Evolution of virulence in facultative pathogens

by

Aakash Pandey

B.Tech., Kathmandu University, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Abstract

In addition to infecting hosts, facultative pathogens can grow independent of hosts in environmental reservoirs. These diverse ecological settings can result in novel life-history trait correlations and trade-offs. The consequence of such trait correlations on virulence evolution is largely understudied, both theoretically and empirically. In this study, I examined how reproduction in reservoirs influences disease dynamics and virulence evolution. I used mathematical models to determine how pathogen growth rate and environmental carrying capacity influence epidemiological outcomes. Further, I used the adaptive dynamics framework to examine the evolutionary dynamics of facultative pathogens under potential trade-offs between transmission and virulence, shedding and virulence, and reservoir persistence and virulence. I then performed critical function analysis to generalize the results independent of specific trade-off assumptions. I found that diverse virulence strategies, sometimes resulting from evolutionary bistability or evolutionary branching conditions, are expected for facultative pathogens. In these pathogens, adaptation to environmental reservoirs can influence virulence within host. I tested this idea using *Stenotrophomonas maltophilia*, an emerging opportunistic human pathogen. Previous research found that a type IV secretion system plays a role in both host-pathogen interactions and competition with other bacteria. I experimentally evolved four strains of *S. maltophilia* (JV3, JCMS, R551-3, and K279a) that vary in their virulence on the host *Caenorhabditis elegans*. I found that these strains also vary in the degree to which they use interference mechanisms in competition with *Escherichia coli* MG1655. Multiple independent populations initiated with highly virulent strains JV3, JCMS, and R551-3 show reductions in both virulence and their competitive ability. These

observations suggest a positive correlation between the virulence and the interspecies competitive ability of the pathogen. Genomic comparison of the ancestral and evolved lines shows that several reduced virulence strains have mutations in their type IV secretion system operon. I also show that both virulence and competitive ability are reduced in a Δ*vir*B10 mutant compared to the wild-type and complemented strains. Similarly, I found that the previously identified effector proteins of type IV secretion system 14245 and 14255 also have dual function in both virulence and interference competition. To test whether the presence of a competitor has any effect on virulence evolution, I evolved the highly virulent JV3 with and without competitor *E. coli* MG1655. Results show that the virulence of population lines evolved without the competitor *E. coli* is reduced to a greater extent than of those population lines evolved in the presence of *E. coli*. These results suggest that coincidental selection for interference competition against heterologous bacteria can contribute to the virulence of *S. maltophilia*. Taken together, this research highlights the importance of environmental adaptation to shaping the consequences that facultative pathogens have on hosts.

Evolution of virulence in facultative pathogens

by

Aakash Pandey

B.Tech in Biotechnology, Kathmandu University, 2016

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Thomas G. Platt

Copyright

Abstract

In addition to infecting hosts, facultative pathogens can grow independent of hosts in environmental reservoirs. These diverse ecological settings can result in novel life-history trait correlations and trade-offs. The consequence of such trait correlations on virulence evolution is largely understudied, both theoretically and empirically. In this study, I examined how reproduction in reservoirs influences disease dynamics and virulence evolution. I used mathematical models to determine how pathogen growth rate and environmental carrying capacity influence epidemiological outcomes. Further, I used the adaptive dynamics framework to examine the evolutionary dynamics of facultative pathogens under potential trade-offs between transmission and virulence, shedding and virulence, and reservoir persistence and virulence. I then performed critical function analysis to generalize the results independent of specific trade-off assumptions. I found that diverse virulence strategies, sometimes resulting from evolutionary bistability or evolutionary branching conditions, are expected for facultative pathogens. In these pathogens, adaptation to environmental reservoirs can influence virulence within host. I tested this idea using *Stenotrophomonas maltophilia*, an emerging opportunistic human pathogen. Previous research found that a type IV secretion system plays a role in both host-pathogen interactions and competition with other bacteria. I experimentally evolved four strains of *S. maltophilia* (JV3, JCMS, R551-3, and K279a) that vary in their virulence on the host *Caenorhabditis elegans*. I found that these strains also vary in the degree to which they use interference mechanisms in competition with *Escherichia coli* MG1655. Multiple independent populations initiated with highly virulent strains JV3, JCMS, and R551-3 show reductions in both virulence and their competitive ability. These

observations suggest a positive correlation between the virulence and the interspecies competitive ability of the pathogen. Genomic comparison of the ancestral and evolved lines shows that several reduced virulence strains have mutations in their type IV secretion system operon. I also show that both virulence and competitive ability are reduced in a Δ*vir*B10 mutant compared to the wild-type and complemented strains. Similarly, I found that the previously identified effector proteins of type IV secretion system 14245 and 14255 also have dual function in both virulence and interference competition. To test whether the presence of a competitor has any effect on virulence evolution, I evolved the highly virulent JV3 with and without competitor *E. coli* MG1655. Results show that the virulence of population lines evolved without the competitor *E. coli* is reduced to a greater extent than of those population lines evolved in the presence of *E. coli*. These results suggest that coincidental selection for interference competition against heterologous bacteria can contribute to the virulence of *S. maltophilia*. Taken together, this research highlights the importance of environmental adaptation to shaping the consequences that facultative pathogens have on hosts.

Table of Contents

List of Figures

List of Tables

Acknowledgements

When I started my graduate school journey in 2017, I knew little about evolution or infectious diseases, let alone about the evolution of infectious diseases. Yet, Tom saw something in me and gave me a chance to be a part of this wonderful community that studies how evolution shapes pathogen traits. I am forever grateful for his patience, dedication, and support throughout this journey.

I would like to thank my committee members Anna Zinovyeva, Michi Tobler, and Majid Jaberi-Douraki. I am fortunate to be under the guidance of these diverse committee members. Anna has always been helpful with the worms related work. Michi guided me on evolutionary questions and Majid helped me with the mathematics. Without the combined effort of all committee members, it would have been difficult to work on this interdisciplinary topic.

Science is a collective pursuit. This thesis would have been impossible without the support of many collaborators. Nicole Mideo had an integral part in theory development. She has also been kind in helping me build my professional network at the EEID meetings. The experimental work would not have been possible without the materials provided by Michael Herman, Leah Radeke, and Nicholas P. Cianciotto. I have a wide range of interests in science. But given the current structure of graduate programs and expectations therein, getting involved in widely varying topics is frowned upon. Fortunately, I had a rare opportunity to get involved in many projects in different topics. Thanks to Tom for allowing me to wander in these projects as I believe this has helped be a better scientist. Thanks to Walter Dodds, Flavia Tromboni, Brad Olson, Paige Hansen, and Ben Sikes for giving me different topics to think about.

Any international graduate student knows the pain of being away from home and family for an extended period. Add pandemic to the mix and it gets worse. It would have been impossible to bear the isolation and distance had I not met great friends in Manhattan. I would like to thank all my friends, Nepalese and otherwise, who made me felt like home. Special thanks to the members of G-18. Thank you for all the love, food, beers, and memories. Thank you to my special friend Meow for bearing me since high school and for being there on my highs and lows.

Lab environment determines a lot on how one's time goes during the graduate school. In Tom, I found not just a great scientist but also a great human being. I cannot say the same about many of my childhood science heroes. I am also fortunate to have wonderful lab members. Priscila and I shared many ups and downs and now she feels like a part of my family. Courtney, Carson, and Tyler have helped a great deal in setting up and running the experimental work. Otherwise, it would have been a tough time in the lab.

I am truly grateful to my family for their love, patience, and support throughout the graduate school journey. I am the first in my extended family to go to a graduate school for a PhD. Despite not knowing what goes into a graduate school life, all the family members have been open, understanding, and supportive of my decisions. Thank you, mom, dad, my brother, and sister-in-law. I love you. The last year of my graduate school has been kind to me as I found a loving and supporting partner. Thank you, Shona, for all the love, food, chai, and support.

Dedication

This thesis is primarily dedicated to my family. Despite several hardships, they have done everything to help me succeed. I would also like to dedicate this thesis to my high school science teacher Mr. Surendra for instilling inside me a sense of wonder about nature.

Preface

"*Disease is no longer the mysterious, personified entity of the past. It has been brought within the domain of the laws which govern all life upon earth*." – Theobald Smith, 1887

Why do pathogens kill their host? Why do some pathogens kill their hosts quickly whereas others only give slight discomforts? These are more than century old questions. All the 'why' questions in biology ultimately fall under the interests of evolutionary biology that aims to explain the variations in life. Ever since the idea of evolution by natural selection was put forward by Darwin and Wallace, people have attempted to explain the variation observed in the harm that pathogens cause towards hosts, i.e., virulence, using the principles of natural selection. The question of why and how virulence changes is also of great public interest given that we are living in an ongoing pandemic due to corona virus.

There is great variation in pathogen's life-history. Some pathogens are obligatory, some need a vector to complete their lifecycle, some are specialists, some are generalists, and some are opportunistic or facultative. This diversity in life-history may contribute to the variation observed in virulence. While there are many studies focusing on obligatory pathogens, very few studies focus on facultative pathogens-those that can live and reproduce in environmental reservoirs. In this dissertation, I examine how different life-history trade-offs arise, and how they influence virulence evolution in facultative pathogens. To do that, I use mathematical modeling and empirical approaches. Mathematical models have been central to the studies of life-history evolution. The goal of modeling is to either explain the observations of nature or to predict future scenarios. In the end, we must confront our models with the reality of data so that

they may not lead us astray. Thus, I combine both modeling and experiments to answer how virulence may evolve in facultative pathogens.

There are no Darwinian demons. Therefore, trade-offs must exist. Traits, including virulence, are usually coupled with other traits resulting in trade-offs. These trade-offs influence fitness of organism, and hence, determine the trajectory of trait evolution. In chapter 1, I deal with different pathogen life-histories and how trade-offs can arise with these life-histories. I focus on the mechanisms of possible trade-offs in pathogens and how these influence different fitness components of pathogens.

Chapter 2 deals with the theoretical exploration of how virulence can evolve in facultative pathogens under potential trade-offs of transmission-virulence, shedding-virulence, and environmental persistence-virulence trade-offs. I look at the effect of environmental growth in the epidemiology and evolutionary dynamics of facultative pathogens.

Chapter 3 focuses on virulence factors in *Stenotrophomonas maltophilia*. Different strains of *S. maltophilia* show varying levels of virulence against *Caenorhabditis elegans*. I look at the potential causes of the variation with the help of experimental evolution, genetic manipulation, and whole genome sequencing. Particularly, I focus on the roles of type IV secretion system and its effectors in virulence.

Understanding the cause and maintenance of biodiversity is one of the central aims of ecology. Diversity of microorganisms, including pathogen diversity, can be shaped by both deterministic and stochastic factors. In the fourth chapter, I take a step back to look at different assembly processes in soil microbiomes. Using the neutral and

null models, I breakdown the relative importance of stochastic and deterministic factors in shaping the soil microbiomes for samples collected across Kansas.

In the final chapter, I look back at the findings of the dissertation and discuss how it helps in moving the field forward. I discuss some of the outstanding questions in the field of virulence evolution and propose new directions. The chapters of this dissertation are either already published or will soon be submitted for publication. As such, the reference formatting for each chapter will vary depending on where it was submitted or will soon be submitted.

Chapter 1 - Trade-offs and heterogeneous selection pressure in virulence evolution

Aakash Pandey[1] and Thomas G. Platt[1]

[1]Division of Biology, Manhattan, KS, USA

Introduction

Fitness tradeoffs are central to the evolution of infectious diseases and the virulence stemming from host-pathogen interactions (Alizon and Michalakis 2015). Virulence is often a key determinant of the pathogen fitness. Virulence evolution can be constrained by trade-offs with other traits that influence pathogen fitness. Pathogens have diverse life-histories and the context in which trade-offs arise varies depending on pathogen ecology and life history. Accordingly, diverse trade-offs and trade-off strengths are assumed in models of pathogen evolution. Understanding the ecological and physiological mechanisms underlying these trade-offs will aid the development of predictive models of pathogen evolution. In this review, we examine how several types of pathogen life-histories give rise to different evolutionary trade-offs and how these trade-offs influence the fitness of the pathogen.

Natural selection results in trait evolution either because of the trait directly influences fitness or because it is correlated with a trait that is linked to individual fitness (Lande and Arnold 1983). Most models of virulence evolution assume a direct influence of virulence on pathogen fitness. The fitness of a pathogen is often approximated by its basic reproduction number (R_0). This metric has several limitations and consequently an alternative metric, invasion fitness, is often used to predict long term evolutionary outcomes (Lion and Metz 2018). Each of these fitness measures describes fitness as a function of virulence trait values, and thereby enables the prediction of responses to selection. Correlation between virulence and other traits can result in conflicting selective pressures acting on different components of pathogen fitness. For example, the virulence-transmission trade-off stemming from a positive correlation between virulence and

transmission (Alizon et al. 2009) results in conflicting selective pressures at within and between host levels. Balancing these conflicting selection pressures results in intermediate virulence (Mideo et al. 2008). Correlation with other pathogen traits can similarly result in conflicting pressures that can shape evolutionary outcomes (Pandey et al. 2022). While these models assume different trade-offs, exact mechanisms of how these trade-offs arise are often overlooked in virulence evolution studies.

Trade-offs and heterogeneous selection pressure

Population structure can result in heterogeneous selection pressures (Ohtsuki et al. 2020). Several factors shape the structure of host and pathogen populations. Host populations can be structured based on the age, sex, immune status, and other attributes of hosts. Host age distribution can shape both epidemiological and evolutionary dynamics of pathogens. For example, young and old populations may show different disease dynamics due to differences in the immune status of individuals in the two populations (Izhar and Ben-Ami 2015; Clark et al. 2017; Davies et al. 2020; Glynn and Moss 2020). These distinct ecological dynamics my result in distinct age/stage specific parasite virulence strategies (Iritani et al. 2019; Hamley and Koella 2021). Immune response can also vary with sex thereby resulting in differential susceptibility of males and females to infectious diseases (Klein and Flanagan 2016). This differential response results in evolution of sex specific virulence (Úbeda and Jansen 2016; Hall and Mideo 2018). Additionally, sex and age can interact to modify the transmission-virulence trade-off leading to complex patterns of age and sex specific virulence (Gipson and Hall 2018). Prior infection and vaccination also alter immune status and consequently many pathogens evolve in response to vaccines (Gandon and Day 2008; Mackinnon et al.

2008). Both theoretical and empirical results show that imperfect vaccines, those that reduce mortality but allow transmission, can select for higher virulence (Read et al. 2015; Miller and Jessica Metcalf 2019; Miller et al. 2022). Likewise, immune mediated selection for antigenic variant is a well-known problem in influenza (Hannoun 2014; Kim et al. 2018; Oidtman et al. 2021). Repeated epidemics due to antigen escape can in turn select for higher virulence (Mideo and Kamiya 2021; Sasaki et al. 2021).

Heterogeneity can also result due to variation in host genetics. Pathogens interacting with a heterogenous host population may have diverse evolutionary outcomes depending on the mechanisms by which host variation impacts pathogen fitness. Some scenarios may result in specialization on specific host types while others may result in a polymorphic pathogen populations or an overall reduction in virulence (Regoes et al. 2000; Osnas and Dobson 2012; Fleming-Davies et al. 2015; Lievens et al. 2018; González et al. 2019; Signe White et al. 2020; Ekroth et al. 2021).

The life-history of many pathogens results in heterogeneous selective pressures on pathogens (Figure 1.1). In vector borne pathogens, the distinct ecology and physiology of vectors and hosts present the pathogen with different environments, likely resulting in heterogeneous selective pressures. For example, in the malarial parasite *Plasmodium falciparum*, within-host selection favors asexual replication, but these cells cannot reproduce within the vector mosquito. Gametocytes are required for vector infection and onward transmission of the pathogen. However, the asexual reproduction trades off with gametocyte production resulting in heterogeneous selective pressures (Greischar et al. 2016; Rono et al. 2018). Likewise, pathogens infecting multi-hosts can experience heterogeneous selective pressures in different species they infect, especially if they

mount distinct immune responses. Consequently, evolutionary virulence outcomes will depend upon the frequency with which the pathogen encounters each host type and the relative importance of each interaction to pathogen fitness (Gandon 2004). Potential trade-offs in pathogen's performance between host types can also influence virulence evolution (Gandon 2004; Osnas and Dobson 2012). While empirical studies on these trade-offs are scant, serial passage experiments demonstrate that adaptation to a new host can lead to virulence attenuation in the original host (Rafaluk et al. 2015).

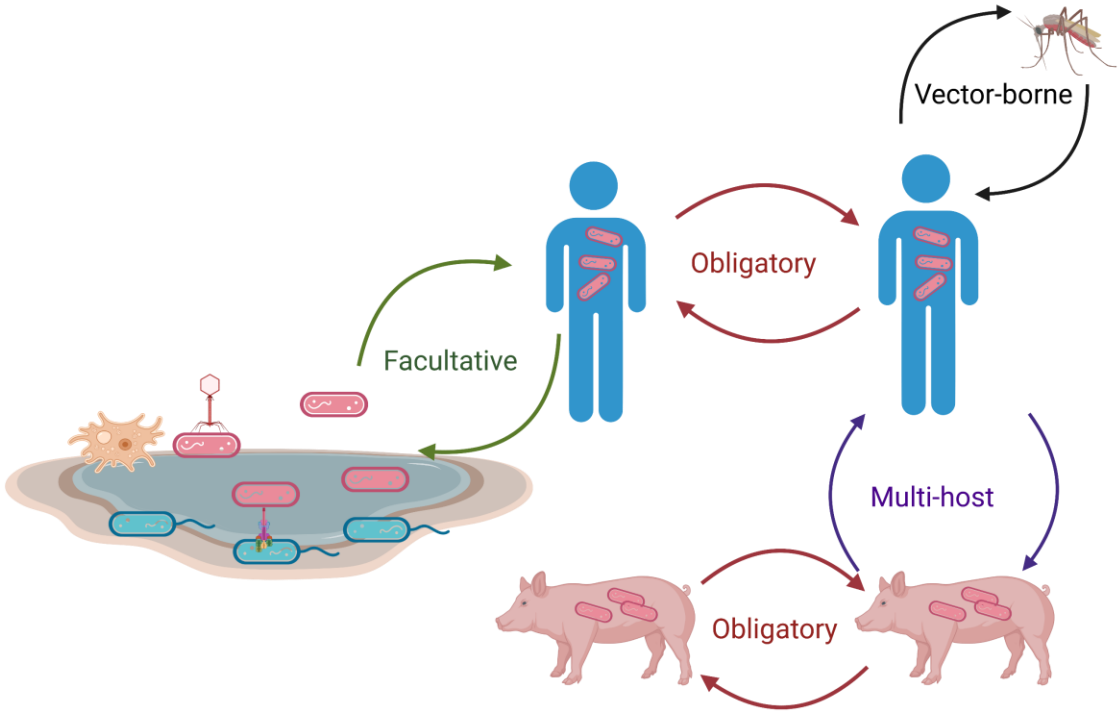

Figure 1.1. Diversity in pathogen's life-histories. Heterogeneous selection pressure can arise in each of these pathogens' classes because of the structure of both hosts and pathogen populations. Illustration created with BioRender.com.

Facultative and opportunistic pathogens maintain populations in environmental reservoirs or in commensal sites within hosts (Brown et al. 2012). These pathogens experience widely varying environments, where disease is manifested in some environments (e.g., those that are host-associated) but not all (e.g., environmental reservoirs or commensal sites within a host) (Figure 1.2). In microbes where populations alternate between two different environments, fitness is optimized across the lifecycle (Bansept et al. 2021; Obeng et al. 2021; Pandey et al. 2022). The structured populations of facultative pathogens exhibit different population dynamics within host/disease sites and reservoir/commensal sites thereby resulting in heterogeneous selective pressures. Virulence evolution in these populations in part depends on correlations between virulence and traits that influence fitness in non-disease contexts (Brown et al. 2012; Pandey et al. 2022). If virulence is positively correlated with survival in the environment, selection in the reservoir populations results in higher virulence, whereas a negative correlation favors the evolution of low virulence or avirulence. Further, there may be multiple trade-offs impacting the evolution of facultative pathogen virulence. Such scenarios can lead to diverse outcomes such as evolutionary bi-stability and evolutionary diversification in virulence trait (Pandey et al. 2022).

Trade-offs depend on genetic correlations due to pleiotropy

Pleiotropy occurs when a gene affects multiple phenotypes (Stearns 2010). Trade-offs or trait correlations can result when two traits share underlying molecular mechanisms. Fitness trade-offs reflect instances where a change in one trait that positively affects fitness results in a change in a correlated trait that negatively affects fitness. Such antagonistic pleiotropy has long been recognized as a major source of

trade-offs in life-history evolution (Williams 1957; Stearns 1989; Roff and Fairbairn 2007). Virulence factors can affect multiple pathogen traits, sometimes resulting in opposing fitness consequences depending on the context of infection. That is, antagonistic pleiotropy can underlie pathogen fitness trade-offs.

Pleiotropy may reflect a single regulatory protein controlling the expression of multiple genes. The transcriptional repressor CodY pleiotropically controls the expression of virulence factors in multiple Gram-positive pathogens (Stenz et al. 2011). In *Bacillus anthracis*, the causative agent of anthrax, CodY influences the expression of three genes (*pagA*, *lef*, and *cya*) that encode the tripartite toxin, a key virulence factor (van Schaik et al. 2009). CodY represses many genes during exponential growth phase thereby promoting cell replication. However, CodY also activates toxin synthesis during stationary phase resulting in context dependent phenotypic consequences. In contrast, CodY represses toxin production in *Clostridium difficile* during both exponential and stationary phase (Dineen et al. 2007).

Fitness trade-offs can result from antagonistic pleiotropy during niche expansion, a phenomenon also known as the cost of generalism. Support for this hypothesis comes from experimental evolution (Elena and Lenski 2003; Duffy et al. 2006; Bono et al. 2020), where adaptation to a new host or environment comes at the cost of reduced performance in the original host or environment. This principle is exploited to produce live attenuated vaccines by serial passaging the pathogen in atypical cells/animals, resulting in reduced virulence in the typical host (Hanley 2011; Galinski et al. 2015). Empirical support for pleiotropic cost of niche expansion also comes from experiments on the RNA bacteriophage φ6 that infects *Pseudomonas* species, where many mutations that allow

niche expansion lead to reduced fitness in the original host (Duffy et al. 2006, 2007; Bono et al. 2020). However, a subset of mutations for host-range expansion either had neutral or positive fitness effects on the original host suggesting that not all host-range expansion mutations lead to antagonistic pleiotropy.

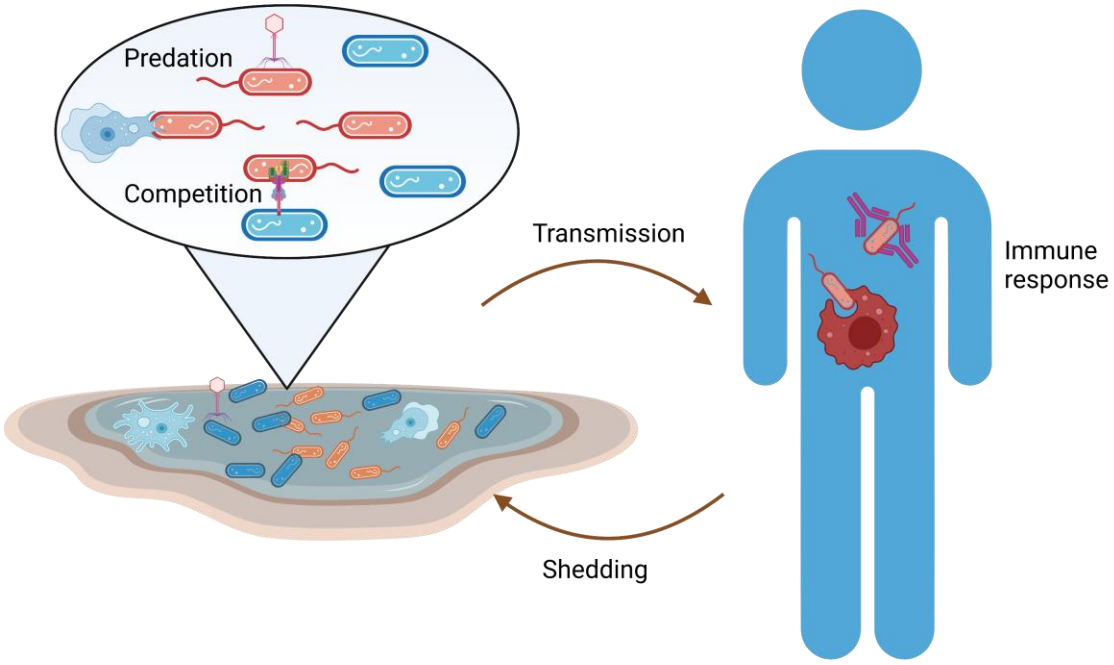

Figure 1.2. Generalized facultative pathogen life-history. Both host and environmental factors can influence pathogen traits. Illustration created with BioRender.com.

Mutation accumulation can also result in fitness trade-offs where neutral mutations accumulate in one selective environment but confer negative fitness consequences in a different environment (Kassen 2002; MacLean and Bell 2002; MacLean et al. 2004). Antagonistic pleiotropy, together with mutation accumulation, can be the source of fitness trade-offs in vector borne pathogens. For example, in vesicular stomatitis virus, adaptation to the vector cells during persistent infection leads to fitness reduction in

mammalian cells through a combination of antagonistic pleiotropy and mutation accumulation (Novella et al. 1995; Presloid et al. 2008). In some cases, a single mutation is enough to cause pleiotropic effects. A point mutation in the NS4B protein of dengue virus results in the virus having different replication rates in mosquito vector cells and mammalian cells (Hanley et al. 2003). The pleotropic mutations can sometimes be compensated by other mutations, potentially resolving the trade-off or even causing positive fitness effects in both environments (Bono et al. 2020). Virulence evolution in such cases is then driven by a positive fitness correlation across the environments rather than a fitness trade-off.

Co-incidental selection occurs when virulence factors are byproducts of adaptation in non-disease contexts, which may result in context dependent trade-offs with virulence (Levin and Edén 1990; Brown et al. 2012). Selection in environmental reservoirs may promote the expression of Shiga toxin in *E. coli*, however, its expression as a virulence factor in mammalian hosts impacts its within host fitness. In environmental reservoirs, *E. coli* is under predation pressure from protist grazers. As an antipredator response, *E. coli* infected with λ coliphage harboring the *stx* loci produces Shiga toxin (Steinberg and Levin 2007; Lainhart et al. 2009; Mauro and Koudelka 2011; Koudelka et al. 2018). Coincidently, this toxin also acts as a virulence factor within host, potentially reducing its within-host fitness. Another potential example of coincidental selection comes from the facultative pathogen *Stenotrophomonas maltophilia*. *S. maltophilia* is found in a broad range of environments including soil, wastewater, and plants. This opportunist pathogen can also infect several hosts including humans and *Caenorhabditis elegans* (White et al. 2016; White and Herman 2018). In its environmental reservoirs, *S. maltopilia* is in

competition with heterologous bacteria. To counter the competition, *S. maltophilia* uses a type IV secretion system to deliver toxins into its competitors (Bayer-Santos et al. 2019; Nas et al. 2019, 2021). The type IV secretion system and its effector proteins that promote interference competition can also influence virulence within *C. elegans* (Pandey et. al., 2022, Chapter 3). This suggests that the effectors and type IV secretion system are selected to provide competitive advantage and coincidentally involve in host-pathogen interactions. Similarly, in multi-host pathogens, selection within a host type can lead to the coincidental increase in virulence in a different host. For example, a variant of West Nile virus with high virulence in corvids may have resulted from evolution for efficient replication in American robins (Geoghegan and Holmes 2018). Understanding how virulence will be impacted because of selection in pathogen reservoirs has great importance to the spillover and epidemic potential of many zoonotic diseases (Becker et al. 2019; Guth et al. 2019).

The transmission-virulence trade-off is often conceptualized as reflecting the effects of pathogen growth rate on both virulence and transmission. Pathogen growth rate is shaped by the effects of many genes with a complex genotype to phenotype mapping. Consequently, the generality of transmission-virulence trade-off is difficult to assess, leading some to question the widespread acceptance of this trade-off (Ebert and Bull 2003). Consistent with this, empirical support for the transmission-virulence trade-off is equivocal (Acevedo et al. 2019) and pathogen growth rate and virulence can even show decoupled evolutionary dynamics (Tardy et al. 2019). Since a complete genotype to phenotype map for pathogen growth rate and other complex traits is difficult to achieve,

we must rely on empirical estimation of the genetic covariance of transmission and virulence to predict how evolution will proceed.

Context dependent fitness benefits

Though virulence factors are essential to pathogen establishment within a host, the expression of virulence factors can also impose fitness costs on pathogens (Sturm et al. 2011; Platt et al. 2012*a*). There can be strong selection against the maintenance of costly virulence factors especially in environments where the benefits are not realized (Barton et al. 2018). In addition, the benefits of virulence factor expression can change during infection. For example, *Pseudomonas aeruginosa* evolves within chronically infected hosts to reduce the fitness cost associated with the expression of virulence factors (Sousa and Pereira 2014). This is mainly due to mutations in pleiotropic regulatory genes such as *mucA*, *lasR*, *rpoL*, and *phuR* (Diard and Hardt 2017). Similar adaptation is seen with *Staphylococcus aureus* in cystic fibrosis lungs where two phenotypically different subpopulations are dominant during the acute and chronic phases of infection. The planktonic type produces toxin and is dominant during acute infection whereas biofilm forming cells dominate during chronic infection. While switching between these two phenotypes is under the control of *agr* quorum sensing system, Agr mutants are frequently isolated from chronic patients suggesting that these cells avoid the fitness cost of virulence factor expression during the chronic infection (García-Betancur et al. 2017; Davis 2020). Together with mutation accumulation and tight regulation of expression, the fitness cost stemming from virulence factor expression can also be limited by heterogeneous expression within a population. In this case, only a fraction of the

population incurs the fitness reduction (Sturm et al. 2011; Diard et al. 2013; Diard and Hardt 2017; Davis 2020).

