(9):3303–3324, 2012.
- Z.C. Wang and J. Wu. Travelling waves of a diffusive Kermack–Mckendrick epidemic model with non–local delayed transmission. *Proceedings of the Royal Society A*, 466(2113):237–261, 2010.

5. **Chapter four: A counter-intuitive relationship between the temporal and spatial spread of diseases**

This chapter is a prepared manuscript for submission to the journal *Emerging Infectious Diseases*

A counter-intuitive relationship between the temporal and spatial spread of diseases

Abdou M. Fofana, Amy Hurford

Abstract

Spatial epidemic models predict a positive relationship between the temporal, r , and the spatial, c , spread rates of diseases because when more infectious individuals are produced per unit of time, more infected hosts can disperse long distances resulting in a larger area covered by the epidemic, however, this prediction has yet to be validated with empirical data. In this paper, we tested whether infectious diseases that produce more infectious individuals per week also spread more kilometres per week using 42 outbreaks caused by 10 infectious organisms with different transmission routes that are reported in the United States of America between 1996 and 2017 by the Centers for Disease Control and Prevention. In contrast to the predictions of spatially explicit disease spread models, we find that infectious diseases that produce more cases per week spread less kilometres per week. This inverse relationship between r and c is unexpected but may reflect a trade-off between parasite dispersal and disease transmission. Our results suggest that assumed relationships between spatial and temporal spread require further investigations, and may inform best approaches for spatially explicit quarantine and vaccination strategies.

Keywords: Epidemic growth rate, spatial spread rate, movement, GAM, CDC.

5.1. Introduction

During the early phase of many infectious disease outbreaks the number of infected individuals in the host population increases exponentially, and the rate of exponential growth is a key epidemiological quantity for the estimation of disease transmission

potential and the evaluation of the effectiveness of control measures (Anderson et al. 1992; Heesterbeek and Dietz 1996; Heffernan et al. 2005; Nishiura 2010; Chowell and Nishiura 2014; Delamater et al. 2019). After the successful introduction of an infectious disease in a host population, the disease often spreads from one location to another in a wavelike pattern and the estimation of the speed of disease propagation is crucial for planning interventions to limit the spatial propagation of the disease (Thieme 1980; Murray et al. 1986; Van den Bosch et al. 1990; Smith et al. 2002). Mathematical formulations have been derived for the initial epidemic growth rate r and the spatial spread rate c of infectious diseases and there is a substantial theoretical support for a positive relationship between c and r (Murray 1993; Murray et al. 2001; Diekmann et al. 2012c).

In a non-spatial epidemic modelling framework, the host population is subdivided into classes with different epidemiological status, and the temporal dynamic of the different classes is investigated whereas the spatial dynamics of the classes is ignored (Anderson et al. 1992; Allen et al. 2008; Diekmann et al. 2012a). Spatial epidemic models, however, account for the spatial and temporal dynamics of the epidemiological classes often by introducing host movement in the model (Fofana and Hurford 2017). The basic reproduction number R_0 and the spatial spread rate c are key epidemiological quantities that are often derived from spatial models (Allen et al. 2008; Diekmann et al. 2012a). The rate of spatial spread c measures the speed of spatial propagation of an infectious disease, and methods have been developed to estimate c using infection cases reported at different locations during an outbreak (Moore 1999; Farnsworth and Ward 2009; Pioz et al. 2011; Mercier et al. 2018; Tisseuil et al. 2016; Goldstein et al. 2019). The basic reproduction number R_0 measures the generational

growth rate of an epidemic and is often difficult to estimate from infection cases data (Diekmann et al. 2012b; Park et al. 2019). A practical alternative way to quantify the growth of an epidemic is to measure its instantaneous growth rate r during the early phase of the outbreak, using reported infection cases data (Lipsitch et al. 2003; Heffernan et al. 2005; Nishiura et al. 2010; Li and Blakeley 2011; Park et al. 2019; Delamater et al. 2019).

Estimated values of the epidemic growth rate r and the spatial spread rate c are useful for disease control and interventions. The epidemic growth rate r is often used to calculate the transmission potential of infectious diseases, and a transmission potential lower than one indicates that the ongoing epidemic is under control (Roberts and Heesterbeek 2007; Chowell et al. 2016; Park et al. 2019). For example, during the 2001 foot-and-mouth disease outbreak in Great Britain mathematical models and early incidence data have been used to guide the implementation of foot-and-mouth control measures in livestock animals (Ferguson et al. 2001a; Keeling et al. 2001; Kao 2002). Many studies showed that following the implementation of culling policies in addition to livestock movement ban in Europe, the transmission potential of foot-and-mouth dropped below one which means that the implemented control measures were effective (Ferguson et al. 2001b; Woolhouse et al. 2001; Ferguson et al. 2001a; Haydon et al. 2004). Similarly, spatial epidemic models have been extensively used to understand the spatial propagation of the 1940 rabies outbreak in fox populations in Europe, and different studies have reported spread rates between 30 and 60 km/year (Van den Bosch et al. 1990; Murray et al. 1986; Alanazi et al. 2019). The estimated spread rates have been used to determine the size of the region within which vaccination strategies must be deployed to prevent further spatial propagation of rabies

(Murray et al. 1986; Evans and Pritchard 2001). While r can be estimated from early disease outbreak data, the estimation of c requires infection case data with some level of spatial resolution which are not available at the beginning of an outbreak. Understanding the relationship between r and c can help in predicting how far an infectious disease can spread and mobilizing earlier the resources that will be needed to control the outbreak.

The relationship between the epidemic growth rate r and the spatial spread rate c has been previously investigated, and diffusion models suggest that c depends on r and the diffusion coefficient D which measures the dispersal distance of host individuals (Källén et al. 1985; Murray et al. 1986; Diekmann et al. 2012c; Osnas et al. 2015). The mathematical expression for c depends on the assumptions of the model, and in a simple case where hosts move randomly and disease transmission is local $c = 2\sqrt{\beta S_0 D}$ which suggest that the relationship between c and r is positive ($r = \beta S_0$ where β and S_0 are disease transmissibility and initial susceptible host density respectively) (Källén et al. 1985; Murray 1993; Murray et al. 2001; Osnas et al. 2015). Similar models have been developed in invasion ecology where many studies reported a positive relationship between the speed of invasion (which is analogous to c) and the intrinsic growth rate of the invading population (which is analogous to r) (Skellam 1951; Andow et al. 1990; Holmes et al. 1994; Kot et al. 1996; Wang and Kot 2001; Neubert et al. 2000; Okubo and Levin 2013; Hastings et al. 2005; Shea et al. 2010). Theoretical models often predict a positive correlation between c and r because when more infections occur per unit time (high r) there are more infected individuals available who can potentially disperse long distances, and as such the area covered by infected hosts and c can be larger.

Many studies have estimated the intrinsic growth rate and the speed of invasion of the invading organism from species range expansion data in invasion ecology (see review Hastings et al. 2005), and tested whether the invasion speed predicted by diffusion model agrees with the observed invasion speed for the muskrat (Skellam 1951), the collared dove, starling (Van den Bosch et al. 1990), the cereal leaf beetle (Andow et al. 1990), the California sea otter (Lubina and Levin 1988) and wolves (Hurford et al. 2006). In disease ecology, a few empirical studies have investigated the pattern and the rate of disease propagation for some infectious diseases. For example, McCallum et al. (2003) compared c in marine and terrestrial pathogens and showed that marine pathogens spread faster than terrestrial pathogens because there is no barrier to dispersal in marine environments.

In this paper, we test the expected positive relationship between disease spread rate c and the epidemic growth rate r using human infectious disease outbreaks data. We collected infection cases data that were reported in the United States of America from 1996 to June 2019 by the Centers for Disease Control and Prevention (CDC) which is publicly available on CDC website, and we estimated the spatial spread rate c (kilometres per week) and the epidemic growth rate r (cases per week) of 42 outbreaks caused by 10 infectious organisms with different transmission routes. In contrast to the predicted positive relationship between the spatial spread rate c and the epidemic growth rate r , we find that c and r are inversely related which suggests that disease outbreaks where the number of cases are rapidly increasing over time spread fewer kilometres per week.

5.2. Method

We investigated the relationship between the initial epidemic growth rate r and the spatial spread rate c of disease outbreaks across different outbreak years and infectious diseases. We collected weekly cases data of infectious diseases that are reported on the website of the Centers for Disease Control and Prevention (CDC), and we estimated c and r for 42 infectious disease outbreaks that are reported in the United States of America, excluding US territories, between 1996 and 2017. The processed CDC disease outbreak data, the estimated c and r data and all the codes that we used are available as electronic supplementary materials S5.1-S5.2 and are publicly available at Figshare doi:10.6084/m9.figshare.11389584.v2. Details on the data, the estimation of c and r and statistical method are presented in appendix C as electronic supplementary materials.

5.2.1. Disease outbreak data

We retrieved infectious disease cases reported in the United States of America by the Centers for Disease Control and Prevention (CDC) through the National Notifiable Diseases Surveillance System (NNDSS) which is an information sharing system for infectious diseases. Cases of notifiable infectious diseases in the different states and territories of the U.S.A are reported and updated weekly, and the data reported from 1995 to present is publicly available on the website of the CDC via CDC Wide-Ranging Online Data for Epidemiologic Research for public health professionals and the general public.

We used web scraping techniques in Matlab to extract the weekly tables which report the infectious diseases, the number of cases notified each epidemiological week within a year and the cumulative year-to-date cases in each state and territory in

the U.S.A. The CDC data is suitable for investigating the spatio-temporal dynamics of infectious diseases because the weekly tables contain information on when (the epidemiological week), where (the state) and the number of notified cases for about 122 human diseases and conditions.

We retrieved weekly and cumulative year-to-date cases for all infectious diseases reported in each state from 1995 to June 2019. We cleaned and processed the data to ensure that the state and disease names are consistent throughout the data. We visualized the data and selected a total of 28 infectious diseases that exhibit an exponential growth profile, which is a signature of an outbreak and are therefore suitable for the estimation of the initial epidemic growth rate r (Chowell et al. 2016). We excluded the US territories data and restricted our analysis to the spread of the infectious diseases from one state to another, and we used data that have been verified and validated by the CDC. Finally, we replaced the names of the states by their longitude and latitude in decimal degrees which we retrieved from the website of the National Ocean and Atmospheric Administration (NOAA) of the U.S.A.

Ethics statement

The data that we collected from the CDC website were anonymous, publicly available, and represent the total number of cases reported within the different states during a week.

5.2.2. Estimation of disease spatial spread rate c

To estimate the spatial spread rate c of a given outbreak, we fit a generalized additive model with thin plate regression splines to the reported infections data, where the response variable is the week of first reported cases in a state and the latitude and

the longitude of the state are the predictors (Example in Figure 5.1). We calculated c as the sum of the inverse of the estimated slopes for each predictor (latitude and longitude in number of weeks per distance) that we obtained from the generalized additive model fit (Mercier et al. 2018; Tisseuil et al. 2016; Farnsworth and Ward 2009; Lizarazo et al. 2019). Finally, we converted c from distance in decimal degrees per week into kilometres per week using map projection methods (Bugayevskiy and Snyder 2013). We implemented the generalized additive model in the *mgcv* R package (Wood 2004; 2011; Marra and Wood 2012; Wood 2012; 2013; Wood et al. 2016; Wood 2017).

A generalized additive model is a generalized linear model where the relationship between each predictor and the response variable is modelled by a smooth function, and finding the form of the function is part of the data fitting procedure (Hastie and R. 1986). We used a generalized additive model because it can effectively capture non-linear disease spread patterns and no prior knowledge of the relationship between the response and the predictors is required for data fitting. Our generalized additive model can be written as,

$$Y_i = f(\text{Latitude}_i, \text{Longitude}_i) + \epsilon_i, \quad (5.1)$$

where Y_i is the week the first cases were reported in a location i , $f(\text{Latitude}, \text{Longitude})$ is an unknown function that describes the change in the response variable as a function of the predictors, and ϵ_i is the error term which we assumed to be normally distributed.

The change of Y with respect to the locations is,

$$\frac{\partial Y}{\partial \text{Latitude} \partial \text{Longitude}} = \frac{\partial f}{\partial \text{Latitude}} + \frac{\partial f}{\partial \text{Longitude}}, \quad (5.2)$$

where $\frac{\partial f}{\partial \text{Latitude}}$ and $\frac{\partial f}{\partial \text{Longitude}}$ are the partial derivatives of the function f with respect

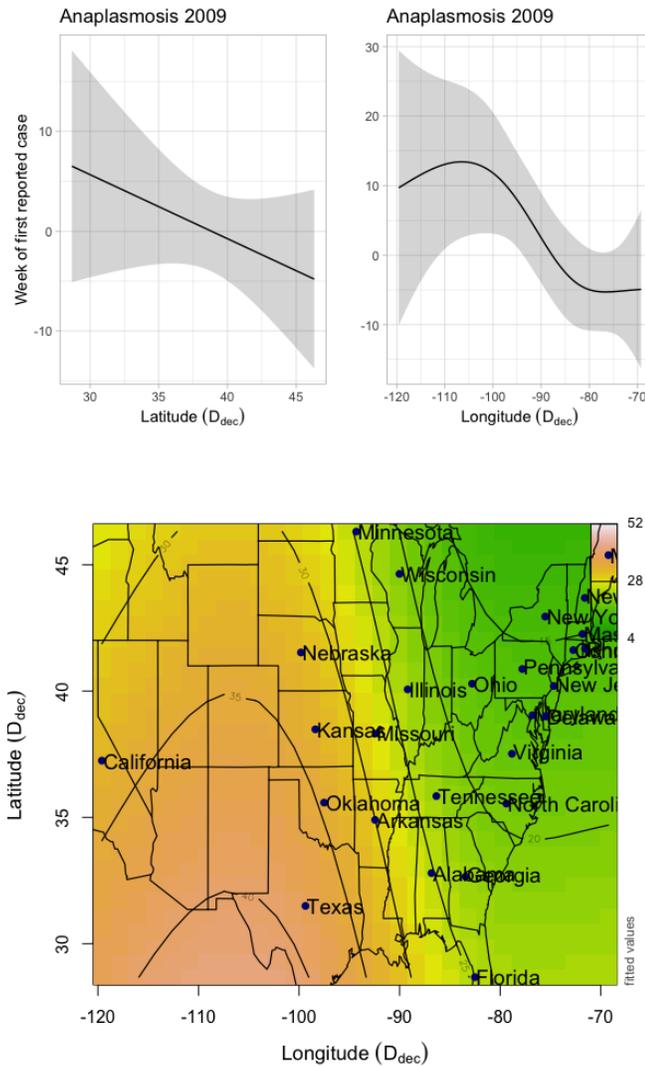


Figure 5.1: The week of first reported cases as a function of the latitude and the longitude (top) and contour plot describing the spatial spread of the 2009 *Anaplasmosis* infections in the U.S.A (bottom). The top graph shows that the disease is travelling at constant speed from North to South (Latitude) and from East to West (Longitude) the speed of disease spread increases rapidly initially and saturates toward the end of the outbreak. The bottom graph shows that the first cases were reported in Northeastern states (Maryland and New Hampshire in green) and the disease spread quickly to the Southern and Western states (Florida, Texas and California in red). The lines of the contour plot are the predicted week of first reported cases. The predicted weeks of first reported cases as a function of the latitude and the longitude are obtained by fitting a generalized additive model to the data (see equation 5.1). The estimated average spatial spread rate c is 1.4 Km/week, and to calculate c we estimated the slope of the curves in (top) using generalized additive model fit to the data, calculated the inverse of the estimated slopes and the sum is c (see section 5.2.2). The data are *Anaplasmosis* infection cases reported in 2009 in different states in the U.S.A by the CDC.

to the Latitude and the Longitude respectively, which are the regression of time on distance for the latitude and Longitude respectively. If the relationship between the time of reported cases Y and the locations (Latitude, Longitude) is linear then equation (5.1) becomes,

$$Y_i = \beta_0 + \beta_1 \text{Latitude}_i + \beta_2 \text{Longitude}_i + \epsilon_i, \quad (5.3)$$

where $\beta_1 = \frac{\partial f}{\partial \text{Latitude}}$ and $\beta_2 = \frac{\partial f}{\partial \text{Longitude}}$ are the slope of the regression of the week of reported case on distance for latitude and longitude respectively (β_1 and β_2 are in weeks per distance). Thus, we can calculate the rate of spatial spread, as the sum of the inverse of the slope of the predictors,

$$c = \frac{1}{\beta_1} + \frac{1}{\beta_2}, \quad (5.4)$$

where c is distance (in kilometres) per week.