Bacterial plasmids often encode genes that provide context dependent fitness benefits to its bacterial cell. Several enteric pathogens including *Shigella* sp. and *Salmonella* sp. harbor virulence plasmids that carry the genes required during pathogenesis (Pilla and Tang 2018). Virulence plasmids provide benefits to bacteria during infection while conferring some cost of carriage and expression. The cost of plasmid carriage depends greatly on the genetic background and can change with time (Carroll and Wong 2018). When the cost is due to genetic conflicts, single compensatory mutations are sometimes enough to reduce the cost of harboring plasmid (Hall et al. 2021). In the plant pathogen *Agrobacterium tumefaciens*, the cost of expressing virulence genes is greater than the cost of maintaining the Ti plasmid (Platt et al. 2012a). The cost of expressing virulence genes can lead reduced proliferation resulting in trade-offs during pathogenesis (Peyraud et al. 2016; Barton et al. 2018). Ti plasmids also harbor catabolic genes that can metabolize opines from plants, thereby offsetting some of the cost (Platt et al. 2012*b*). Free-loaders or cheaters that do not pay the cost of expression while reaping the benefit of opine catabolism can outcompete those that do express virulence genes. While these cheaters may have fitness advantage within host, they cannot initiate new infections leading to low between host fitness. Thus, the genetic linkage between virulence and catabolic genes can help maintain the trade-off if the cells expressing virulence genes are not driven to extinction because of fitness cost. How plasmids are maintained despite their fitness cost (a.k.a. the plasmid paradox) is an interesting research avenue that can shed light to virulence trade-offs and their evolution.

Concluding remarks

The ‘trade-off revolution’ centered trade-offs in evolutionary studies of virulence evolution. Most models of virulence evolution assume a virulence-transmission trade-off, sometimes with little justification. The nature and strength of trade-offs can differ for pathogens with different life-histories and with the context of infection. This is reflected by the fact that only a modest number of empirical observations support the trade-off hypothesis. In addition, the trade-offs themselves can evolve, thereby changing their strength and nature. Therefore, it is critical to understand i) what trait correlations and trade-offs are present, ii) how they arise, and iii) how they might evolve to better understand how virulence will evolve in different host-pathogen systems. In this review, we examined some of the mechanisms by which trade-offs arise. We discussed how pleiotropy, mutation accumulation, and genetic linkage can result in trait correlations and trade-offs. Understanding the molecular and genetic basis of trade-offs will help further the development of virulence management strategies.

References

Acevedo, M. A., F. P. Dillemuth, A. J. Flick, M. J. Faldyn, and B. D. Elderd. 2019. Virulence-driven trade-offs in disease transmission: A meta-analysis. Evolution 73:636–647.

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

Alizon, S., and Y. Michalakis. 2015. Adaptive virulence evolution: The good old fitness-based approach. Trends in Ecology and Evolution 30:248–254.

Bansept, F., N. Obeng, H. Schulenburg, and A. Traulsen. 2021. Modeling host-associating microbes under selection. The ISME Journal 2021 15:12 15:3648–3656.

Barton, I. S., C. Fuqua, and T. G. Platt. 2018. Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. Environmental Microbiology 20:16–29.

Bayer-Santos, E., W. Cenens, B. Y. Matsuyama, G. U. Oka, G. di Sessa, I. D. V. Mininel, T. L. Alves, et al. 2019. The opportunistic pathogen *Stenotrophomonas maltophilia* utilizes a type IV secretion system for interbacterial killing. PLOS Pathogens 15:e1007651.

Becker, D. J., A. D. Washburne, C. L. Faust, J. R. C. Pulliam, E. A. Mordecai, J. O. Lloyd-Smith, and R. K. Plowright. 2019. Dynamic and integrative approaches to understanding pathogen spillover. Philosophical Transactions of the Royal Society B 374.

Bono, L. M., J. A. Draghi, and P. E. Turner. 2020. Evolvability costs of niche expansion. Trends in Genetics 36:14–23.

Brown, S. P., D. M. Cornforth, and N. Mideo. 2012. Evolution of virulence in opportunistic pathogens: Generalism, plasticity, and control. Trends in Microbiology 20:336–342.

Carroll, A. C., and A. Wong. 2018. Plasmid persistence: costs, benefits, and the plasmid paradox. Canadian Journal of Microbiology 64:293–304.

Clark, J., J. S. Garbutt, L. McNally, and T. J. Little. 2017. Disease spread in age structured populations with maternal age effects. Ecology Letters 20:445–451.

Davies, N. G., P. Klepac, Y. Liu, K. Prem, M. Jit, C. A. B. Pearson, B. J. Quilty, et al. 2020. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine 2020 26:8 26:1205–1211.

Davis, K. M. 2020. For the greater (bacterial) good: Heterogeneous expression of energetically costly virulence factors. Infection and Immunity 88.

Diard, M., V. Garcia, L. Maier, M. N. P. Remus-Emsermann, R. R. Regoes, M. Ackermann, and W. D. Hardt. 2013. Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature 494:353–356.

Diard, M., and W. D. Hardt. 2017. Evolution of bacterial virulence. FEMS Microbiology Reviews 41:679–697.

Dineen, S. S., A. C. Villapakkam, J. T. Nordman, and A. L. Sonenshein. 2007. Repression of Clostridium difficile toxin gene expression by CodY. Molecular Microbiology 66:206–219.

Duffy, S., C. L. Burch, and P. E. Turner. 2007. Evolution of host specificity drives reproductive isolation among RNA viruses. Evolution 61:2614–2622.

Duffy, S., P. E. Turner, and C. L. Burch. 2006. Pleiotropic costs of niche expansion in the RNA bacteriophage Φ6. Genetics 172:751–757.

Ebert, D., and J. J. Bull. 2003. Challenging the trade-off model for the evolution of virulence. Trends in Microbiology 11:15–20.

Ekroth, A. K. E., M. Gerth, E. J. Stevens, S. A. Ford, and K. C. King. 2021. Host genotype and genetic diversity shape the evolution of a novel bacterial infection. The ISME Journal 2021 15:7 15:2146–2157.

Elena, S. F., and R. E. Lenski. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nature Reviews Genetics 2003 4:6 4:457–469.

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

Galinski, M. S., K. Sra, J. I. Haynes, and J. Naspinski. 2015. Live attenuated viral vaccines. Pages 1–44 *in* B. Nunnally, V. Turula, and R. Sitrin, eds. Vaccine Analysis: Strategies, Principles, and Control. Springer, Berlin, Heidelberg.

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

Gandon, S., and T. Day. 2008. Evidences of parasite evolution after vaccination. Vaccine 26:C4–C7.

García-Betancur, J. C., A. G. Moreno, T. Horger, M. Schott, M. Sharan, J. Eikmeier, B. Wohlmuth, et al. 2017. Cell differentiation defines acute and chronic infection cell types in *Staphylococcus aureus*. eLife 6.

Geoghegan, J. L., and E. C. Holmes. 2018. The phylogenomics of evolving virus virulence. Nature Reviews Genetics 2018 19:12 19:756–769.

Gipson, S. A. Y., and M. D. Hall. 2018. Interactions between host sex and age of exposure modify the virulence–transmission trade-off. Journal of Evolutionary Biology 31:428–437.

Glynn, J. R., and P. A. H. Moss. 2020. Systematic analysis of infectious disease outcomes by age shows lowest severity in school-age children. Scientific Data 2020 7:1 7:1–13.

González, R., A. Butković, and S. F. Elena. 2019. Role of host genetic diversity for susceptibility-to-infection in the evolution of virulence of a plant virus. Virus Evolution 5.

Greischar, M. A., N. Mideo, A. F. Read, and O. N. Bjørnstad. 2016. Predicting optimal transmission investment in malaria parasites. Evolution 70:1542–1558.

Guth, S., E. Visher, M. Boots, and C. E. Brook. 2019. Host phylogenetic distance drives trends in virus virulence and transmissibility across the animal–human interface. Philosophical Transactions of the Royal Society B 374.

Hall, J. P. J., R. C. T. Wright, E. Harrison, K. J. Muddiman, A. J. Wood, S. Paterson, and M. A. Brockhurst. 2021. Plasmid fitness costs are caused by specific genetic conflicts enabling resolution by compensatory mutation. PLOS Biology 19:e3001225.

Hall, M. D., and N. Mideo. 2018. Linking sex differences to the evolution of infectious disease life-histories. Philosophical Transactions of the Royal Society B: Biological Sciences 373.

Hamley, J. I. D., and J. C. Koella. 2021. Parasite evolution in an age-structured population. Journal of Theoretical Biology 527:110732.

Hanley, K. A. 2011. The double-edged sword: How evolution can make or break a live-attenuated virus vaccine. Evolution 4:635.

Hanley, K. A., L. R. Manlucu, L. E. Gilmore, J. E. Blaney, C. T. Hanson, B. R. Murphy, and S. S. Whitehead. 2003. A trade-off in replication in mosquito versus mammalian

systems conferred by a point mutation in the NS4B protein of dengue virus type 4. Virology 312:222–232.

Hannoun, C. 2014. The evolving history of influenza viruses and influenza vaccines. Expert Review of Vaccines 12:1085–1094.

Iritani, R., E. Visher, and M. Boots. 2019. The evolution of stage-specific virulence: Differential selection of parasites in juveniles. Evolution Letters 3:162–172.

Izhar, R., and F. Ben-Ami. 2015. Host age modulates parasite infectivity, virulence and reproduction. Journal of Animal Ecology 84:1018–1028.

Kassen, R. 2002. The experimental evolution of specialists, generalists, and the maintenance of diversity. Journal of Evolutionary Biology 15:173–190.

Kim, H., R. G. Webster, and R. J. Webby. 2018. Influenza virus: Dealing with a drifting and shifting pathogen. Viral Immunology 31:174–183.

Klein, S. L., and K. L. Flanagan. 2016. Sex differences in immune responses. Nature Reviews Immunology 2016 16:10 16:626–638.

Koudelka, G. B., J. W. Arnold, and D. Chakraborty. 2018. Evolution of STEC virulence: Insights from the antipredator activities of Shiga toxin producing *E. coli*. International Journal of Medical Microbiology 308:956–961.

Lainhart, W., G. Stolfa, and G. B. Koudelka. 2009. Shiga toxin as a bacterial defense against a eukaryotic predator, *Tetrahymena thermophila*. Journal of Bacteriology 191:5116–5122.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Levin, B. R., and C. S. Edén. 1990. Selection and evolution of virulence in bacteria: an ecumenical excursion and modest suggestion. Parasitology 100:S103–S115.

Lievens, E. J. P., J. Perreau, P. Agnew, Y. Michalakis, and T. Lenormand. 2018. Decomposing parasite fitness reveals the basis of specialization in a two-host, two-parasite system. Evolution Letters 2:390–405.

Lion, S., and J. A. J. Metz. 2018. Beyond R_0 maximisation: on pathogen evolution and environmental dimensions. Trends in Ecology & Evolution 33:458–473.

Mackinnon, M. J., S. Gandon, and A. F. Read. 2008. Virulence evolution in response to vaccination: The case of malaria. Vaccine 26:C42–C52.

MacLean, R. C., and G. Bell. 2002. Experimental adaptive radiation in *Pseudomonas*. American Naturalist 160:569–581.

MacLean, R. C., G. Bell, and P. B. Rainey. 2004. The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. Proceedings of the National Academy of Sciences of the United States of America 101:8072–8077.

Mauro, S. A., and G. B. Koudelka. 2011. Shiga toxin: Expression, distribution, and its role in the rnvironment. toxins 3:608–625.

Mideo, N., S. Alizon, and T. Day. 2008. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. Trends in Ecology and Evolution 23:511–517.

Mideo, N., and T. Kamiya. 2021. Antigenic evolution can drive virulence evolution. Nature Ecology & Evolution 2021 6:1 6:24–25.

Miller, I. F., C. Jessica, and E. Metcalf. 2022. Assessing the risk of vaccine-driven virulence evolution in SARS-CoV-2. Royal Society Open Science 9.

Miller, I. F., and C. Jessica Metcalf. 2019. Vaccine-driven virulence evolution: consequences of unbalanced reductions in mortality and transmission and implications for pertussis vaccines. Journal of the Royal Society Interface 16.

Nas, M. Y., J. Gabell, and N. P. Cianciotto. 2021. Effectors of the *Stenotrophomonas maltophilia* type IV secretion system mediate killing of clinical isolates of *Pseudomonas aeruginosa*. mBio 12:e01502-21.

Nas, M. Y., R. C. White, A. L. DuMont, A. E. Lopez, and N. P. Cianciotto. 2019. *Stenotrophomonas maltophilia* encodes a VirB/VirD4 Type IV Secretion System that modulates apoptosis in human cells and promotes competition against heterologous bacteria, including *Pseudomonas aeruginosa*. Infection and Immunity 87:10–12.

Novella, I. S., D. K. Clarke, † Josep Quer, E. A. Duarte, C. H. Lee, S. C. Weaver, § Santiago, et al. 1995. Extreme fitness differences in mammalian and insect hosts after continuous replication of vesicular stomatitis virus in sandfly cells. Journal of Virology 69:6805–6809.

Obeng, N., F. Bansept, M. Sieber, A. Traulsen, and H. Schulenburg. 2021. Evolution of microbiota–host associations: The microbe's perspective. Trends in Microbiology 29:779–787.

Ohtsuki, H., C. Rueffler, J. Y. Wakano, K. Parvinen, and L. Lehmann. 2020. The components of directional and disruptive selection in heterogeneous group-structured populations. Journal of Theoretical Biology 507:110449.

Oidtman, R. J., P. Arevalo, Q. Bi, L. McGough, C. J. Russo, D. Vera Cruz, M. Costa Vieira, et al. 2021. Influenza immune escape under heterogeneous host immune histories. Trends in Microbiology 29:1072–1082.

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

Pandey, A., N. Mideo, and T. G. Platt. 2022. Virulence evolution of pathogens that can grow in reservoir environments. The American Naturalist 199:141–158.

Peyraud, R., L. Cottret, L. Marmiesse, J. Gouzy, and S. Genin. 2016. A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. PLOS Pathogens 12:e1005939.

Pilla, G., and C. M. Tang. 2018. Going around in circles: virulence plasmids in enteric pathogens. Nature Reviews Microbiology 2018 16:8 16:484–495.

Platt, T. G., J. D. Bever, and C. Fuqua. 2012*a*. A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. Proceedings of the Royal Society B: Biological Sciences 279:1691–1699.

Platt, T. G., C. Fuqua, and J. D. Bever. 2012*b*. Resource and competitive dynamics shape the benefits of public goods cooperation in a plant pathogen. Evolution; international journal of organic evolution 66:1953.

Presloid, J. B., B. E. Ebendick-Corpus, S. Zárate, and I. S. Novella. 2008. Antagonistic pleiotropy involving promoter sequences in a virus. Journal of Molecular Biology 382:342–352.

Rafaluk, C., G. Jansen, H. Schulenburg, and G. Joop. 2015. When experimental selection for virulence leads to loss of virulence. Trends in Parasitology 31:426–434.

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

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

Roff, D. A., and D. J. Fairbairn. 2007. The evolution of trade-offs: where are we? Journal of evolutionary biology 20:433–47.

Rono, M. K., M. A. Nyonda, J. J. Simam, J. M. Ngoi, S. Mok, M. M. Kortok, A. S. Abdullah, et al. 2018. Adaptation of *Plasmodium falciparum* to its transmission environment. Nature Ecology and Evolution 2:377–387.

Sasaki, A., S. Lion, and M. Boots. 2021. Antigenic escape selects for the evolution of higher pathogen transmission and virulence. Nature Ecology & Evolution 2021 6:1 6:51–62.

Signe White, P., A. Choi, R. Pandey, A. Menezes, M. Penley, A. K. Gibson, J. de Roode, et al. 2020. Host heterogeneity mitigates virulence evolution. Biology Letters 16.

Sousa, A. M., and M. O. Pereira. 2014. *Pseudomonas aeruginosa* diversification during infection development in cystic fibrosis lungs—A review. Pathogens 2014, Vol. 3, Pages 680-703 3:680–703.

Stearns, F. W. 2010. One hundred years of pleiotropy: A retrospective. Genetics 186:767–773.

Stearns, S. C. C. 1989. Trade-offs in life-history evolution. Functional Ecology 3:259–268.

Steinberg, K. M., and B. R. Levin. 2007. Grazing protozoa and the evolution of the *Escherichia coli* O157:H7 Shiga toxin-encoding prophage. Proceedings of the Royal Society B: 274:1921–1929.

Stenz, L., P. Francois, K. Whiteson, C. Wolz, P. Linder, and J. Schrenzel. 2011. The CodY pleiotropic repressor controls virulence in gram-positive pathogens. FEMS Immunology & Medical Microbiology 62:123–139.

Sturm, A., M. Heinemann, M. Arnoldini, A. Benecke, M. Ackermann, M. Benz, J. Dormann, et al. 2011. The cost of virulence: Retarded growth of *Salmonella typhimurium* cells expressing Type III secretion system 1. PLoS Pathogens 7:e1002143.

Tardy, L., M. Giraudeau, G. E. Hill, K. J. McGraw, and C. Bonneaud. 2019. Contrasting evolution of virulence and replication rate in an emerging bacterial pathogen. Proceedings of the National Academy of Sciences of the United States of America 116:16927–16932.

Úbeda, F., and V. A. A. Jansen. 2016. The evolution of sex-specific virulence in infectious diseases. Nature Communications 2016 7:1 7:1–9.

van Schaik, W., A. Château, M. A. Dillies, J. Y. Coppée, A. L. Sonenshein, and A. Fouet. 2009. The global regulator CodY regulates toxin gene expression in Bacillus anthracis and is required for full virulence. Infection and Immunity 77:4437–4445.

White, C. v, B. J. Darby, R. J. Breeden, and M. A. Herman. 2016. A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. Infection and Immunity 84:524–536.

White, C. v., and M. A. Herman. 2018. Transcriptomic, functional, and network analyses reveal novel genes involved in the interaction between *Caenorhabditis elegans* and *Stenotrophomonas maltophilia*. Frontiers in Cellular and Infection Microbiology 8:266.

Williams, G. C. 1957. Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411.

Chapter 2 - Virulence evolution of pathogens that can grow in environmental reservoirs

Aakash Pandey[1*], Nicole Mideo[2], Thomas G. Platt[1].

1Division of Biology, Kansas State University, Manhattan, KS, USA

2Department of Ecology and Evolution, University of Toronto, Toronto, ON, Canada

*Corresponding author: apandey@ksu.edu

Published as: Pandey, A., N. Mideo, and T. G. Platt. 2022. Virulence evolution of pathogens that can grow in reservoir environments. *The American Naturalist* 199:141–158.

Abstract

Many pathogens reside in environmental reservoirs within which they can reproduce and from which they can infect hosts. These facultative pathogens experience different selective pressures in host-associated environments and reservoir environments. Heterogeneous selective pressures have the potential to influence the virulence evolution of these pathogens. Previous research has examined how environmental transmission influences the selective pressures shaping the virulence of pathogens that cannot reproduce in environmental reservoirs, yet many pathogens of human, crop plants, and livestock can reproduce in these environments. We build on this work to examine how reproduction in reservoirs influences disease dynamics and virulence evolution in a simple facultative pathogen model. We use adaptive dynamics to examine the evolutionary dynamics of facultative pathogens under potential trade-offs between transmission and virulence, shedding and virulence, and reservoir persistence and virulence. We then perform critical function analysis to generalize the results independent of specific trade-off assumptions. We determine that diverse virulence strategies, sometimes resulting from evolutionary bistability or evolutionary branching conditions, are expected for facultative pathogens. Our findings motivate research establishing which trade-offs most strongly influence the virulence evolution of facultative pathogens.

Introduction

Life history trade-offs reflect constraints that ensure that changes in one trait result in fitness costs associated with another trait (Charnov 1989; Stearns 1989; Roff 2000). Pathogens can experience trade-offs associated with various traits that influence their fitness. Many pathogens reproduce within hosts and in environmental reservoirs and consequently experience heterogeneous selective pressures. These facultative pathogens include environmental opportunists like Vibrio cholerae and Bacillus anthracis, the aetiological agents of cholera and anthrax, and commensal opportunists like Candida albicans and Streptococcus pneumoniae, the aetiological agents of candidiasis and pneumococcal diseases (Brown et al. 2012). Most theoretical studies of virulence evolution focus on trade-offs experienced by obligate pathogens (Anderson and May 1978; Ewald 1983; van Baalen and Sabelis 1995; Frank 1996; Alizon et al. 2009; Cressler et al. 2016). The heterogeneous selective pressures faced by pathogens that can infect from environmental reservoir populations are thought to relax the virulence-transmission trade-off, resulting in higher overall virulence (Ewald 1983, 1991a, 1991b; Bonhoeffer et al. 1996; Gandon 1998; Day 2002a; Walther and Ewald 2004; Roche et al. 2011; Boldin and Kisdi 2012). Selection in environments where pathogens do not cause disease can influence the virulence evolution of opportunistic pathogens when there are correlations between the pathogen's population dynamics in these environments and in environments where they do cause disease (Brown et al. 2012). However, no theoretical study of virulence evolution has allowed for pathogen replication in the environmental reservoir, though some studies have examined the epidemiological dynamics of facultative

pathogens (Godfray et al. 1999; Bani-Yaghoub et al. 2012; Kaitala et al. 2017; Lanzas et al. 2019).

The environment in which facultative pathogens live and reproduce varies over space and time—with host and reservoir environments imposing different requirements for survival and reproduction. Consequently, these organisms routinely experience heterogeneous selective pressures (Caraco and Wang 2008; Mikonranta et al. 2012, 2015; Barton et al. 2018). The importance of heterogeneous selection pressures on virulence evolution has already been highlighted in the context of multi-host pathogen systems (Regoes et al. 2000; Gandon 2004). Owing to their life history, with a population stage structured in host and reservoir environments, novel trait correlations and trade-offs can impact the course of virulence evolution in these pathogens (Sokurenko et al. 2006; Brown et al. 2012; Barton et al. 2018). For example, traits influencing the environmental persistence of free-living pathogens can be correlated with virulence, potentially resulting in a trade-off between these functions. Consistent with a relationship between virulence and anti-predation traits that promote environmental persistence, the experimental evolution of the opportunistic pathogen Serratia marcescens in the presence of a protozoan predator results in reduced virulence (Friman et al. 2009; Mikonranta et al. 2012). Similarly, the evolution of a facultative fish pathogen Flavobacterium columnarae in the presence of lytic phages similarly reduces virulence (Laanto et al. 2012). Due to associations between virulence and traits that affect functions in reservoir environments, selection in these environments likely influences the degree of virulence facultative pathogens exhibit.

The importance of reservoir dynamics to the evolution of virulence may depend on the context-dependent fitness consequences of virulence factors. Costs stemming from the expression of virulence factors or the maintenance of genes encoding virulence factors (Sturm et al. 2011; Platt et al. 2012; Peyraud et al. 2016; Pilla et al. 2017) may adversely affect the fitness of facultative pathogens in reservoir environments—where the benefits of pathogenesis are absent. For this reason, costs associated with virulence factors may result in a negative correlation between virulence and environmental persistence—a relationship that may influence the virulence evolution of facultative pathogens. The virulence factors of some facultative pathogens may alternatively have a function promoting the persistence of the pathogen in the environmental reservoir. For example, Shiga toxin produced by E. coli and Shigella sp. pathogens, may function in both promoting virulence and resistance to grazing by protozoan predators (Steinberg and Levin 2007; Adiba et al. 2010). Similarly, virulence factors may mediate interspecific competition in the reservoir environments in addition to mediating the virulence of facultative pathogens (Anttila et al. 2013). For example, in a facultative pathogen Stenotrophomonas maltophilia, a type 4 secretion system is involved in competition against heterologous bacteria and is also responsible for apoptosis of human macrophages such that a shared mechanism links virulence and competitive ability (Bayer-Santos et al. 2019; Nas et al. 2019). The dual function of these virulence factors, reflecting either pre-adaptation or coincidental selection of traits impacting virulence (Levin 1996; Brown et al. 2012), may result in a positive correlation between virulence and environmental persistence of these facultative pathogens.

One main challenge for studying the virulence evolution in facultative pathogens is adequately defining pathogen fitness, since infecting a new host is no longer the only way for a pathogen to reproduce. Although the basic reproduction number is used as a fitness proxy in most theoretical studies of virulence evolution, it is not a suitable fitness proxy in many epidemiological settings (Lion and Metz 2018). In the case of facultative pathogens, the basic reproduction number is an inadequate fitness metric because it omits the consequences of reservoir dynamics. In this study, we consider a simple epidemiological model of a facultative pathogen where we highlight the challenges associated with deriving and interpreting the basic reproduction number. We then derive the invasion fitness for the system of facultative pathogens using the adaptive dynamics framework to examine the evolutionary outcomes for virulence, influenced by potential trade-offs between that trait and transmission, shedding, or persistence. We begin by considering individual trade-offs and later explore evolutionary outcomes when there are multiple trade-offs. We then perform critical function analysis which allows us to identify evolutionary outcomes independent of specific trade-off assumptions. Evolution under individual trade-offs can result in intermediate or high virulence if the fitness gain from increasing the correlated trait (transmission, shedding, or persistence) more than compensates the fitness loss due to increased virulence. Otherwise, low virulence or avirulence is expected. More diverse evolutionary outcomes are predicted when we consider multiple, simultaneous trade-offs influencing virulence evolution. We observed bistability conditions under a wide range of parameters. Similarly, evolutionary branching points occur in multiple scenarios suggesting that diverse virulence strategies are

possible in facultative pathogens. Hence our work lays the foundation for a more system specific exploration of virulence evolution in facultative pathogens.

Model

We extended the modeling approach described by Boldin & Kisdi (2012) to capture a system in which a pathogen can be transmitted both directly from hosts and indirectly through the environment by allowing the pathogen to reproduce in the environment. In so doing, our model describes the population dynamics of facultative pathogens such as *V. cholerae*, *B. anthracis*, and *S. maltophilia*. This model describes a system composed of susceptible hosts (*S*), infected hosts (*I*), and a pathogen reservoir (*P*). The pathogen is present in both infected hosts and an environmental reservoir where it reproduces. Susceptible hosts can become infected following contact with either infected hosts or the environmental reservoir. The relative contribution of each transmission route to the disease dynamics likely differs across disease systems. In some cases, both transmission routes can be significant (e.g., cholera, Tien et al. 2011) whereas in many other water-borne pathogens, environmental transmission is likely more important compared to direct transmission (Julian 2016). Hence, we include both transmission routes to keep the model general. The flow diagram in Figure 2.1 illustrates the relationships among the *S*, *I*, and *P* compartments.

The following system of ordinary differential equations describes the system illustrated by Figure 2.1,

$$\frac{dS}{dt} = b - \beta SI - \gamma S\frac{\phi P}{P+\tau} - dS$$

$$\frac{dI}{dt} = \beta SI + \gamma S\frac{\phi P}{P+\tau} - (d + \rho + \alpha)I$$

$$\frac{dP}{dt} = \theta I - \sigma P + rP\left(1 - \frac{P}{K}\right) \quad (1).$$

We assume that susceptible hosts are born at a constant rate (*b*) and die at a per capita rate (*d*), with direct infection following mass action incidence with rate constant β. The recovery rate is given by ρ and we assume life-long immunity once the host is recovered. Although we assume ρ to be zero in simulation results, we have not dropped it from analytic expressions to make them applicable to broader scenarios. We assume that the contact rate between susceptible hosts and pathogens in the environmental reservoir is an increasing but saturating function of *P*, described by $\frac{\phi P}{P+\tau}$. Although there is little empirical data for any system with which to define an exact analytical form for environmental transmission, this form of environmental transmission is used by several studies because it incorporates concentration dependent pathogen transmission from environment (Codeço 2001) and is analytically tractable (Roche et al. 2011; Boldin and Kisdi 2012; Breban 2013). For simplicity, we assume $\phi = 1$ and $\tau = 1$. γ represents the probability of environmental transmission upon contact with pathogens in the reservoir. The loss of free-living pathogens from the reservoir due to the uptake by host is assumed to be negligible. Free-living pathogen population dynamics depend on shedding from infected hosts at rate θ, death of free-living pathogen propagules at rate σ, and density dependent (specifically, logistic) reproduction of pathogen propagules at per capita rate r.

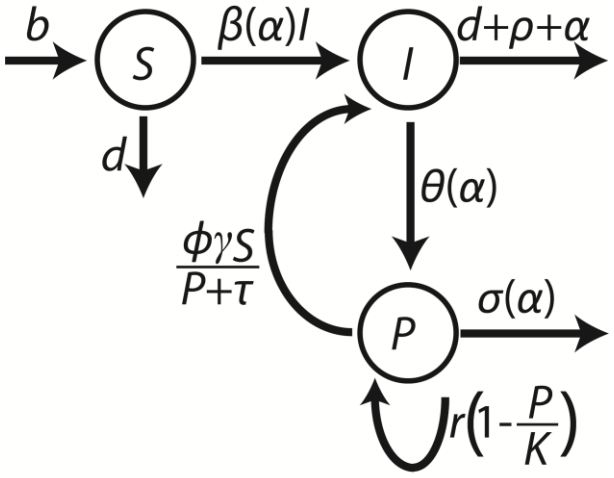

Figure 2.1. Schematic representation of the model described by system (*1*). *S* and *I* represent the number of susceptible and infected hosts, respectively. *P* represents the number of free-living propagules in an environmental reservoir. See Table 1 for variable and parameter definitions.