5.2.3. Estimation of disease epidemic growth rate r

In the early phase of an outbreak the infected class grows exponentially and the rate of exponential growth at the beginning of the outbreak, which is often denoted r , is the epidemic growth rate (Diekmann et al. 2012b; Heffernan et al. 2005; Ma et al. 2014; Chowell et al. 2016). The exponential growth of the infected class during the early phase of an outbreak can be written as,

$$i(t) = I_0 e^{rt}, \quad (5.5)$$

where $i(t)$ is the number of new infections reported at week t , I_0 is a constant and r is the epidemic growth rate (Roberts and Heesterbeek 2007; Nishiura et al. 2010; 2009b;a; Chowell and Nishiura 2014). For each outbreak data we fit equation (5.5) to

the exponentially growing part of the incidence curve, and we estimated r by non-linear least squares approach (Examples in Figure 5.2). We fit equation (5.5) to the incidence data instead of the cumulative number of cases because the cumulative data is often autocorrelated, and as such violate the independence assumptions of the least squares method (King et al. 2015).

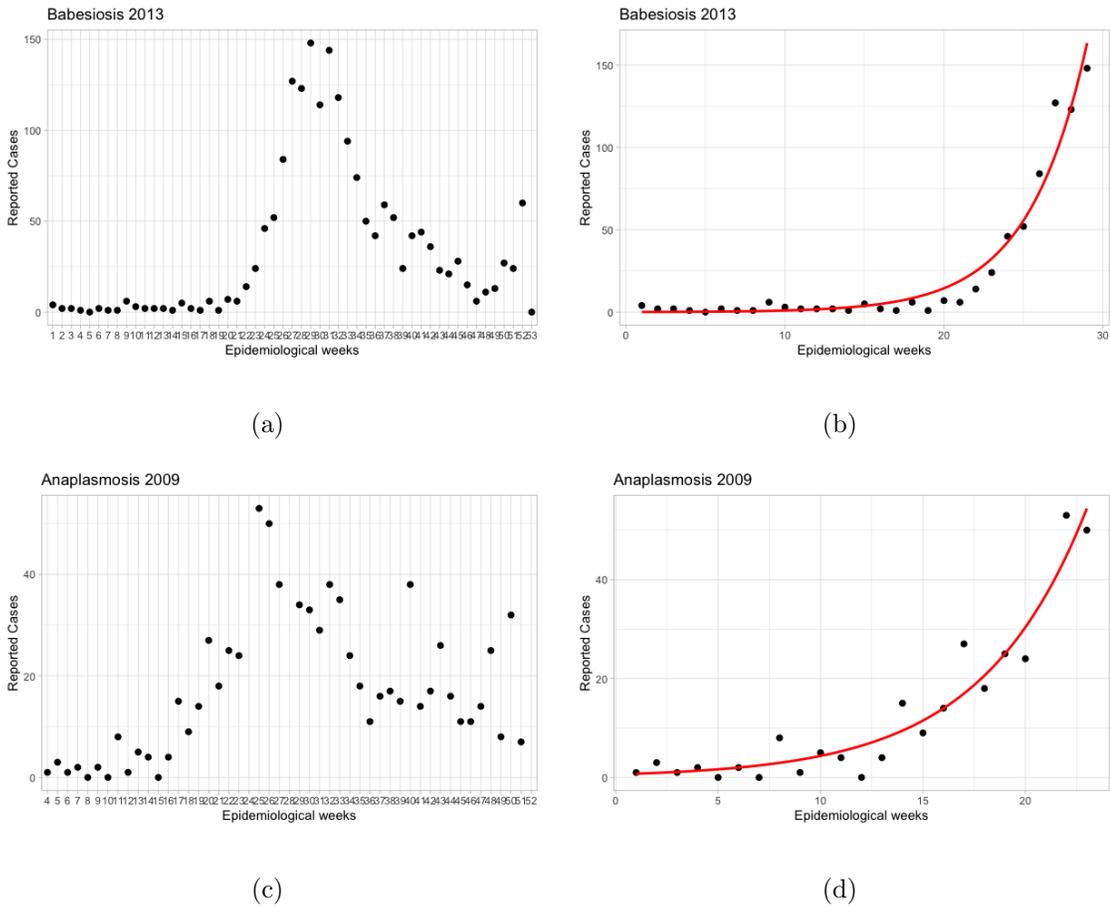


Figure 5.2: The number of cases reported each epidemiological week for *Babesiosis* (a) and *Anaplasmosis* (b), and the non-linear fit of an exponential function to the exponentially growing part of the incidence curve for *Babesiosis* (c) and *Anaplasmosis* (d). For these sample graphs the epidemic growth rate is $r = 0.27$ and 0.42 for (c) and (d) respectively. The data are the 2013 *Babesiosis* and the 2009 *Anaplasmosis* infections reported on the CDC website.

5.2.4. Statistical analysis

In addition to c , r and outbreak year variables, we performed a web search to identify the main transmission route of each infectious disease. We removed the 2016 *Anaplasmosis* outbreak data from some of analysis because it is an influential outlier, it has very high $c = 77$ Km/week for an intermediate $r = 0.2$ relative to the general trend of the relationship between c and r followed by the other observations. Also, we removed the 2011 *Coccidioidomycosis* outbreak data because the incidence data was not exponential and the estimated r is unreliable.

To investigate the relationship between c and r across the years and disease transmission routes we fit a generalized linear model with Gamma distributed response variable (c) to our data because the distribution of c is left-skewed, and as such, a classic linear model is not appropriate. We compared different candidate models using likelihood ratio tests and analysis of deviance, and we selected the models with the lowest deviance, the lowest AIC and the highest likelihood compared to the null model. We selected the link function for our Gamma-generalized linear model by comparing 3 candidate link functions (the inverse, log and identity links) and selecting the link function that captures the general trend of the data. To improve model fit we square root transformed either c or both c and r . We compared c and r in tick-borne and airborne infections only because the other transmission routes have very few observations (sample size is 2 and 1 for direct and waterborne transmissions respectively).

5.3. Results

We tested the predicted positive relationship between the spatial spread rate c and the epidemic growth rate r using human infectious disease outbreaks reported

in the United States of America by the Centers for Disease Control and Prevention (CDC) from 1996 to 2017. We find that infectious diseases that produce more cases per week spread fewer kilometres per week. Also, disease outbreaks caused by tick-borne infections on average have higher epidemic growth rate and spread fewer kilometres per year, whereas disease outbreaks caused by airborne infections spread more kilometres per year and have lower epidemic growth rate.

5.3.1. Epidemic growth rate r and spatial spread rate c are inversely related

We find that r accounts for an important part of the explained variability in c , while the effect of outbreak years is statistically unclear (Total pseudo- $R^2 = 0.31$ and r marginal pseudo- $R^2 = 0.29$). The best model describing c as a function of r and outbreak years does not include interaction terms (Table 5.1). The results show an inverse relationship between c and r which suggests that when the number of cases per week is high, the spatial expansion rate of the infected area is low and the relationship is non-linear (Figure 5.3a). Also, we compared the epidemic growth rate r of 41 infectious disease outbreaks from 1996 to 2017 and find that more recent outbreaks have higher epidemic growth rate (pseudo- $R^2 = 0.17$, and see Figure 5.3b). From 1996 to 2017 the average r of reported infectious diseases in the U.S.A has significantly increased from 0.023 (0.014 0.03, 95 % CI) to 0.16 (0.15 0.18, 95 % CI) cases per week, and this result means that epidemic growth rate r has increased by approximately 85 % in 20 years.

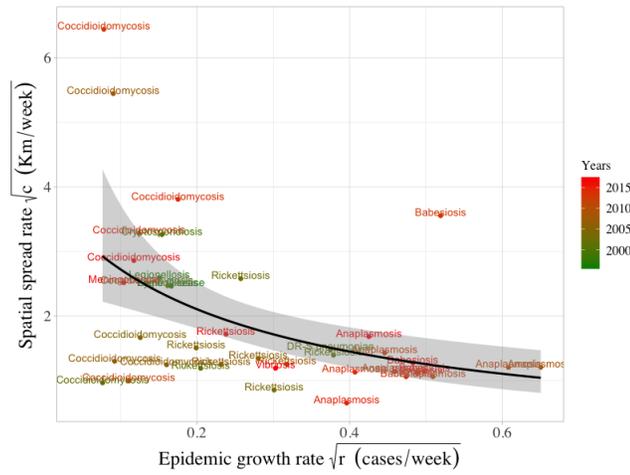
Table 5.1: The model with spatial spread rate c , epidemic growth rate r , and outbreak years T . We fit Gamma-generalized linear models with inverse link function the data. We calculated the ΔAIC , ΔDev ($\Delta\text{Deviance}$), ΔLL ($\Delta\text{Log-likelihood}$) and pseudo- R^2 with respect to the model 7. The pseudo- $R^2 = 1 - \text{Model residual deviance}/\text{deviance of model 7}$. The models are ordered from the best to the worst.

N	Models	K	ΔAIC	ΔDev	ΔLL	Pseudo- R^2
1	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1\sqrt{r} + \beta_2T$	3	11.277	3.660	-7.638	0.31
2	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1\sqrt{r} + \beta_2T + \beta_3\sqrt{r}T$	4	9.286	3.661	-7.642	0.31
3	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1\frac{1}{\sqrt{r}} + \beta_2T + \beta_3\frac{1}{\sqrt{r}}T$	4	7.533	3.305	-6.766	0.28
4	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1\frac{1}{\sqrt{r}} + \beta_2T$	3	9.382	3.273	-6.691	0.27
5	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1r + \beta_2T + \beta_3rT$	4	6.287	3.042	-6.143	0.25
6	$\frac{1}{\sqrt{c}} \sim \beta_0 + \beta_1r + \beta_2T$	3	8.286	3.041	-6.143	0.25
7	$\frac{1}{\sqrt{c}} \sim \beta_0$	1	0	0	0	0

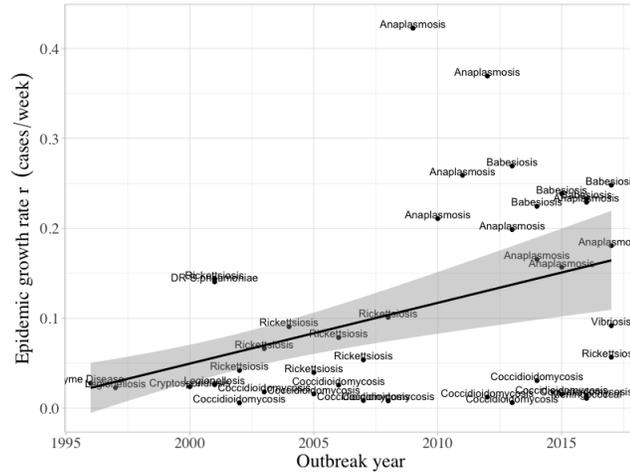
5.3.2. Airborne infections grow slowly and spread faster compared to tick-borne infections

We investigated whether the link between epidemic growth rate r and spatial spread rate c is different in tick-borne and airborne infections, and we find that the difference between the slope of the correlation between c and r in tick-borne and airborne infections is statistically unclear (P-value > 0.5 , pseudo- $R^2 = 0.33$, and see Figures 5.4 and 5.5). As the best model does not include interaction terms (Table 5.2) we compared the spatial spread rate c and the epidemic growth rate r in airborne and tick-borne infections, and we find that on average the spatial spread rate of airborne infections is 1.75 times higher compared to tick-borne infections (P-value < 0.02 , model pseudo- $R^2 = 0.32$, see Figure 5.5a).

Moreover, we compared the epidemic growth rate r from 1996 to 2017 in tick-borne and airborne infections, and we find that more recent outbreaks have higher r in tick-borne infections, but in airborne infections the correlation between r and outbreak years is statistically unclear (Figure 5.4b, and see Table A5.1). From 1996



(a)



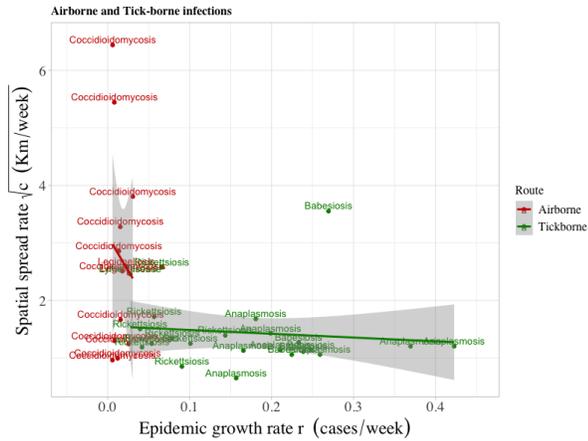
(b)

Figure 5.3: The spatial spread rate (c) and the initial epidemic growth rate (r) are inversely related (a), and more recent outbreaks have higher r (b). We estimated c and r of 42 diseases outbreaks (10 different infectious diseases) that occurred between 1996 and 2017 in the United States of America, excluding U.S. territories and publicly available on the CDC website. We fit Gamma Generalized Linear Models with inverse (a) and log links (b) to the estimated c and r data. The dots are the estimated c and r measures for an outbreak, the colours are the years the outbreaks occurred, the black line through the data is the best fit model, and the grey area is the 95 % confidence interval. For all graphs the correlations are significant (p-values < 0.05), and pseudo- R^2 is 0.30, and 0.17 in (a) and (b) respectively. In (a) the residuals are independent ($\rho = -0.09$, and p-value > 0.5), whereas in (b) the residuals are correlated ($\rho = 0.56$, and p-value < 0.0001).

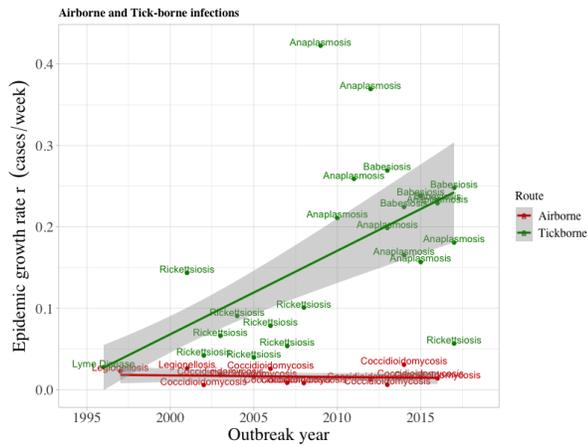
Table 5.2: Spatial spread rate c and epidemic growth rate r in tick-borne and airborne infections, where R is disease transmission route variable. We calculated the ΔAIC , ΔDev ($\Delta\text{Deviance}$), ΔLL ($\Delta\text{Log-likelihood}$) and pseudo- R^2 with respect to the model 3. The pseudo- $R^2 = 1 - \text{Model residual deviance}/\text{deviance of model 3}$. The models are ordered from the best to the worst.