Trade-offs are represented by functional relationships between traits. As in Boldin and Kisdi (2012), we consider the trade-off between (direct) transmission and virulence where transmission is a saturating function of virulence, $\beta(\alpha) = \frac{c\alpha}{c\alpha+h}$. While several studies have documented such a trade-off, a recent meta-analysis identified that there have been insufficient empirical studies evaluating trade-offs between virulence and transmission (Acevedo et al. 2019). Further, we know of no studies directly evaluating this relationship for a facultative pathogen. If virulence is directly correlated with the pathogen population size within a host, then the degree of pathogen shedding by an infected host may depend on virulence. For example, virulence and duration of infection is shown to be correlated with within-host pathogen number in the facultative pathogen *S. maltophilia* (White et al. 2016). High virulence strains cause greater intestinal distension and are potentially shed from hosts at a higher rate, either continuously during

infection or abruptly following death of the host. We consider the case of continuous shedding in our model. As in direct transmission-virulence trade-off, we consider a case where the shedding rate increases monotonically with virulence. Specifically, we assume the following functional relationship between shedding and virulence, $\theta(\alpha) = \frac{g\alpha}{g\alpha+f}$.

We also consider two potential relationships between reservoir persistence (i.e., the inverse of pathogen death rate in the reservoir) and virulence. First, strategies that enhance a facultative pathogen's survival in the reservoir can also result in higher virulence within a host as a result of preadaptation or coincidental evolution. For example, antipredator adaptations in the reservoir environment could function as virulence factors within a host (Adiba et al. 2010; Robino et al. 2019). As a result, pathogen persistence and virulence may trade-off since increased persistence positively impacts pathogen fitness whereas virulence results in a fitness cost due to death of the host, reducing the infectious period. We assume that the pathogen death rate in the reservoir decreases (equivalently, persistence increases) as a function of virulence and is given as, $\sigma(\alpha) = e^{-u\alpha} + v$. Under this relationship, the free-living propagules of a strain that does not harm the host $(\alpha = 0)$ are expected to survive in the environmental reservoir $\frac{1}{1+v}$ units of time, whereas this time period approaches $\frac{1}{v}$ units of time as virulence increases. Caraco and Wang (2008) consider a similar functional relationship where the free-living pathogen persistence increases with increasing virulence. Second, we consider a different functional relationship where increasing virulence is associated with higher free-living pathogen death rate. This negative correlation between virulence and pathogen persistence in the reservoir environment can occur due to energetic constraints where the pathogen's investment in one environment comes at the cost of reduced performance

in another. Specifically, we assume $\sigma(\alpha) = \frac{a}{e^{-u\alpha}+v}$. In all cases, *a, c, f, g, h, u, ϕ, τ* and *v* are constants and are positive. Although our simulation results depend on the functional relationships considered, qualitative interpretation of our analytical results are not restricted to specific functional forms.

Evolutionary outcomes in the adaptive dynamics framework are critically dependent on the assumed trade-off shapes. Hence, to identify general results for trade-offs of arbitrary shapes, we follow Kisdi (2006) to perform critical function analysis. The critical function is derived from the invasion fitness, $W(\alpha, \alpha_m)$. Singular strategies occur when the trait correlation is locally tangent to a solution of the critical function. The nature of a singular strategy is determined by evaluating the following second derivatives:

$$E = \frac{\partial^2}{\partial \alpha_m^2} W(\alpha, \alpha_m)\Big|_{\alpha_m=\alpha=\alpha*}$$

$$M = \frac{\partial^2}{\partial \alpha \partial \alpha_m} W(\alpha, \alpha_m)\Big|_{\alpha_m=\alpha=\alpha^*}.$$

Here, α_m represents the mutant trait value whereas α^* is the singular trait. The singular trait is convergence stable if $E + M < 0$ and evolutionary stable if $E < 0$. Evolutionary branching points occur when the singular strategy is convergence stable but not evolutionary stable. Results from the critical function analysis for a case of multiple trade-offs is presented in the results section (e.g., Figure 2.8) whereas similar analyses for other scenarios are presented in the Supplementary Materials (Section A7, Figures S2.5, and Figure S2.6).

Results

Epidemiological results

System (1) has a disease-free equilibrium (DFE) point (S_0, I_0, P_0) given by ($\frac{b}{d}$,0,0). The stability of this equilibrium is determined by the eigenvalue of the Jacobian evaluated at this DFE. If the eigenvalue is greater than zero, the DFE is unstable, whereas if the eigenvalue is less than zero, the DFE is stable (Appendix A1). Alternatively, the Jacobian can be re-framed into a next generation matrix whose spectral radius (λ) determines the stability of the DFE (Diekmann et al. 2009). We focused on this threshold parameter of the next generation matrix in our description of the stability of DFE because it is easier to interpret and can be related to another important epidemiological threshold parameter, the basic reproduction number (R_0). The spectral radius (λ) of the next generation matrix (*A1.4*) for system 1 evaluated at the DFE is,

$$\lambda = \frac{1}{2}\left(\frac{\beta S_0}{(d+\rho+\alpha)} + \frac{r}{\sigma} + \sqrt{\left(\frac{\beta S_0}{(d+\rho+\alpha)} - \frac{r}{\sigma}\right)^2 + 4\frac{\gamma S_0}{\sigma}\frac{\theta}{(d+\rho+\alpha)}}\right) \qquad (2).$$

The DFE is stable whenever $\lambda \leq 1$ and unstable if $\lambda > 1$. The DFE can only be stable if the free-living pathogen death rate, σ, is greater than its birth rate, r, in the environmental reservoir (Appendix A1). Assuming that the pathogen cannot reproduce in the reservoir (i.e. $r = 0$) and that there is no environmental infection (i.e. $\gamma = 0$), then equation 2 reduces to $\lambda = \beta S_0/(d+\rho+\alpha)$, which is also the expression of R_0 for a simple model of a directly-transmitted obligate pathogen (e.g. Anderson and May 1981).

Although λ is a threshold parameter, we refrain from calling it the basic reproduction number. In the derivation of λ, we include the birth of facultative pathogens in the reservoir together with the epidemiological 'births' in the 'transmission' matrix

(Appendix A1). Thus, λ is not equal to the expected number of infections caused by a single infected individual in an otherwise susceptible population (i.e. R_0). Instead, λ is the demographic reproduction number representing per generational overall growth of the facultative pathogen which is distinct from R_0, the epidemiological basic reproduction number (see Heesterbeek 2002 for review of R_0 and its relation to demography).These parameters, λ and R_0, are interchangeable when the 'transmission' matrix includes only one type of birth (i.e. the epidemiological birth) as is the case in Boldin and Kisdi (2012) (but see Appendix A2). In contrast, λ and R_0 are qualitatively and quantitatively different for models where the pathogen can reproduce independent of focal host (e.g. facultative pathogens).

Table 2.1 The definitions of the symbols used in the model. N represents number. For variables and parameters that vary, please see individual figure legends for values used.

Symbols	**Definition**	**Values (dimensions)**
S	Number of susceptible hosts	- (N)
I	Number of infected hosts	- (N)
P	Number of free pathogens	- (N)
b	Birth rate of hosts	1 (N time^{-1})
d	Per capita death rate of hosts	0.2 (N N^{-1} time^{-1})
α	Per capita disease induced death rate (virulence) of hosts	Varies (N N^{-1} time^{-1})
ρ	Per capita recovery rate of hosts	0 (N N^{-1} time^{-1})

θ	Per host shedding rate	Varies (N N^{-1} $time^{-1}$)
β	Transmission rate	Varies (N^{-1} $time^{-1}$)
γ	Probability of infection on contact between susceptible-free pathogen	0.35 (-)
r	Per capita birth rate of environmental propagules	0.1 (N N^{-1} $time^{-1}$)
K	Carrying capacity of population of pathogen propagules in the reservoir	Varies (N)
σ	Free pathogen death rate	Varies (N N^{-1} $time^{-1}$)

If the birth of different 'types' have non-zero entries in the 'transmission' matrix, then R_0 for such system can be defined as a form of type reproduction number (Roberts and Heesterbeek 2003; Heesterbeek and Roberts 2007; van den Driessche 2017). Following the procedure of deriving the type reproduction number, a general expression for R_0 can be derived as

$$R_0 = R_{0d} + \frac{\gamma S_0}{\sigma} \frac{\theta}{d+\rho+\alpha} \left(\frac{r}{\sigma} + \left(\frac{r}{\sigma} \right)^2 + \cdots \right) \quad (3)$$

where,

$$R_{0d} = \frac{\beta S_0}{d+\rho+\alpha} + \frac{\gamma S_0}{\sigma} \frac{\theta}{d+\rho+\alpha} \quad (4).$$

R_{0d} represents expected number of new infections resulting directly from a single infected individual and the propagules it sheds before they reproduce in the reservoir. Remaining terms on the right-hand side of equation 3 represent the expected number of

infections due to the growth of the shed pathogens in the reservoir. Hence, this partitioning of the basic reproduction number is helpful in highlighting the importance of growth in the reservoir to disease dynamics (Figure 2.2a). In Figure 2.2a, R_{0d} is less than one over a considerable parameter space (gray region), but the disease still spreads, as reflected by λ and R_0 both exceeding one. Here, the disease is endemic because of the frequent introduction of new cases through environmental infection that is dependent on growth of free-living pathogens. In the dark gray region, the disease spreads irrespective of the growth of pathogen in the reservoir because low virulence allows for a relatively long duration of infection. In the light gray region, disease fails to spread in the population because there is insufficient host-host direct transmission (due to a short duration of infection) coupled with insufficient growth of the pathogen in the reservoir.

If $r < \sigma$, the limit of the equation 3 exists and can be written as

$$R_0 = \frac{\beta S_0}{d+\rho+\alpha} + \frac{\gamma S_0}{\sigma}\frac{\theta}{d+\rho+\alpha}\left(\frac{1}{1-\frac{r}{\sigma}}\right) \quad (5).$$

However, if $r > \sigma$ – that is, the rate at which pathogens reproduce in the environmental reservoir is greater than the rate at which they die – then $R_0 = \infty$ (see Appendix A2 for derivation). This is because pathogens shed from a single infected host can cause an unending chain of indirect environmental transmissions from the reservoir population that the pathogen sustains independent of hosts. Although λ (equation 2) and R_0 (equation 3) both are accurate threshold parameters reflecting disease dynamics (Figure 2.2a), R_0 is less informative when $r > \sigma$. When the net growth rate of free-living pathogens is positive, we must consider an infinite number of generations of growth for the free-living pathogens shed from a single infected host and the total environmental

transmissions resulting from those generations to compute R_0. In practice, we are primarily interested in epidemiological dynamics occurring at short timescales or finite generations. Thus, other metrics such as epidemic growth rate and disease prevalence (Figure 2.2b) might be more relevant than R_0.

The reservoir population of facultative pathogens can also influence disease prevalence in the host population. Under the assumption that environmental transmission saturates when the reservoir population is sufficiently dense, we found that the pathogen's birth rate in the reservoir has little effect on the endemic disease prevalence (Figure 2.2b). The carrying capacity of the reservoir, however, can show significant differences in the level of endemic disease prevalence (Figure 2.2b).

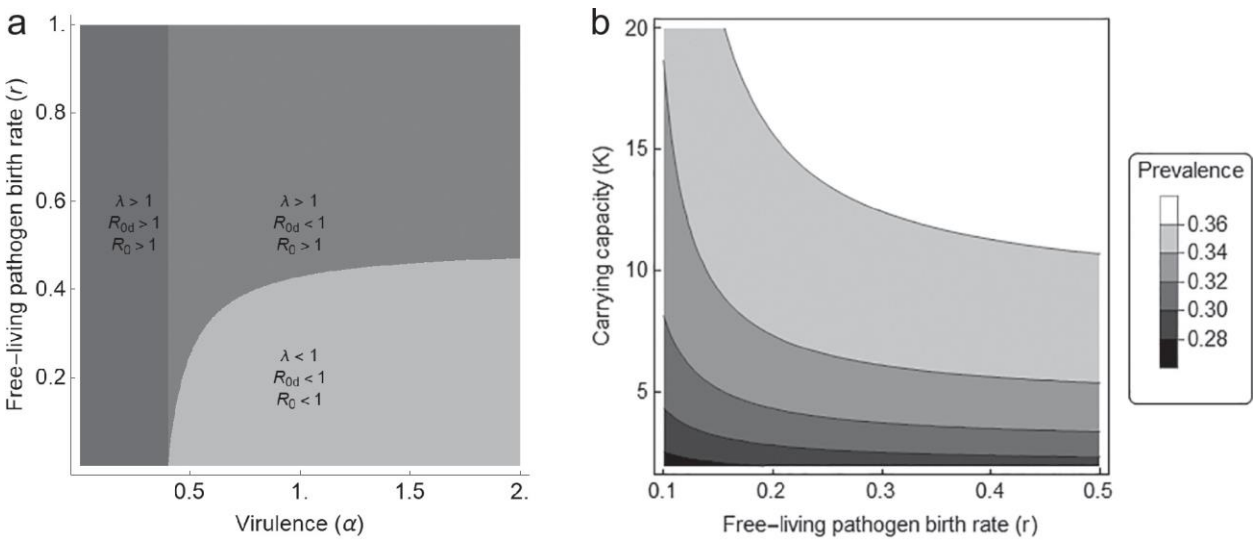

Figure 2.2 . Epidemiological predictions of the facultative pathogen model. (a) Spectral radius (λ) (equation 2), R_0 (equation 3), and R_{0d} (equation 5) contour plots show different epidemiological outcomes. Light gray region corresponds to parameter values where the disease fails to spread at the population level. In gray region, the spread of the disease requires sufficient reproduction in the environmental reservoir. In dark gray region, the disease spreads irrespective of the growth

rate of pathogen in the reservoir. Parameter values: $\gamma = 0.1$, $\beta = 0.1$, $\theta = 0.1$, and $\sigma = 0.5$ (b) Effect of pathogen birth rate in the reservoir and environmental carrying capacity on disease prevalence. Increasing birth rate shows little change in endemic disease prevalence whereas increasing reservoir carrying capacity more strongly impacts prevalence. Parameter values: $\gamma = 0.3$, $\beta = 0.01$, $\theta = 0.1$, $\sigma = 0.1$, $\tau = 1$ and $\alpha = 0.3$.

In our analysis, we have assumed that the disease dynamics start with a single infected host and no pathogens in the environmental reservoir. This may not always be the case. There is another possible disease-free equilibrium which results when $\sigma < r$ and hosts do not contact the pathogen's reservoir population. In this case, the DFE is given as,

$$(S_0, I_0, P_0) = \left(\frac{b}{d}, 0, K\left(1 - \frac{\sigma}{r}\right)\right).$$

If the pathogen reservoir and susceptible host population are spatially structured with little or no interaction between the two populations, then S_0 is determined by the host ecology alone and P_0 is determined by the pathogen reservoir ecology alone. When hosts subsequently contact the environmental reservoir, the stability of this DFE is determined by the eigenvalue at this point (Appendix A6).

Epidemiological R0 does not describe pathogen fitness in facultative pathogens

In the previous section, we described the limitations of epidemiological R_0 for describing disease dynamics. Historically, the epidemiological R_0 is also considered to be a suitable fitness proxy for evolutionary analysis. However, since R_0 only describes the dynamics of the infected hosts, it largely omits free-living pathogen dynamics. Thus, this epidemiological R_0 often cannot be used as a fitness proxy in evolutionary analysis for facultative pathogens. For example, consider the extreme case where there is little

environmental and direct transmission (γ and β are positive but close to 0). Under these conditions, the evolution of the pathogen will be largely determined by reservoir dynamics that do not contribute to R_0 (Equation 3). Though λ (equation 2) incorporates the growth of pathogens in the reservoir, it omits the effect of density dependence in the reservoir and consequently may be an inadequate fitness proxy when the reservoir initially harbors a pathogen population. In this case, a suitable fitness metric for studying long-term evolution is the invasion fitness, defined as the per capita growth rate of the mutant strain in a resident population that has reached its ecological/epidemiological attractor (Metz et al. 1996; Lion and Metz 2018). For facultative pathogens, the invasion fitness should be derived when the system has reached both epidemiological (endemic equilibrium) and ecological attractors (non-zero equilibrium of free-living pathogen population in the reservoir). The next generation matrix approach provides a biologically intuitive framework to derive the invasion fitness for organisms with multiple transmission routes or stage structured life histories (Hurford et al. 2010). Thus, we use this framework to perform evolutionary invasion analysis in the following sections.

Evolutionary Invasion Analysis

Whenever $\lambda > 1$, system 1 goes to the endemic equilibrium (EE) ($\hat{S} > 0, \hat{I} > 0, \hat{P} > 0$). We chose the parameter values such that there exists a locally stable endemic equilibrium (Appendix A4) with positive numbers of individuals (Appendix A3). The individual expressions for $\hat{S}$, $\hat{I}$ and $\hat{P}$ have many terms, and so we do not present explicit expressions for these equilibria. In the following sections, we address the outcome of long-term evolution when virulence trades off with other traits that impact fitness.

Case I: Transmission-Virulence Trade-off

We evaluated how a trade-off between direct transmission and virulence influences the virulence evolution of a facultative pathogen. We consider the case where the direct transmission rate is a saturating function of virulence, $\beta(\alpha) = c\alpha/(c\alpha + h)$. The invasion fitness of a mutant trait α_m in an environment set by the resident trait α is derived as the maximum eigenvalue of the next generation matrix (*A5.6*) for the mutant subsystem (Appendix A5). Assuming a transmission-virulence trade-off, we obtain the following invasion fitness:

$$W(\alpha_m, \alpha) = 1 + \left(s_d \frac{\partial}{\partial \alpha_m} \left(\frac{\beta(\alpha_m)}{d + \rho + \alpha_m} \right) + s_s \theta \frac{\partial}{\partial \alpha_m} \left(\frac{1}{d + \rho + \alpha_m} \right) \right) \Bigg|_{\alpha_m = \alpha} (\alpha_m - \alpha) \quad (6).$$

Here,

$$s_d = \frac{\theta}{d + \rho + \alpha} \frac{\gamma \hat{S}}{\sigma(1 + \hat{S})} \hat{S}$$

and

$$s_s = \left(1 - \frac{\beta \hat{S}}{d + \rho + \alpha} \right) \frac{\gamma \hat{S}}{\sigma(1 + \hat{S})}$$

where s_d describes the lifetime fitness gain from an infected host due to new infections and s_s describes the lifetime fitness gain from an infected host due to shedding of environmental propagules. Both s_s and s_d are proportional to the product of the elements of the left and right eigenvectors of the next generation matrix. The left eigenvector is composed of the reproductive values of an infected host and an environmental propagule. The right eigenvector is composed of the equilibrium number of infected hosts and free-living propagules (Appendix *A5*, (Hurford et al. 2010)). In the case of s_d, $\theta/(d + \rho + \alpha)$ is the scaled lifetime reproductive value of an infected host and $\gamma \hat{S}/(\sigma(1 + \hat{S}))$ is the scaled equilibrium population of infected hosts. Similarly, in the case

of s_s, $\left(1-\beta\hat{S}/(d+\rho+\alpha)\right)$ is the scaled lifetime reproductive value of free-living pathogens. The partial derivatives within equation 6 are the fitness gradients that determine the direction of virulence evolution. A mutant pathogen with a higher virulence compared to the resident pathogen population ($\alpha_m > \alpha$) can invade the resident population if the sum of the scaled fitness gradients is positive. Similarly, if $\alpha_m < \alpha$, then the mutant can invade if the sum of scaled fitness gradients is negative.

In equation 6, the first fitness gradient represents fitness gain due to increased transmission whereas the second fitness gradient

$$\frac{\partial}{\partial \alpha_m}\left(\frac{1}{d+\rho+\alpha_m}\right)$$

is negative and thus represents fitness loss due to increased virulence. Higher virulence reduces the duration of infection which consequently reduces the duration over which pathogen is shed into the reservoir. An intermediate virulence ESS occurs if the fitness gain due to increased transmission can more than compensate the fitness loss due to the reduction in shedding (e.g. strong transmission-virulence relationship, Figure 2.3b). Otherwise, low virulence is expected (Figure 2.3a). Since we are assuming that the virulence does not impact traits relevant to the reservoir, there is no variation in performance of the pathogen in the reservoir for the selection to act. Consequently, the evolutionarily stable strategy (ESS) is independent of shedding rate (θ) and death rate (σ).

Case II: Shedding-Virulence Trade-off

Under the assumption where the shedding rate is a monotonically increasing function of virulence (i.e., $\theta(\alpha) = g\alpha/(g\alpha+f)$), the invasion fitness is given as,

$$W(\alpha_m, \alpha) = 1 + \left(s_d \beta \frac{\partial}{\partial \alpha_m} \left(\frac{1}{d + \rho + \alpha_m} \right) \Big|_{\alpha_m = \alpha} + s_s \frac{\partial}{\partial \alpha_m} \left(\frac{\theta(\alpha_m)}{d + \rho + \alpha_m} \right) \Big|_{\alpha_m = \alpha} \right) (\alpha_m - \alpha) \quad (7).$$

Here, the fitness gradient term has the same functional form as in case I. Since the terms before the fitness gradients are proportional to the product of reproductive values and equilibrium populations, they have positive values. Hence, it also results in an intermediate evolutionarily stable strategy (ESS) if the shedding and virulence are under a strong trade-off, while avirulence or low virulence is the outcome for a weak trade-off relationship.

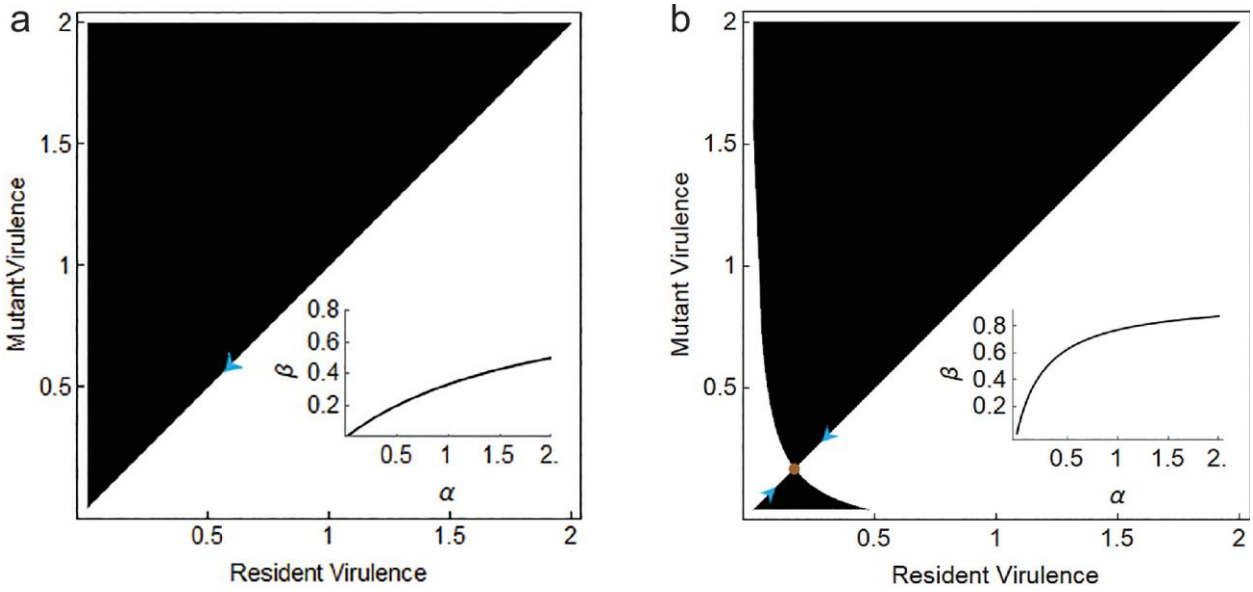

Figure 2.3. Pairwise Invasibility Plots (PIPs) showing virulence evolutionary outcomes where there is a trade-off between transmission and virulence. Shaded region shows the area where the invasion fitness has negative sign (i.e., mutant cannot invade) whereas white region shows area with positive invasion fitness values. Arrows indicate the direction of trait evolution. (a) With a weak tradeoff (shown in the inset, $h = 2$), the pathogen evolves towards avirulence. (b) For a stronger transmission-virulence trade-off ($h = 0.3$), intermediate virulence (brown point) is expected, as shown. Parameter values: $c = 1$, $\theta = 0.2$, $\sigma = 0.05$, $\tau = 1$, $K = 15$.

If we assume that both transmission-virulence and shedding-virulence trade-offs exist for a given pathogen, then the outcome of evolution is determined by the combined fitness gradients from case I and II. Under these trade-off relationships, both fitness gradients point towards intermediate ESS (Figure 2.4). However, if the shedding per generation and transmission per generation are maximized at different levels of virulence, then the outcome will be determined by the relative fitness effects of shedding and transmission.

Case III: Free-living Pathogen Persistence-Virulence Trade-off

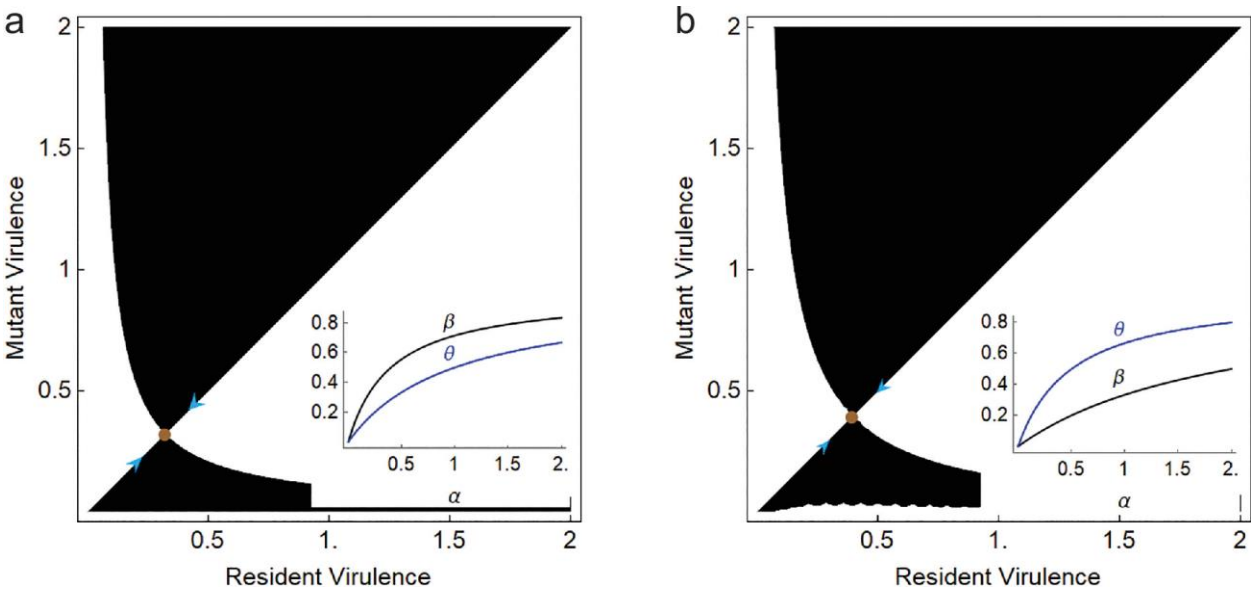

Figure 2.4. Pairwise Invasibility Plots (PIPs) showing virulence outcome for combined transmission-virulence and shedding-virulence tradeoffs. Arrows indicate the direction of trait evolution. (a) Intermediate evolutionarily stable strategy (ESS) results with strong transmission-virulence tradeoff (black line in inset; $h = 0.4$) and weak shedding-virulence trade-off (blue line in inset; $f = 1$). (b) Similar intermediate ESS results with strong shedding-virulence trade-off ($f =$

0.5) and weak transmission-virulence trade-off ($h = 2$). All other parameters are the same as that of Figure 2.3.

The distinctive feature of facultative pathogens is their potential for host-independent growth and survival. We have decomposed the birth and death rate of free-living pathogens in the reservoir in the system (1). Adaptation towards the reservoir environment can come in at least two different forms: increase in birth rate or decrease in death rate (increase in persistence). Because system (1) has density-dependent birth rates, for simplicity, we consider the case of selection to reduce the death rate, thereby increasing the period of time that environmental propagules persist in the reservoir ($1/\sigma$). The invasion fitness in this case can be written as,

$$W(\alpha_m, \alpha) = 1 + \left(s_d \beta \frac{\partial}{\partial \alpha_m}\left(\frac{1}{d+\rho+\alpha_m}\right) + s_s \theta \frac{\partial}{\partial \alpha_m}\left(\frac{1}{d+\rho+\alpha_m}\right) + s_e \frac{\partial}{\partial \alpha_m}\left(\frac{1}{\sigma(\alpha_m)}\right)\right)\Bigg|_{\alpha_m=\alpha} (\alpha_m - \alpha) \quad (8).$$

Here,

$$s_e = \left(1 - \frac{\beta \hat{S}}{d+\rho+\alpha}\right)\frac{\theta}{d+\rho+\alpha}\frac{\gamma \hat{S}}{(1+\hat{S})} + \left(1 - \frac{\beta \hat{S}}{d+\rho+\alpha}\right)^2 r\left(1 - \frac{\hat{P}}{K}\right)$$

and represents lifetime contribution due to free-living pathogens. The first term

$$\left(1 - \frac{\beta \hat{S}}{d+\rho+\alpha}\right)\frac{\theta}{d+\rho+\alpha}\frac{\gamma \hat{S}}{(1+\hat{S})}$$

gives the lifetime contribution of free-living pathogens to new infections whereas the second term

$$\left(1 - \frac{\beta \hat{S}}{d+\rho+\alpha}\right)^2 r\left(1 - \frac{\hat{P}}{K}\right)$$

gives the lifetime contribution of free-living pathogens towards growth in environmental reservoir. Depending upon the exact mechanisms involved, two different functional relationships between pathogen persistence and virulence can be considered. For the first case, we consider a positive correlation between persistence and virulence. The selection gradient associated with s_e represents fitness gain due to increased persistence whereas selection gradients associated with s_d and s_s represent fitness losses due to increased virulence and shorter duration of infection. The result of evolution then depends on the balance of these opposing selection gradients. Bistability is observed for some parameter values and the attained ESS virulence depends on the initial virulence level (Figure 2.5a). If the initial virulence level is low, it pays for the pathogen to have lower virulence because of the benefit of increased duration of infection. However, if the initial virulence starts above the trait value of the repeller point (red dot in Figure 2.5a), increased persistence more than compensates for the fitness loss due to increased virulence. Consequently, virulence evolves to higher levels (Figure 2.5a, Figure S2.5).

We also consider a different functional relationship between persistence and virulence, where the increase in virulence leads to lowering of pathogen persistence. Specifically, we assume the functional form $\sigma(\alpha) = a/(e^{-u\alpha} + v)$, where a, u, and v are positive constants. Here, all three fitness gradients have the same sign such that it is beneficial for the pathogen to have both low death rate in reservoir and low virulence because of the fitness benefits. Consequently, avirulence is the expected outcome for this form of virulence-persistence relationships (Figure 2.5b, Figure S2.5).

Case IV: Multiple Trade-offs

Although the outcomes of evolution can be easily determined for single trade-offs, systems with multiple trade-offs can have more diverse evolutionary outcomes. For some pathogens, it is likely that virulence is correlated with traits that determine fitness both within hosts and in the reservoir. The outcome of virulence evolution then depends on the selection gradients acting in these environments. To explore potential outcomes, we start by considering both transmission-virulence and persistence-virulence trade-offs, followed by a scenario where we also include a relationship between virulence and shedding.

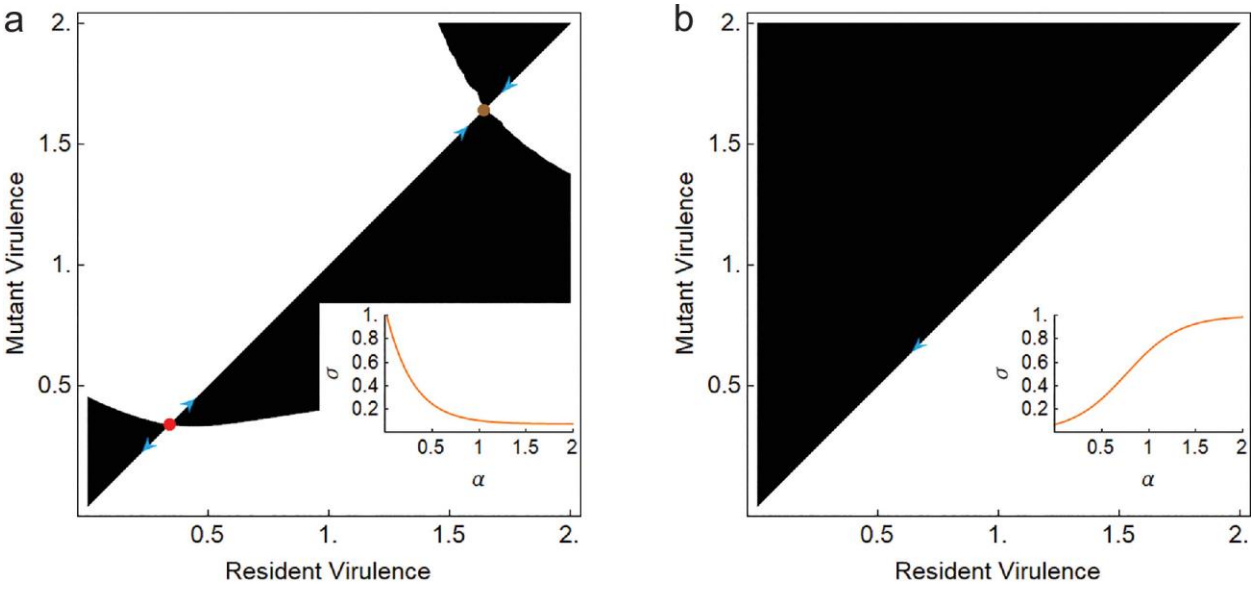

Figure 2.5. Pairwise Invasibility Plots (PIPs) for virulence evolution under persistence-virulence trait correlation. Arrows indicate the direction of trait evolution. (a) Persistence-virulence trade-off ($\sigma(\alpha) = e^{-u\alpha} + v$; i.e., death rate in the environment decreases with increasing virulence) results in bistability conditions separated by a repeller (red point) where the evolutionarily stable strategy (ESS; brown point) depends on the initial condition for virulence. Parameter values: $b = 2$, $\beta = 0.12$, $\theta = 0.3$, $r = 0.3$, $u = 3.5$, $v = 0.07, \tau = 1$. (b) Positive correlation between virulence and death rate ($\sigma(\alpha) = a/(e^{-u\alpha} + v)$) results in evolution towards avirulence. Parameter values: $a = 0.07$, $u = 3.5$, $v = 0.07$, all other parameters are the same as in previous figures.

IV a) Combined transmission-virulence and persistence-virulence tradeoff:

The invasion fitness in this case is given as,

$$W(\alpha_m, \alpha) = 1 + \left(s_d \frac{\partial}{\partial \alpha_m}\left(\frac{\beta(\alpha_m)\hat{S}}{d+\rho+\alpha_m}\right) + s_s\theta \frac{\partial}{\partial \alpha_m}\left(\frac{1}{d+\rho+\alpha_m}\right) + s_e \frac{\partial}{\partial \alpha_m}\left(\frac{1}{\sigma(\alpha_m)}\right)\right)\Bigg|_{\alpha_m=\alpha} (\alpha_m - \alpha) \quad (9).$$

Assuming that both transmission and pathogen persistence are positively correlated with virulence, the selection gradient associated with s_d represents the fitness benefit associated with increased transmission, s_s represents the fitness cost associated with increased virulence (decreased duration for shedding), and s_e represents the fitness benefit associated with increased persistence. When initial virulence levels are low, fitness gradients associated with both transmission and persistence point towards higher virulence. An opposing selection gradient is associated with shedding which points towards lower virulence. ESS virulence occurs whenever the three weighted selection gradients are balanced. If the virulence is too high, the fitness benefit due to increased transmission and persistence does not compensate for the fitness loss due to increased virulence.

IV b) Virulence-transmission, virulence-shedding, and virulence-persistence:

In this section, we consider scenarios where all three trade-offs explored above are present. The invasion fitness for this case can be written as,

$$W(\alpha_m, \alpha) = 1 + \left(s_d \frac{\partial}{\partial \alpha_m}\left(\frac{\beta(\alpha_m)}{d+\rho+\alpha_m}\right) + s_s \frac{\partial}{\partial \alpha_m}\left(\frac{\theta(\alpha_m)}{d+\rho+\alpha_m}\right) + s_e \frac{\partial}{\partial \alpha_m}\left(\frac{1}{\sigma(\alpha_m)}\right)\right)\Bigg|_{\alpha_m=\alpha} (\alpha_m - \alpha) \quad (10).$$

The three weighted fitness gradients determine the direction of selection on virulence. Singular points occur whenever the sum of all fitness gradients weighted by their respective reproductive values is zero. This can be achieved because of the trade-offs, such that selection gradients act in opposite directions with respect to the trait, and fitness contribution obtained by changing the trait in one term is cancelled by the fitness change in another term.

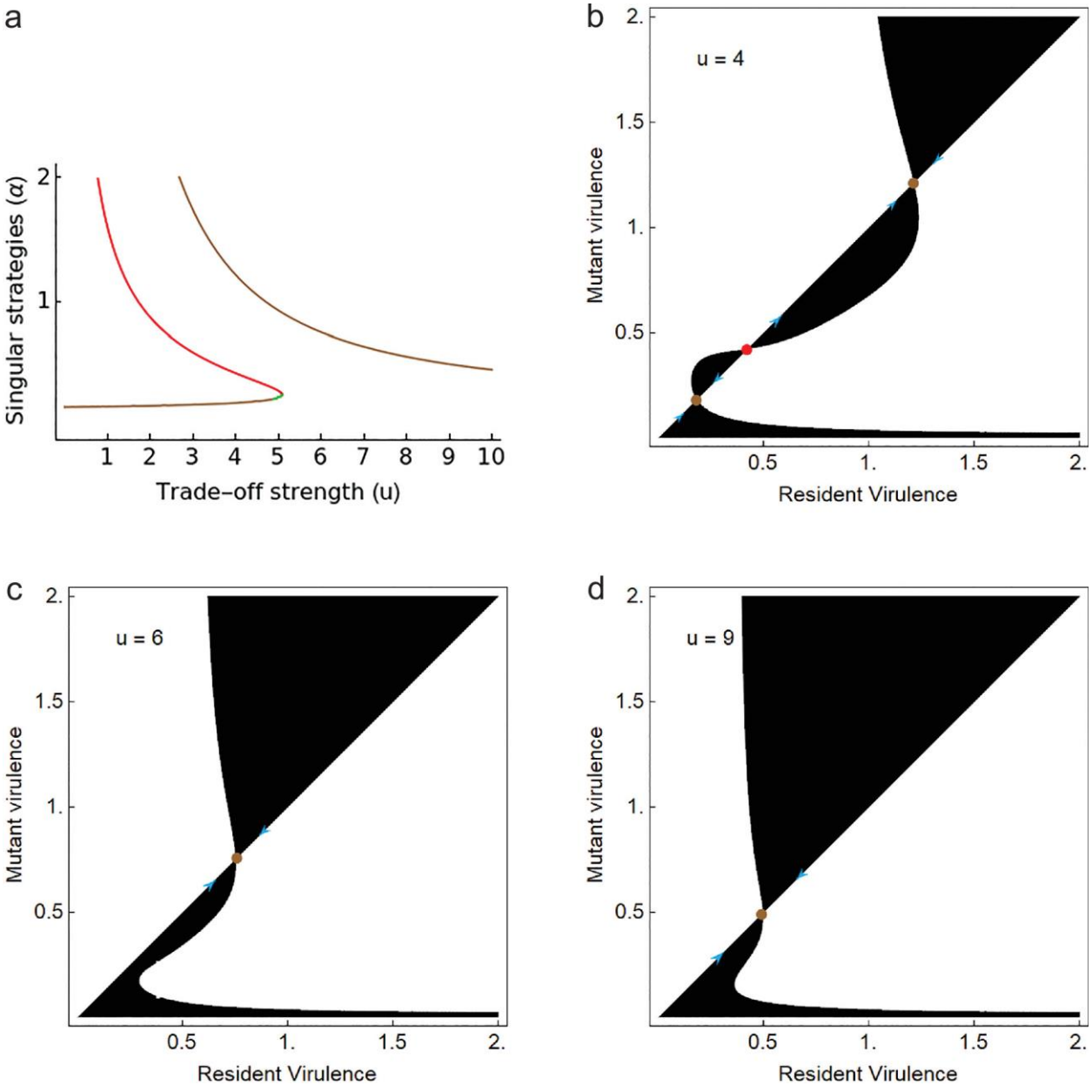

Figure 2.6. Virulence evolution with multiple trade-offs. (a) Bifurcation diagram with singular strategies as a function of u which determines the shape of the relationship between environmental persistence and virulence. Brown lines indicate evolutionarily stable strategies (ESSs) whereas the red and green line represent repellers and evolutionary branching points, respectively. (b) - (d). Pairwise Invasibility Plots (PIPs) with different values of u. Brown points represent evolutionarily stable strategies and red points indicate repeller points. Arrows indicate the direction of trait evolution. Parameter values: $v = 0.02$, $g = 1$, $f = 1$, $c = 1$, $h = 0.1$, $K = 15$, $r = 0.07$, $\tau = 1$.

If we assume a positive correlation between virulence and persistence, then

$$\frac{\partial}{\partial \alpha_m}\left(\frac{1}{\sigma(\alpha_m)}\right) \geq 0.$$

In this case, no singular strategies exist where virulence increases both transmission and shedding. Singular strategies can exist when there is a positive relationship between virulence and persistence if

$$s_d \frac{\partial}{\partial \alpha_m}\left(\frac{\beta(\alpha_m)}{d+\rho+\alpha_m}\right) + s_s \frac{\partial}{\partial \alpha_m}\left(\frac{\theta(\alpha_m)}{d+\rho+\alpha_m}\right) < 0.$$

We can explore different possibilities by changing the strength of trade-offs. For example, if we keep the strength of transmission-virulence and shedding-virulence trade-offs constant, then increasing the strength of the trade-off between persistence and virulence leads to lower virulence (Figure 2.6). A simple comparison between the results in the figure 2.6 and that in Boldin and Kisdi (2012) shows that the addition of growth in reservoir increases the value of singular strategies. Bistability is observed for a wide range of (u) values (Figure 2.6a) where the final ESS attained is dependent on the virulence of the initial strain. Depending on the strength of trade-offs in the system,

singular strategies can become evolutionary branching points where pathogens with both low and high virulence can co-exist (compare Figure 2.7c and Figure 2.8).

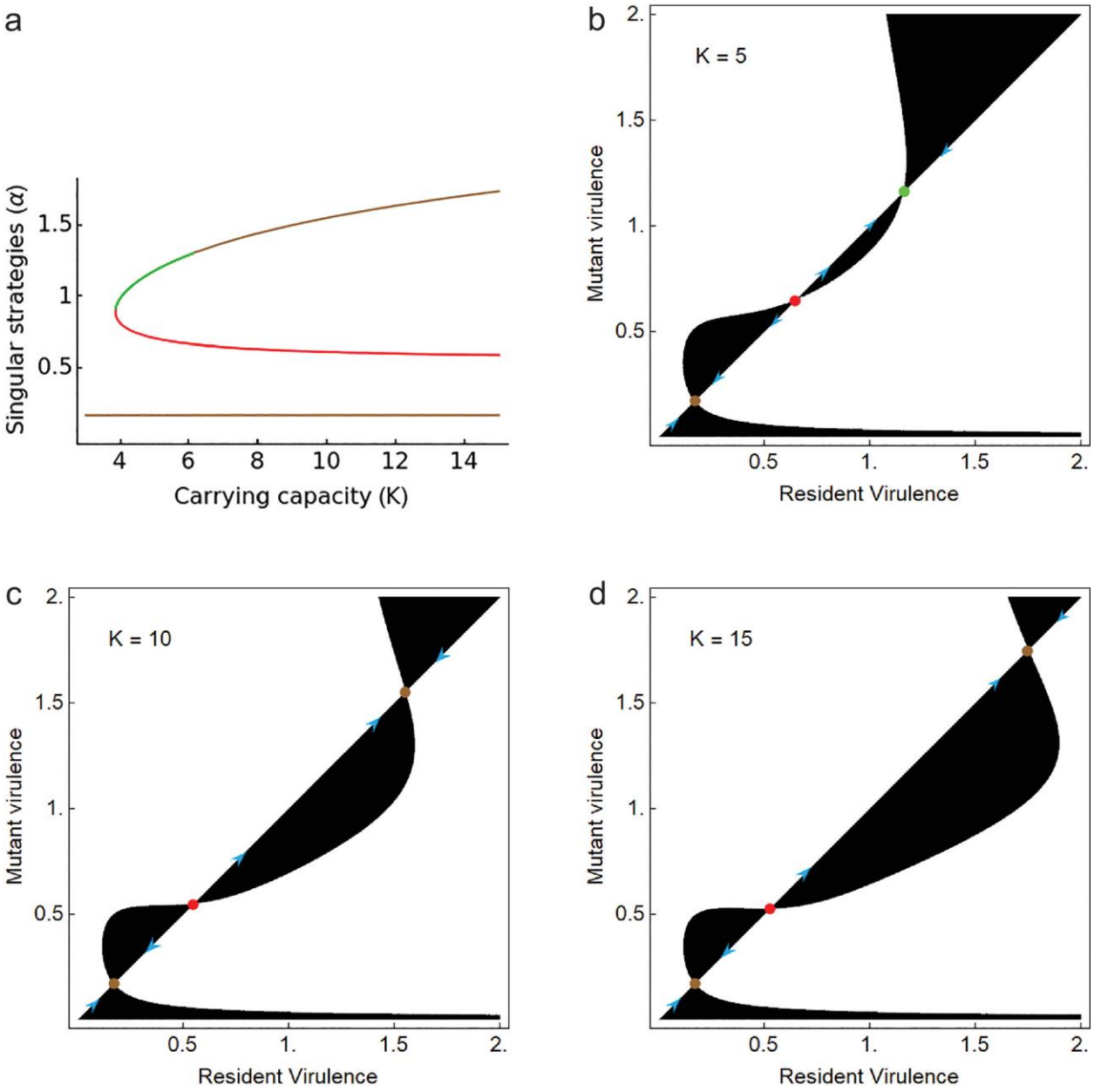

Figure 2.7. Effect of reservoir carrying capacity on virulence evolutionary outcomes (a). Bifurcation plot with singular strategies as a function of carrying capacity (*K*). Brown lines indicate evolutionarily stable strategies whereas the red and green lines represent repellers and evolutionary branching points, respectively. The upper evolutionarily stable strategies increase

with increasing *K*. (b) - (d). Pairwise Invasibility Plots (PIPs) with different values of *K*. Brown points indicate evolutionarily stable strategies, red points represent repellers, and the green point represents an evolutionary branching point. Arrows indicate the direction of trait evolution. Parameter values: $v = 0.02$, $g = 1$, $f = 1$, $c = 1$, $h = 0.1$, $\alpha = 1$, $r = 0.1$, $u = 3$, $\tau = 1$.

IV c) Effect of Carrying Capacity

Several factors, such as a change in environmental conditions or competitor populations, can change the carrying capacity of the pathogen's reservoir populations. While we have not assumed any correlation between virulence and the reservoir carrying capacity, a change in carrying capacity can still impact the direction of virulence evolution by changing the scaling factor s_e in equation 8. Our results suggest that as the reservoir becomes more favorable for the pathogens by increasing the carrying capacity we expect the evolution of higher virulence. Similarly, increased carrying capacity also increases the disease prevalence (Figure 2.2b). High carrying capacity can lead to high equilibrium pathogen population size in the reservoir, which can then increase the number of infections.

Critical Function Analysis

The simulation results presented so far assume specific trade-off forms (e.g. Figure 2.3). Critical function analysis (de Mazancourt and Dieckmann 2004; Kisdi 2006) provides the framework to generalize the results that are independent of specific trade-off assumptions. Scenarios involving a single trade-off are straightforward (see Appendix A7). Here, we focus on multiple trade-offs considered in Figure 2.6 (Figure 2.8), with additional scenarios presented in the Supplementary Materials (Section A7, Figures S2.5, and Figure S2.6).

From equation 10, singular strategies are obtained when

$$\left(s_d \frac{\partial}{\partial \alpha_m}\left(\frac{\beta(\alpha_m)}{d+\rho+\alpha_m}\right) + s_s \frac{\partial}{\partial \alpha_m}\left(\frac{\theta(\alpha_m)}{d+\rho+\alpha_m}\right) + s_e \frac{\partial}{\partial \alpha_m}\left(\frac{1}{\sigma(\alpha_m)}\right)\right)\Bigg|_{\alpha_m=\alpha} = 0 \qquad (11).$$

If the shapes of all three trade-offs are unknown, equation 11 is underdetermined. Building on the analysis presented in section IVb, we fix the shape of transmission-virulence and shedding virulence trade-offs and assume a general relationship for persistence-virulence. This gives us the following differential equation,

$$\sigma'_{crit}(\alpha) = \left(\frac{\sigma^2_{crit}(\alpha)}{s_e}\right)\frac{\left(s_d\big((d+\rho+\alpha)\beta'(\alpha)-\beta(\alpha)\big) + s_s\big((d+\rho+\alpha)\theta'(\alpha)-\theta(\alpha)\big)\right)}{(d+\rho+\alpha)^2} \quad (12),$$

where, $\sigma_{crit}(\alpha)$ is a critical function. The solutions of the equation 12 with different initial conditions result in a family of critical functions (e.g. Figure 2.8a).

Singular points are obtained whenever the trade-off is tangent to the critical function. Convergence stability of the singular points is determined by the local shape of the trade-off. If the trade-off is more concave than the critical function, the singular point is convergence stable (see Figure 2.8a and Appendix A7). Convergence stability along with the nature of evolutionary stability determine the type of singular strategies possible. For example, in Figure 2.8b, the singular strategy is convergence stable but evolutionarily unstable and thus results in an evolutionary branching point (Figure 2.8b).

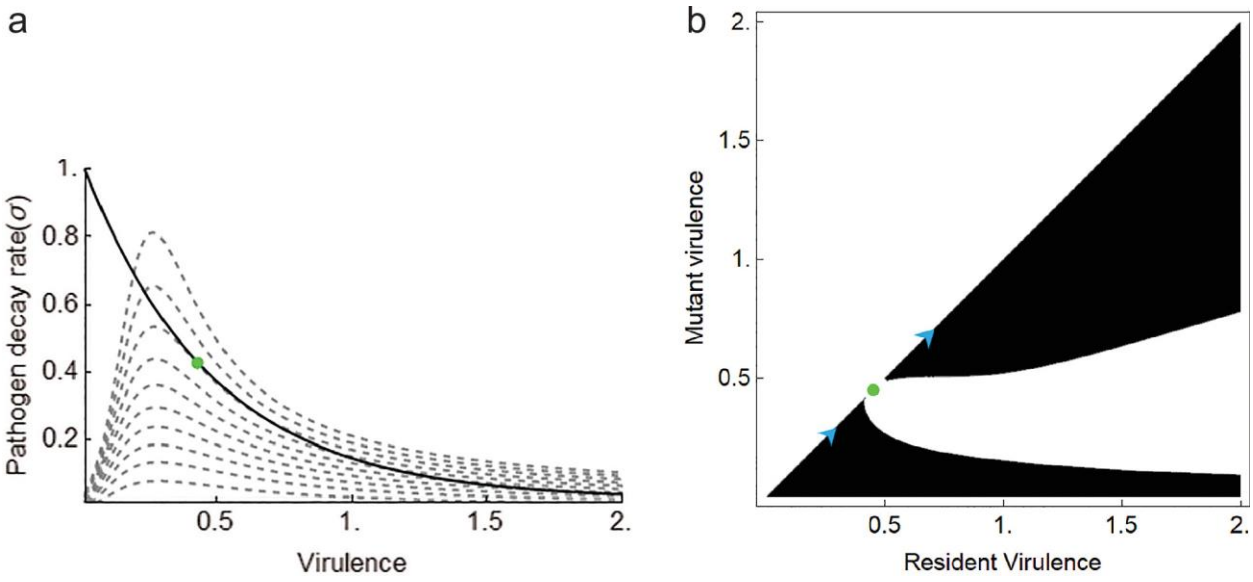

Figure 2.8. Critical function analysis and the nature of singular strategies. (a). Dashed curves represent critical functions that are numerical solutions of equation 12 with different initial conditions. The black solid line represents the persistence-virulence trade-off ($u = 2.1$). The green dot represents the singular strategy where the trade-off is tangent to the critical function. (b). The singular strategy obtained in the (a) is convergence stable but evolutionarily unstable thus leading to an evolutionary branching point (green dot in pairwise-invasibility plot). Arrows indicate the direction of trait evolution. Parameter values are the same as in Figure 2.7c, except $h = 0.3$.

Discussion

The selective pressures shaping the evolution of virulence depend on the pathogen's life history and the fitness trade-offs the pathogen faces. Many pathogens can be environmentally transmitted from and reproduce within reservoir environments. The fitness of these pathogens depends on how well they perform functions associated with infecting hosts, as well as selective pressures imposed in environmental reservoirs. Evolutionary responses to these selective pressures can have significant consequences for how facultative pathogens interact with hosts, particularly when virulence factors have additional functions in reservoir environments.

Stage-structured life histories of pathogens may diminish the role of a trade-off between virulence and transmission but may pose alternative fitness trade-offs shaping the evolution of virulence. The importance of different types of constraints on virulence evolution due to heterogeneous environments has already studied in the context of multi-host parasite systems (Regoes et al. 2000; Gandon 2004). Environmental transmission of dormant environmental propagules (e.g. spores or virions) may allow pathogens to partly circumvent a potential transmission-virulence trade-off allowing for higher virulence (Gandon 1998; Day 2002*b*; Boldin and Kisdi 2012). Consistent with these theoretical predictions, waterborne pathogens which primarily infect hosts from environmental reservoirs are highly virulent (Ewald 1991*b*; Walther and Ewald 2004; Pulkkinen et al. 2010; Breban 2013; Kinnula et al. 2017).

Pathogen growth in environmental reservoirs provides abundant opportunities for ecological trade-offs stemming from traits that are beneficial in one environment being detrimental in another (Sokurenko et al. 2006; Caraco and Wang 2008; Brown et al. 2012; Friman et al. 2013; Martínez 2014; Mikonranta et al. 2015; Barton et al. 2018). We explored the consequences of a potential trade-off between virulence and persistence, virulence and transmission, and virulence and shedding on the evolution of facultative pathogen virulence. When considered in isolation, each trade-off predicts intermediate to high virulence provided that the fitness gain from increased virulence in correlated traits more than compensates for the fitness loss due to reduced duration of infection (increased virulence; Figure 2.3a, 2.4a, and 2.5a). This occurs when there are strong trade-offs, such that a slight increase in virulence leads to a large change in the correlated trait (transmission, shedding, or persistence).

The life history of facultative pathogens and its inherent heterogeneous selection may result in multiple, simultaneous trade-offs associated with virulence. Based on qualitative analysis of the conditions under which the invasion fitness (equation 10) equals 1, at least one singular strategy is possible only when one fitness gradient associated with direct transmission, shedding, or environment persistence is of opposite sign relative to the others. With multiple trade-offs, bistability and evolutionary branching points occur over a wide range of parameter values (Figures 2.6a, 2.7a). Evolutionary bistability occurs when a monomorphic population can evolve to either low virulence or high virulence depending on the virulence of the initial population. Bistability occurs even with a single trade-off between persistence and virulence (Figure 2.5a). Branching points occur when any monomorphic population starting on either side of the branching point will converge towards this point. However, once the population reaches the branching point, disruptive selection results in a dimorphic population. The long-term fate of each subpopulation depends on the invasion fitness of each phenotype in the dimorphic population (Geritz et al. 1998). We observe that evolutionary branching is possible for facultative pathogens (Figures 2.6a, 2.7a), in part depending on the strength of the trade-offs in the system (e.g. compare Figure 2.7c and Figure 2.8). Consistent with the idea that diverse virulence strategies may be likely in pathogens with heterogeneous selection and environmental transmission, Boldin and Kisdi (2012) also identified bistability and branching outcomes for a wide parameter range in their model.

The relationship between the virulence and growth rate of pathogens in the reservoir is likely context-dependent and variable among facultative pathogens. In some cases, there may be pre-adaptation or coincidental selection such that the pathogen's

growth rate in the reservoir is positively correlated with virulence (Levin 1996; Brown et al. 2012; Sundberg et al. 2016). This may be true when virulence factors also have a role in promoting survival in the environmental reservoir. For example, in an emerging opportunistic human pathogen *S. maltophilia,* a type IV secretion system (T4SS) is involved in both competition against heterologous bacteria and apoptosis of human macrophages (Bayer-Santos et al. 2019; Nas et al. 2019). Competition in environmental reservoirs can impose strong selection pressure to maintain these molecular machineries, which can then be co-opted for other functions such as secretion of virulence factors. Similar correlations between pathogen survival in reservoirs and virulence occurs in Shiga toxin producing pathogens, where Shiga toxin provides resistance against grazers and increases host damage (Steinberg and Levin 2007; Adiba et al. 2010). Consistent with these observations, our model predicts that a positive correlation between environmental persistence and virulence results in high ESS virulence (Figure 2.5a).

Alternatively, specialized mechanisms may be required for survival and resource acquisition of the pathogen in host and reservoir environments. This may result in a negative correlation between growth/survival in the reservoir and virulence (Kaitala et al. 2017). As an example, evolution of phage resistance – essential to environmental persistence – has been shown to reduce virulence in *Listeria monocytogenes* (Sumrall et al. 2019). In this study, phage resistance mediated by modifications of *L. monocytogenes* O-antigen results in lower invasiveness of mammalian cells and *in vivo* virulence attenuation in mice. When increased virulence is associated with lower environmental persistence, our model predicts a decrease in virulence (Figure 2.5b). This is consistent with the empirical observation where the evolution of increased environmental

persistence via anti-predator adaptation results in reduced virulence of *Serratia marcescens* (Mikonranta et al. 2012). Similarly, the evolution of phage resistance also resulted in the reduction of virulence in a fish pathogen, *Flavobacterium columnarae* (Laanto et al. 2012). For these reasons, the functional relationship between virulence and traits affecting performance in the reservoir are required to predict the evolution of virulence.