N	Models	K	ΔAIC	ΔDev	ΔLL	Pseudo- R^2
1	$\sqrt{c} \sim \beta_0 + \beta_1 r + \beta_2 R$	3	10.592	3.620	-7.296	0.32
2	$\sqrt{c} \sim \beta_0 + \beta_1 r + \beta_2 R + \beta_3 r R$	4	9.328	3.770	-7.664	0.33
3	$\sqrt{c} \sim \beta_0$	1	0	0	0	0

to 2017 the average r of reported tick-borne infections has increased from 0.027 (0.017 0.037, 95 % CI) to 0.24 (0.22 0.26, 95 % CI) cases per week in 2017 which means that the average epidemic growth rate of reported infectious diseases has increased by 88 % in 20 years. Finally, we find that on average the epidemic growth rate r is 9 times higher in tick-borne than airborne infections, and this result suggests that disease outbreaks produced by tick-borne infections grow more faster compared to airborne infections ($r = 0.019$ (0.015 0.024, 95 % CI) and $r = 0.17$ (0.13 0.21, 95 % CI) in airborne and tick-borne infection respectively).

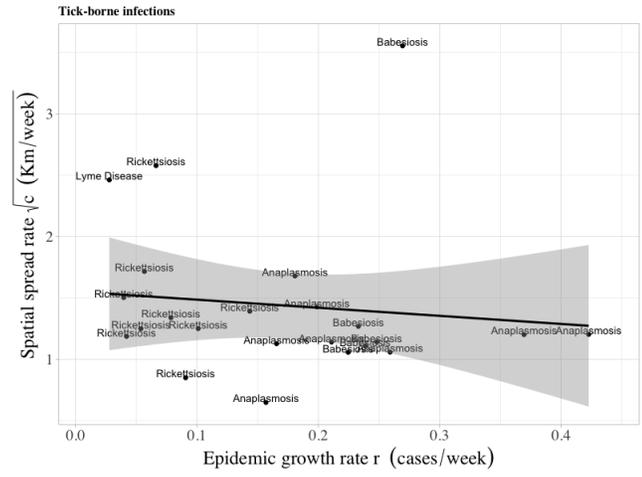


(a)

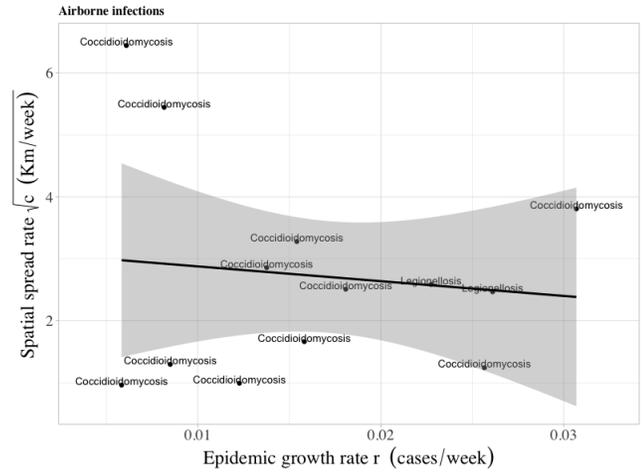


(b)

Figure 5.4: Airborne infections have higher c and lower r , and tick-borne infections have lower c and higher r (a), and more recent outbreaks have higher r in tick-borne infections (b). We fit Gamma Generalized Linear Models with identity links to the estimated c and r data (23 tick-borne and 13 airborne infection outbreaks). The dots are the estimated c and r measures for an outbreak, the line through the data is the best fit model, the colours are the different transmission routes (red and green for airborne and tick-borne infections respectively) and the grey area is the 95 % confidence interval. In (a) c is significantly higher in airborne compared to tick-borne (p-value < 0.01 and pseudo- $R^2 = 0.32$), but the correlation between c and r is not significant for tick-borne and airborne infections (p-value > 0.3 and pseudo- $R^2 = 0.33$). In (b), r for tick-borne and airborne infections are statistically different (p-value < 0.05), the correlation between r and outbreak year is significant in tick-borne (p-value < 0.0001) and statistically unclear in airborne infections (p-value > 0.3), and model pseudo- $R^2 = 0.79$. For both graphs the residuals are independent (p-value < 0.05), the correlation coefficients are $\rho = -0.15$ and $\rho = 0.13$ for (a) and (b) respectively.



(a)



(b)

Figure 5.5: The relationship between c and r is statistically unclear for tick-borne (a) and airborne infections (b). We estimated c and r for 23 tick-borne and 13 airborne infections that occurred between 1996 and 2017 in the United States of America, excluding U.S. territories and publicly available on the CDC website. We fit Gamma Generalized Linear Models with identity links to the estimated c and r data. The dots are the estimated c and r measures for an outbreak, the line through the data is the best fit model and the grey area is the 95 % confidence interval. In (a) the estimated slope $\rho = -0.6$, p-value 0.6, and pseudo- $R^2 = 0.01$, and in (b) the estimated slope $\rho = -23.84$, p-value is 0.7, and pseudo- $R^2 = 0.01$.

5.4. Discussion

A positive relationship between the spatial spread rate c and the epidemic growth rate r has been derived using spatial epidemic models, but this prediction has yet to be validated with empirical data. In this paper, we estimated c (kilometres per week) and r (cases per week) for 42 outbreaks caused by 10 human infectious diseases with different transmission routes that have been reported in the United States of America by the Centers for Disease Control and Prevention (CDC) from 1996 to 2017, and we investigated whether infectious disease outbreaks that produce more cases per week spread more kilometres per week. We find an inverse relationship between c and r , which suggests that infectious disease outbreaks where the number of cases are rapidly increasing over time spread fewer kilometres per week. Also, we find that disease outbreaks caused by tick-borne infections produce on average more cases per week and spread fewer kilometres per week, whereas disease outbreaks caused by airborne infections produce on average fewer cases per week and spread more kilometres per week.

5.4.1. *The inverse relationship between c and r can reflect a link between the spatial dispersal and the transmission of parasites*

The inverse relationship between spatial spread rate c and epidemic growth rate r is unexpected because spatial epidemic models suggest that when more cases are produced there are more infected individuals who can potentially disperse long distances and the spatial region covered by infected hosts is larger, and as such the spatial spread rate will be higher (Källén et al. 1985; Murray et al. 2001; Diekmann et al. 2012c). An inverse relationship between c and r can reflect a relationship between parasite transmission and the dispersal of infectious organisms (Figures 5.3a

and 5.4a).

The density of infected humans do not contribute to the growth of the epidemic in the majority of the infectious diseases that we have investigated because humans are dead-end hosts for these parasites, thus human-to-human transmission does not occur. One explanation for why some infectious disease outbreaks generate more new cases per week is that the likelihood of infection is high due to higher density of the infectious organisms. In theory, the infectious disease outbreaks that generate more infections per week will spread more kilometres per week, however, we observe the opposite which can be explained by a possible trade-off between dispersal and disease parasite transmission. When more infections occur per week the spatial spread rate of the disease c is higher because there are more infected hosts that can disperse long distances, but higher spread rate of infectious diseases can also be achieved by particular transmission routes: wind, water or animal. As such, some infectious disease outbreaks spread more kilometres per week because the infectious organisms can disperse long distances by wind or water which can result in larger area covered by the disease. For example, *Coccidioides* fungi spread more kilometres per week because the infectious aerosols which contribute directly to the growth of the epidemic can disperse long distances by wind.

The positive relationship between c and r predicted by spatial epidemic models may hold because the new cases generated per week can disperse longer distances and directly contribute to the growth of the epidemic by infecting more susceptible hosts. One possible explanation for why the positive relationship between c and r does not hold in airborne and tick-borne infections is that new cases that are generated each week do not generate more cases and contribute directly to the growth of the epidemic

and the main factor driving the spatial propagation of the disease is the dispersal of the infectious organisms by wind, water and animals. An alternative explanation is that long distance dispersal is associated with higher parasite mortality or limited access to susceptible hosts, which results in fewer cases per week. Infectious organisms that disperse long distances by air or water have a higher risk of dying or dispersing away from susceptible hosts, whereas infectious organisms that disperse locally, by vectors for example, have a stable access to susceptible hosts and the risk of death due to long distance dispersal is lower. As such, infectious organisms that spread locally produce more cases per week compared to infectious organisms that disperse long distances.

5.4.2. Disease spread rate c and epidemic growth rate r in tick-borne and airborne infections

Our analysis suggests that on average tick-borne infections spread fewer kilometres per week than airborne infections, and infectious disease outbreaks caused by tick-borne infections have higher epidemic growth rates compared to airborne infections (Figure 5.4a). Tick-borne infections have lower spatial spread rates because ticks, which are the vectors of the diseases, are absent from many states and rely on migratory birds for long distance dispersal (Gabriel et al. 2009; Rikihisa 2011). For example the main vectors of Lyme disease, *Ixodes scapularis* and *I. pacificus*, are present in the eastern and western states in the U.S.A respectively (Dennis et al. 1998; Eisen et al. 2016; Owen et al. 2019), disperse locally few kilometres per day attached to rodents and deers and occasionally more than 250 kilometres per day attached to migratory birds (Ogden et al. 2008; Leighton et al. 2012; Tonelli and Dearborn 2019). The geographic range within which infectious contacts and infections can occur is relatively smaller for tick-borne infections due to their limited spatial distribution and move-

ment. As such the reported tick-borne cases are within the small geographic range of the ticks, and over a period of time the disease spread kilometres per week.

In contrast, airborne infectious organisms have higher spatial spread rates because the infectious aerosols can disperse longer distances by wind. For example the *Coccidioides* which are the fungi that cause Coccidioidomycosis are endemic to the southwestern United States (California, Arizona, New Mexico and Texas States), and some studies found that the geographic range of *Coccidioides* species is expanding to Utah and Washington states (Brown et al. 2013; Litvintseva et al. 2014; Johnson et al. 2014; Engelthaler et al. 2016; Barker et al. 2019). Cases of Coccidioidomycosis are reported from states that are far from the endemic regions because the *Coccidioides* spores can disperse longer distances by wind. Thus, the area covered by humans infected with *Coccidioides* is larger and the estimated spatial spread rate is higher. In summary, tick-borne infections spread fewer kilometres per week than airborne infections because the infected ticks that transmit the infections disperse shorter distances attached to their hosts, whereas the infectious aerosols that cause airborne infections can disperse longer distances by wind. Similar ideas have been proposed to explain why tick-borne infections have lower spatial spread rates c compared to mosquito-borne infections (Leiby 2019).

Moreover, infectious disease outbreaks caused by tick-borne infections produce on average more infections per week than airborne infectious disease outbreak (Figure 5.4a). One possible explanation for this result is that more infectious contacts per week occur between humans and ticks than humans and the infectious aerosols that cause airborne infections. Either the overall density of infected ticks is larger than the density of the infected aerosols or some human activities like hunting and walking in the forest

increases the probability of infectious contacts between infected ticks and humans, and as such more cases per week due to tick-borne infections are reported than cases due to airborne infections. However, this result must be taken with caution because the incidence curve of most of our airborne infections do not exhibit a well defined exponential growth profile, and as such the estimation of r for airborne infections may be unreliable.

Finally, recent outbreaks have higher epidemic growth rates r , and this result is supported in tick-borne infections and unclear in airborne infections (Figure 5.3b and 5.4b). It is clear that the observed increased epidemic growth rate during the past 20 years is driven by tick-borne infections. A potential explanation for increased epidemic growth rate in tick-borne infections is that the density and/or the geographic range of ticks has increased during the past 20 years. As reported by previous studies, the geographic range of the ticks has increased during the past 50 years due the expansion of suitable habitat like forests and climate change has been reported as important factor (Barbour and Fish 1993; Brownstein et al. 2003; Eisen and Eisen 2018; VanAcker et al. 2019). Higher ticks population density will increase the probability of infectious contacts between ticks and humans and the incidence of tick-borne infections, and as such, more recent outbreaks due to tick-borne infections will produce more infections per week.

5.5. Concluding remarks

Spatial epidemic models predict a positive relationship between the epidemic growth rate r and the spatial spread rate c , and the empirical clarification of the link between c and r can help in predicting how far an infectious disease will propagate. In this paper, we test whether infectious disease outbreaks where the number of

cases are rapidly increasing over time spread more kilometres per week using human infectious disease cases reported in the U.S.A by the CDC during the past 20 years. In contrast to the theoretical predictions we found an inverse relationship between c and r . Our work is a step forward in clarifying the link between c and r , and we formulate three recommendations for future works. Firstly, to achieve the goal of predicting the spatial spread rate c early during an outbreak we need to clarify the contribution of disease transmission routes to the epidemic growth rate r and the spatial spread rate c , which can clarify whether the predicted positive relationship between c and r holds across transmission routes. Secondly, future works will clarify the impact of host movement on the spatial spread rate of infectious diseases. In simple spatial epidemic models hosts movement affect the rate of spatial spread via the diffusion rate, but in complex models the effect of hosts movement is unclear because other movement parameters (e.g. the advection rate, the coefficient of correlation between movement directions) come into play. Thirdly, the data for investigating the link between the epidemic growth rate and the spatial spread rate is available (not public sometimes), but the challenge is to organize the outbreak data in a format that can be used directly in statistical softwares. As such, more work is needed to get open, high spatial resolution and useful data to achieve the goal of predicting how far an infectious disease will propagate using early case data.

Literature cited

K.M. Alanazi, Z. Jackiewicz, and H.R. Thieme. Numerical simulations of spread of rabies in a spatially distributed fox population. *Mathematics and Computers in Simulation*, 159: 161–182, 2019.

- L.J. Allen, F. Brauer, P. Van den Driessche, and J. Wu. *Mathematical epidemiology*, volume 1945. Springer, 2008.
- R.M. Anderson, B. Anderson, and R.M. May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.
- D.A. Andow, P.M. Kareiva, S.A. Levin, and A. Okubo. Spread of invading organisms. *Landscape Ecology*, 4(2):177–188, 1990.
- A.G. Barbour and D. Fish. The biological and social phenomenon of Lyme disease. *Science*, 260(5114):1610–1616, 1993.
- B.M. Barker, A.P. Litvintseva, M. Riquelme, and L. Vargas-Gastélum. Coccidioides ecology and genomics. *Medical mycology*, 57(Supplement_1):S21–S29, 2019.
- J. Brown, K. Benedict, B.J. Park, and G.R. Thompson III. Coccidioidomycosis: epidemiology. *Clinical epidemiology*, 5:185, 2013.
- J.S. Brownstein, T.R. Holford, and D. Fish. A climate-based model predicts the spatial distribution of the Lyme disease vector *Ixodes scapularis* in the United States. *Environmental health perspectives*, 111(9):1152–1157, 2003.
- L.M. Bugayevskiy and J. Snyder. *Map projections: A reference manual*. CRC Press, 2013.
- G. Chowell and H. Nishiura. Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC medicine*, 12(1):196, 2014.
- G. Chowell, L. Sattenspiel, S. Bansal, and C. Viboud. Mathematical models to characterize early epidemic growth: A review. *Physics of life reviews*, 18:66–97, 2016.
- P.L. Delamater, E.J. Street, T.F. Leslie, Y.T. Yang, and K.H. Jacobsen. Complexity of the basic reproduction number (R_0). *Emerging infectious diseases*, 25(1):1, 2019.
- D.T. Dennis, T.S. Nekomoto, J.C. Victor, W.S. Paul, and J. Piesman. Reported distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the United States. *Journal of medical entomology*, 35(5):629–638, 1998.

- O. Diekmann, H. Heesterbeek, and T. Britton. *Mathematical tools for understanding infectious disease dynamics*, volume 7. Princeton University Press, 2012a.
- O. Diekmann, H. Heesterbeek, and T. Britton. The epidemic in a closed population. In *Mathematical tools for understanding infectious diseases dynamics*, chapter 1, pages 3–36. Princeton University Press, 2012b.
- O. Diekmann, H. Heesterbeek, and T. Britton. Spatial spread. In *Mathematical tools for understanding infectious disease dynamics*, chapter 10, pages 275–288. Princeton University Press, 2012c.
- R.J. Eisen and L. Eisen. The blacklegged tick, *Ixodes scapularis*: an increasing public health concern. *Trends in parasitology*, 34(4):295–309, 2018.
- R.J. Eisen, L. Eisen, and C.B. Beard. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United states. *Journal of Medical Entomology*, 53(2):349–386, 2016.
- D.M. Engelthaler, C.C. Roe, C.M. Hepp, M. Teixeira, E.M. Driebe, J.M. Schupp, L. Gade, V. Waddell, K. Komatsu, E. Arathoon, et al. Local population structure and patterns of western hemisphere dispersal for *Coccidioides* spp., the fungal cause of valley fever. *MBio*, 7(2):e00550–16, 2016.
- N.D. Evans and A.J. Pritchard. A control theoretic approach to containing the spread of rabies. *Mathematical Medicine and Biology: A Journal of the IMA*, 18(1):1–23, 2001.
- M.L. Farnsworth and M.P. Ward. Identifying spatio-temporal patterns of transboundary disease spread: examples using avian influenza H5N1 outbreaks. *Veterinary research*, 40(3):1–14, 2009.
- N.M. Ferguson, C.A. Donnelly, and R.M. Anderson. The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science*, 292(5519):1155–1160, 2001a.