In addition to the effects of varying trade-offs, other environmental parameters can potentially change both epidemiological and evolutionary outcomes. Our results suggest that the endemic prevalence of disease is more sensitive to reservoir carrying capacity than to the effect of free-living pathogen birth rate (Figure 2.2b). Reservoir carrying capacity depends on abiotic (e.g. nutrients, temperature) and biotic (e.g. competition, parasitism, predation) factors and has been shown to influence epidemiological outcomes for environmentally growing, opportunistic pathogens (Anttila et al. 2013, 2016; Kaitala et al. 2017; Merikanto et al. 2018). We also explored evolutionary consequences for virulence of varying reservoir carrying capacities while keeping the strength of trade-offs constant. In general, increasing the carrying capacity in the environmental reservoir can increase a facultative pathogen's evolutionarily stable virulence strategy (Figure 2.7), suggesting that the higher carrying capacity can offset the fitness cost of increased virulence. This result is similar to the prediction from multi-host systems that increased fitness in one host can compensate for the fitness cost associated with increased virulence affecting a second host (Gandon 2004). When carrying capacity is relatively low, our results suggest possible evolutionary diversification indicated by the evolutionary branching points (Figure 2.7a, 2.7b). Interestingly, a long-term evolution experiment of *F.*

columnarae under low nutrient conditions resulted in the diversification of morphotypes with varying virulence (Sundberg et al. 2014). In contrast, the evolution of *F. columnarae* under high nutrient conditions in the reservoir results in higher virulence (Kinnula et al. 2017), which is consistent with our predictions for increasing the carrying capacity (Figure 2.7b-d).

The adaptive dynamics framework relies on the assumption of specific trade-off shapes. However, the empirical determination of the exact shape of trade-off is often difficult (Ebert and Bull 2003; Alizon et al. 2009; Alizon and Michalakis 2015). Though we have assumed qualitatively different trait correlations in some of the results (e.g., Figure 2.5a and 2.5b), more general results are obtained through geometric analyses that are independent of specific trade-off shapes (de Mazancourt and Dieckmann 2004; Kisdi 2006). We used critical function analysis to assess the location and nature of singular strategies in the trait space (Figure 2.8, Figure S2.5, Figure S2.6). A singular strategy occurs when the critical function is tangent to the trade-off function. With multiple trade-offs, it is possible to get a singularity of an arbitrary type given enough environmental feedbacks (Kisdi 2015). To avoid this, we fix some of the trade-off shapes and explore outcomes for others. For example, we fixed the shape of transmission-virulence and shedding-virulence trade-offs to get an explicit differential equation of the critical function related to the trait correlation of pathogen decay rate and virulence (equation 12). Numerical solutions of this equation and invasion analysis show that stronger negative correlations between pathogen decay rate and virulence can result in lower virulence singular traits than weaker negative correlations (Figure 2.6, Figure 2.8a). However, unlike the case of a single trait correlation (Figure S2.5), non-zero singular trait is possible

even for positive correlation between pathogen decay rate and virulence. Similar analysis can be performed for transmission-virulence or shedding-virulence trait correlations if the shapes of remaining two trait correlations are fixed (see Appendix A7).

Facultative pathogens pose a challenge to the determination and interpretation of the basic reproduction number (R_0)—an important threshold parameter used to evaluate potential epidemiological and evolutionary outcomes. The next generation matrix method provides an intuitive approach for deriving R_0 of different systems. There are different ways to construct the next generation matrix for the same system, but how we structure the next generation matrix changes the biological interpretation of the threshold parameter obtained. Although many of these have the same threshold properties as R_0 (e.g., Bani-Yaghoub et al. 2012), not all threshold parameters obtained from the next generation matrices follow the strict definition of the epidemiological R_0. If the objective is to purely look at the disease outcomes by deciding whether the disease dies out or leads to an outbreak, then all the R_0-like threshold parameters derived from different next generation matrix considerations behave equivalently (Appendix A2). However, if the objective is to quantify the contribution of a single infected individual to secondary infections, then the R_0 is needed. The conventional method of deriving the R_0 as the maximum eigenvalue of the next generation matrix leads to a threshold metric quantitatively different than the epidemiological R_0. Unlike Boldin and Kisdi (2012), our model has two 'birth' types in the 'transmission matrix'. When there is more than one type of birth in the 'transmission' matrix component of the next generation matrix, there is an important distinction between the demographic reproduction number (which we refer to as λ in the results section) and the epidemiological reproduction number (R_0) as these

parameters represent two fundamentally different biological quantities and hence, can have different magnitudes. Thus, we provide a general recipe for R_0 derivation and explain its difference with the demographic threshold parameter λ. The advantage of using this technique to determine the R_0 is that it is easier to interpret than other threshold expressions, and allows us to decompose the contribution of each transmission route to new infections.

For many pathogens, the epidemiological R_0 can be a suitable fitness proxy to study virulence evolution. However, since the epidemiological R_0, by definition, is focused on the disease transmission to the hosts, it omits the free-living pathogen growth dynamics in the reservoir. A better fitness metric should ideally close the life-cycle of pathogens (Alizon and Michalakis 2015). For this reason, the demographic reproduction number (λ) better describes facultative pathogen dynamics than R_0. However, because λ is derived under the assumption of no initial free-living pathogens in the reservoir, it is only a suitable fitness metric when the pathogen is initially absent in the reservoir. Consequently, because the invasion fitness is derived when both the host and the free-living pathogen population reach equilibrium it is a better fitness metric than λ and R_0 since it incorporates the effect of density dependence in the reservoir.

The life history of facultative pathogens features distinct population dynamics occurring in environmental reservoirs and during transmission to hosts and can involve multiple transmission routes. Our model extends the model from Boldin & Kisdi (2012) in order to explore consequences of pre-adaptation and coincidental selection in reservoirs for facultative pathogens. We further explored the consequences of parameters relevant to reservoir dynamics (e.g., reservoir carrying capacity, free-living pathogen growth rate)

to disease dynamics and virulence evolution. However, we deliberately kept the model simple to maintain tractability and facilitate interpretation. Although a positive correlation between the pathogen's birth rate in the reservoir and virulence is possible, density dependence with birth rate makes calculations required to explore this trade-off cumbersome. Thus, we restricted our analysis to a simpler correlation between the pathogen's reservoir death rate and virulence. We also neglected potential interactions of free-living propagules with other organisms in the reservoir. Bacterial competitors and predators can potentially influence the trait dynamics in the reservoir (Godfray et al. 1999; Friman et al. 2009).

Many facultative pathogens are of public health, economic, and ecological importance. Despite this, the consequences of such life-histories on virulence evolution have often been overlooked in previous studies. Although our model is a simplification of facultative pathogen life-history, the alignment of qualitative model predictions with several empirical observations is encouraging. These empirical agreements are however merely correlative suggestions, and hence, future work connecting modeling and experiment is required to establish the selective pressures determining the evolution of facultative pathogen virulence. Empirical work evaluating the trade-offs faced by facultative pathogens is required for a comprehensive theory on virulence evolution capable of predicting how virulence evolves in these systems.

Acknowledgements

We thank Jim Bever, Megan Greischar, Spencer Hall, Majid Jaberi-Douraki, and Tsukushi Kamiya for helpful discussions and the Kansas National Science Foundation EPSCoR (OIA 1656006) and the Natural Sciences and Engineering Research Council

(RGPIN-2018-06017) of Canada for funding. Research reported in this publication was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health under award number P20GM130448. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Acevedo, M. A., F. P. Dillemuth, A. J. Flick, M. J. Faldyn, and B. D. Elderd. 2019. Virulence-driven trade-offs in disease transmission: A meta-analysis. Evolution 73:636–647.

Adiba, S., C. Nizak, M. van Baalen, E. Denamur, and F. Depaulis. 2010. From grazing resistance to pathogenesis: The coincidental evolution of virulence factors. PLoS ONE 5:e11882.

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

Alizon, S., and Y. Michalakis. 2015. Adaptive virulence evolution: The good old fitness-based approach. Trends in Ecology and Evolution 30:248–254.

Anderson, R. M., and R. M. May. 1978. Regulation and stability of host-parasite population interactions: I. Regulatory processes. Journal of Animal Ecology 47:219–247.

Anderson, R. M., and R. M. May. 1981. The population dynamics of microparasites and their invertebrate hosts. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 291:451–524.

Anttila, J., J. Laakso, V. Kaitala, and L. Ruokolainen. 2016. Environmental variation enables invasions of environmental opportunist pathogens. Oikos 125:1144–1152.

Anttila, J., L. Ruokolainen, V. Kaitala, and J. Laakso. 2013. Loss of competition in the outside host environment generates outbreaks of environmental opportunist pathogens. PLoS ONE 8:e71621.

Bani-Yaghoub, M., R. Gautam, Z. Shuai, P. van den Driessche, and R. Ivanek. 2012. Reproduction numbers for infections with free-living pathogens growing in the environment. Journal of Biological Dynamics 6:923–940.

Barton, I. S., C. Fuqua, and T. G. Platt. 2018. Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. Environmental Microbiology 20:16–29.

Bayer-Santos, E., W. Cenens, B. Y. Matsuyama, G. U. Oka, G. Di Sessa, I. D. V. Mininel, T. L. Alves, et al. 2019. The opportunistic pathogen *Stenotrophomonas maltophilia* utilizes a type IV secretion system for interbacterial killing. PLOS Pathogens 15:e1007651.

Boldin, B., and É. Kisdi. 2012. On the evolutionary dynamics of pathogens with direct and environmental transmission. Evolution 66:2514–2527.

Bonhoeffer, S., R. E. Lenski, and D. Ebert. 1996. The curse of the pharaoh: the evolution of virulence in pathogens with long living propagules. Proceedings of the Royal Society B: 263:715–721.

Breban, R. 2013. Role of environmental persistence in pathogen transmission: a mathematical modeling approach. Journal of Mathematical Biology 66:535–546.

Brown, S. P., D. M. Cornforth, and N. Mideo. 2012. Evolution of virulence in opportunistic pathogens: Generalism, plasticity, and control. Trends in Microbiology 20:336–342.

Caraco, T., and I. N. Wang. 2008. Free-living pathogens: Life-history constraints and strain competition. Journal of Theoretical Biology 250:569–579.

Charnov, E. 1989. Phenotypic evolution under Fisher's fundamental theorem of natural selection. Heredity 113–116.

Codeço, C. T. 2001. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. BMC infectious diseases 1:1.

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

Day, T. 2002*a*. The evolution of virulence in vector-borne and directly transmitted parasites. Theoretical Population Biology 62:199–213.

———. 2002*b*. Virulence evolution via host exploitation and toxin production in spore-producing pathogens. Ecology Letters 5:471–476.

de Mazancourt, C., and U. Dieckmann. 2004. Trade-off geometries and frequency-dependent selection. American Naturalist 164:765–778.

Diekmann, O., J. A. P. Heesterbeek, and M. G. Roberts. 2009. The construction of next-generation matrices for compartmental epidemic models. Journal of The Royal Society Interface 7:873–885.

Ebert, D., and J. J. Bull. 2003. Challenging the trade-off model for the evolution of virulence. Trends in Microbiology 11:15–20.

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

Ewald, P. W. 1991*a*. Transmission modes and the evolution of virulence - With special reference to cholera, influenza, and AIDS. Human Nature 2:1–30.

Ewald, P. W. 1991*b*. Waterborne transmission and the evolution of virulence among gastrointestinal bacteria. Epidemiology and Infection 106:83–119.

Frank, S. A. 1996. Models of parasite virulence. The Quarterly Review of Biology 71:37–78.

Friman, V. P., M. Ghoul, S. Molin, H. K. Johansen, and A. Buckling. 2013. *Pseudomonas aeruginosa* adaptation to lungs of cystic fibrosis patients leads to lowered resistance to phage and protist enemies. PLoS ONE 8:e75380.

Friman, V. P., C. Lindstedt, T. Hiltunen, J. Laakso, and J. Mappes. 2009. Predation on multiple trophic levels shapes the evolution of pathogen virulence. PLoS ONE 4:e6761.

Gandon, S. 1998. The curse of the pharaoh hypothesis. Proceedings of the Royal Society B 265:1545–1552.

Gandon, S. 2004. Evolution of Multihost Parasites. Evolution 58:455–469.

Geritz, S. A. H., É. Kisdi, G. Meszena, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology 12:35–57.

Godfray, H. C. J., C. J. Briggs, N. D. Barlow, M. O'Callaghan, T. R. Glare, and T. A. Jackson. 1999. A model of insect-pathogen dynamics in which a pathogenic bacterium can also reproduce saprophytically. Proceedings of the Royal Society B: 266:233–240.

Heesterbeek, J. A. P. 2002. A brief history of R_0 and a recipe for its calculation. Acta Biotheoretica 50:189–204.

Heesterbeek, J. A. P., and M. G. Roberts. 2007. The type-reproduction number *T* in models for infectious disease control. Mathematical Biosciences 206:3–10.

Hurford, A., D. Cownden, and T. Day. 2010. Next-generation tools for evolutionary invasion analyses. Journal of The Royal Society Interface 7:561–571.

Julian, T. R. 2016. Environmental transmission of diarrheal pathogens in low and middle income countries. Environmental Science: Processes and Impacts. Royal Society of Chemistry.

Kaitala, V., L. Ruokolainen, R. D. Holt, and J. K. Blackburn. 2017. Population dynamics, invasion, and biological control of environmentally growing opportunistic pathogen. Pages 213–245 *in*Modeling the Transmission and Prevention of Infectious Disease, Advances in Environmental Microbiology (Vol. 4). Springer International Publishing, Cham.

Kinnula, H., J. Mappes, J. K. Valkonen, K. Pulkkinen, and L.-R. Sundberg. 2017. Higher resource level promotes virulence in an environmentally transmitted bacterial fish pathogen. Evolutionary Applications 10:462–470.

Kisdi, É. 2006. Trade-off geometries and the adaptive dynamics of two co-evolving species. Evolutionary Ecology Research 8:959–973.

———. 2015. Construction of multiple trade-offs to obtain arbitrary singularities of adaptive dynamics. Journal of Mathematical Biology 70:1093–1117.

Laanto, E., J. K. H. Bamford, J. Laakso, and L.-R. Sundberg. 2012. Phage-driven loss of virulence in a fish pathogenic bacterium. PLoS ONE 7:e53157.

Lanzas, C., K. Davies, S. Erwin, and D. Dawson. 2019. On modelling environmentally transmitted pathogens. Interface Focus 10:20190056.

Levin, B. R. 1996. The evolution and maintenance of virulence in microparasites. Emerging Infectious Diseases 2:93–102.

Lion, S., and J. A. J. Metz. 2018. Beyond R_0 maximisation: on pathogen evolution and environmental dimensions. Trends in Ecology & Evolution 33:458–473.

Martínez, J. L. 2014. Short-sighted evolution of bacterial opportunistic pathogens with an environmental origin. Frontiers in Microbiology 5:239.

Merikanto, I., J. T. Laakso, and V. Kaitala. 2018. Outside-host phage therapy as a biological control against environmental infectious diseases. Theoretical Biology and Medical Modelling 15:7.

Metz, J. A. J., S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden. 1996. Adaptive Dynamics: A geometrical study of the consequences of nearly faithful reproduction. Stochastic and Spatial Structures of Dynamical Systems 45:183–231.

Mikonranta, L., V. P. Friman, and J. Laakso. 2012. Life history trade-offs and relaxed selection can decrease bacterial virulence in environmental reservoirs. PLoS ONE 7:e43801.

Mikonranta, L., J. Mappes, J. Laakso, and T. Ketola. 2015. Within-host evolution decreases virulence in an opportunistic bacterial pathogen. BMC Evolutionary Biology 15:165.

Nas, M. Y., R. C. White, A. L. DuMont, A. E. Lopez, and N. P. Cianciotto. 2019. *Stenotrophomonas maltophilia* encodes a VirB/VirD4 Type IV Secretion System that

modulates apoptosis in human cells and promotes competition against heterologous bacteria, including *Pseudomonas aeruginosa*. Infection and Immunity 87:10–12.

Peyraud, R., L. Cottret, L. Marmiesse, J. Gouzy, and S. Genin. 2016. A resource allocation trade-off between virulence and proliferation drives metabolic versatility in the plant pathogen *Ralstonia solanacearum*. PLoS Pathogens 12:e1005939.

Pilla, G., G. McVicker, and C. M. Tang. 2017. Genetic plasticity of the *Shigella* virulence plasmid is mediated by intra- and inter-molecular events between insertion sequences. (M. Blokesch, ed.)PLoS Genetics 13:e1007014.

Platt, T. G., J. D. Bever, and C. Fuqua. 2012. A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. Proceedings of the Royal Society B: Biological Sciences 279:1691–1699.

Pulkkinen, K., L.-R. Suomalainen, A. F. Read, D. Ebert, P. Rintamäki, and E. T. Valtonen. 2010. Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. Proceedings of the Royal Society B: 277:593–600.

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

Roberts, M. G., and J. A. P. Heesterbeek. 2003. A new method for estimating the effort required to control an infectious disease. Proceedings of the Royal Society B: 270:1359–1364.

Robino, E., A. C. Poirier, H. Amraoui, S. Le Bissonnais, A. Perret, C. Lopez-Joven, J. Auguet, et al. 2019. Resistance of the oyster pathogen *Vibrio tasmaniensis* LGP32 against grazing by *Vannella* sp. marine amoeba involves Vsm and CopA virulence factors. Environmental Microbiology 1462-2920.14770.

Roche, B., J. M. Drake, and P. Rohani. 2011. The curse of the Pharaoh revisited: evolutionary bi-stability in environmentally transmitted pathogens. Ecology Letters 14:569–575.

Roff, D. A. 2000. Trade-offs between growth and reproduction : an analysis of the quantitative genetic evidence. Journal of Evolutionary Biology 13:434–445.

Sokurenko, E. V., R. Gomulkiewicz, and D. E. Dykhuizen. 2006. Source–sink dynamics of virulence evolution. Nature Reviews Microbiology 4:548–555.

Stearns, S. C. 1989. Trade-offs in life-history evolution. Functional Ecology 3:259–268.

Steinberg, K. M., and B. R. Levin. 2007. Grazing protozoa and the evolution of the *Escherichia coli* O157:H7 Shiga toxin-encoding prophage. Proceedings of the Royal Society B: 274:1921–1929.

Sturm, A., M. Heinemann, M. Arnoldini, A. Benecke, M. Ackermann, M. Benz, J. Dormann, et al. 2011. The cost of virulence: Retarded growth of *Salmonella typhimurium* cells expressing Type III secretion system 1. PLoS Pathogens 7:e1002143.

Sumrall, E. T., Y. Shen, A. P. Keller, J. Rismondo, M. Pavlou, M. R. Eugster, S. Boulos, et al. 2019. Phage resistance at the cost of virulence: *Listeria monocytogenes* serovar 4b requires galactosylated teichoic acids for InlB-mediated invasion. PLOS Pathogens 15:e1008032.

Sundberg, L.-R., T. Ketola, E. Laanto, H. Kinnula, J. K. H. Bamford, R. Penttinen, and J. Mappes. 2016. Intensive aquaculture selects for increased virulence and interference competition in bacteria. Proceedings of the Royal Society B: 283:20153069.

Sundberg, L.-R., H. M. T. Kunttu, and E. Valtonen. 2014. Starvation can diversify the population structure and virulence strategies of an environmentally transmitting fish pathogen. BMC Microbiology 14:67.

Tien, J. H., H. N. Poinar, D. N. Fisman, and D. J. D. Earn. 2011. Herald waves of cholera in nineteenth century London. Journal of The Royal Society Interface 8:756–760.

van Baalen, M., and M. W. Sabelis. 1995. The dynamics of multiple infection and the evolution of virulence. The American Naturalist 146:881–910.

van den Driessche, P. 2017. Reproduction numbers of infectious disease models. Infectious Disease Modelling 2:288–303.

Walther, B. A., and P. W. Ewald. 2004. Pathogen survival in the external environment and the evolution of virulence. Biological Reviews 79:849–869.

White, C. V, B. J. Darby, R. J. Breeden, and M. A. Herman. 2016. A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. Infection and Immunity 84:524–536.

Chapter 3 - Coincidental selection stemming from interspecific competition promotes higher virulence in *Stenotrophomonas maltophilia*

Aakash Pandey[1*], Brandi Lynn Cobe[2], Tyler Hanson[1], Nicholas P. Cianciotto[2], Thomas G. Platt[1]

[1]Division of Biology, Kansas State University, Manhattan, KS, USA

[2]Department of Microbiology-Immunology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

*Corresponding author: apandey@ksu.edu

Abstract

Selection occurring in the environmental reservoirs of free-living pathogens can have important consequences for the evolution of pathogen traits, including those that shape virulence. *Stenotrophomonas maltophilia* is an emerging opportunistic human pathogen found in soil and aquatic reservoirs. Although several functions that contribute to its colonization in host tissues have been identified, the ecological drivers that shape the evolution of these potential virulence factors are not well understood. In this study, we examine how the heterogeneous selective pressures acting on the type IV secretion system and its associated effector proteins shape virulence in *Caenorhabditis elegans-S. maltophilia* interactions. We found that highly virulent strains are better competitors against *Escherichia coli* MG1655 compared to low virulent strains. Experimental evolution of *S. maltophilia* strains in an unstructured environment shows that relaxed selection can result in both low virulence and low competitive ability. Genomic comparison of ancestral and evolved lines derived from a highly virulent strain JV3 show identical mutations in the type IV secretion system operon. In addition, populations derived from the experimental evolution of JV3 in the presence of competitor *E. coli* MG1655 retained higher degree of virulence compared to population evolving without the competitor. Taken together, these results suggest that coincidental selection for interference competition against heterologous bacteria can promote elevated virulence in *S. maltophilia*.

Introduction

Traits, such as virulence, that can influence the fitness of a pathogen are shaped by the ecology of the host and the pathogen. In free-living environmental pathogens, selective pressures occurring in environmental reservoirs can be important drivers of pathogen traits [1–4]. Strong selection pressure can result from predation and competition in environment. The factors that contribute to environmental survival of these pathogens can also involve in within-host interactions leading to coincidental virulence. For example, protozoal predation influences virulence and within-host survival in environmental pathogens such as *Pseudomonas aeruginosa* [5] and *Vibrio cholerae* [6]. Virulence in environmental pathogens can thus evolve due to coincidental selection for other traits [7,8].

Stenotrophomonas maltophilia is an environmental pathogen that is often found in diverse settings such as rhizosphere[9], plant tissues[10], and water reservoirs[11]. Hence, the selection pressures and the consequent adaptations to these environments can potentially shape the virulence of *S. maltophilia*. *S. maltophila* is an emerging opportunistic human pathogen [12–14], whose incidence and prevalence of multidrug resistant strains are increasing in recent decades [15–18]. It is also one of the most common bacteria associated with severe and critical cases of covid-19 disease [19]. *S. maltophilia* infections are often problematic because this pathogen is often resistant to many antibiotics [20,21]. Although several potential virulence factors of *S. maltophilia* that contribute to its colonization in host tissues are identified [20,22–24], the ecological drivers that shape the evolution of these virulence factors are not well understood.

S. maltophilia is often associated with *Caenorhabditis elegans* in nature and *S. maltophilia* strains vary considerably in their virulence on *C. elegans* [25]. This variation may

partially reflect differences in how these strains impact host gene expression profiles [26,27]. How this virulence polymorphism is evolutionarily maintained is not well understood. Previous research has shown that *S. maltophilia* employs a type IV secretion system that plays a dual role in host-pathogen interaction and interference competition against heterologous bacteria [28]. Since *S. maltophilia* likely faces considerable competition in reservoir environments like the rhizosphere, in this study we examine if selection for interference competition influences the virulence evolution in *S. maltophilia*.

In this study, we test the role of the type IV secretion system in virulence against *C. elegans* and show that the type IV secretion system and its associated effector proteins [29] contribute to virulence against *C. elegans*, in addition to being involved in interference competition. We further show that the four different *S. maltophilia* strains (K279a, R551-3, JCMS, and JV3) that vary in virulence against *C. elegans* also vary in their competitive ability against *E. coli* MG1655. JV3 shows the highest virulence whereas K279a is the least virulent strain among the four strains. R551-3 and JCMS show intermediate level of virulence. We found that high virulence is associated with greater interspecific competitive ability against *E. coli* MG1655. Experimental evolution of JV3, JCMS, and R551-3 in unstructured environments results in reduced virulence and reduced competitive ability. In contrast, experimental evolution of the lowest virulence strain, K279a, did not show any significant changes in virulence. Populations derived from K279a, however, did have reduced competitive ability against *E. coli* MG1655. Variant analysis of the independent evolution lines derived from the JV3 strain show the same mutation in type IV secretion system locus. To test whether the presence of competitor has any effect on virulence evolution, we evolved JV3 with and without the competitor *E. coli* MG1655. Both

treatments show reduction in virulence; however, virulence was reduced to a greater extent in populations evolving without a competitor than in populations evolving with a competitor present. Taken together, we show that virulence and interference competition are correlated in *S. maltophilia* with both being dependent on a type IV secretion system, and that selection for interference competition can help maintain high virulence.

Materials and methods

Strains and growth conditions:

All strains used in this study are listed in Table S1. *E. coli* OP50 was grown overnight in LB broth media at 30^{o}C and used as a food source for *C. elegans*. *C. elegans* was grown on nematode growth media (NGM) at 25 ^{o}C. Virulence assays were performed at 25 ^{o}C and competition experiments were performed at 30 ^{o}C. *S. maltophilia* strains were grown overnight in LB broth at 30^{o}C and diluted to the 0.2 OD_{600nm} to perform virulence and 0.3 OD_{600nm} for competition assays.

Virulence and survival assay:

To measure virulence, we used the nematode survival assay described by White et al.[25] with some minor modifications. In short, *S. maltophilia* strains were grown overnight in LB broth at 30 ^{o}C. The culture was then washed with MQ water (3X) and diluted to an OD_{600nm} of 0.2. 200 µl of the diluted culture was then spread on a NGM plate to make a lawn of bacteria. For experiments that include Δ*virB10*/*virB10*+, 14245/14245+, and 14255/14255+, we included IPTG (100 µM) in the NGM media. *C. elegans* were grown to the larval 4 (L4) stage. We transferred 10 L4 stage worms using a worm-picker to a NGM plate inoculated with *S. maltophilia*. The plates were then incubated at 25 ^{o}C. The plates were checked daily for the next generation *C. elegans* larvae (< L4 stage), and

the lavae were removed if present. Surviving *C. elegans* were transferred to a fresh plate containing *S. maltophilia* every other day. The number of dead and alive worms were counted using a Leica S7 E Stereomicroscope either every day, every other day, or after 72 hours depending on the experiment. The worms were poked by worm picker to test if they were alive. If there was no response to touch, they were considered dead.

Competition assays

The bacterial interspecific competition assay was performed following established protocols [28,30] with slight modifications. In short, both *S. maltophilia* and *E. coli* MG1655 were grown overnight in LB broth at 30^{o}C. The cultures were diluted 1:100 into fresh LB broth and grown to mid-log phase at 30^{o}C. The cultures were then diluted to an OD_{600nm} of 0.3 and equal volumes of each diluted culture were mixed. We spotted 5 µl of the cell mixture onto LB agar and incubated at 30^{o}C for 2 hours. The mixture was also diluted and plated on LB agar plates with X-gal (40µg/ml) and IPTG (100 µM) to quantify the initial density of each strain present. We recovered the cells from the spotted area after 2 hours of incubation and resuspended the cells in MQ water. We then serially diluted this cell suspension and plated on LB agar plates with X-gal (40µg/ml) and IPTG (100 µM) to quantify the final density of each strain present. The plates were then incubated at 30^{o}C for 48 hours before the numbers of white (*S. maltophilia*) and blue (*E. coli*) colonies on each dilution plate were recorded. Competitive ability was calculated as the ratio of *S. maltophilia* to *E. coli* at the end of the experiment divided by this ratio at the start of the experiment.

Evolution experiment

The evolution experiment involved serial passage of *S. maltophilia* cultures in 96 deep well plates containing LB media. The experiment had 10 replicate populations of each of 4 strains (JV3, JCMS, R551-3 and K279a) with populations being randomly assigned to a position in the 96-well plate. For initial inoculation, each population were grown to mid-log phase, diluted, and inoculated into the wells to make the final OD_{600nm} of 0.005. The plates were incubated at 30^{o}C for 24 hours. After 24 hours each population was then diluted 1:100 (10 µl /1000 µl) into a new deep well plate with 1ml of fresh LB media. This was repeated for 30 transfers. After 30 transfers, samples were frozen in 25 % glycerol and kept at -80 ^{o}C for use in subsequent virulence and competition assays. Some of the evolution lines were lost due to power failure to the -80^{o}C freezer and hence, different numbers of evolution lines are shown in the results.

To evaluate the effect of interspecies competition on virulence evolution, 8 replicate JV3 populations were serially passaged in three different conditions. In one treatment, JV3 was passaged alone following the procedures same as above. In the second treatment, JV3 was mixed with *E. coli* MG1655 in a 1:1 ratio which was then diluted and serially passaged as a mixture in 1:100 ratio (10µl mixed culture/1000µl LB media). The third treatment also started wtih JV3 and *E. coli* MG1655 in a 1:1 ratio but also received 5µl of fresh MG1655 cells (OD_{600nm} of 0.5) at every transfer. The third treatment was a precaution for the possibility of MG1655 extinction due to the competitive advantage of JV3. All other details were the same as for the evolution experiment described above. All three treatments were serially passaged for 15 transfers after which samples were frozen and kept at -80 ^{o}C.

Whole genome sequencing

We extracted genomic DNA of both ancestral and evolved strains using the Qiagen Blood and Tissue Kit on overnight cultures overnight. Library preparation and whole genome sequencing was done by Microbial Genome Sequencing Center (https://www.migscenter.com/) that uses Illumina DNA library preparations and NextSeq 2000 platform for sequencing. Ancestral genomes were sequenced to 300 Mbs depth whereas derived lines were sequenced to 150 Mbs depth.

Variant analysis

FASTQC (version 0.11.5) was used to monitor quality of sequence reads before and after trimming [26]. Trimmomatic (version 0.38) was used to remove low quality bases from 3' end version[27]. SPAdes (version 3.11.1) was used to assemble short reads into contigs [28]. The genome was annotated by Prokka pipeline (version 1.13) [29]. Read mapping to reference ancestral strain and variant calling was done by GATK pipeline (version v4.0.5.1) [30]. The resulting VCF and BAM files were visualized using online version of IGV (https://igv.org/app/).

Results

S. maltophilia strains vary in virulence and competitive ability.

Consistent with previous findings [25], we observed that different *S. maltophilia* strains vary in their virulence towards *C. elegans*. Among the strains tested, JV3 has the highest virulence, JCMS and R551-3 have intermediate virulence whereas, and K279a is the least virulent strain (Figure 3.1a). The mean survival time of *C. elegans* interacting with JV3 is significantly lower than than of worms interacting with JCMS, R551-3, and K279a (Log-rank test: Table S2). Similarly, *C. elegans* interacting with JCMS has significantly lower mean survival time than those fed on K279a, whereas the mean

survival time of *C. elegans* with interacting with R551-3 and K279a are not significantly different. *C. elegans* interacting with JV3 and JCMS have significantly higher *S. maltophilia* load (within-host CFUs) than worms interacting with R551-3 and K279a (Figure S1). JV3 strain also shows the highest levels of interference effects as shown by the change in the ratio of *S. maltophilia* to *E. coli* densities (Figure 1b, $p<0.01$). R551-3 and JCMS strains also strongly antagonize MG1655. K27a shows relatively weak but still has statistically significant interference competition effects on *E. coli* MG1655 ($p<0.05$).

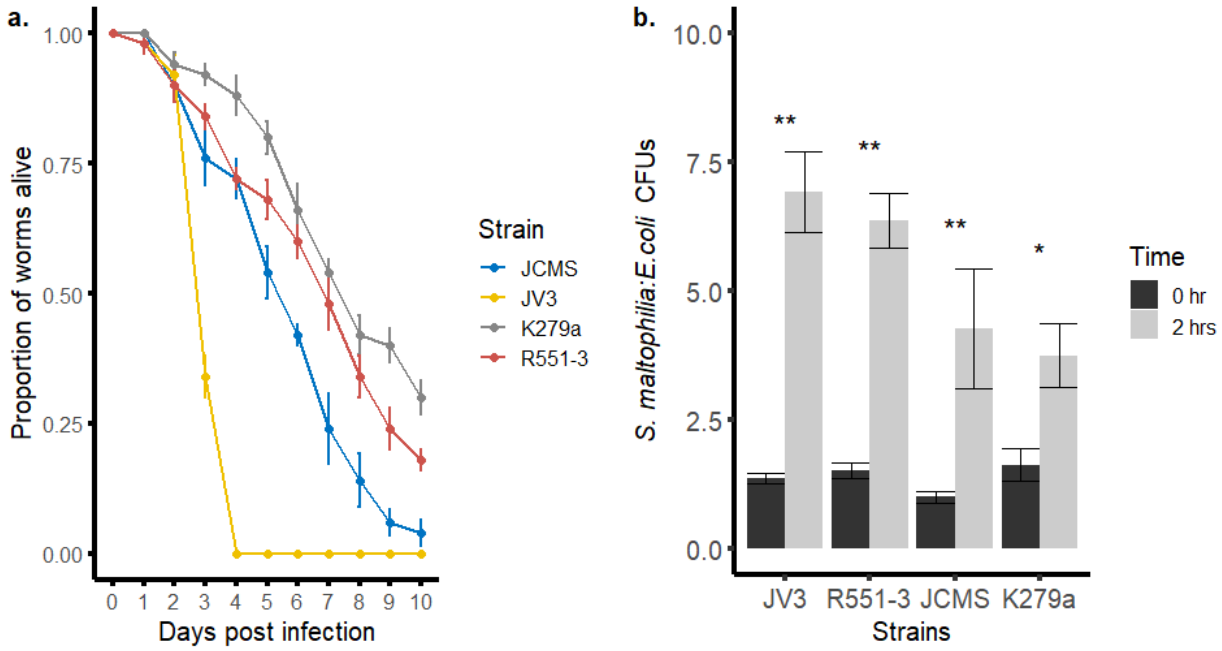

Figure 3.1. *S. maltophilia* strains vary in their virulence in *C. elegans* and their interference antagonism of *E. coli* MG1655. a. The mean proportion of worms alive over the course of 10 days of interaction with different *S. maltophilia* strains. Error bars show standard error (n=5 each). b. The effect of several *S. maltophilia* strains on the recovery of *E. coli* MG1655 cells. Y-axis represents CFU ratio of *S. maltophilia* to *E. coli* MG1655. Data are represented as means and

standard errors (n=6 each). The asterisks represent significant differences between the two time points as calculated from Wilcoxon test (*, p<0.05; **, p<0.01; ***p, <0.001).

Highly virulent strains evolve reduced virulence in unstructured environments.

To determine how virulence evolves in an unstructured environment, we serially transferred populations derived from all four strains JV3, JCMS, R551-3, and K279a in LB media. We measured the survival of *C. elegans* after 72 hours of interaction as a measure of virulence for both ancestral and evolution lines. We observed a significant reduction in virulence in all populations initiated with the highly virulent strain JV3 (Figure 2a, p<0.05). Similarly, JCMS and R551-3 populations also evolved significant lower virulence compared (Figure 2b, 2c). In contrast, populations of the lowest virulence strain, K279a, did not significantly change in virulence (Figure 2d).

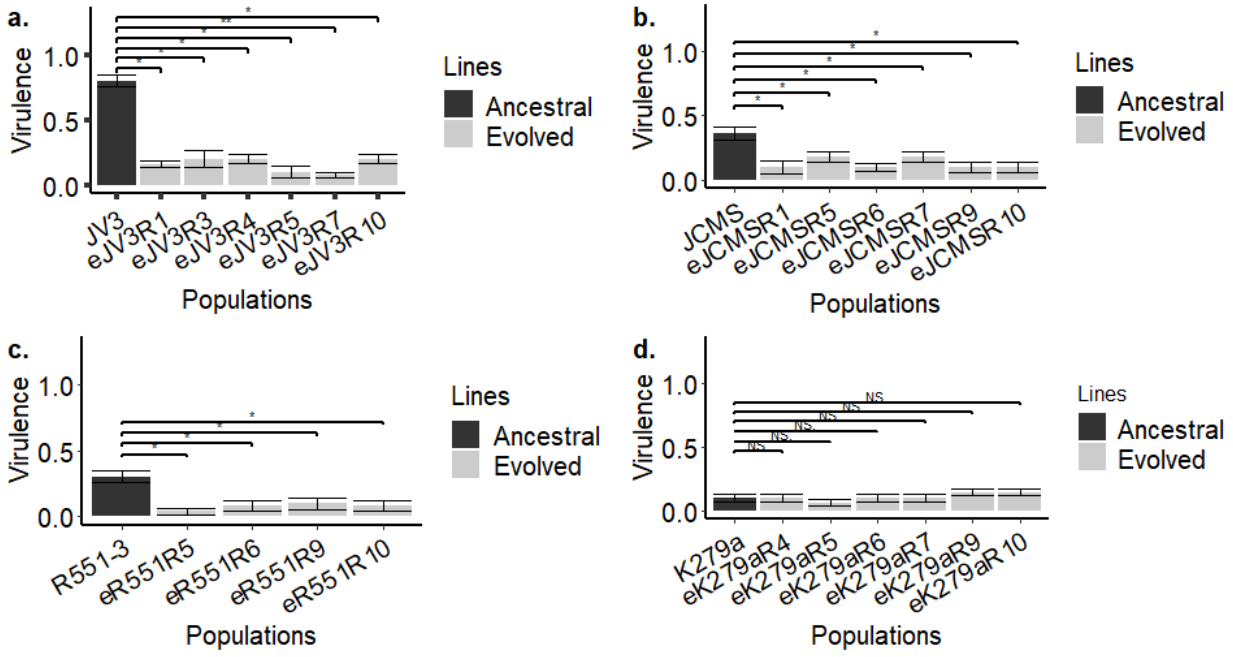

Figure 3.2. Virulence in evolution lines derived from the ancestral strains JV3 (a.), JCMS (b.), R551-3 (c.), and K279a (d.). Virulence in this case is the proportion of dead worms after 72 hours

of infection. Data represents mean and standard error (n=5 each), and significance is calculated using the Wilcoxon test.

Evolved populations show reduced competitive ability in independent evolution lines.

Similar to the observed changes in virulence, we observed that JV3 populations evolved significantly lower interference competition effects on MG1655 (Figure 3a). One of these populations (eJV3R5) retained significant, albeit reduced, antagonism of MG1655, while the remaining JV3 populations had no effect the CFU ratio over the course of two hours. The degree of interference was similarly reduced in populations derived from JCMS and R551-3 (Figure 3b, 3c), with 4 populations (eJCMSR6, eJCMSR9, eR551-R9, eR551-3R10) still causing significant differences between the initial and final CFU ratios. All evolved K279a populations lost the ability to antagonize MG1655 (Figure 3d).

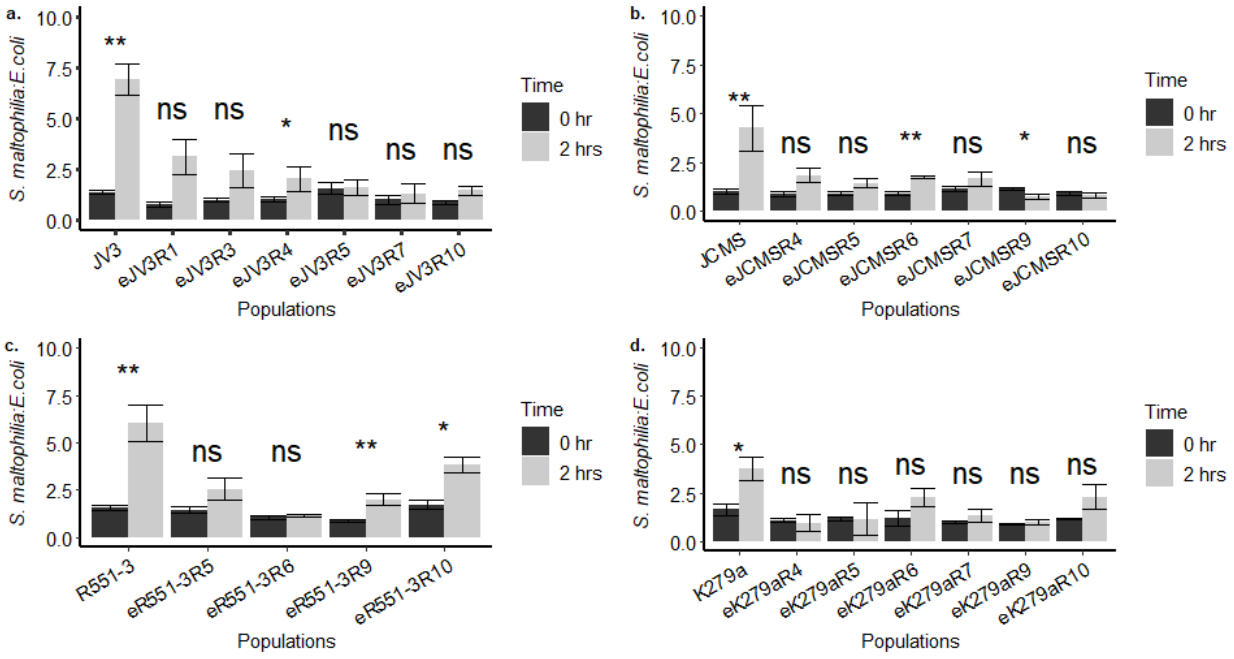

Figure 3.3. Competitive ability in evolution lines derived from the ancestral strains JV3 (a.), JCMS (b.), R551-3 (c.), and K279a (d.). All the bars whose labels start with “e” represent evolution lines.

Data represents mean and standard error (n=5 each), and significance is calculated using the Wilcoxon test.

Low virulence evolved derivatives of the high virulence strain have a mutation in the type IV secretion system operon.

We obtained the whole genome sequences of 5 evolution lines of JV3 in addition to the whole genome sequence of all 4 ancestral strains. Variant analysis revealed numerous SNPs and indel mutations throughout the genome sequences of evolution lines. We found an identical substitution mutation in the type IV secretion system operon of all JV3 evolved isolates (Figure 4). We also observed extensive mutations in the putative coding region of A-type flagellin in all sequenced evolution lines.

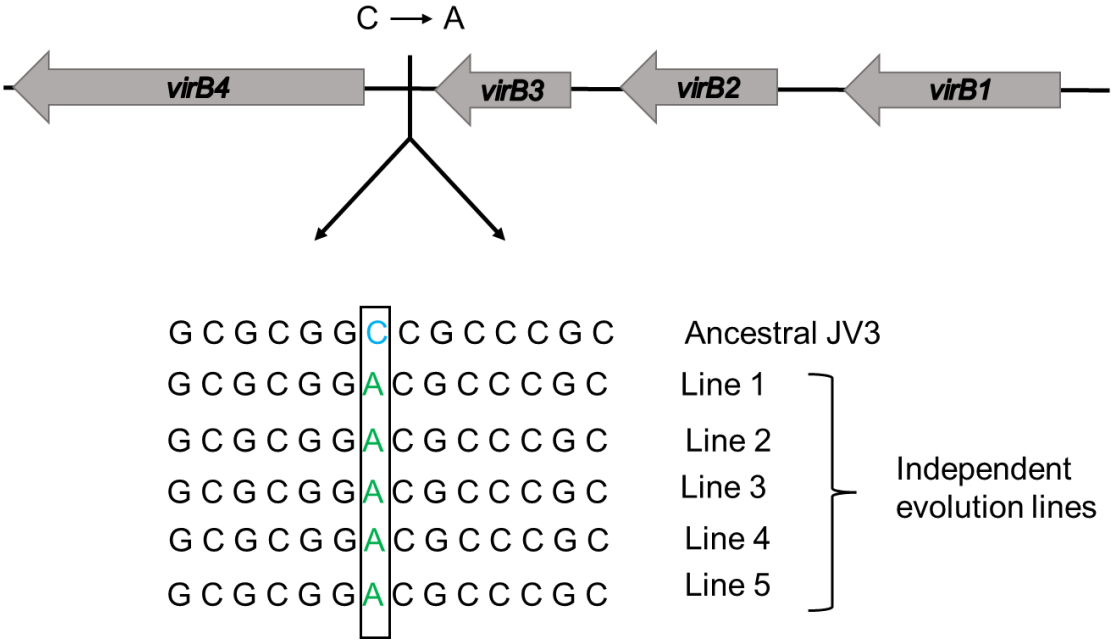

Figure 3.4. Variant analysis comparing ancestral and JV3 evolved isolates identifies a mutation in the type IV secretion system operon of the evolved strains. All 5 sequenced evolved JV3 isolates have the same substitution mutation in the type IV secretion system locus.

Type IV secretion system and its associated effector proteins influence both virulence and competitive ability.

To understand the role of the type IV secretion system in virulence, we measured the survival of *C. elegans* interacting with K279a derivatives where a gene required for type IV secretion system function (*virB10*) and two genes encoding key effector proteins (*14245* and *14255*) were deleted [29]. We also measured the interference competitive ability of these strains against *E. coli* MG1655. We observed that *C. elegans* have greater survival when interacting with K279a Δ*virB10* gene than when interacting with K279a (Figure 5). Further, complementation of K279a Δ*virB10* with a plasmid-borne expression construct (pP*lac*::*virB10*) restores *C. elegans* survival to a level that is not significantly different from the survival with the wild-type K279a strain. The survival of *C. elegans* interacting with the effector mutants (Δ*14245*, Δ*14255*, and Δ*14245*Δ*14255*) is significantly higher than survival when interacting with the K279a wild-type. The survival on the complemented strain 14245/14245+ and 14255/14255+ is not significantly different from the survival on the wild-type strain. Unlike the K279a parental strain, each of these deletion strains (Δ*virB10*, Δ*14245*, Δ*14255*, Δ*14245*Δ*14255*) do not antagonize MG1655. Complementation of each single mutant with the with the appropriate expression plasmid (i.e., Δ *virB10* pP*lac*::*virB10*, Δ14245 pP*lac*::*14245*, Δ14255 pP*lac*::*14255*) (Figure S3) restores the ability to antagonize MG1655.

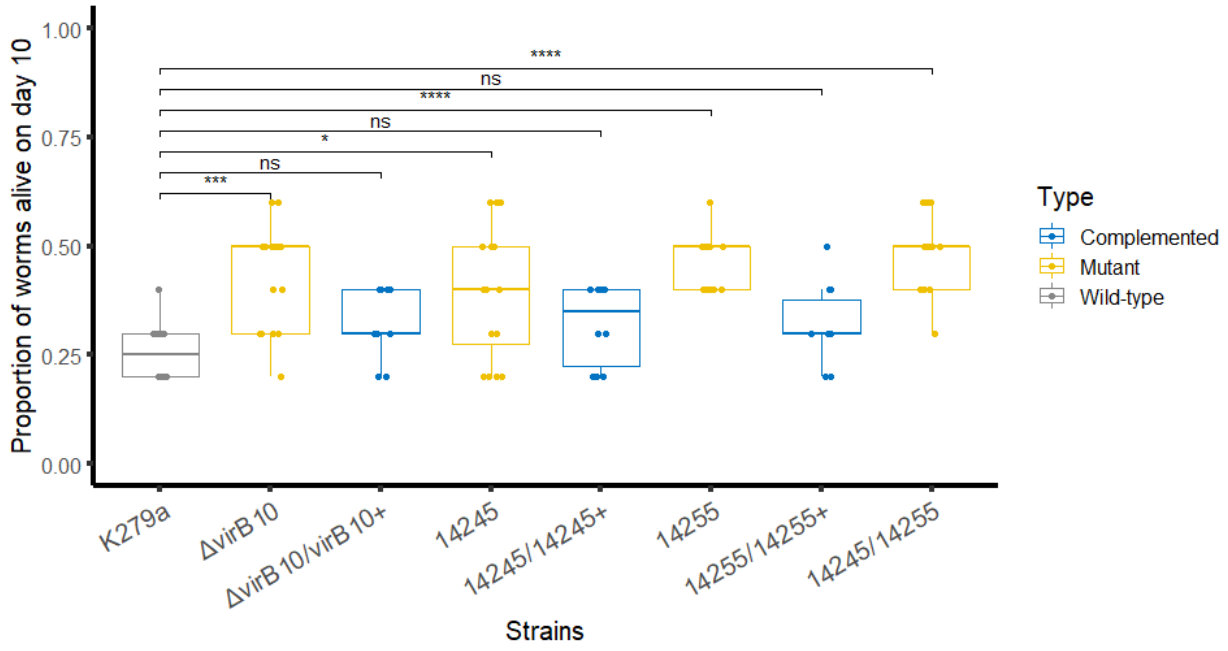

Figure 3.5. Effect of type IV secretion system and its associated effector proteins in survival of C. elgans. Significance is evaluated using the Wilcoxon test from the data combining 3 independent experiments (n=4 plates, each containing 10 worms). * indicates p<0.05, *** indicates p<0.001, ns indicates non-significant.

Virulence is reduced to a lesser extent when the high virulence strain evolves in the presence of a competitor

To test whether the presence of a competitor influences virulence evolution, we evolved replicate populations of the highly virulent JV3 strain under 3 treatment regimens– without *E. coli*, with *E. coli*, with *E. coli* and addition of *E. coli* cells at every transfer. In all treatments, host survival is higher than the ancestral strain (Figure 6; Figure S4). However, mean host survival when interacting with isolates from the populations evolving without *E. coli* is significantly greater than mean host survival when interacting with isolates from the populations that evolved with a competitor (Table S3). Addition of

fresh *E. coli* at every transfer did not show any effect compared to the treatment where *E. coli* was only introduced at the start of serial passaging (χ^2=0.93, p=0.33, Table S3).

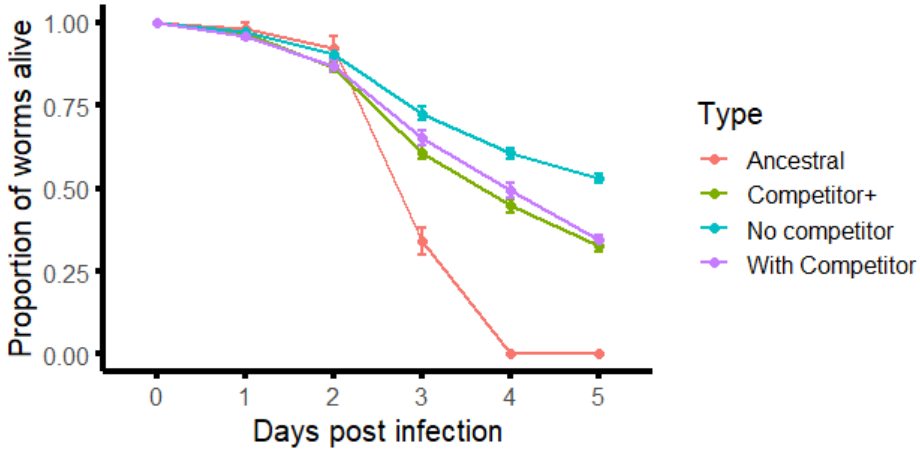

Figure 3.6. Survival curve of *C. elegans* interacting with the JV3 ancestral strain (red) and derivatives from JV3 populations that evolved with (purple and green) and without (blue) the competitor *E. coli* MG1655. The purple line indicates evolution with the competitor only introduced at the start of the experiment, while the green line indicates evolution with additional introduction of *E. coli* at each passage. Data represents mean and standard error for 8 population lines (each replicated 5 times with 10 worms on each).

Discussion

Facultative pathogens like *S. maltophilia* navigate widely varying environments within host and in environmental reservoirs. Selection pressure influencing environmental survival can influence consequences of within-host interactions. Coincidental hypothesis posits that virulence in these pathogens may be a coincidental result of adaptation to non-host conditions. In this study, we show that the degree of interference antagonism of *E. coli* MG1655 competitors and virulence in *C. elegans* hosts are both influenced by a type IV secretion system and its associated effector proteins. Further evolution under relaxed selection results in the evolution of low virulence and low competitive ability in populations

derived from highly competitive and highly virulent strains. In addition, we also show that presence of a competitor can help maintain relatively high virulence in the highly virulent strain of *S. maltophilia*.

Type IV secretion is widespread among Gram-negative bacteria and is often studied in the context of host-pathogen interactions. Several pathogens use the type IV secretion systems to deliver effector proteins into host cells, resulting in either host manipulation or cell death [36,37]. *S. maltophilia* uses a type IV secretion system to modulate apoptosis of macrophages [29]. In addition to being involved in host-pathogen interactions, the many members of the Xanthomonas family, including *S. maltophilia*, use the type IV secretion system to deliver toxins that antagonize heterologous competitor bacteria [28–30]. Here we show that the same effector proteins (those encoded by the *14245* and *14255* loci) can influence both virulence in *C. elegans* and interference competition against *E. coli* MG1655. Since strains differ in the degree of interference competition and virulence they show, we hypothesize that expression levels impacting effector protein delivery likely differ between low and high virulence strains.

Potential fitness costs associated with virulence factors within a host can be offset by the fitness benefits they provide in the non-disease context, such as in environmental reservoirs [7,38]. Biotic interactions such as predation can contribute to virulence maintenance in *E. coli* [39], *Pseudomonas aeruginosa* [40], *Vibrio cholerae* [6] and other *Vibrio* spp [41,42]. Contrasting outcomes can also occur when virulence is attenuated by predation [5,43]. Whether high virulence is tolerated depends on the mechanisms shaping fitness in environmental reservoirs and host-pathogen interactions [3]. In some cases, there may be positive pleiotropy such that shared mechanisms increase both environmental survival

and virulence [44]. In contrast, in other cases, there may be antagonistic pleiotropy resulting in virulence attenuation when increased environmental survival comes at the expense of fitness within the host (e.g. due to reduced ability to evade host immune responses).

In the case of *S. maltophilia*, we have shown that virulence and interference competitive antagonism are positively correlated (Figure S2), and both depend on the same molecular mechanism-- specifically the function of a type IV secretion system. The ability to antagonize competitors can increase fitness in environmental reservoirs. Because the same mechanisms that mediate interference competition are also virulence factors mediating high virulence, selective pressures associated with interspecific competition in environmental reservoirs may coincidentally select increased virulence.

Our study has focused on the role of the type IV secretion system, but other factors such as diffusible signal factors [45] and the effectors of type II secretion systems [46,47] also contribute to *S. maltophilia* virulence. These molecules are not known to contribute to interference competition and are unlikely to explain concurrent evolution of reduced virulence and competitive ability we observed. Similarly, flagella can contribute to surface attachment [48] and may thus contribute to virulence in *C. elegans* by facilitating gut surface colonization. We did observe several substitution mutations in the *fliC1* gene of our evolved JV3 isolates. These mutations may contribute to the reduced virulence we observed in those evolution lines, but it is less clear how they can account for the reduced interference antagonism we observed. Thus, the concurrent evolution towards low virulence and low interference competition we observed most likely stems from the mutation we observed in the JV3 type IV secretion system locus.

Facultative pathogens experience diverse selective pressures in environmental reservoirs and when associating with hosts. Consequently, the virulence of these pathogens may be analogous to the spandrels of San Marco [49] rather than an adaptive response to the host. In accordance with the coincidental selection hypothesis, we have shown that selection for interference competition antagonism can help maintain high virulence in *S. maltophilia*. This can have important consequences for chronic polymicrobial infections such as those of cystic fibrosis patients where strong selective pressures for competitive advantages could also select for more virulent strains. Thus, further research into specific effector proteins that yield a competitive edge to *S. maltophilia* strains such as JV3 and JCMS is required as they may also contribute to higher virulence in these strains. For this reason, our study motivates integrative approaches examining how virulence factors evolve in different contexts as a key element of optimal virulence management strategies.

Acknowledgements

We would like to thank Mike Herman, Leah Radeke, and Anna Zinovyeva for providing materials. Research reported in this publication was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health under award P20GM130448. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1 Caraco T, Wang IN. Free-living pathogens: Life-history constraints and strain competition. *Journal of Theoretical Biology* 2008; **250**: 569–579.

2 Brown SP, Cornforth DM, Mideo N. Evolution of virulence in opportunistic pathogens: Generalism, plasticity, and control. *Trends in Microbiology* 2012; **20**: 336–342.

3 Pandey A, Mideo N, Platt TG. Virulence evolution of pathogens that can grow in reservoir environments. *The American Naturalist* 2022; **199**: 141–158.

4 Barton IS, Fuqua C, Platt TG. Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. *Environmental Microbiology* 2018; **20**: 16–29.

5 Leong W, Poh WH, Williams J, Lutz C, Hoque MM, Poh YH *et al.* Adaptation to an amoeba host leads to *Pseudomonas aeruginosa* isolates with attenuated virulence. *Applied and Environmental Microbiology* 2022; **88**: aem-02322.

6 Hoque MM, Noorian P, Espinoza-Vergara G, Manuneedhi Cholan P, Kim M, Rahman MH *et al.* Adaptation to an amoeba host drives selection of virulence-associated traits in *Vibrio cholerae*. *The ISME Journal 2021* 2021; **16**: 1–12.

7 Levin BR. The evolution and maintenance of virulence in microparasites. *Emerging Infectious Diseases* 1996; **2**: 93–102.

8 Levin BR, Edén CS. Selection and evolution of virulence in bacteria: an ecumenical excursion and modest suggestion. *Parasitology* 1990; **100**: S103–S115.

9 Berg G, Eberl L, Hartmann A. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environmental Microbiology* 2005; **7**: 1673–1685.

10 Ryan RP, Monchy S, Cardinale M, Taghavi S, Crossman L, Avison MB *et al.* The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nature Reviews Microbiology* 2009; **7**: 514–525.

11 Verweij PE, Meis JFGM, Christmann V, van der Bor M, Melchers WJG, Hilderink BGM *et al.* Nosocomial outbreak of colonization and infection with *Stenotrophomonas maltophilia* in preterm infants associated with contaminated tap water. *Epidemiology & Infection* 1998; **120**: 251–256.

12 Abbott IJ, Slavin MA, Turnidge JD, Thursky KA, Worth LJ. *Stenotrophomonas maltophilia* : emerging disease patterns and challenges for treatment. *Expert Review of Anti-infective Therapy* 2011; **9**: 471–488.

13 Brooke JS. New strategies against *Stenotrophomonas maltophilia* : a serious worldwide intrinsically drug-resistant opportunistic pathogen. *Expert Review of Anti-infective Therapy* 2014; **12**: 1–4.

14 Brooke JS. *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clinical Microbiology Reviews* 2012; **25**: 2–41.

15 Denton M, Kerr KG. Microbiological and clinical aspects of infection associated with *Stenotrophomonas maltophilia*. *Clinical Microbiology Reviews* 1998; **11**: 57–80.