- N.M. Ferguson, C.A. Donnelly, and R.M. Anderson. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature*, 413(6855):542, 2001b.
- A.M. Fofana and A. Hurford. Mechanistic movement models to understand epidemic spread. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719):20160086, 2017.
- M.W. Gabriel, R.N. Brown, J.E. Foley, J.M. Higley, and R.G. Botzler. Ecology of *Anaplasma phagocytophilum* infection in gray foxes (*Urocyon cinereoargenteus*) in northwestern California. *Journal of wildlife diseases*, 45(2):344–354, 2009.
- J. Goldstein, J. Park, M. Haran, A. Liebhold, and O.N. Bjørnstad. Quantifying spatio-temporal variation of invasion spread. *Proceedings of the Royal Society B*, 286(1894):20182294, 2019.
- T. Hastie and Tibshirani R. Generalized additive models. *Statistical Science*, 1(3):297–318, 1986.
- A. Hastings, K. Cuddington, K.F. Davies, C.J. Dugaw, S. Elmendorf, A. Freestone, S. Harrison, M. Holland, J. Lambrinos, U. Malvadkar, et al. The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, 8(1):91–101, 2005.
- D.T. Haydon, R.R. Kao, and R.P. Kitching. The UK foot-and-mouth disease outbreak—the aftermath. *Nature Reviews Microbiology*, 2(8):675, 2004.
- J.A.P Heesterbeek and K. Dietz. The concept of R_0 in epidemic theory. *Statistica neerlandica*, 50(1):89–110, 1996.
- J.M. Heffernan, R.J. Smith, and L.M. Wahl. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4):281–293, 2005.
- E.E. Holmes, M.A. Lewis, J.E. Banks, and R.R. Veit. Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology*, 75(1):17–29, 1994.

- Amy Hurford, Mark Hebblewhite, and Mark A Lewis. A spatially explicit model for an Allee effect: why wolves recolonize so slowly in Greater Yellowstone. *Theoretical population biology*, 70(3):244–254, 2006.
- S.M. Johnson, E.L. Carlson, F.S. Fisher, and D. Pappagianis. Demonstration of *Coccidioides immitis* and *Coccidioides posadasii* DNA in soil samples collected from Dinosaur National Monument, Utah. *Medical Mycology*, 52(6):610–617, 2014.
- A. Källén, P. Arcuri, and J.D. Murray. A simple model for the spatial spread and control of rabies. *Journal of theoretical biology*, 116(3):377–393, 1985.
- R.R. Kao. The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. *Trends in microbiology*, 10(6):279–286, 2002.
- M.J. Keeling, M.E. Woolhouse, D.J. Shaw, L. Matthews, M. Chase-Topping, D.T. Haydon, S.J. Cornell, J. Kappey, J. Wilesmith, and B.T. Grenfell. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, 294(5543):813–817, 2001.
- A.A. King, M.D. de Celles, F. Magpantay, and P. Rohani. Avoidable errors in modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. B*, 282(1806):20150347, 2015.
- M. Kot, M.A. Lewis, and P. van den Driessche. Dispersal data and the spread of invading organisms. *Ecology*, 77(7):2027–2042, 1996.
- D.A. Leiby. Tick-borne infections: Beware the tortoises among us. In *Blood Safety*, pages 207–221. Springer, 2019.
- P.A. Leighton, J.K. Koffi, Y. Pelcat, L.R. Lindsay, and N.H. Ogden. Predicting the speed of tick invasion: an empirical model of range expansion for the Lyme disease vector *Ixodes scapularis* in Canada. *Journal of Applied Ecology*, 49(2):457–464, 2012.
- J. Li and D. Blakeley. The failure of R_0 . *Computational and mathematical methods in*

- medicine*, 2011, 2011.
- M. Lipsitch, T. Cohen, B. Cooper, J.M. Robins, S. Ma, L. James, G. Gopalakrishna, S.K. Chew, C.C. Tan, M.H. Samore, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970, 2003.
- A.P. Litvintseva, N. Marsden-Haug, S. Hurst, H. Hill, L. Gade, E.M. Driebe, C. Ralston, C. Roe, B.M. Barker, M. Goldoft, et al. Valley fever: finding new places for an old disease: *Coccidioides immitis* found in Washington State soil associated with recent human infection. *Clinical Infectious Diseases*, 60(1):e1–e3, 2014.
- E. Lizarazo, M. Vincenti-Gonzalez, M.E. Grillet, S. Bethencourt, O. Diaz, N. Ojeda, H. Ochoa, M. A. Rangel, and A. Tami. Spatial dynamics of Chikungunya virus, Venezuela, 2014. *Emerging infectious diseases*, 25(4):672, 2019.
- J.A. Lubina and S.A. Levin. The spread of a reinvading species: range expansion in the California sea otter. *The American Naturalist*, 131(4):526–543, 1988.
- J. Ma, J. Dushoff, B.M. Bolker, and D.J. Earn. Estimating initial epidemic growth rates. *Bulletin of mathematical biology*, 76(1):245–260, 2014.
- G. Marra and S.N. Wood. Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, 39(1):53–74, 2012.
- H. McCallum, D. Harvell, and A. Dobson. Rates of spread of marine pathogens. *Ecology Letters*, 6(12):1062–1067, 2003.
- A. Mercier, E. Arsevska, L. Bournez, A. Bronner, D. Calavas, J. Cauchard, S. Falala, P. Caufour, C. Tisseuil, T. Lefrançois, et al. Spread rate of lumpy skin disease in the Balkans, 2015–2016. *Transboundary and emerging diseases*, 65(1):240–243, 2018.
- D.A. Moore. Spatial diffusion of raccoon rabies in Pennsylvania, USA. *Preventive veterinary medicine*, 40(1):19–32, 1999.
- A.G. Murray, M. O’Callaghan, and B. Jones. Simple models of massive epidemics of her-

- pesvirus in Australian (and New Zealand) pilchards. *Environment international*, 27(2-3): 243–248, 2001.
- J.D. Murray. *Mathematical biology. Second corrected edition*. Springer-Verlag, Berlin, 1993.
- J.D. Murray, E.A. Stanley, and D.L. Brown. On the spatial spread of rabies among foxes. *Proceedings of the Royal society of London. Series B. Biological sciences*, 229(1255):111–150, 1986.
- M.G. Neubert, M. Kot, and M.A. Lewis. Invasion speeds in fluctuating environments. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1453): 1603–1610, 2000.
- H. Nishiura. Correcting the actual reproduction number: a simple method to estimate R_0 from early epidemic growth data. *International journal of environmental research and public health*, 7(1):291–302, 2010.
- H. Nishiura, C. Castillo-Chavez, M. Safan, and G. Chowell. Transmission potential of the new influenza A (H1N1) virus and its age-specificity in Japan. *Eurosurveillance*, 14(22): 19227, 2009a.
- H. Nishiura, N. Wilson, and M.G. Backer. Estimating the reproduction number of the novel influenza A virus (H1N1) in a Southern Hemisphere setting: preliminary estimate in New Zealand. *The New Zealand Medical Journal (Online)*, 122(1299), 2009b.
- H. Nishiura, G. Chowell, M. Safan, and C. Castillo-Chavez. Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theoretical Biology and Medical Modelling*, 7(1):1, 2010.
- N.H. Ogden, L.R. Lindsay, K. Hanincová, I.K. Barker, M. Bigras-Poulin, D.F. Charron, A. Heagy, C.M. Francis, C.J. O’Callaghan, I. Schwartz, et al. Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. *Appl. Environ. Microbiol.*, 74(6):1780–1790,

- 2008.
- A. Okubo and S.A. Levin. *Diffusion and ecological problems: modern perspectives*, volume 14. Springer Science & Business Media, 2013.
- E.E. Osnas, P.J. Hurtado, and A.P. Dobson. Evolution of pathogen virulence across space during an epidemic. *The American Naturalist*, 185(3):332–342, 2015.
- H. Owen, S. Lisowski, C. Schaefer, T. Yao, J. Allen, N. Goetz, J. Hernandez, R. Kreisler, J.K. Lee, M. Quinlan, et al. Variation in the geographic distribution and Rickettsial infection rates of *Rhipicephalus sanguineus* contributes to the spread of RMSF in Arizona and Mexico. *The FASEB Journal*, 33(1-supplement):662–48, 2019.
- S.W. Park, D. Champredon, J.S. Weitz, and J. Dushoff. A practical generation-interval-based approach to inferring the strength of epidemics from their speed. *Epidemics*, 27:12–18, 2019.
- M. Pioz, H. Guis, D. Calavas, B. Durand, D. Abrial, and C. Ducrot. Estimating front-wave velocity of infectious diseases: a simple, efficient method applied to bluetongue. *Veterinary research*, 42(1):60, 2011.
- Y. Rikihisa. Mechanisms of obligatory intracellular infection with *Anaplasma phagocytophilum*. *Clinical microbiology reviews*, 24(3):469–489, 2011.
- M.G. Roberts and J.A.P. Heesterbeek. Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. *Journal of mathematical biology*, 55(5-6):803, 2007.
- K. Shea, E. Jongejans, O. Skarpaas, D. Kelly, and A.W. Sheppard. Optimal management strategies to control local population growth or population spread may not be the same. *Ecological applications*, 20(4):1148–1161, 2010.
- J.G. Skellam. Random dispersal in theoretical populations. *Biometrika*, 38(1/2):196–218, 1951.

- D.L. Smith, B. Lucey, L.A. Waller, J.E. Childs, and Leslie A. Real. Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences*, 99(6):3668–3672, 2002.
- H.R. Thieme. Some mathematical considerations of how to stop the spatial spread of a rabies epidemic. In *Biological Growth and Spread*, pages 310–319. Springer, 1980.
- C. Tisseuil, A. Gryspeirt, R. Lancelot, M. Pioz, A. Liebhold, and M. Gilbert. Evaluating methods to quantify spatial variation in the velocity of biological invasions. *Ecography*, 39(5):409–418, 2016.
- B.A. Tonelli and D.C. Dearborn. An individual-based model for the dispersal of *Ixodes scapularis* by ovenbirds and wood thrushes during fall migration. *Ticks and tick-borne diseases*, 10(5):1096–1104, 2019.
- F. Van den Bosch, J.A.J Metz, and O. Diekmann. The velocity of spatial population expansion. *Journal of Mathematical Biology*, 28(5):529–565, 1990.
- M.C. VanAcker, E.A. Little, G. Molaei, W.I. Bajwa, and M.A. Diuk-Wasser. Enhancement of risk for Lyme disease by landscape connectivity, New York, New York, USA. *Emerging infectious diseases*, 25(6):1136, 2019.
- M.H Wang and M. Kot. Speeds of invasion in a model with strong or weak Allee effects. *Mathematical biosciences*, 171(1):83–97, 2001.
- Simon N Wood. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC, 2017.
- S.N. Wood. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467):673–686, 2004.
- S.N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(1):3–36, 2011.

- S.N. Wood. On p-values for smooth components of an extended generalized additive model. *Biometrika*, 100(1):221–228, 2012.
- S.N. Wood. A simple test for random effects in regression models. *Biometrika*, 100(4):1005–1010, 2013.
- S.N. Wood, N. Pya, and B. Säfken. Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, 111(516):1548–1563, 2016.
- M. Woolhouse, M. Chase-Topping, D. Haydon, J. Friar, L. Matthews, G. Hughes, D. Shaw, J. Wilesmith, A. Donaldson, S. Cornell, et al. Foot-and-mouth disease under control in the UK. *Nature*, 411(6835):258, 2001.

6. Summary

The adaptive virulence evolution hypothesis is widely accepted as an explanation for why some parasites kill their hosts, but has been recently criticized for its limited applicability and the lack of spatial aspects of disease spread in the theoretical framework. This thesis is an attempt to spatially extend the mathematical framework of the adaptive virulence hypothesis by explicitly accounting for why and how hosts move towards food, conspecifics and mates and address the broad applicability problems.

In summary, I show that the empirical investigation of the adaptive virulence hypothesis at the cross-species level is feasible with parasite species that are ecologically similar. Our simulation data show that virulence can be adaptive at the species level, and as such, the adaptive virulence hypothesis may be broadly applicable and can be invoked to explain why some parasite species cause higher virulence than others (Figures 2.5 and 2.6). Also, I show that evolutionary bistable virulence can emerge when parasites induce lethargy and death to their hosts. I explain how parasites with low- and high-virulence can be maintained in transient coexistence in host populations (Figure 3.3). Moreover, I reviewed epidemic models that account for host movement to investigate how different types of host movement affect epidemic spread, I find that the main aspect that require further focus is how spatially heterogenous epidemiological parameters (e.g., parasite transmission rate) can emerge from the spatial structure of host population or other spatiotemporal processes (Table 4.2). Finally, I investigated whether the relationship between the growth rate r and the spread rate c of infectious diseases is positive, I find an inverse relationship between c and r and discuss how this result can reflect a trade-off between parasite dispersal and transmission (Figure

5.3).

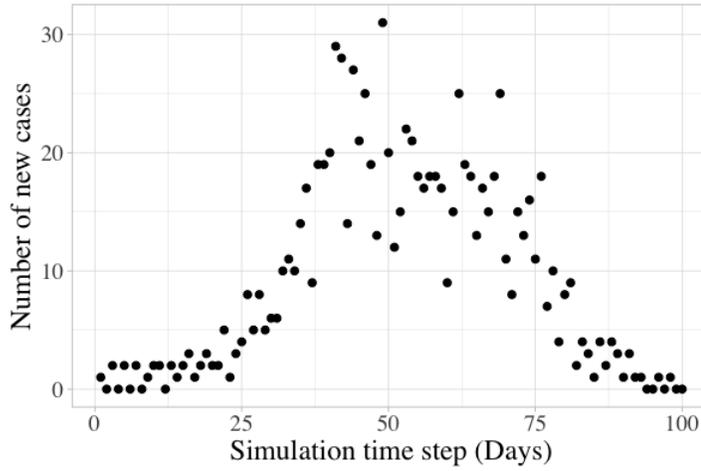
Diffusion models to describe epidemic spread can effectively capture the transmission and the spatial spread of infectious diseases, but the derived basic reproduction number R_0 , which is a measure of parasite fitness, is often independent of host movement parameters. In chapter three, I show that for spatial epidemic models with spatially and/or temporally heterogeneous epidemiological parameters (e.g., disease transmission rate) R_0 and host movement are linked via the diffusion coefficient of the hosts. As such, spatial epidemic models with spatially heterogeneous disease transmission rates can be appropriate for the spatial extension of the adaptive virulence evolution hypothesis. However, the mathematical analysis of spatial epidemic models with spatially heterogeneous disease transmission rate can be difficult and only very few studies have derived an explicit expression of R_0 for such models. Future works will investigate how spatially heterogeneous disease transmission rate can emerge from the spatiotemporal dynamics of host-parasite interactions, formulate the results as trade-offs between parasite traits and spatial aspects of disease spread like host movement and explicitly incorporate those trade-offs directly in the mathematical framework of the adaptive virulence hypothesis. Moreover, the spatial spread rate c , which is the speed of spatial propagation of infectious diseases, can be an important component of parasite fitness because a parasite can achieve higher transmission by spreading globally in the host population. As such, the empirical investigation of the relationship between the temporal and the spatial spread of infectious diseases is critical for understanding the contribution of c to the lifetime transmission success of parasites and the implications for the evolution of virulence.

Appendix A.

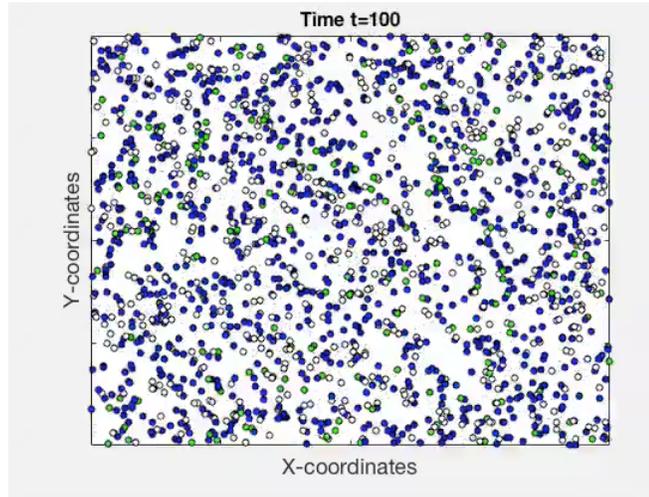
Chapter one: Is virulence adaptive? A numerical investigation with cross-species disease outbreak data

Parasite species and Disease outbreak simulation

We generated numerically a total of 1500 parasites, 50 species and 30 strains within each species, that are phylogenetically related and each parasite has an incubation period (σ_d), the time before the host recovers from an infection (γ_d), the time before the host dies due to infection (ν_d), the probability of disease transmission given a contact (β) and the infected host movement step length reduction factor (s). We generated outbreaks from the evolved epidemiological parameters in host populations where the size of the population is selected randomly. The epidemic model is spatially explicit, stochastic, Susceptible-Exposed-Infectious-Removed (SEIR) disease progression and with random walk host movement (Figures A1 and A2).



(a)



(b)

Figure A.1: Disease outbreak caused by a parasite in a population where hosts move randomly in two-dimensional spatial domain with periodic boundaries, we graph the epidemic curve in (a) and the movie in (b) illustrates disease spread in the population (Video is available at Figshare doi:10.6084/m9.figshare.11392626). Parasite incubation period, the time from infection to host death, the time from infection to host recovery are gamma-distributed with mean $\sigma_d = 1.5$, $\gamma_d = 4.3$ and $\nu_d = 3.6$ days respectively. The probability of parasite transmission given a contact is $p = 0.6$ and the movement step length reduction factor is $s = 0.3$ and host population size $N = 1869$ individuals. The epidemic model is a spatially explicit stochastic Susceptible-Exposed-Infectious-Removed (SEIR) model, each dot is a host individual, it's location at a given time and epidemiological status. Susceptible hosts are blue, exposed and infectious hosts are red, recovered hosts are green and dead hosts are empty dots. The total number of secondary cases by the primary case is $R_0 = 4$.

The relationship between virulence and parasite fitness measures: within-species data

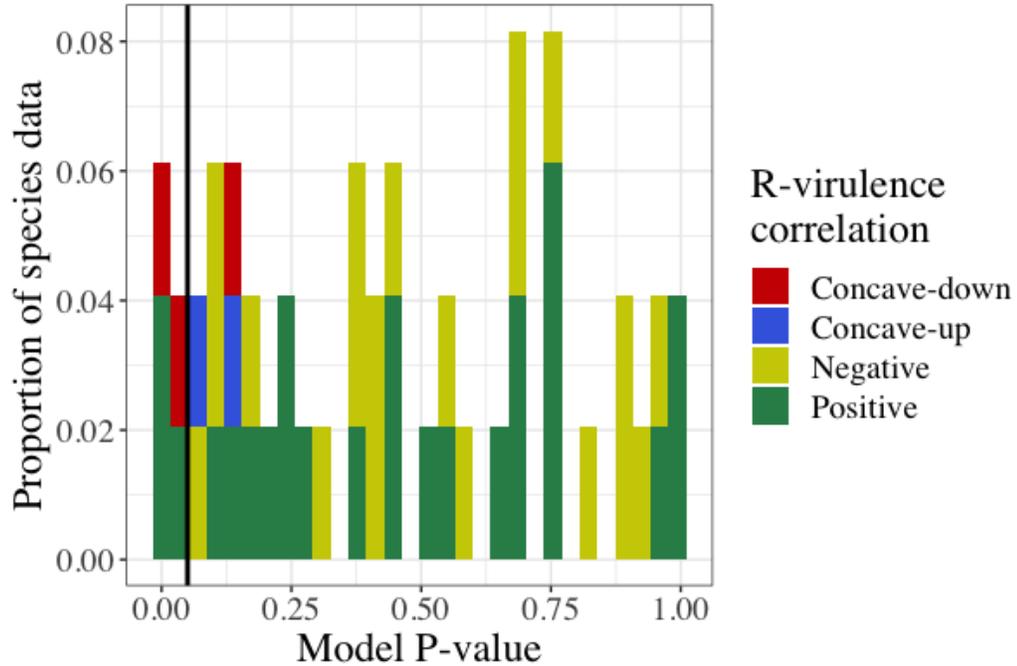


Figure A.2: For most of the species' data, the relationship between virulence and the effective reproduction number R is statistically unclear. We graph the proportion of model fit to within-species data as a function of model p-value for which the relationship between virulence and R is positive (green bar), negative (yellow bar), concave-up (blue bar) or concave-down (red bar). We fit linear and parabolic models to each species' data (sample sizes are 26-30 observations for the different species) and we did likelihood ratio tests to select the best fit model. The vertical black line is the 0.05 significance level.

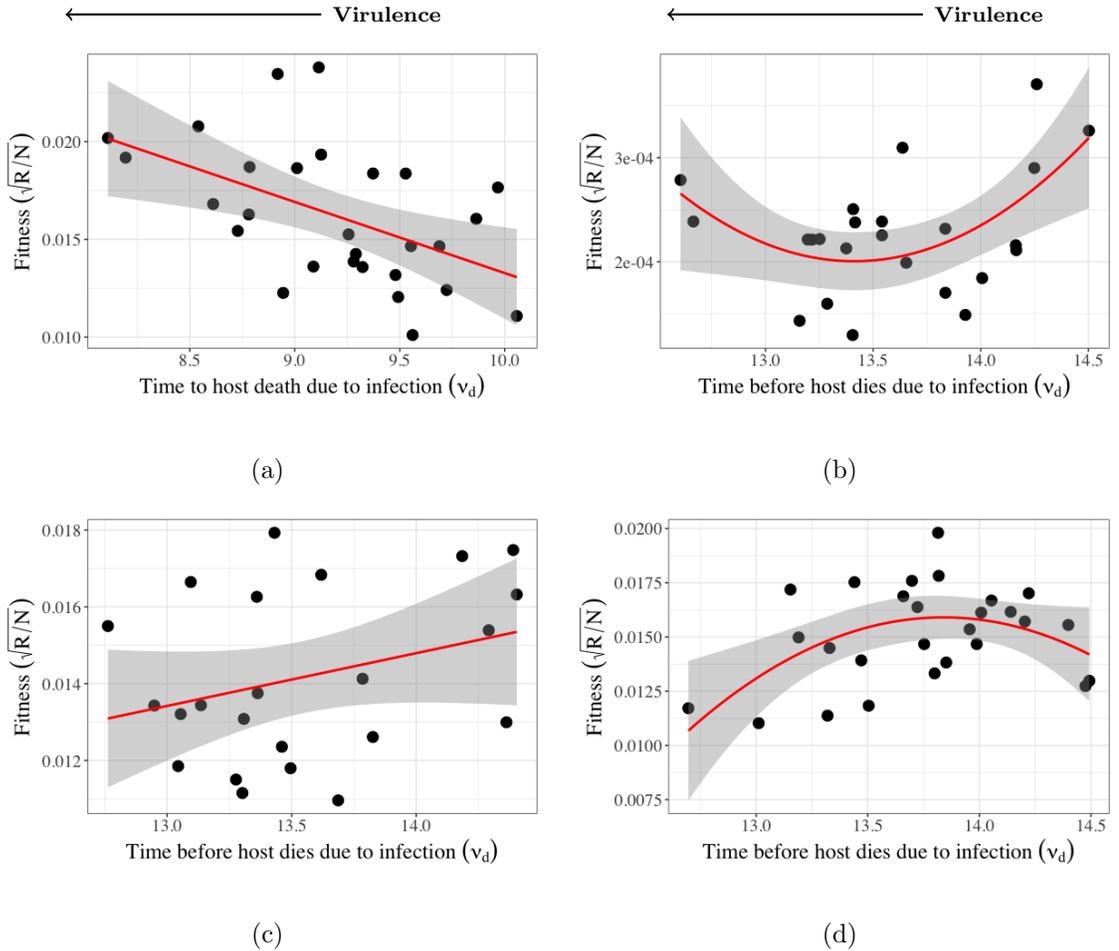


Figure A.3: The correlation between virulence and the effective reproduction number R can be positive (a), concave-up (b), negative (c) and concave-down (d), and we graph cases where the relationship is significant (except graph (c)). The dots are the estimated R for each strain (averaged over 30 outbreaks simulation runs), the line through the data is the best fit model, and the grey area is the 95% confidence interval. Model p-values are less than 0.05 for all graphs except graph (c) where p-value is 0.1, and adjusted $R^2 = 0.23, 0.2, 0.05$ and 0.2 for (a), (b), (c) and (d) respectively. We divided R_0 by host population density (N) to correct for the effect of host population density and we square root transformed the response variables to meet the assumptions of the linear model.

The relationship between virulence and parasite fitness measures: cross-species data

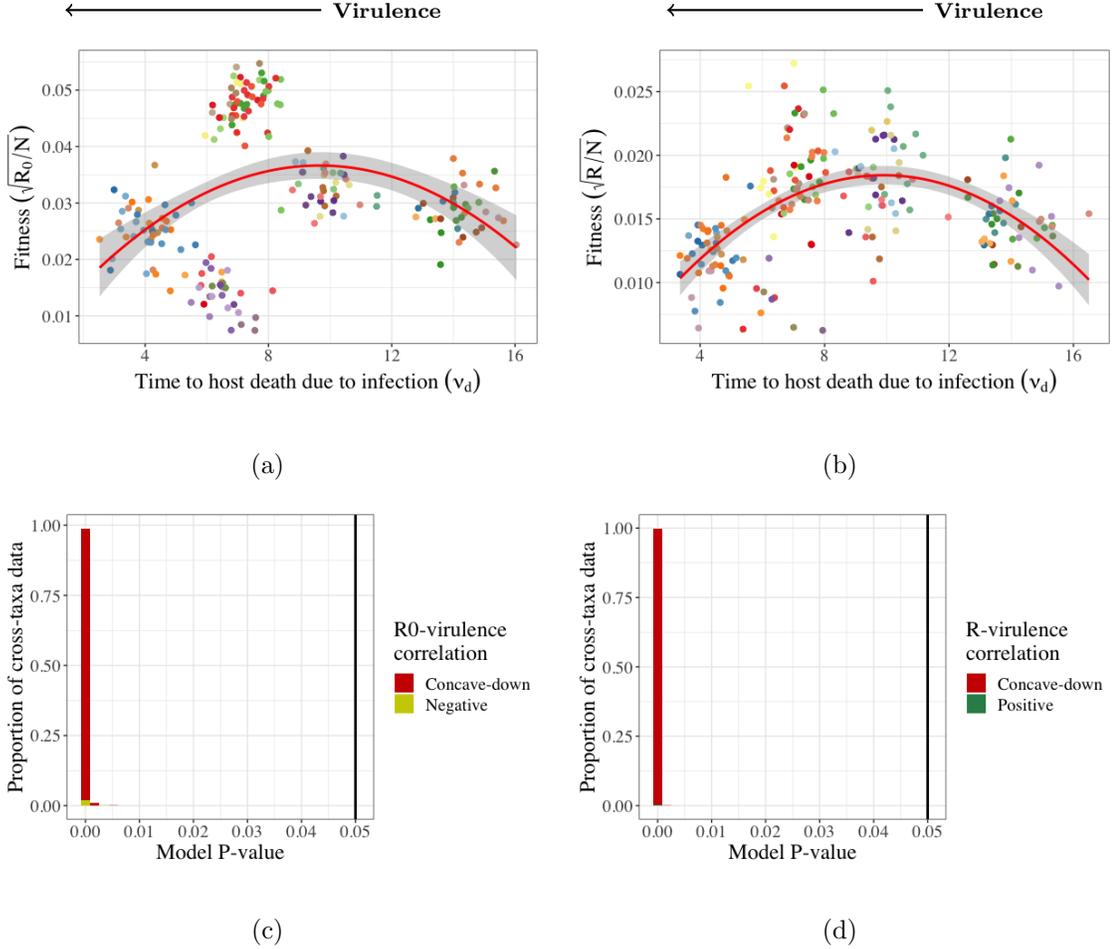


Figure A.4: The correlation between virulence and the basic reproduction number R_0 in (a) and the effective reproduction number R in (b) is concave down, and in (c) and (d) the proportion of random cross-species data samples that supports the R_0 and the R results is 97.7 % and 99.5 % respectively. The dots are the estimated fitness measures for each parasite over 30 outbreaks (the colour indicates strains of the same species), the line through the data is the best fit model, and the grey area is the 95% confidence interval. In (a) and (b) the best fit polynomial models are significant (Model p-value < 0.0001) and the adjusted $R^2 = 0.5$ and $R^2 = 0.38$ for (a) and (b) respectively. To select the best fit model, we fit phylogenetically corrected linear and parabolic models to 1000 randomly sampled cross-species data (sample size is 200 observations per randomly sampled data), we did likelihood ratio tests to select the best fit model, and the proportion of samples for each best fit model and their corresponding model p-values are presented in (c) and (d). The red, yellow and green bars are the best fit polynomial, linear-negative and linear-positive models respectively, and the vertical black line is the 0.05 significance level.