16 Rutter WC, Burgess DR, Burgess DS. Increasing incidence of multidrug resistance among cystic fibrosis respiratory bacterial isolates. *https://home.liebertpub.com/mdr* 2017; **23**: 51–55.

17 Valdezate S, Vindel A, Martín-Dávila P, del Saz BS, Baquero F, Cantón R. High genetic diversity among *Stenotrophomonas maltophilia* strains despite their originating at a single hospital. *Journal of Clinical Microbiology* 2004; **42**: 693–699.

18 Gajdács M, Urbán E. Prevalence and antibiotic resistance of *Stenotrophomonas maltophilia* in respiratory tract samples: A 10-year epidemiological snapshot. *Health services research and managerial epidemiology* 2019; **6**: 2333392819870774.

19 Yang S, Hua M, Liu X, Du C, Pu L, Xiang P *et al.* Bacterial and fungal co-infections among COVID-19 patients in intensive care unit. *Microbes and Infection* 2021; **23**: 104806.

20 Trifonova A, Strateva T. *Stenotrophomonas maltophilia* – a low-grade pathogen with numerous virulence factors. *Infectious Diseases* 2019; **51**: 168–178.

21 Falagas ME, Kastoris AC, Vouloumanou EK, Rafailidis PI, Kapaskelis AM, Dimopoulos G. Attributable mortality of *Stenotrophomonas maltophilia* infections: A systematic review of the literature. *Future Microbiology* 2009; **4**: 1103–1109.

22 Adamek M, Linke B, Schwartz T. Virulence genes in clinical and environmental *Stenotrophomas maltophilia* isolates: A genome sequencing and gene expression approach. *Microbial Pathogenesis* 2014; **67–68**: 20–30.

23 Kalidasan V, Joseph N, Kumar S, Awang Hamat R, Neela VK. Iron and virulence in *Stenotrophomonas maltophilia*: All we know so far. *Frontiers in Cellular and Infection Microbiology* 2018; **8**: 401.

24 di Bonaventura G, Prosseda G, del Chierico F, Cannavacciuolo S, Cipriani P, Petrucca A *et al.* Molecular characterization of virulence determinants of *Stenotrophomonas maltophilia* strains isolated from patients affected by cystic fibrosis. *International Journal of Immunopathology and Pharmacology* 2007; **20**: 529–537.

25 White C v, Darby BJ, Breeden RJ, Herman MA. A *Stenotrophomonas maltophilia* strain evades a major *Caenorhabditis elegans* defense pathway. *Infection and Immunity* 2016; **84**: 524–536.

26 White C v., Herman MA. Transcriptomic, functional, and network analyses reveal novel genes involved in the interaction between *Caenorhabditis elegans* and *Stenotrophomonas maltophilia*. *Frontiers in Cellular and Infection Microbiology* 2018; **8**: 266.

27 Radeke LJ, Herman MA. Identification and characterization of differentially expressed genes in *Caenorhabditis elegans* in response to pathogenic and nonpathogenic *Stenotrophomonas maltophilia*. *BMC Microbiology* 2020; **20**: 1–20.

28 Nas MY, White RC, DuMont AL, Lopez AE, Cianciotto NP. *Stenotrophomonas maltophilia* encodes a VirB/VirD4 Type IV Secretion System that modulates apoptosis in human cells and promotes competition against heterologous bacteria, including *Pseudomonas aeruginosa*. *Infection and Immunity* 2019; **87**: 10–12.

29 Nas MY, Gabell J, Cianciotto NP. Effectors of the *Stenotrophomonas maltophilia* type IV secretion system mediate killing of clinical isolates of *Pseudomonas aeruginosa*. *mBio* 2021; **12**: e01502-21.

30 Bayer-Santos E, Cenens W, Matsuyama BY, Oka GU, di Sessa G, Mininel IDV *et al.* The opportunistic pathogen *Stenotrophomonas maltophilia* utilizes a type IV secretion system for interbacterial killing. *PLOS Pathogens* 2019; **15**: e1007651.

31 Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed 5 Feb2022).

32 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–2120.

33 Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology* 2012; **19**: 455–477.

34 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.

35 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 2010; **20**: 1297–1303.

36 Cascales E, Christie PJ. The versatile bacterial type IV secretion systems. *Nature Reviews Microbiology 2003 1:2* 2003; **1**: 137–149.

37 Voth DE, Broederdorf LJ, Graham JG. Bacterial type IV secretion systems: versatile virulence machines. *Future Microbiology* 2012; **7**: 241–257.

38 Read AF. The evolution of virulence. *Trends in Microbiology* 1994; **2**: 73–76.

39 Adiba S, Nizak C, van Baalen M, Denamur E, Depaulis F. From grazing resistance to pathogenesis: The coincidental evolution of virulence factors. *PLoS ONE* 2010; **5**: e11882.

40 Hosseinidoust Z, van de Ven TGM, Tufenkji N. Evolution of *Pseudomonas aeruginosa* virulence as a result of phage predation. *Applied and Environmental Microbiology* 2013; **79**: 6110–6116.

41 Robino E, Poirier AC, Amraoui H, le Bissonnais S, Perret A, Lopez-Joven C *et al.* Resistance of the oyster pathogen *Vibrio tasmaniensis* LGP32 against grazing by *Vannella* sp. marine amoeba involves Vsm and CopA virulence factors. *Environmental Microbiology* 2019; : 1462-2920.14770.

42 Espinoza-Vergara G, Hoque MM, McDougald D, Noorian P. The impact of protozoan predation on the pathogenicity of *Vibrio cholerae*. *Frontiers in Microbiology* 2020; **11**: 17.

43 Kortright KE, Done RE, Chan BK, Souza V, Turner PE. Selection for phage resistance reduces virulence of *Shigella flexneri*. *Applied and Environmental Microbiology* 2022; **88**: AEM-01514.

44 Sun S, Noorian P, McDougald D. Dual role of mechanisms involved in resistance to predation by protozoa and virulence to humans. *Frontiers in Microbiology* 2018; **9**: 1017.

45 Alcaraz E, García C, Friedman L, de Rossi BP. The *rpf* /DSF signalling system of *Stenotrophomonas maltophilia* positively regulates biofilm formation, production of virulence-associated factors and β-lactamase induction. *FEMS Microbiology Letters* 2019; **366**: fnz069.

46 DuMont AL, Cianciotto NP. *Stenotrophomonas maltophilia* serine protease StmPr1 induces matrilysis, anoikis, and protease-activated receptor 2 activation in human lung epithelial cells. *Infection and immunity* 2017; **85**: e00544-17.

47 Karaba SM, White RC, Cianciotto NP. *Stenotrophomonas maltophilia* encodes a type II protein secretion system that promotes detrimental effects on lung epithelial cells. *Infection and immunity* 2013; **81**: 3210–3219.

48 Pompilio A, Crocetta V, Confalone P, Nicoletti M, Petrucca A, Guarnieri S *et al.* Adhesion to and biofilm formation on IB3-1 bronchial cells by *Stenotrophomonas maltophilia* isolates from cystic fibrosis patients. *BMC Microbiology* 2010; **10**: 1–15.

49 Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London Series B Biological Sciences* 1979; **205**: 581–598.

Chapter 4 - Neutral and niche-based assembly of microbiome communities in agricultural and prairie soils

Aakash Pandey[1*], Paige Hansen[2, 3], Willow Kessler[2], Marcos Sarto[4], Ligia Souza[2], Austin Yoder[5], James D. Bever[2], Sharon Billings[2], Matthew Kirk[6], Terry Loecke[5,7], Charles W. Rice[4], Benjamin A. Sikes[2,7], Thomas G. Platt[1]

[1]Division of Biology, Kansas State University, Manhattan, KS, USA

[2]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA

[3]Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, USA

[4]Department of Agronomy, Kansas State University, Manhattan, KS, USA

[5]Kansas Biological Survey, University of Kansas, Lawrence, KS, USA

[6]Department of Geology, Kansas State University, Manhattan, KS, USA

[7]Department of Environmental Studies, University of Kansas, Lawrence, KS, USA

*** Correspondence:**

Aakash Pandey

evomathbio@gmail.com

Abstract

Neutral and niche-based processes both influence bacterial and fungal community structure. The relative importance of these processes, however, is poorly understood—including how their contributions may differ across microbiomes from distinct areas or over time. Here we evaluate how well neutral and niche-based processes account for microbiome composition observed in agricultural and prairie soils over two seasons. We compare observed soil community composition to neutral and null model predictions, and evaluate a wide range of abiotic factors (e.g., depth, pH, land use history) as potential environmental filters causing deviations from the neutral model. We find that the neutral model better accounts for the composition of bacterial communities than it did for fungal communities. Further, the neutral model best accounted for soil bacterial community composition at intermediate soil depth (15 – 30 cm) at all sites. This pattern was, however, not observed in fungal samples. Nearest taxon index (NTI) shows significant phylogenetic clustering for most bacterial and some fungal samples. The null models comparing samples between shallow and deeper depths suggest that variable selection influences these communities. In contrast, differences in communities from similar depths were mostly explained by either homogeneous selection or stochastic processes. Correlations between βNTI and environmental factors such as pH, moisture, and enzymatic activities show that pH influences bacterial but not fungal communities, and the degree of influence varies between seasons and across sites. Together, these results show that the degree to which neutral and niche-based processes shape bacterial and fungal community assembly is context dependent, with specific factors like soil depth, and pH influencing differences in their relative importance.

Introduction

Microorganisms have diverse ecological roles and contribute to several ecosystem services (McKenney et al., 2018; Cavicchioli et al., 2019; Dodds et al., 2020). Understanding how microbial communities assemble has implications in many areas including agriculture, land management, wastewater treatment, human and animal health, and biotechnology (Ofiţeru et al., 2010; D'Haens and Jobin, 2019; Arif et al., 2020; Duncker et al., 2021; Estrela et al., 2021). The combination of deterministic and stochastic processes influence community assembly and the degree of influence can differ between communities in different environments. Recent technological and methodological advancements have made it possible to assess the relative importance of these assembly processes in shaping microbial communities.

The diverse processes that contribute to the assembly of communities can be broadly categorized into four types: selection, drift, dispersal, and diversification (Vellend, 2010; Nemergut et al., 2013). The relative importance of each of these processes during community assembly is often debated. Abiotic conditions (e.g., temperature, pH, and moisture, etc.) can act as environmental filters that determine community composition (Kraft et al., 2015; Cadotte and Tucker, 2017). Similarly, biotic interactions (e.g., competition, predation, mutualism, facilitation, etc.) can result in fitness differences among organisms of different types shaping community structure(Yan et al. 2021; Wang et al. 2020). Homogeneous biotic and abiotic factors lead to similar pressures driving assembly thereby resulting in communities having similar structures. Conversely, heterogeneous driving factors (e.g., variable selection) lead to differences in community structures (Stegen et al., 2013; Zhou and Ning, 2017). Besides these deterministic

processes, stochastic events (e.g., birth, death, emigration, and immigration) can also change community composition (Hubbell, 2006; Woodcock et al., 2007; Vellend et al., 2014; Sieber et al., 2019). Drift plays an important role if the selection is weak, and the community size is small (Orrock and Watling, 2010; Chase and Myers, 2011). The metacommunity species pool is composed of all taxa in the region, with dispersal from this species pool shaping local communities (Chave, 2004; Nemergut et al., 2013). High (homogenizing) dispersal leads to communities having similar structures whereas dispersal limitation together with drift and diversification leads to the differences between communities (Stegen et al., 2012, 2015; Dini-Andreote et al., 2015; Zhou and Ning, 2017). A key objective of this paper is to assess how these factors jointly shape microbiome community structure and which of these processes dominate the assembly of soil microbial communities across a range of land use and spatial (e.g., soil depth) contexts.

Soil bacterial and fungal communities are important drivers of nutrient cycling and plant health (Saleem et al., 2019; Naylor et al., 2020). Soil provides diverse and heterogeneous environments. Consequently, several factors can influence community assembly of these microbiomes. Abiotic factors such as pH (Lauber et al., 2009; Tripathi et al., 2018), temperature (Jiao and Lu, 2020), salinity (Zhang et al., 2019), moisture and precipitation (Chase, 2007) affect their community structure. Similarly, land use practices can also significantly alter microbiome communities (Jiao and Lu, 2020; Jiao et al., 2020; Köberl et al., 2020). Neutral processes of random demographic events along with these deterministic factors combine to shape microbial community structure (Wang et al., 2020; Dong et al., 2021; Yan et al., 2021). However, fungal and bacterial communities can show distinct patterns of community assembly (Bahram et al., 2018). Understanding the factors

governing the assembly of these microbiomes is broadly relevant to their importance to soil function and plant health.

Several techniques are routinely used to tease apart the influence of deterministic and stochastic processes during community assembly. Most of these techniques fall into three broad categories – multivariate analysis, neutral model analysis, and null model analysis (Zhou and Ning, 2017). Multivariate analysis incorporates several statistical approaches such as ordination-based analysis, correlation-based analysis, and variation partitioning. The neutral model assumes that the demographic events of birth, death, and dispersal are stochastic with respect to the taxa and relates the observed relative taxa abundance to their frequency in the global species pool. Null models work by comparing the expected results obtained from randomization with the observed results. Patterns in ecology are often multi-causal. Since different processes can produce the same pattern, we perform multiple analyses with different underlying assumptions in a complimentary fashion to confidently infer which factors underlie the observed patterns (Zhou and Ning, 2017).

In this manuscript, we use the neutral community model developed by Sloan et. al., that is specifically suited to microbial communities (Sloan et al., 2006; Woodcock et al., 2007) to fit to the soil microbiome data collected across Kansas. We use goodness-of-fit measures to evaluate the strength of the fit to the available data. We find that stochastic processes often influence soil microbiome assembly at all land uses and sites, more so in bacterial communities compared to fungal communities. In addition, we use two different types of null models – composition-based and phylogeny-based null models to disentangle different assembly processes. We use the Raup-Crick metric (RC_{bray})

(Chase et al., 2011) modified to incorporate taxa's relative abundance (Stegen et al., 2013) as a community composition based null model. This method uses the taxa co-occurrence data for a given pair of samples to infer whether stochastic, homogenizing, or diversifying forces can explain differences between community compositions. We also used the phylogeny-based null model that takes phylogenetic relationships between taxa into consideration (Stegen et al., 2012; Wang et al., 2013; Dini-Andreote et al., 2015). We show that the community differences observed in the null models correlate with soil depth profiles and pH differences between the samples for bacterial samples. However, we do not see this pattern in fungal communities. We also do not see the effect of moisture content on both bacterial and fungal communities. Taken together, we show that distinct processes are dominant in soil bacterial and fungal communities and the dominance is context dependent.

Materials and Methods

Microbiome samples and sequences

Site selection

The state of Kansas has considerable variation in precipitation and land use history. Kansas has a precipitation gradient that ranges from 350 mm per year in the West to 1162 mm per year in the East (Figure 4.1A). The state is also comprised of a patchwork of land uses, including remnant, unplowed tall- and mixed grass prairies, active crop lands, and abandoned agricultural fields, some of which are now active restorations. We took advantage of Kansas's precipitation gradient and variable land use histories to assess how they and soil depth influence microbial community assembly dynamics. We selected sites whose historical precipitation regime represents relatively wet (Site 1, 1045

mm/y) intermediate (Site 2, 865 mm/y), and dry (Site 3, 533 mm/y), amounts of annual precipitation (Figure 4.1A). While the timing of seasonal rainfall events differs among these sites, precipitation is greater in the spring than in the fall in all three (Figure S4.1). At each location, we sampled three land use histories: a remnant native grassland, a post-agricultural restoration, and an active agricultural field (Figure 4.1B). Because relatively little land in Kansas represents remnant grassland, these land use histories are pseudoreplicated within site. To maintain consistency with availability of remnant prairie, we sampled from only one replicate of each land use type at each site. While broad, these three land use categories encompass the expected variation in land management and crops, as well as in soil property variables. For example, all three restorations (Sites 1, 2, and 3) were seeded with native forbs and grasses, but vary in management (e.g., mowing, burning) and in restoration age. Likewise, while all sampled agricultural fields are regularly tilled and planted with cover crops, they are rarely planted at the same time and with the same crops, including the year we sampled. Information about individual site location, land use history, annual precipitation, soil type, and plant community at time of sampling is in Table S1. At each site x land use history location, we established four plots to sample soils.

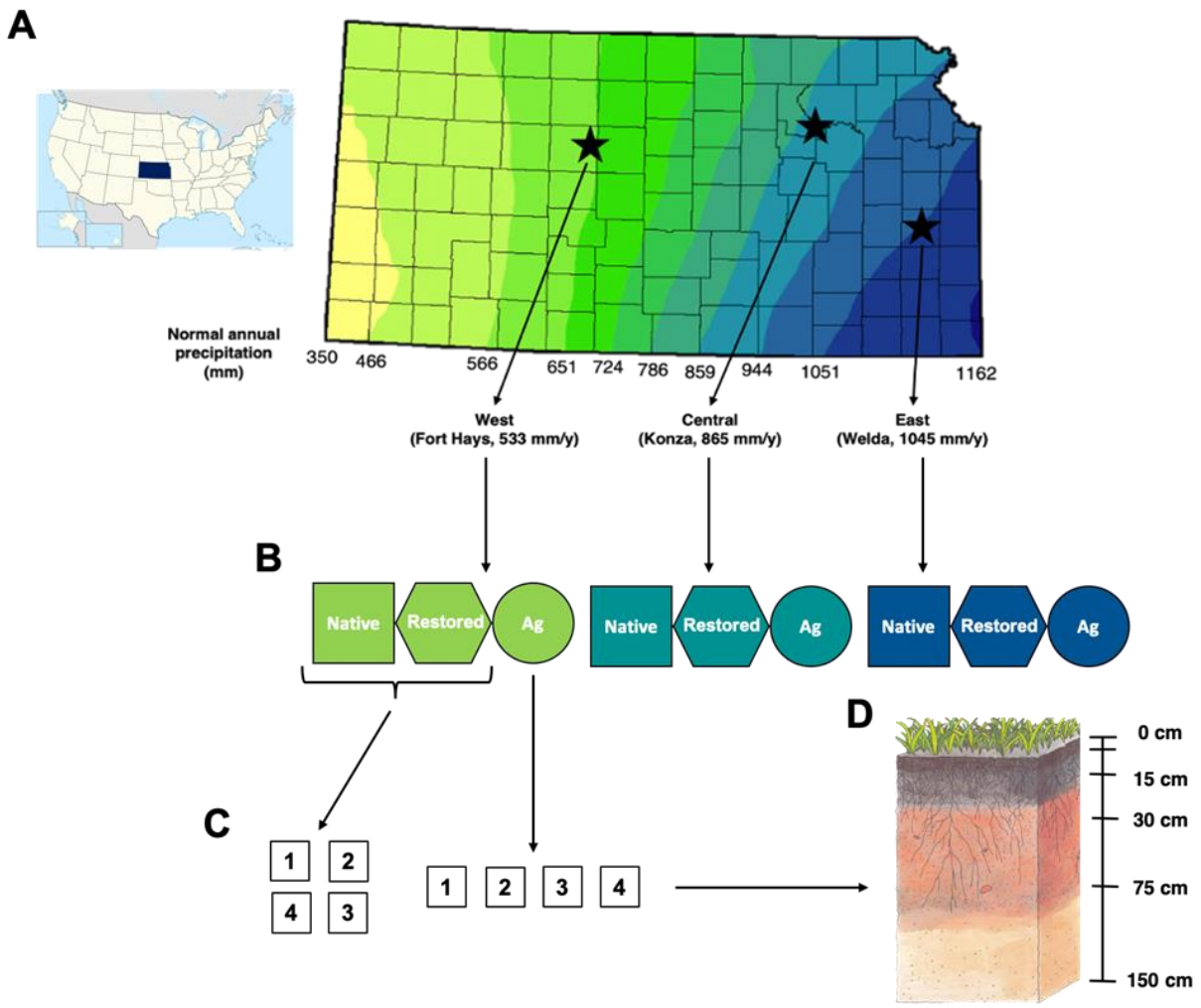

Figure 4.1. Complete sampling design. (A) Kansas's location in the United States, and the state's normal annual precipitation from 1981-2010. Lighter and darker colors represent lower and higher amounts of annual precipitation, respectively. Yearly precipitation estimates in mm are indicated along the bottom of the map. Stars represent sites selected for sampling. (B) At each site, we sampled from a native prairie, a restored prairie, and an agricultural field. (C) We sampled from 4 plots in each land use history. Native and restored plots were arranged quadratically, and agricultural plots were arranged linearly. (D) In each plot, 3 replicate soil cores were taken to 150 cm at the West site, and 120 cm at the Central and East sites. These cores were divided into 0-5 cm, 5-15 cm, 15-30 cm, 30-75 cm, and 75-150 or 75-120 cm increments. The 3 replicates from each plot were combined to create 1 single sample per plot. Figure adapted from Wikipedia

Commons, the Kansas Office of the State Climatologist, and the USDA-NRCS. Copyright (2020) by Paige Hansen. Reprinted with permission.

Soil sampling and processing

In each plot, we sampled soils at multiple depths using a Giddings probe (Giddings Machine Company, Windsor, CO, USA). Individual plots had adjacent subplots for spring (June 2018) and fall (September 2018) sampling (Figure 4.1C). At both time points, cores were taken to a maximum depth of 150 cm, and then divided into five depth increments: 0-5 cm; 5-15 cm; 15-30 cm; 30-75 cm; and 75-120 (Site 1 and 2) or 75-150 cm (Site 3; Figure 4.1D); maximum depth sampled was limited by equipment and rock layers at Site 1 and 2. Additionally, at the Site 1 native prairie site, a rock layer in the fall subplot prohibited sampling past 75 cm. At each sampling time, three soil cores (4 cm diameter) from each subplot were collected, then split by depth increment and pooled. For both spring and fall, this produced four pooled replicate samples for each depth, from each land use history, at each site. All soil samples were transported on ice to the nearest university (either University of Kansas or Kansas State University), stored overnight at 4°C, then thoroughly homogenized and subsampled. We subsampled 25 g of soil for community sequencing, and stored these at -80°C until further processing.

Generation and processing and 16S and ITS amplicon sequencing libraries

Immediately prior to DNA extraction, we once again homogenized and subsampled our soils to 5 g (Kang and Mills, 2006; Song et al., 2015). After performing a series of test DNA extractions aimed at identifying an extraction kit that performed optimally across all our sample types (see SI), we extracted total genomic DNA from 0.25 g of soil and from 1 blank (molecular-grade water) control per season using the DNeasy PowerSoil Pro Kit

(Qiagen, Germantown, MD, USA), following the manufacturer's protocol. Extract quality and DNA concentration were assessed using the Qubit 2.0 Flourometer (Thermo Fisher Scientific, Waltham, MA, USA) High-Sensitivity dsDNA assay.

We amplified the V4 region of the 16S small subunit ribosomal gene using the Earth Microbiome Project (Thompson et al., 2017) primers 515F-Y (Parada et al., 2016) and 806R (Apprill et al., 2015) to generate bacterial community data, and amplified the ITS2 region using primers ITS7 (Ihrmark et al., 2012) and ITS4 (White et. al. 1990) to generate fungal community data. After this first amplification step, we separately purified both 16S and ITS amplicons using the NucleoMag® PCR prep kit (Macherey-Nagel Gmbh & Co., Düren, Germany), then ligated Illumina® Nextera™ (Illumina, Inc., San Diego, CA, USA) indices to the purified amplicons in a second PCR reaction. Amplicons were purified a second time, then pooled in equimolar concentrations in two groups that corresponded to samples taken in the spring and fall. Pooled 16S and ITS libraries were sequenced at the University of Kansas Genomic Sequencing Core (Lawrence, KS, USA) and the Kansas State University Genomic Sequencing Core (Manhattan, KS, USA), respectively, using Illumina 2x300 bp MiSeq v2 chemistry. This resulted in 71,148 ± 20,223 and 86,905 ± 55,127 (mean ± sd) high-quality reads per sample in the spring and fall 16S pools, and 77,779 ± 19,124 and 86,365 ± 30,201 reads per sample in the spring and fall ITS pools. See SI for more information on both 16S and ITS library preparation.

All 16S and ITS sequences were processed separately in QIIME2.10 (Bolyen et al., 2019). Demultiplexed sequences from spring and fall runs were denoised using the QIIME2 implementation of DADA2 (Callahan et al., 2016), and combined. We grouped all both 16S and ITS sequences into amplicon sequence variants (ASVs) and open

reference operational taxonomic units (OTUs) at 97% sequence similarity using the UNITE database (Nilsson et al., 2019). We then created a 16S phylogenetic tree using the QIIME2 implementation of MAFFT fasttree (Price et al., 2010). Bacterial and fungal taxonomy was assigned by alignment to the SILVA (Quast et al., 2013) and UNITE databases, respectively, using DADA2 through R version 3.6.2 (R Development Core Team, 2019.). For both our bacterial and fungal data sets, we conservatively removed all samples containing less then 1000 reads, as well as all non-reproducible ASVs/OTUs (i.e., those that were observed less than 20 times and in under 5 samples). After sequence processing and quality filtering, our final 16S data set comprised a total of 7836 ASVs across 165 spring and 169 fall samples, and our final ITS data set included 1925 OTUs across 168 spring and 166 fall samples.

Neutral model

To assess the role of neutral processes during community assembly of microbiomes, we used the neutral community model modified for a large population (Sloan et al., 2006; Woodcock et al., 2007; Sieber et al., 2019). Briefly, we consider a community saturated with a total of N_T individuals. Individuals die at a constant rate, irrespective of the taxa they belong to. The open space is then filled either by dispersal of an individual from a source pool with probability m, or by reproduction from the local community with probability $1 - m$. Since the total number of individuals and the number of taxa in a microbiome is large, it is reasonable to assume that the relative abundance of a taxa i, ($x_i = N_i/N_T$), is continuous. The probability density function of x_i, $\phi_i = \phi_i(x_i, t)$, is then given by the forward Kolmogorov equation (a.k.a Fokker-Plank equation) whose equilibrium solution is approximated by the following beta distribution

$$\phi_i(x_i;, N_T, p_i, m) = c(1 - x_i)^{N_T m(1-p_i)-1} x_i^{N_T m p_i - 1} \quad 1.$$

Here, $c = \Gamma(N_T m)/(\Gamma\big(N_T m(1-p_i)\big)\Gamma(N_T m p_i))m$, and $p_i$represents the relative abundance of taxa i in the source pool community. We can empirically estimate p_i by combining the sequence data from many microbiome samples. This provides a connection between equation 1 and the empirical observations (16S rRNA and ITS sequences). If d is a relative threshold abundance below which a taxon will not be detected, the probability of observing an i^{th} taxon in the local community is given as,

$$P(taxon\ i\ is\ observed\ in\ the\ sample) = \int_d^1 \phi_i(x_i;, N_T, p_i, m)dx \quad 2.$$

Here, N_T is the number of sequence reads per sample and d is set to $1/N_T$. The immigration parameter (m) is the only free parameter which can be estimated by fitting the equation 2 to a given data set.

Fit to the neutral model

Amplicon sequence variant (ASV) table and OTU table for both bacterial and fungal samples were rarefied to an even depth. Equation 2 was fitted to the sequence table using *minpack.lm* package in R software (https://rdocumentation.org/packages/minpack.lm/) (Elzhov et al., 2016). We used the coefficient of determination (equation 3) as a goodness of fit measure test of the neutral model prediction.

$$R^2 = 1 - \frac{\sum_i (f_i - \Phi_i)^2}{\sum_i (f_i - \overline{p_i})^2} \quad 3.$$

Here, f_i is occurrence frequency observed in the rarefied OTU table, $\overline{p_i}$ is mean relative abundance, and Φ_i is expected occurrence frequency obtained from best-fit neutral model. 95% bootstrapped confidence intervals were calculated by resampling the

samples 100 times with replacement and following the fitting procedure described above. Statistical tests were performed between the confidence interval using non-parametric test (Wilcoxon) in R software using *ggpubr* package (Kassambara, 2020).

Null models

To complement the neutral model and to detect any potential influence of niche-based processes, we performed null model analyses. First, we calculated the mean nearest taxon distance (MNTD) which is the weighted mean of phylogenetic distance to a nearest taxon for each taxon within a sample (Webb et al., 2002; Stegen et al., 2012).

$$MNTD = \sum_{i_k=1}^{n_k} f_{i_k} \min \Delta_{i_k j_k} \quad 4,$$

where, f_{i_k} is the relative abundance of taxon i in community k, $\min(\Delta_{i_k j_k})$ is the minimum phylogenetic distance between a taxon i and all other taxa j in the sample, n_k is the total number of taxa (ASVs or OTUs) in the sample. The observed MNTD is then compared to the null distribution obtained by randomization of taxa and their relative abundance across the tips of phylogeny. Nearest taxon index (NTI) is then used to measure the deviation of the observed MNTD from the mean of the null distribution. If the $NTI > +2$ for a sample, the co-occuring taxa in the sample are more closely related than expected by chance and hence represents clustering. Likewise, $NTI < -2$ represents overdispersion indicating co-occuring taxa are more distantly related than expected by chance. Means of NTIs were calculated for samples based on location, land-use, and depth.