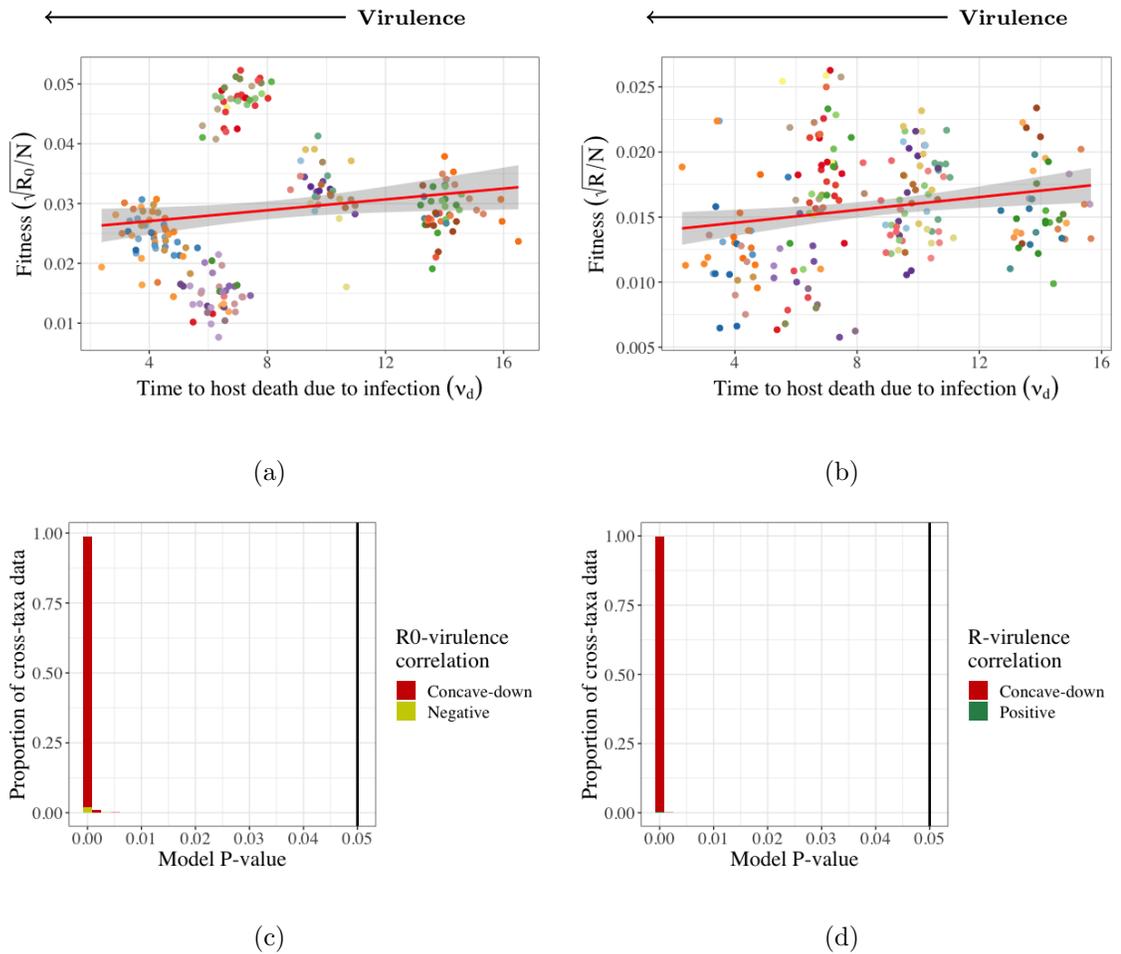


Figure A.5: The correlation between virulence and the basic reproduction number R_0 is negative in (a) and the effective reproduction number R is positive in (b), and in (c) and (d) the proportion of random cross-species data samples that supports the R_0 and the R results is 2.2 % and 0.5 % respectively. In (a) and (b) the dots are the estimated fitness measures for each parasite over 30 outbreaks (the colour indicates strains of the same species), the line through the data is the best fit model to the sampled data, and the grey area is the 95% confidence interval. The model fit is poor, adjusted $R^2 = 0.09$ for both graphs (a) and (b), the model diagnostics performed using the *gvlma* R package, show that the residuals are correlated and the linear relationship assumption is not satisfied (p-value < 0.05). We divided R_0 and R by host population density (N) to correct for the effect of host population density and we square root transformed the response variables to meet the assumptions of the linear model. We fit phylogenetically corrected linear and parabolic models to 1000 randomly sampled cross-species data (sample size is 200 observations per randomly sampled data), we performed likelihood ratio tests to select the best fit model, and the proportion of samples for each best fit model and their corresponding model p-values are presented in (c) and (d). The red, yellow and green bars are the best fit polynomial, linear-negative and linear-positive models respectively, and the vertical black line is the 0.05 significance level.

Appendix B.

Chapter two: Parasite-induced shifts in host movement may explain the transient coexistence of high- and low-pathogenic disease strains

Epidemiological and evolution models presented in the main text

Epidemiological dynamics

$$\frac{dS_M}{dt} = \theta + \gamma (I_M + I_R) - S_M (\Lambda + d) \quad (\text{B.1})$$

$$\frac{dI_M}{dt} = \Lambda S_M - I_M [d + \gamma + \psi(\alpha)] \quad (\text{B.2})$$

$$\frac{dI_R}{dt} = \psi(\alpha) I_M - I_R [d + \gamma + \nu(\alpha)], \quad (\text{B.3})$$

where $\Lambda = \alpha (c_m I_M + c_r I_R)$ represents the force of infection. The system of equations B.1-B.3 (system 3.1-3.3 in the main text) has two equilibria. A disease-free equilibrium (E_{DF}),

$$E_{DF} = \left(S_M^* = \frac{\theta}{d}, I_M^* = 0, I_R^* = 0, \right)$$

and an endemic equilibrium (E_E),

$$E_E = \left(\begin{array}{l} S_M^* = \frac{[d + \gamma + \nu(\alpha)] [d + \gamma + \psi(\alpha)]}{\alpha (c_m [d + \gamma + \nu(\alpha)] + \alpha c_r \psi(\alpha))}, \\ I_M^* = \frac{[d + \gamma + \nu(\alpha)] \left([d + \gamma + \nu(\alpha)] [\alpha c_m \theta - d(d + \gamma)] + [\alpha c_r \theta - d(d + \gamma + \nu(\alpha))] \psi(\alpha) \right)}{\alpha \left(c_m [d + \gamma + \nu(\alpha)] + c_r \psi(\alpha) \right) \left(d [d + \gamma + \nu(\alpha)] + [d + \nu(\alpha)] \psi(\alpha) \right)}, \\ I_R^* = \frac{\psi(\alpha)}{d + \gamma + \nu(\alpha)} I_M^*. \end{array} \right)$$

If both

$$\frac{\theta}{d} > \frac{d + \gamma}{\alpha c_m},$$

and

$$\frac{\theta}{d} > \frac{d + \gamma + \nu(\alpha)}{\alpha c_r},$$

then I_M^* and I_R^* are non-negative and the endemic equilibrium is biologically feasible.

To investigate the stability of disease-free equilibrium (E_{DF}) we use the next-generation matrix method (see van den Driessche and Watmough 2002), and we compute the basic reproduction number (R_0) of the system B.1-B.3. We write the Jacobian matrix of the system B.1-B.3 as $J_{eco} = F - V$ where,

$$F = \begin{bmatrix} \alpha c_m S_M^* & \alpha c_r S_M^* \\ 0 & 0 \end{bmatrix},$$

and

$$V = \begin{bmatrix} d + \gamma + \psi(\alpha) & 0 \\ -\psi(\alpha) & d + \gamma + \nu(\alpha) \end{bmatrix}.$$

According to the next-generation theorem, R_0 is given by the dominant eigenvalue of the next-generation matrix which is,

$$FV^{-1} = \begin{bmatrix} \left(\frac{\alpha c_m}{d + \gamma + \psi(\alpha)} + \frac{\alpha c_r \psi(\alpha)}{[d + \gamma + \nu(\alpha)][d + \gamma + \psi(\alpha)]} \right) S_M^* & \frac{\alpha c_r}{d + \gamma + \nu(\alpha)} S_M^* \\ 0 & 0 \end{bmatrix},$$

and the dominant eigenvalue of FV^{-1} is,

$$\rho(FV^{-1}) = R_0 = \left[\frac{\alpha c_m}{d + \gamma + \psi(\alpha)} + \frac{\alpha c_r}{d + \gamma + \nu(\alpha)} \times \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)} \right] S_M^*, \quad (\text{B.4})$$

where $S_M^* = \theta/d$ is the size of the susceptible host population at the disease-free equilibrium. If $R_0 < 1$ then E_{DF} is stable and no outbreak occurs, in contrast, if $R_0 > 1$ then E_{DF} is unstable and an outbreak occurs. Following an outbreak the system reaches a stable endemic equilibrium as long as there is a permanent input of susceptible hosts through recovery and immigration.

Evolutionary dynamics

$$\frac{dS_M}{dt} = \theta + \gamma (I_{M1} + I_{M2} + I_{R1} + I_{R2}) - S_M (\Lambda_1 + \Lambda_2 + d) \quad (\text{B.5})$$

$$\frac{dI_{M1}}{dt} = \Lambda_1 S_M - I_{M1} [d + \gamma + \psi(\alpha_1)] \quad (\text{B.6})$$

$$\frac{dI_{M2}}{dt} = \Lambda_2 S_M - I_{M2} [d + \gamma + \psi(\alpha_2)] \quad (\text{B.7})$$

$$\frac{dI_{R1}}{dt} = \psi(\alpha_1) I_{M1} - I_{R1} [d + \gamma + \nu(\alpha_1)] \quad (\text{B.8})$$

$$\frac{dI_{R2}}{dt} = \psi(\alpha_2) I_{M2} - I_{R2} [d + \gamma + \nu(\alpha_2)], \quad (\text{B.9})$$

where $\Lambda_1 = \alpha_1 (c_m I_{M1} + c_r I_{R1})$ and $\Lambda_2 = \alpha_2 (c_m I_{M2} + c_r I_{R2})$ are the force of infections of the resident strain (α_1) and the mutant strain (α_2) respectively. The symbols α_1 and α_2 are within-host net replication rates of the resident and the mutant strains respectively. The system of equations B.5-B.9 has 3 equilibria: the disease-free, the resident-free and the mutant-free equilibria. For the purposes of the evolutionary invasion analysis we are interested in the mutant-free equilibrium (E_{MF}) which is

$$E_{MF} = \left(\begin{array}{l} S_M^* = \frac{[d + \gamma + \nu(\alpha_1)] [d + \gamma + \psi(\alpha_1)]}{\alpha_1 (c_m [d + \gamma + \nu(\alpha_1)] + \alpha_1 c_r \psi(\alpha_1))}, \\ I_{M1}^* = \frac{[d + \gamma + \nu(\alpha_1)] \left([d + \gamma + \nu(\alpha_1)] [\alpha_1 c_m \theta - d(d + \gamma)] + [\alpha_1 c_r \theta - d(d + \gamma + \nu(\alpha_1))] \psi(\alpha_1) \right)}{\alpha_1 \left(c_m [d + \gamma + \nu(\alpha_1)] + c_r \psi(\alpha_1) \right) \left(d [d + \gamma + \nu(\alpha_1)] + [d + \nu(\alpha_1)] \psi(\alpha_1) \right)}, \\ I_{R1}^* = \frac{\psi(\alpha_1)}{d + \gamma + \nu(\alpha_1)} I_{M1}^*, \\ I_{M2}^* = 0, \\ I_{R2}^* = 0. \end{array} \right)$$

To investigate the stability of the mutant-free equilibrium (E_{MF}) we write the Jacobian matrix of the system B.5-B.9 (J_{evo}) and we evaluate J_{evo} at the mutant-free equilibrium.

$$J_{evo} = \left[\begin{array}{c|c} J_{res} & U \\ \hline 0 & J_{mut} \end{array} \right], \quad (\text{B.10})$$

where

$$U = \begin{bmatrix} -\alpha_2 c_m S_M^* + \gamma & -\alpha_2 c_r S_M^* + \gamma \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$J_{res} = \begin{bmatrix} -d - \alpha_1(c_m I_{M1}^* + c_r I_{R1}^*) & -\alpha_1 c_m S_M^* + \gamma & -\alpha_1 c_r S_M^* + \gamma \\ \alpha_1(c_m I_{M1}^* + c_r I_{R1}^*) & -d - \gamma - \psi(\alpha_1) + \alpha_1 c_m S_M^* & \alpha_1 c_r S_M^* \\ 0 & \psi(\alpha_1) & -d - \gamma - \nu(\alpha_1) \end{bmatrix},$$

and

$$J_{mut} = \begin{bmatrix} -d - \gamma - \psi(\alpha_2) + \alpha_2 c_m S_M^* & \alpha_2 c_r S_M^* \\ \psi(\alpha_2) & -d - \gamma - \nu(\alpha_2) \end{bmatrix}.$$

First, we assume that the resident strain is established in the host population, meaning that an epidemic occurred ($R_0 > 1$) and the system reaches a stable endemic equilibrium (J_{res} is locally stable). Then a rare mutant strain arises in the population. We investigate the conditions for the rare mutant strain to invade and replace the dominant resident strain, by analyzing the stability of the system of equation B.5-B.9 at the mutant-free equilibrium. The dynamics of the system of equation B.5-B.9 are governed by the stability of the sub-matrices J_{res} and J_{mut} . We assumed that J_{res} is locally stable, thus the dynamics of J_{evo} are governed by the stability of J_{mut} . If J_{mut} is unstable then J_{evo} is unstable and the rare mutant strain replaces the resident strain, and if J_{mut} is stable then J_{evo} is stable and the rare mutant strain goes extinct.

To investigate the stability of J_{mut} , we use the next-generation theorem for the evolutionary invasion analysis (see, Hurford et al. 2010). We write $J_{mut} = F - V$ and we compute the leading eigenvalue ($\rho(FV^{-1})$) of the J_{mut} sub-matrix, which is given

by,

$$\rho(FV^{-1}) = R(\alpha_2, \alpha_1) = \frac{\alpha_2 \left(c_m [d + \gamma + \nu(\alpha_2)] + c_r \psi(\alpha_2) \right)}{\left[d + \gamma + \nu(\alpha_2) \right] \left[d + \gamma + \psi(\alpha_2) \right]} S_M^*. \quad (\text{B.11})$$

where

$$S_M^* = \frac{\left[d + \gamma + \nu(\alpha_1) \right] \left[d + \gamma + \psi(\alpha_1) \right]}{\alpha_1 \left(c_m [d + \gamma + \nu(\alpha_1)] + c_r \psi(\alpha_1) \right)}. \quad (\text{B.12})$$

Equation B.11 is known as the invasion fitness of a rare mutant strain in a resident population at endemic equilibrium. Replacing equation B.12 in equation B.11 we have,

$$R(\alpha_2, \alpha_1) = \frac{\alpha_2 \left(c_m [d + \gamma + \nu(\alpha_2)] + c_r \psi(\alpha_2) \right)}{\left[d + \gamma + \nu(\alpha_2) \right] \left[d + \gamma + \psi(\alpha_2) \right]} \times \frac{\left[d + \gamma + \nu(\alpha_1) \right] \left[d + \gamma + \psi(\alpha_1) \right]}{\alpha_1 \left(c_m [d + \gamma + \nu(\alpha_1)] + c_r \psi(\alpha_1) \right)}. \quad (\text{B.13})$$

It can be noticed that

$$R(\alpha_2, \alpha_1) = \frac{R(\alpha_2)}{R(\alpha_1)},$$

with $i = 1$ and 2 and

$$R(\alpha_i) = \frac{\alpha_i \left(c_m [d + \gamma + \nu(\alpha_i)] + c_r \psi(\alpha_i) \right)}{\left[d + \gamma + \nu(\alpha_i) \right] \left[d + \gamma + \psi(\alpha_i) \right]}. \quad (\text{B.14})$$

According to the Next-generation theorem, (see Hurford et al. 2010), J_{mut} sub-matrix is unstable if

$$\rho(FV^{-1}) = R(\alpha_2, \alpha_1) > 1.$$

Therefore a rare mutant strain invades the host population dominated by the resident

strain if,

$$\frac{\alpha_2 \left(c_m [d + \gamma + \nu(\alpha_2)] + c_r \psi(\alpha_2) \right)}{\left[d + \gamma + \nu(\alpha_2) \right] \left[d + \gamma + \psi(\alpha_2) \right]} > \frac{\alpha_1 \left(c_m [d + \gamma + \nu(\alpha_1)] + c_r \psi(\alpha_1) \right)}{\left[d + \gamma + \nu(\alpha_1) \right] \left[d + \gamma + \psi(\alpha_1) \right]}. \quad (\text{B.15})$$

We discuss the evolutionary implications of this result in the main text.

The evolutionarily stable within-host parasite net replication rate (ESS α^*)

First we recall that the invasion fitness is,

$$R(\alpha_2, \alpha_1) = R_0(\alpha_2) \times \frac{1}{R_0(\alpha_1)},$$

Assuming that mutants are slightly different from the resident strain, a net replication rate that is evolutionarily stable (denoted α^*) must satisfy:

$$\left. \frac{\partial R(\alpha_2, \alpha_1)}{\partial \alpha_2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} = \frac{1}{R_0(\alpha^*)} \left. \frac{\partial R_0(\alpha_2)}{\partial \alpha_2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} = 0 \quad (\text{B.16})$$

and

$$\left. \frac{\partial^2 R(\alpha_2, \alpha_1)}{\partial \alpha_2^2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} = \frac{1}{R_0(\alpha^*)} \left. \frac{\partial^2 R_0(\alpha_2)}{\partial \alpha_2^2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} \leq 0. \quad (\text{B.17})$$

The condition B.16 is the first partial derivative of the invasion fitness with respect to α_2 evaluated at $\alpha_2 = \alpha_1 = \alpha^*$ and the condition B.17 is the second partial derivative of the invasion fitness with respect to α_2 evaluated at $\alpha_2 = \alpha_1 = \alpha^*$. From equation B.14, we know that

$$R(\alpha_2) = \frac{\alpha_2 \left(c_m [d + \gamma + \nu(\alpha_2)] + c_r \psi(\alpha_2) \right)}{\left[d + \gamma + \nu(\alpha_2) \right] \left[d + \gamma + \psi(\alpha_2) \right]}.$$

The first and the second derivatives of $R(\alpha_2)$ with respect to α_2 evaluated at $\alpha_2 = \alpha_1 = \alpha^*$ are respectively,

$$\left. \frac{\partial R(\alpha_2)}{\partial \alpha_2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} = R(\alpha^*) \left[\frac{1}{\alpha^*} + \left(\frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} - \frac{1}{[d+\gamma+\psi]} \right) \psi' - \frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \nu' \right], \quad (\text{B.18})$$

and

$$\begin{aligned} \left. \frac{\partial^2 R(\alpha_2)}{\partial \alpha_2^2} \right|_{\substack{\alpha_2=\alpha^* \\ \alpha_1=\alpha^*}} &= R(\alpha^*) \left[-\frac{1}{\alpha^{*2}} - \left(\frac{(c_m - c_r)(d+\gamma) + c_m\nu}{[d+\gamma+\psi][c_m(d+\gamma+\nu) + c_r\psi]} \right) \psi'' - \right. \\ &\quad \left. \frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \nu'' - \left[\frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \psi' \right]^2 + \left[\frac{1}{[d+\gamma+\psi]} \psi' \right]^2 \right. \\ &\quad \left. \frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \left(\frac{2c_m(d+\gamma+\nu)}{[c_m(d+\gamma+\nu) + c_r\psi][d+\gamma+\nu]} \right) \psi' \nu' + \right. \\ &\quad \left. \frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \left(\frac{2c_m(d+\gamma+\nu) + c_r\psi}{[c_m(d+\gamma+\nu) + c_r\psi][d+\gamma+\nu]} \right) \nu'^2 \right], \quad (\text{B.19}) \end{aligned}$$

where, ψ , ψ' and ψ'' are used in place of $\psi(\alpha^*)$, $\psi'(\alpha^*)$ and $\psi''(\alpha^*)$ respectively, and ν , ν' and ν'' are used in place of $\nu(\alpha^*)$, $\nu'(\alpha^*)$ and $\nu''(\alpha^*)$ respectively for notational brevity. Also, ψ' and ψ'' are respectively the first and the second derivatives of $\psi(\alpha_2)$ with respect α_2 evaluated at α^* , whereas ν and ν'' are respectively the first and the second derivatives of $\nu(\alpha_2)$ with respect α_2 evaluated at α^* . We substitute equation B.18 in the ESS condition B.16 and after few simplifications we found that if

$$\frac{1}{\alpha^*} = \frac{(c_m - c_r)(d+\gamma) + c_m\nu}{[d+\gamma+\psi][c_m(d+\gamma+\nu) + c_r\psi]} \psi' + \frac{c_r}{[c_m(d+\gamma+\nu) + c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \nu', \quad (\text{B.20})$$

then the condition B.16 is satisfied. From equation B.20 we solve for α^* , and it is

given by

$$\alpha^* = \frac{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \psi][d + \gamma + \nu]}{[(c_m - c_r)(d + \gamma) + c_m\nu][d + \gamma + \nu]\psi' + [d + \gamma + \psi]c_r\psi\nu'}. \quad (\text{B.21})$$

For equation B.21 to make sense biologically α^* must be non-negative. In the model formulation we assume that $c_m > c_r$, thus if both ψ' and ν' are positive then α^* is non-negative. Similarly, we substitute equation B.19 in the ESS condition B.17 and after few simplifications we found that if

$$\begin{aligned} & \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \left(\frac{2c_m(d + \gamma + \nu) + c_r\psi}{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \nu]} \right) \nu'^2 + \left[\frac{1}{d + \gamma + \psi} \psi' \right]^2 \leq \\ & \frac{1}{\alpha^{*2}} + \left(\frac{(c_m - c_r)(d + \gamma) + c_m\nu}{[d + \gamma + \psi][c_m(d + \gamma + \nu) + c_r\psi]} \right) \psi'' + \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \nu'' + \left[\frac{c_r}{c_m(d + \gamma + \nu) + c_r\psi} \psi' \right]^2 \\ & \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \left(\frac{2c_m(d + \gamma + \nu)}{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \nu]} \right) \psi'\nu', \end{aligned} \quad (\text{B.22})$$

then the condition B.17 is satisfied. We replace the expression of α^* (equation B.21) in inequality B.22 and we have,

$$\begin{aligned} & \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \left(\frac{2c_m(d + \gamma + \nu) + c_r\psi}{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \nu]} \right) \nu'^2 + \left[\frac{1}{d + \gamma + \psi} \psi' \right]^2 \leq \\ & \left(\frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \nu' \right)^2 + \left(\frac{(c_m - c_r)(d + \gamma) + c_m\nu}{[d + \gamma + \psi][c_m(d + \gamma + \nu) + c_r\psi]} \psi' \right)^2 + \\ & 2 \left(\frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \right) \left(\frac{(c_m - c_r)(d + \gamma) + c_m\nu}{[d + \gamma + \psi][c_m(d + \gamma + \nu) + c_r\psi]} \right) \psi'\nu' + \\ & \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \left(\frac{2c_m(d + \gamma + \nu)}{[c_m(d + \gamma + \nu) + c_r\psi][d + \gamma + \nu]} \right) \psi'\nu' + \left[\left(\frac{c_r}{c_m(d + \gamma + \nu)} - \frac{1}{d + \gamma + \psi} \right) \psi' \right]^2 + \\ & \left(\frac{(c_m - c_r)(d + \gamma) + c_m\nu}{[d + \gamma + \psi][c_m(d + \gamma + \nu) + c_r\psi]} \right) \psi'' + \frac{c_r}{[c_m(d + \gamma + \nu) + c_r\psi]} \frac{\psi}{[d + \gamma + \nu]} \nu''. \end{aligned} \quad (\text{B.23})$$

It can be shown that if both ψ'' and ν'' are positive or if ψ'' is positive and $\nu'' = 0$ then inequality B.23 holds, and α^* satisfies both conditions B.16 and B.17. Thus, if both parasite-induced host resting rate ($\psi(\alpha)$) and parasite-induced host mortality rate ($\nu(\alpha)$) increase at an increasing rate as within-host parasite net replication rate

(α) increases (meaning that both $\psi(\alpha)$ and $\nu(\alpha)$ have a concave-up form) then α^* is a biologically feasible evolutionarily stable within-host parasite net replication rate. Also, if $\psi(\alpha)$ has a concave up form whereas $\nu(\alpha)$ is linear then equation B.21 is a biologically feasible evolutionarily stable within-host parasite net replication rate. In contrast, when both $\psi(\alpha)$ and $\nu(\alpha)$ have a linear form then no evolutionarily stable parasite net replication rate is possible. In the main paper we focus on the case where both $\psi(\alpha)$ and $\nu(\alpha)$ have a concave up form.

The convergence stable within-host parasite net replication rate (CSS)

An ESS, if it exists, is also convergence stable if

$$\frac{d}{d\alpha_1} \left\{ \left. \frac{\partial R(\alpha_2, \alpha_1)}{\partial \alpha_2} \right|_{\alpha_2=\alpha_1} \right\}_{\alpha_1=\alpha^*} < 0. \quad (\text{B.24})$$

The CSS condition (equation B.24) and condition B.17 are similar except the inequality sign. We found that if

$$\begin{aligned} & \frac{c_r}{[c_m(d+\gamma+\nu)+c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \left(\frac{2c_m(d+\gamma+\nu)+c_r\psi}{[c_m(d+\gamma+\nu)+c_r\psi][d+\gamma+\nu]} \right) \nu'^2 + \left[\frac{1}{d+\gamma+\psi} \psi' \right]^2 < \\ & \left(\frac{c_r}{[c_m(d+\gamma+\nu)+c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \nu' \right)^2 + \left(\frac{(c_m-c_r)(d+\gamma)+c_m\nu}{[d+\gamma+\psi][c_m(d+\gamma+\nu)+c_r\psi]} \psi' \right)^2 + \\ & 2 \left(\frac{c_r}{[c_m(d+\gamma+\nu)+c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \right) \left(\frac{(c_m-c_r)(d+\gamma)+c_m\nu}{[d+\gamma+\psi][c_m(d+\gamma+\nu)+c_r\psi]} \right) \psi' \nu' + \\ & \frac{c_r}{[c_m(d+\gamma+\nu)+c_r\psi]} \left(\frac{2c_m(d+\gamma+\nu)}{[c_m(d+\gamma+\nu)+c_r\psi][d+\gamma+\nu]} \right) \psi' \nu' + \left[\left(\frac{c_r}{c_m(d+\gamma+\nu)} - \frac{1}{d+\gamma+\psi} \right) \psi' \right]^2 + \\ & \left(\frac{(c_m-c_r)(d+\gamma)+c_m\nu}{[d+\gamma+\psi][c_m(d+\gamma+\nu)+c_r\psi]} \right) \psi'' + \frac{c_r}{[c_m(d+\gamma+\nu)+c_r\psi]} \frac{\psi}{[d+\gamma+\nu]} \nu'' \end{aligned} \quad (\text{B.25})$$

then equation B.24 is satisfied. If $c_r = 0$ then inequality B.25 becomes

$$0 < \left[\frac{1}{d+\gamma+\psi} \psi' \right]^2 + \left[\frac{1}{d+\gamma+\psi} \right] \psi''. \quad (\text{B.26})$$

From the ESS conditions we know that both ψ'' and ψ' are positive. Thus, we conclude that if $c_r = 0$ then an evolutionarily stable within-host parasite net replication rate (ESS α^*) is also convergence stable (CSS α^*).

Evolutionary dynamics when parasite infection is non-lethal

The model is similar to the system B.1-B.3, with no disease-induced host death ($\nu(\alpha) = 0$). We substitute $\nu(\alpha) = 0$ in equation B.4 and we obtain the basic reproduction number which is given by

$$R_0 = \left[\frac{\alpha c_m}{d + \gamma + \psi(\alpha)} + \frac{\alpha c_r}{(d + \gamma)} \times \frac{\psi(\alpha)}{d + \gamma + \psi(\alpha)} \right] S_M^*, \quad (\text{B.27})$$

where $S_M^* = \theta/d$ is the size of susceptible host population at disease-free equilibrium. Similarly we substitute $\nu(\alpha) = 0$ in equation B.13 and we obtain the invasion fitness which is given by

$$R(\alpha_2, \alpha_1) = \frac{\alpha_2 \left[c_m (d + \gamma) + c_r \psi(\alpha_2) \right]}{\left[d + \gamma \right] \left[d + \gamma + \psi(\alpha_2) \right]} \times \frac{\left[d + \gamma \right] \left[d + \gamma + \psi(\alpha_1) \right]}{\alpha_1 \left[c_m (d + \gamma) + c_r \psi(\alpha_1) \right]}. \quad (\text{B.28})$$

The conditions for an ESS net replication rate to exist are the same as those provided in B.16 and B.17. We substitute $\nu(\alpha) = 0$ in equation B.21 and we obtain the expression of the within-host net replication rate that is evolutionarily stable. It is given by

$$\alpha^* = \frac{\left[c_m (d + \gamma) + c_r \psi \right] \left[d + \gamma + \psi \right]}{\left[(c_m - c_r) (d + \gamma) \right] \psi'}. \quad (\text{B.29})$$

For α^* to be non-negative, thus biologically meaningful, ψ' must be positive. The ESS condition (B.17) is satisfied if

$$-\frac{1}{\alpha^{*2}} - \left(\frac{(c_m - c_r)(d + \gamma)}{\left[d + \gamma + \psi \right] \left[c_m (d + \gamma) + c_r \psi \right]} \right) \psi'' - \left[\frac{c_r}{c_m (d + \gamma) + c_r \psi} \psi' \right]^2 + \left[\frac{1}{d + \gamma + \psi} \psi' \right]^2 \leq 0. \quad (\text{B.30})$$

We replace α^* (equation B.29) in equation B.30 and after few simplifications we have

$$2c_r\psi'^2 - \left[c_m(d + \gamma) + c_r\psi \right] \psi'' \leq 0. \quad (\text{B.31})$$

As in the case where parasite infection is potentially lethal ψ'' must be positive for α^* (equation B.29) to be biologically feasible. Therefore, the trade-off between parasite-induced host lethargy rate ($\psi(\alpha)$) and within-host net parasite replication rate (α) is concave-up.

To derive the condition for the ESS to be a CSS, we apply the condition B.24, and we find that if

$$2c_r\psi'^2 - \left[c_m(d + \gamma) + c_r\psi \right] \psi'' < 0. \quad (\text{B.32})$$

then equation (B.29) is also a CSS. It can be noticed that if $c_r = 0$ then inequality B.32 holds. Therefore, similarly to the case where parasite infection is potentially lethal, if $c_r = 0$ then whenever α^* is an ESS it is also a CSS.

Dynamical simulation

To simulate the evolution of the within-host parasite net replication rate (α), we solve the system of ordinary differential equations (ODEs) describing the epidemiological dynamics (B.1-B.3), where only the resident strain (α_1) is present in the host population. We set the parameter values such that an epidemic occurs ($R_0 > 1$) and the system reaches a stable endemic equilibrium (which is reached within 500 time steps maximum).

For the evolutionary dynamics, we set the initial within-host net replication rate $\alpha_i = \alpha_1$ as the dominant strain for the first generation. At the end of each generation,

we produce 20 different mutant strains from uniformly distributed α values, with the centre of the distribution being the α value of the current dominant strain. The lower and the upper bounds of the distribution are chosen to reflect the magnitude of the effect of mutation. We set bounds to $\alpha_1 \pm 0.1$ and $\alpha_1 \pm 0.55$ for small- and large-effect mutations respectively. We calculate the fitness for all parasite strains present in the population using equation S14, and we compare the fitness of mutants to the fitness of the current resident strain. For the following generation, the new dominant resident strain is the strain with the highest fitness. We assume that all the other strains go extinct. We iterate this evolution process for 300 generations (evolutionary equilibrium is reached in all simulations before 300 generations). We repeat the evolution simulation 100 times, but we plot only one sample evolutionary path to illustrate the PIP.

For simulations in Figures 3.3d and 3.3e, we run the simulations with initial α values below (dotted lines) and above (dashed lines) the *invasible repellor* which is ≈ 0.7 . For all simulations we model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = 0.01\alpha^2$, and we set $c_m = 0.8$, $c_r = 0.08$, $d = 0.0001$ and $\gamma = 0.065$ except Figure 3.3b where we set $c_r = 0$.

Multimedia materials

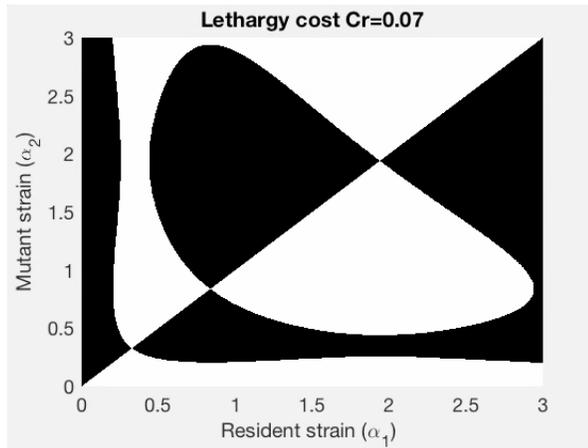


Figure B.1: Movie of Pairwise Invasibility Plots (PIP) illustrating the effect of increasing the contact rate in the resting state (c_r) on the evolutionary dynamics, Video is available at Figshare doi:10.6084/m9.figshare.11392617. We set $c_m = 0.8$, $b = 0.01$, $d = 0.0001$, $\gamma = 0.065$, and we vary c_r values from 0 to 0.25. The colours on the PIPs represent the fate of a rare mutant strain in a host population where the resident strain is at endemic equilibrium for different combinations of mutant-resident α values (α_1 on the x-axis and α_2 on the y-axis). For a given combination (α_1, α_2) , white indicates that the rare mutant goes extinct (equation 3.13, in the main text, is negative), and black indicates that the rare mutant replaces the resident (equation 3.13, in the main text, is positive). The transitions between black and white occur where equation 3.13, in the main text, equals zero, and the intersections are evolutionary equilibria. The intersections are either one ESS that is convergence stable or 2 ESS separated by an *invasible repeller*. We model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = b\alpha^2$.

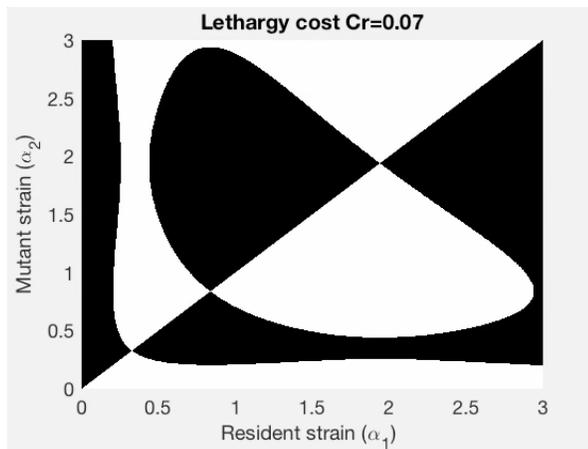


Figure B.2: Movie of Pairwise Invasibility Plots (PIP) illustrating the effect of increasing increasing the ratio of host mortality to lethargy rates (b) on the evolutionary dynamics, Video is available at Figshare doi:10.6084/m9.figshare.11393319. We set we $c_m = 0.8$, $c_r = 0.08$, $d = 0.0001$, $\gamma = 0.065$, and we vary b values from 0 to 0.05. We model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = b\alpha^2$. See the caption of Figure B.1 for how to read a PIP.

Throughout the paper, we assumed that the probability of disease transmission given an infectious contact, which is proportional to the within-host parasite net replication rate (α), is the same in the moving and the resting states, but the probability of disease transmission given an infectious contact may be higher in the resting state because of a higher parasite load. We investigated the case where the probability of disease transmission given an infectious contact (α) is higher in the resting state than the moving state ($\alpha_m > \alpha_r$, where α_m and α_r are the within-host parasite net replication rates in the moving and the resting state respectively). To formalize this idea, we assume that α_m is lower by a factor of c than α_r . For example, if $c = 0.5$ and the probability of disease transmission given an infectious contact in the resting state is $\alpha_r = 1$ then the probability of disease transmission given an infectious contact in the moving state is $\alpha_m = 0.5$. We found that the results are qualitatively similar to the case where α is the same in the moving and the resting states. When the contribution of one state (moving or resting) to the expected number of secondary infections per susceptible host (equation 15 in the main text) is not substantial then only one ESS is possible. In contrast, when both states can substantially contribute to the expected number of secondary infections per susceptible host then a bistability occurs.

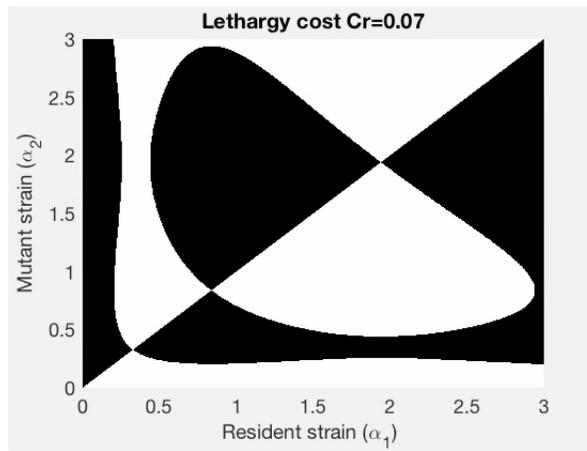


Figure B.3: The results are qualitatively similar when we assume that the probability of disease transmission given an infectious contact is higher in the resting than the moving state, Video is available at Figshare doi:10.6084/m9.figshare.11393343. We set $c_m = 0.8$, $c_r = 0.08$, $b = 0.01$, $d = 0.0001$, $\gamma = 0.065$, and the movie shows the PIPs for $c = \alpha_m/\alpha_r$ values from 0 to 1 (α_m and α_r are the within-host parasite net replication rates in the moving and the resting states respectively). We model the concave-up trade-offs using a power function $\psi(\alpha) = \alpha^2$ and $\nu(\alpha) = b\alpha^2$. See the caption of Figure B.1 for how to read a PIP.

Literature cited

- A. Hurford, D. Cownden, and T. Day. Next-generation tools for evolutionary invasion analyses. *J. Royal Soc. Interface*, 7(45):561–571, 2010.
- P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.*, 180(1):29–48, 2002.

Appendix C.

Chapter four: A counter-intuitive relationship between the temporal and spatial spread of diseases

Description of the CDC data

We retrieved infectious disease cases reported in the United States of America by the Centers for Disease Control and Prevention (CDC) through the National Notifiable Diseases Surveillance System (NNDSS) which is an information sharing system on infectious diseases. We cleaned and processed the data to make it directly usable in any statistical softwares or program, and the processed data are publicly available at Figshare doi:. The data have 3754064 observations of 11 variables which are: the reporting years (ReportingYears) is from 1996 to 2019, the reporting week (ReportingWeeks) is 1-52 or 1-53 within each year, the locations (Locations) are the names of the states as reported on the CDC websites, the names of the diseases (DiseaseNames) as reported on the CDC website, the number of cases reported the current epidemiological week (CurrentWeekInf), the cumulative number of cases to date reported for a disease at a location during the current year (CumYearInf1), the cumulative number of cases to date reported for a disease at a location during the previous year (CumYearInf2), the processed disease names (DiseaseNamesProc), the processed location names (LocationsProc), latitudes (Lat), and longitudes (Long). We processed the disease and the state names because the reporting was not consistent from one year to another. To make the disease names consistent across the years we reviewed the the case definition and the updates for each infectious diseases that is provided by the Council of State and Territorial Epidemiologists (CSTE). We used the latitude

and longitude coordinates (in decimal degrees) which we retrieved from the website of the National Ocean and Atmospheric Administration (NOAA) of the U.S.A. The CDC processed data is publicly available at doi:.

Preliminary statistical analysis

We explored the data to investigate the distribution of the spatial spread rate c and the epidemic growth rate r . We used box plots to identify extreme observations that can be outliers and potential influential observations.

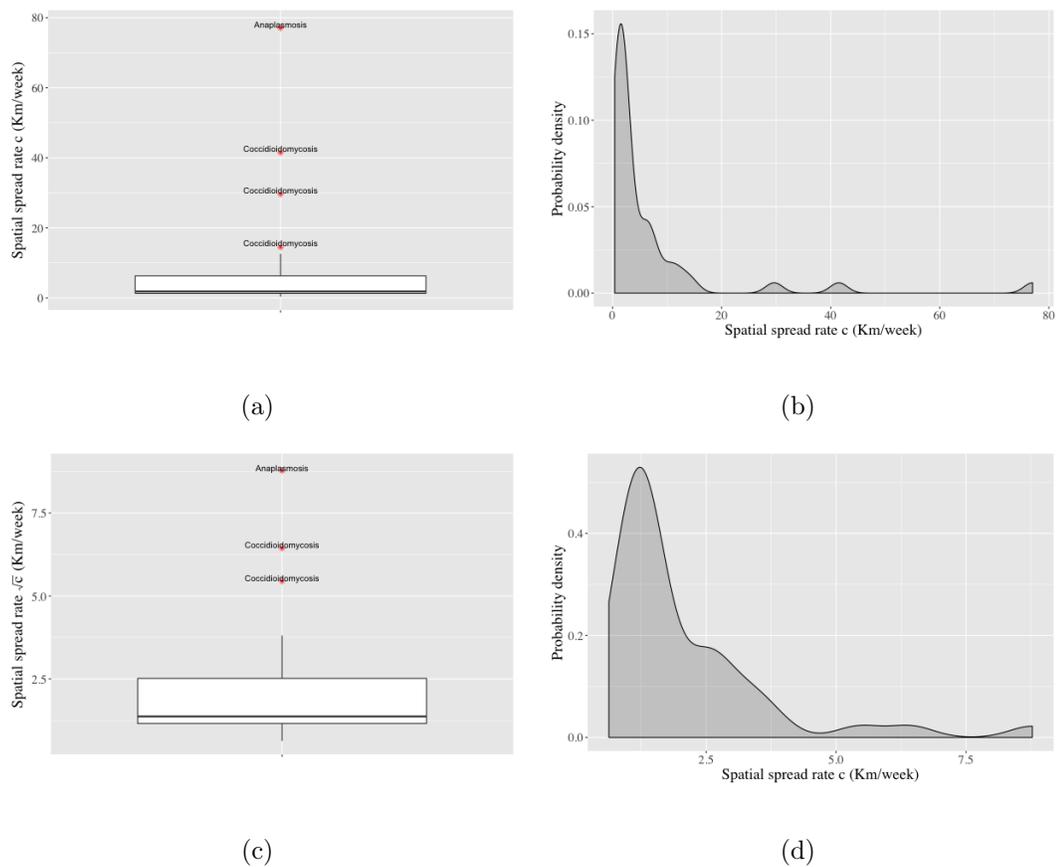
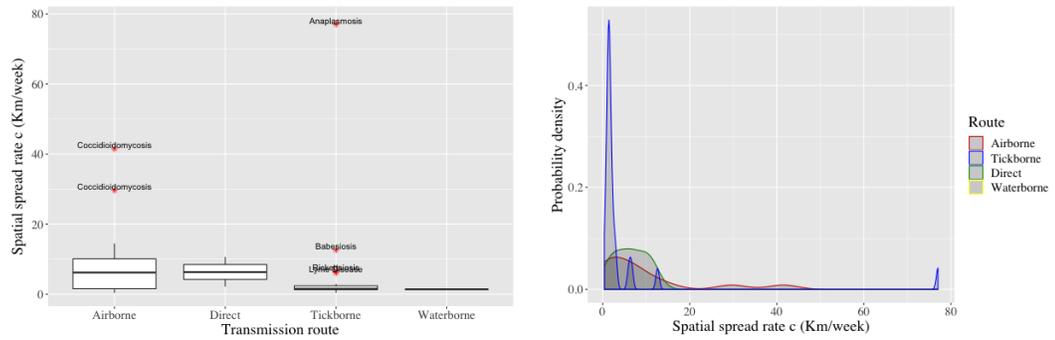
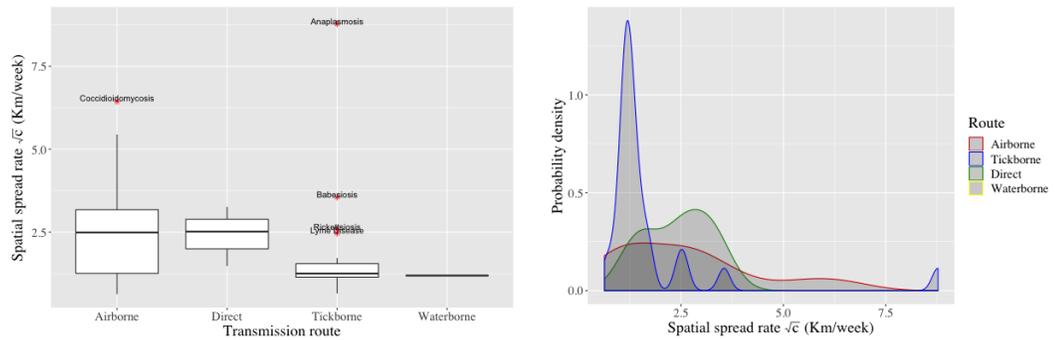


Figure C.1: The distribution of the spatial spread rate c for the untransformed data (a and b) and the square root transformed data (c and d). With these graphs, we started thinking about data transformation and dealing with possible outliers, which was decided after model fit and diagnostics.



(a)

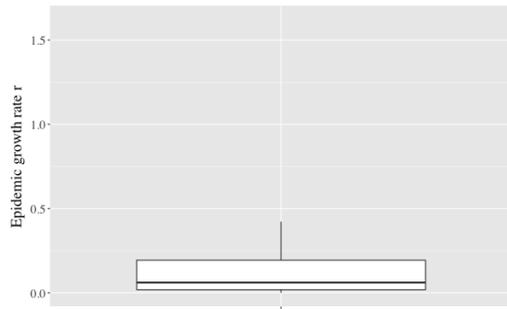
(b)



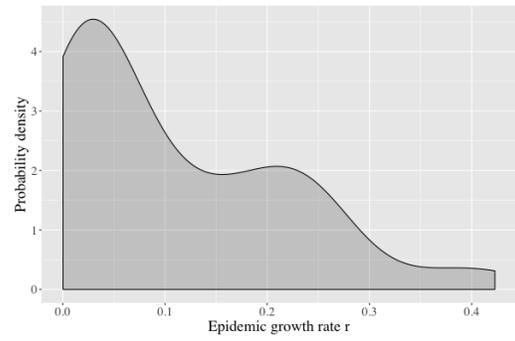
(c)

(d)

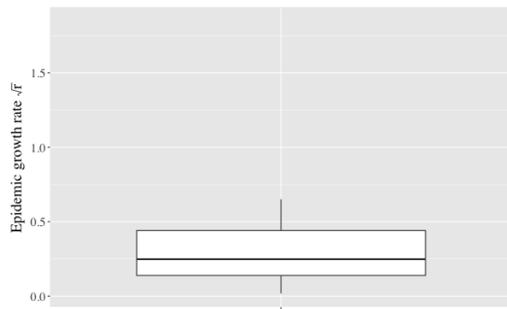
Figure C.2: The distribution of the spatial spread rate c for the untransformed data (a and b) and the square root transformed data (c and d) for each transmission route data.



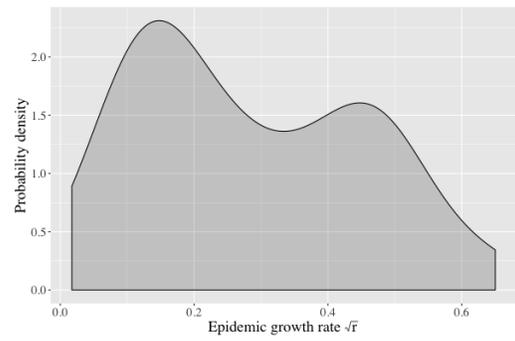
(a)



(b)



(c)



(d)

Figure C.3: The distribution of the epidemic growth rate r for the untransformed data (a and b) and the square root transformed data (c and d). For r measures the square root transformation is sometimes not necessary to get a good model fit.

Supplementary tables

Table C.1: Epidemic growth rate r as a function of outbreak years T in tick-borne and airborne infections, where R is disease transmission route variable. We calculated the ΔAIC , ΔDev ($\Delta\text{Deviance}$), ΔLL (Δ Log-likelihood) and pseudo- R^2 with respect to the model 3. The pseudo- $R^2 = 1 - \text{Model residual deviance}/\text{deviance of model 3}$. The models are ordered from the best to the worst.

N	Models	K	ΔAIC	ΔDev	ΔLL	Pseudo- R^2
1	$\frac{1}{r} \sim \beta_0 + \beta_1 T + \beta_2 R + \beta_3 TR$	4	56.380	32.738	-31.190	0.79
2	$\frac{1}{r} \sim \beta_0 + \beta_1 T + \beta_3 R$	3	42.165	28.223	-23.082	0.68
3	$\frac{1}{r} \sim \beta_0$	1	0	0	0	0