The phylogenetic differences between communities were estimated by the beta mean nearest taxon distance ($\beta MNTD$) metric (Kembel, 2009; Fine and Kembel, 2011; Stegen et al., 2012; Dini-Andreote et al., 2015). $\beta MNTD$ quantifies the phylogenetic distance between a taxon in one community (k) with its closest relative in another community (m):

$$\beta MNTD = \frac{1}{2}\left(\sum_{i_k=1}^{n_k} f_{i_k} \min\left(\Delta_{i_k j_m}\right) + \sum_{i_m=1}^{n_m} f_{i_m} \min\left(\Delta_{i_m j_k}\right)\right) \quad 5.$$

Here, n_k and n_m are the number of taxa in community k and m respectively. f_{i_k} and f_{i_m} represent the relative abundance of taxon i in community k and m respectively. $\min\left(\Delta_{i_k j_m}\right)$ is the minimum phylogenetic distance between a taxon i in community k and all other taxa in community m. β-nearest taxon index (βNTI) is then used to compute the deviation of the observed $\beta MNTD$ from expected null $\beta MNTD$ distribution obtained by randomization. Values with $|\beta NTI| > 2$ are statistically significant and is interpreted as niche-based processes being dominant in shaping community structure. If $\beta NTI > 2$, the two communities are phylogenetically more different than expected by chance, whereas, if $\beta NTI < -2$, the communities are more similar than expected by chance. Additionally, if the mean βNTI of all pairwise comparisons is significantly different from 0, the phylogenetic differences between communities is interpreted to be influenced by niche-based processes.

In addition to the phylogeny-based null models, we also used a composition-based null model to distinguish between potential homogenizing and diversifying forces that relate to community similarities and differences. For this, we used the Raup-Crick metric

(RC_{bray}) (Chase et al., 2011) modified to incorporate taxa's relative abundance (Stegen et al., 2013). Unlike $\beta MNTD$, RC_{bray} does not consider the phylogenetic relationship between taxa and is solely based on presence-absence data. RC_{bray} metric ranges from -1 to 1. If the $RC_{bray} > 0.95$, the two communities under comparisons are significantly different from null expectations. We interpret this scenario as the diversifying forces such as variable selection being dominant. Similarly, if $RC_{bray} < -0.95$, are more compositionally more similar than expected by chance. In this case, we consider homogenizing forces such as homogenizing selection and/or high dispersal between communities to be dominant. The values in between are assumed to be due to stochastic processes.

Results

Neutral model fit to bacterial and fungal samples

We fit the neutral model to both bacterial and fungal samples at two different scales – one at the regional scale combining all samples collected across Kansas and another at the scale of individual sites. At the regional scale, the neutral model showed a better fit to the bacterial samples ($R^2 = 0.80$, Figure 4.2a) compared to the fungal samples ($R^2 = 0.46$, Figure 4.3a). Similar trend was observed for spring samples (Figure S4.1 and S4.2). Neutral model fits to Site 3 samples show better fit compared to Site 1 and Site 2 samples for both bacterial (Figure 4.2, Figure S4.1) and fungal samples (Figure 4.3, Figure S4.2). Overall bacterial and fungal samples also have better fit compared to those at Site 1 and Site 2 samples (Figures 4.2, 4.3, S4.1, S4.2). We also tested whether the choice between OTUs and ASVs had any effect on these analyses. Goodness of fit for both bacterial and fungal samples using OTUs was better than with ASVs but the site-wise patterns largely

hold regardless of the choice of OTUs or ASVs (Figure 4.2, Figure 4.3, Figure S4.1 and Figure S4.2 vs Figure S4.3, S4.4, S4.5, and S4.6).

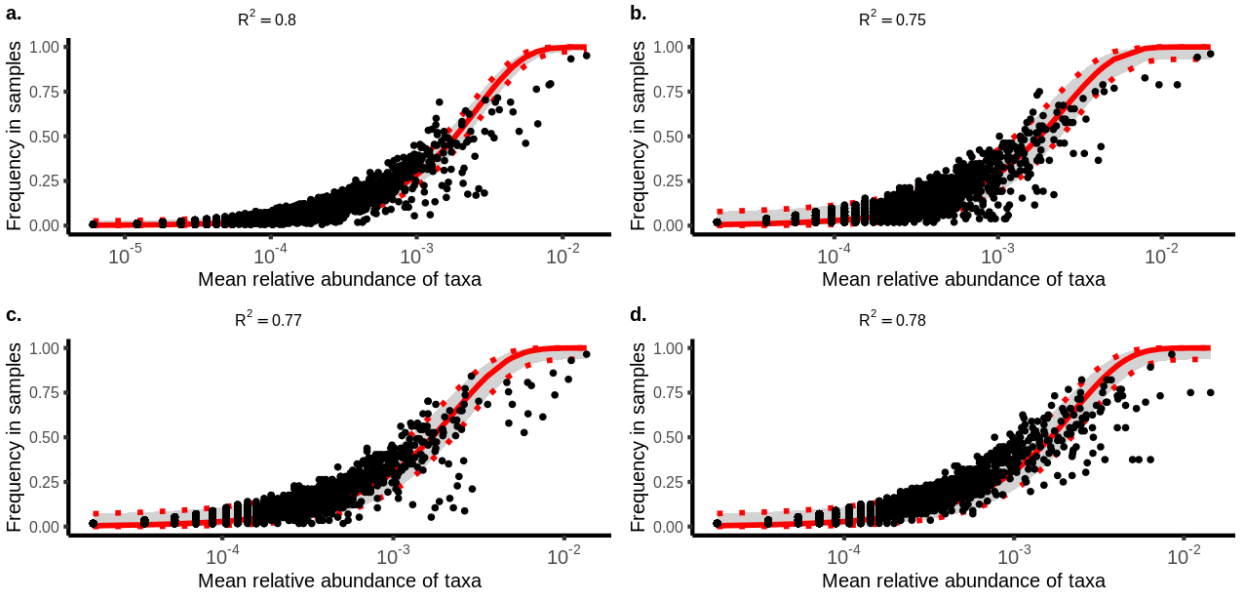

Figure 4.2. Neutral model fit for bacterial (ASVs) samples at different scales. a. shows fit for all samples collected in fall 2018. b. shows fit for the samples from site 1, c. shows fit for the samples from site 2, and d. shows fit for the samples from site 3. Red line represents prediction from the neutral model, gray shading within red dashed lines represents 95% confidence interval, and black dot represents the observed taxa. Site 1 is in eastern Kansas, site 2 is in Konza Prairie, and site 3 is near Fort Hays.

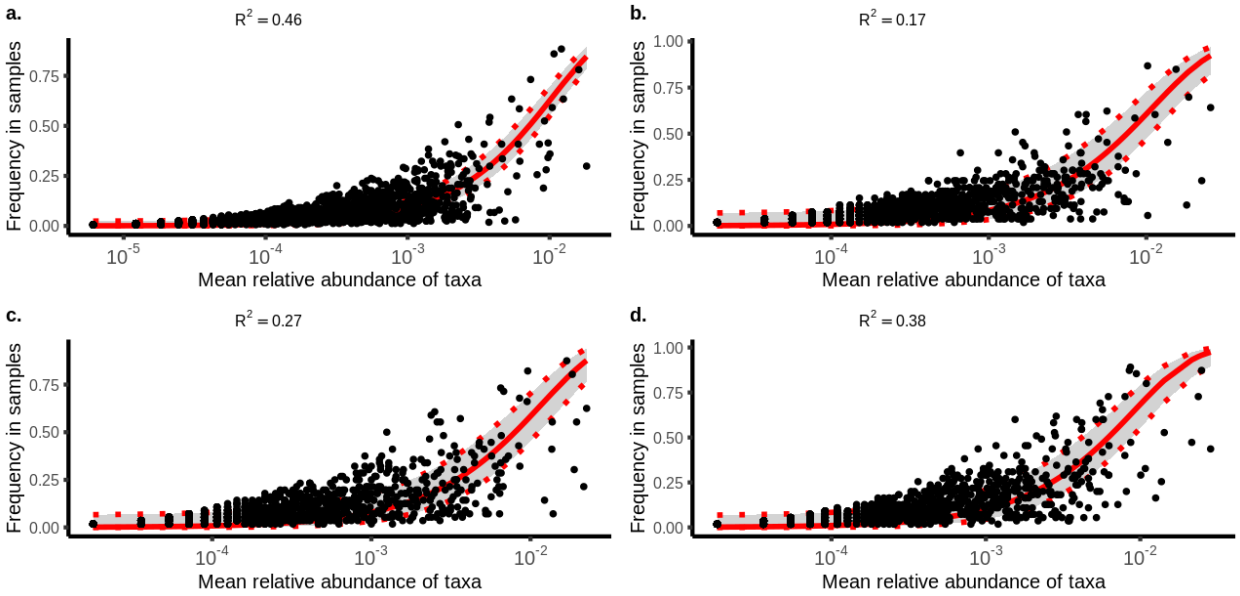

Figure 4.3. Neutral model fit for fungal samples (ASVs) at different scales. a. shows fit for all samples collected in fall 2018. b. shows fit for the samples from site 1, c. shows fit for the samples from site 2, and d. shows fit for the samples from site 3. Red line represents prediction from the neutral model, gray shading within red dashed lines represents 95% confidence interval, and black dot represents the observed taxa.

Neutral model shows best fit for bacterial samples obtained from intermediate depths across all sites.

To understand the role of stochastic processes in shaping the communities at different depths, we categorized both bacterial and fungal samples based on the depth they were sampled from and fit with the neutral model. For bacterial communities, the neutral model fit was best at intermediate depth (30 cm) compared to other depths. The pattern is repeated for all three sites as shown by the bootstrapped R^2 values (Figures 4.3 and S4.7). For fungal samples, however, samples from >75 cm of depth showed the best fit compared to samples from shallower depths based on bootstrapped R^2 values (Figure S4.8).

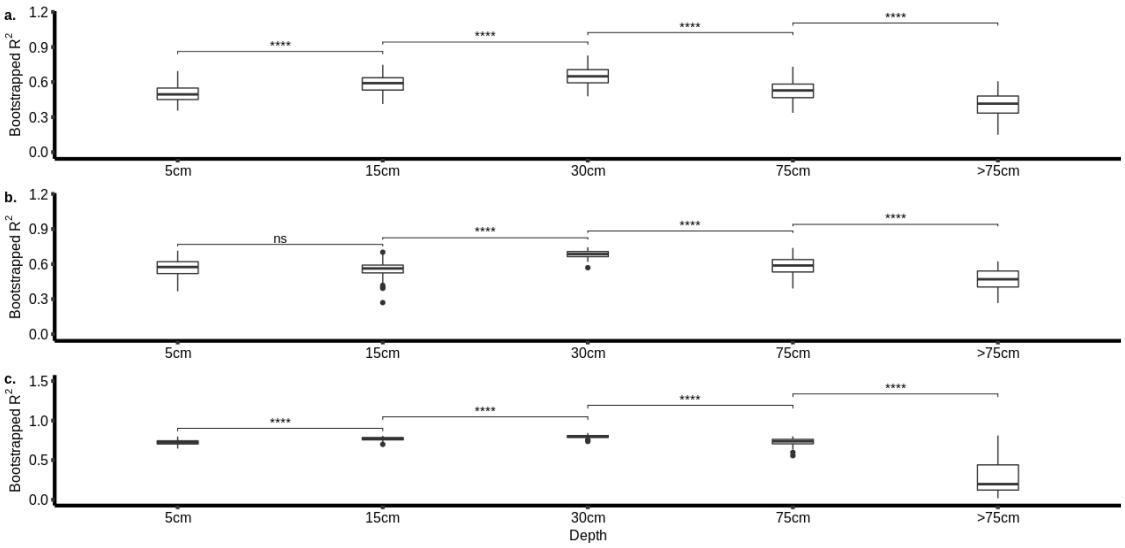

Figure 4.4. Bootstrapped goodness of fit of neutral model for bacterial samples (OTUs) obtained from different depths a. shows fit for the samples from site 1, b. shows fit for the samples from site 2, and c. shows fit for the samples from site 3.

The relative contributions of deterministic and stochastic processes vary with depth and land use

We estimated the phylogenetic differences among communities using βNTI values. This metric allows assessment of the roles of deterministic and stochastic processes in shaping community structure. This was done between samples at different depths and land-use history within a site. Comparisons of bacterial samples between different depths show that the communities from nearby depths are more similar to each other than are communities from distant depths across all sites (Figures 4.5a, 4.5c, 4.5e). This is indicated by more negative βNTI values for nearby depths and the gradual increase of βNTI values with increasing distance between depth comparisons in Figure 4.5. This pattern, however, was not observed for fungal communities except at site 3 (Figures 4.6a, 4.6c, 6e). In addition, we also compared the samples from within the same land-use and between different land-use histories. Same site and land-use comparisons

for both bacterial and fungal samples show that the taxa in a given community are more closely related to each other than by random chance (i.e., phylogenetic clustering). This is shown by positive NTI values in the figures for all depths (Figures S4.9-S4.14). For between-land use comparisons, the results show that bacterial communities obtained from the native land are more similar to the communities obtained from the restored sites (shown as Post-Ag in Figures 4.5 and 4.6) compared to those obtained from agricultural land in all sites. The pattern is similar for fungal communities except in site 2 (Figure 4.5d).

In addition to the phylogeny-based analysis, we calculated RC_{bray} based on compositional data using both ASVs and OTUs for fungal and bacterial samples. Bacterial samples based on depth show that samples on nearby depths are mostly explained by stochastic processes whereas sample comparisons from varying depths are mostly explained by diversifying forces (i.e. communities are more dissimilar than expected by chance, Figures S4.19-S4.21). With regards to samples based on land use management, community comparisons between samples from native and restored lands are mostly explained by stochastic processes whereas other two comparisons (native and agricultural plus agricultural and restored lands) are mostly explained by diversifying forces (e.g., variable selection). On the other hand, only a few comparisons of fungal communities lie within stochastic ranges whereas most comparisons are explained by diversifying forces (Figures S4.22-S4.25). This is consistent with neutral model analysis where we see low fit to neutral model from fungal samples.

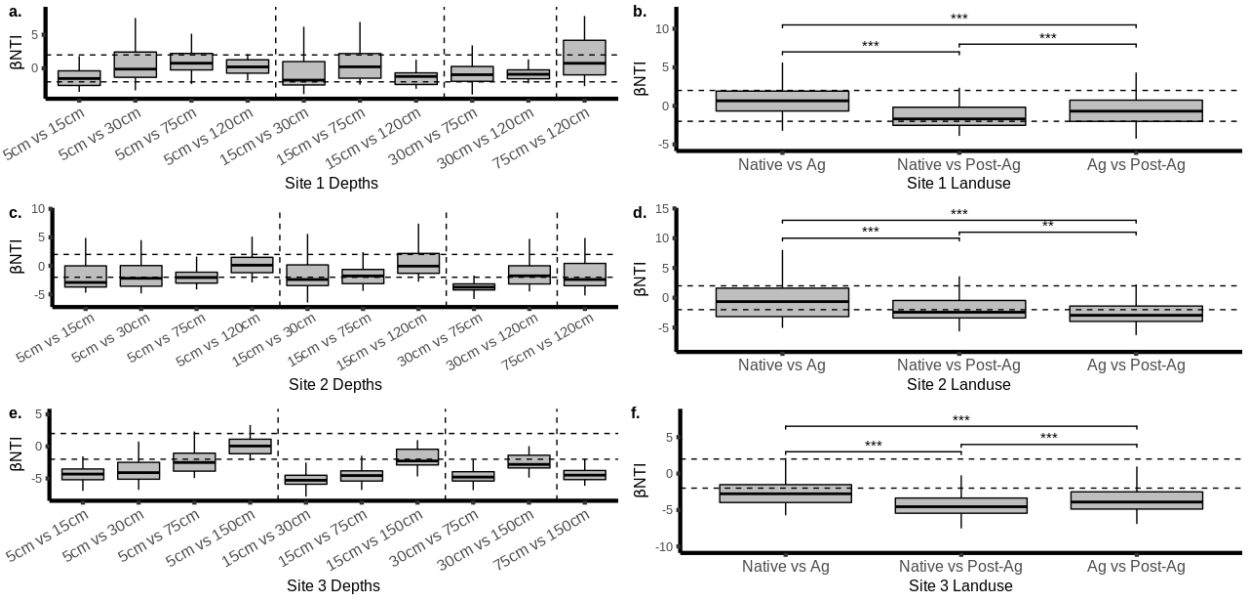

Figure 4.5. Community differences between bacterial samples (ASVs) obtained from different depths and land use history. All samples are from 2018 fall. In figures b., d., and f., Ag represents samples from agricultural land and Post-Ag represents sample from post-agricultural land. Horizontal dashed lines show upper and lower significance threshold values at $\beta NTI = +2$ and $\beta NTI = -2$, respectively.

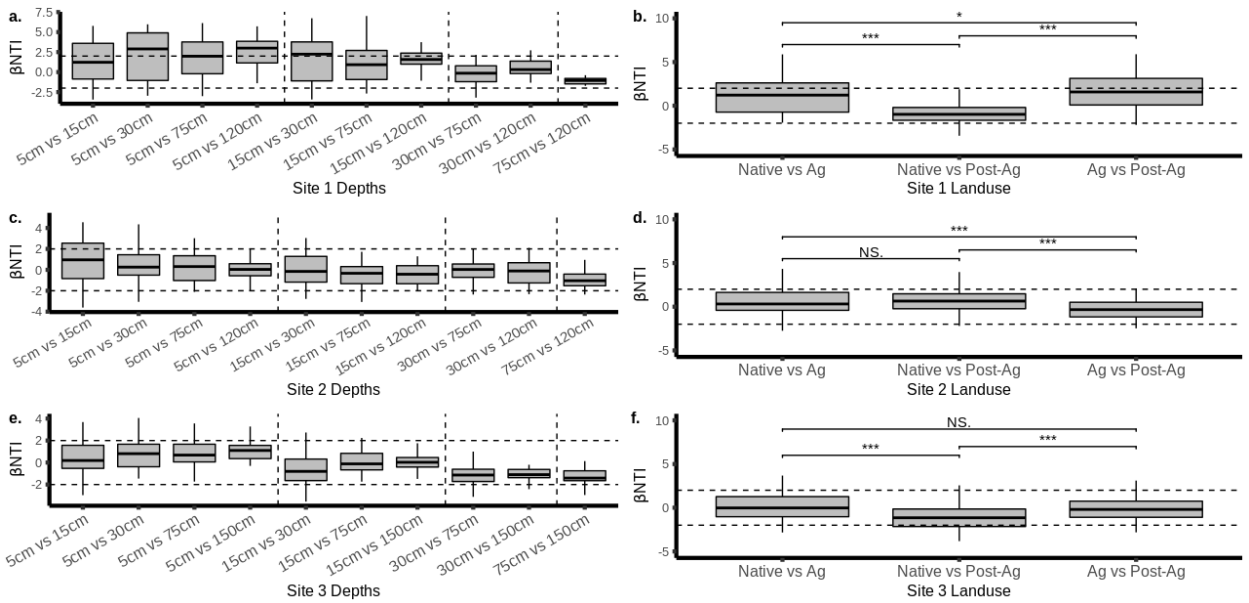

Figure 4.6. Community differences between fungal samples (ASVs) obtained from different depths and land use history. All samples are from 2018 fall. In figures b., d., and f., Ag represents

samples from agricultural land and Post-Ag represents sample from post-agricultural land. Horizontal dashed lines show upper and lower significance threshold values at $\beta NTI = +2$ and $\beta NTI = -2$, respectively.

The degree of bacterial community differences positively correlates with pH differences, but the strength of the relationship is season dependent.

Differences in communities was correlated with different environmental variables measured for each site. Community differences based on differences in soil moisture content show either weak positive, weak negative, or no significant correlations based on sites and season (Figures S4.29-S4.31). The pattern is similar for both fungal and bacterial samples. We also looked at the correlation of community differences on enzymatic activity differences of BGAse, NAGase, and phosphatase enzymes. For each site and enzyme, bacterial βNTI was positively correlated with differences in the log activity of the enzyme for samples obtained in the spring season (Figures S4.32-S4.34). For fungal samples, only the site 3 showed a positive correlation (Figures S4.37). All other samples either showed weak positive, weak negative, or no correlation (Figures S4.35-S4.37). pH was the only environmental factor that showed a consistently significant correlation for bacterial samples (Figure 4.7). The correlation was strongest for samples from Site 2 ($R^2 = 0.51, p < 0.01$) and Site 3 ($R^2 = 0.43, p < 0.01$) collected in the spring (Figures 4.7c and 4.7e). The strength of the correlation decreased for samples from fall season (Figures 4.7b, 4.7d, and 4.7f). Fungal samples did not show consistent correlation patterns based on pH differences between the samples. The choice of OTUs vs ASVs shows some difference in the corresponding strength of correlation (e.g., compare Figure 4.7 and Figure S4.26).

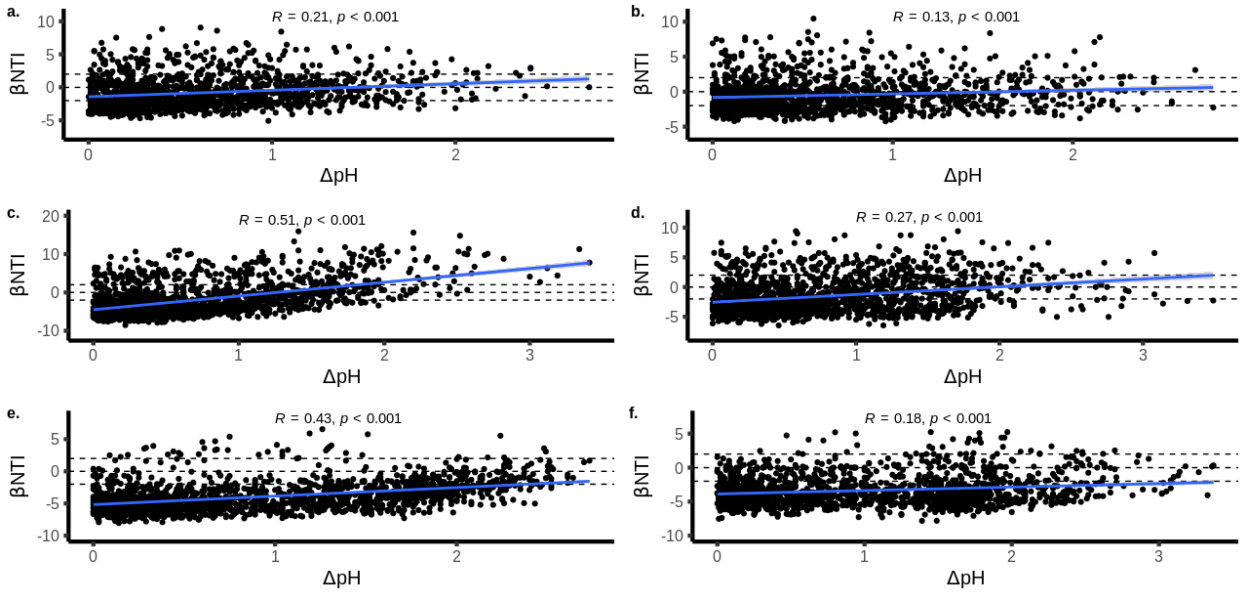

Figure 4.7. βNTI values plotted against the difference in pH values between bacterial for 3 sites across two seasons. Left panel represents samples from spring 2018 whereas right panel shows samples from fall 2018. Each dot represents βNTI values between samples based on phylogenetic tree constructed for ASVs. Top panel shows sample from Site 1, middle panel from Site 2, and bottom panel from Site 3. Blue line shows the best fit regression lines and their associated R^2 and p-values are reported within each figure.

Quantitative estimates of the assembly processes.

We quantified the proportions of community differences explained by different assembly processes for both bacterial and fungal samples (Figure 4.8) using RC_{bray} values. For most sample pairings, the differences in sample communities were mostly explained by a combination of stochastic and diversifying forces (e.g., variable selection) for both bacterial and fungal samples. Based on ASVs of bacterial samples, variable selection was more dominant in fall season compared to spring (54.7% fall vs 45.4% spring for site 1, 59.1% fall vs 53.2% spring for site 2, and 53.1% fall vs 50.9% spring for site 3). This pattern was similar for fungal samples except at site 1 (63.8% fall vs 65.3%

spring for Site 1, 71.8% fall vs 61.6% spring for Site 2, 70.8% fall vs 67.2% spring for site 3). The sample pairings that are mostly explained by variable selection are mostly derived from different depths (Figure S4.19-S4.25). No significant homogeneous selection was observed for spring fungal samples (5% or less). On the other hand, about 25% of bacterial community differences in fall of site 3 can be explained by homogeneous selection.

To understand whether the choice of ASVs vs OTUs made any differences, we also quantified RC_{bray} values based on OTUs table for both bacterial and fungal samples. Like the ASVs results, community differences in most sample pairings are explained by a combination of variable and stochastic forces (Figure S4.38). However, the seasonal trend is reversed for both bacterial and fungal samples. For example, variable selection accounted for less proportion of sample pairings in fall compared to spring for bacterial samples (40% fall vs 49.5% spring for site 1, 50.5% fall vs 55.1% spring for site 2, and 41.4% fall vs 48.1% spring for site 3). The trend is similar for fungal sample except at site 1 (61.9 % fall vs 70% spring for site 1, 73.3% fall vs 65.2% spring for site 2, and 75.9% fall vs 62.2% spring for site 3).

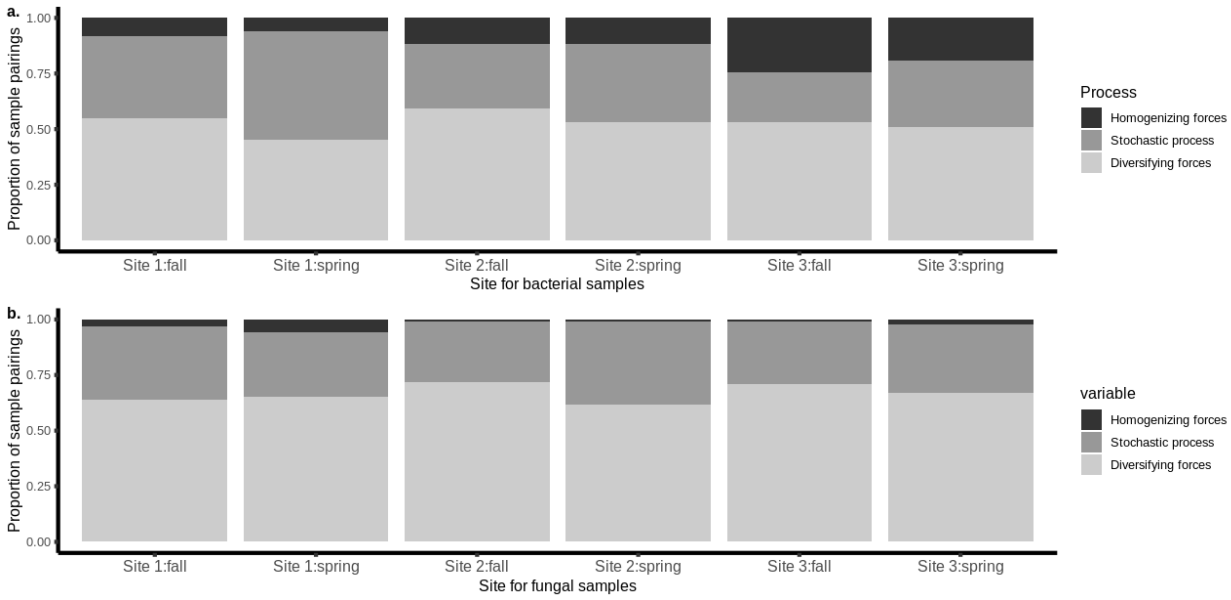

Figure 4.8. Proportion of community turnovers explained by homogeneous selection, variable selection, and stochastic processes for bacterial (a) and fungal (b) samples.

Discussion

The relative contribution of stochastic and deterministic processes to community assembly of microbiomes varies across environments (Stegen et al., 2015). In this study, we used neutral and null models to examine the contribution of both stochastic and deterministic processes during the community assembly of microbiomes obtained from agricultural and prairie soil samples collected across Kansas. Using the neutral community model, we find that the stochastic processes are prevalent in their influence on soil communities of all land uses and sites and that bacterial communities more strongly show evidence of neutral/ stochastic assembly processes than do fungal communities. A null model based on compositional data (RC_{bray}) supports this inference where most of the differences between pairwise community comparisons are explained by a combination of stochastic and diversifying forces.

Of several possible factors we tested, soil depth profiles show distinct effects on bacterial communities. Results show that the samples from intermediate depths (30cm) have a better fit to the model compared to the samples from shallow and deeper depths. Shallow depth is temporally more dynamic as it is influenced by plant communities (Emilia Hannula et al., 2019; Hao et al., 2021). Likewise, abiotic climatic factors are more pronounced at shallow depth influencing microbial communities (Dove et al., 2021). Thus, deterministic processes due to biotic and abiotic factors may have contributed to the lower model fit. Biotic and abiotic factors change rapidly across soil depth profiles that has consequences in microbial diversity (Mundra et al., 2021). Both fungal and bacterial diversity decreases with soil depth (Jiao et al., 2018). The deeper soil profile may provide consistent selection pressures thereby limiting the type of taxa that are present. It is also likely that the rate of dispersal is low in deeper depths. These factors can jointly explain the low fit of the neutral model at deeper depths. A null model based on phylogeny shows that greater the differences in depth profile of the samples is associated with less similarity in the phylogenetic structure of bacterial communities. The composition based null model also shows that samples from nearby depth intervals had communities that were more similar to one another than samples from intervals with greater separation. Together, our results show that soil depth profile strongly influences assembly processes in bacterial communities.

Fungal samples do not follow the trends observed in bacterial samples. The overall lower fit to the neutral model may be attributed to low dispersal rate in fungal communities. Consistent with previous findings, we observed about 10X low dispersal rate (not shown) in fungal samples compared to the bacterial samples(Chen et al., 2020). In addition, plant

communities can strongly influence fungal community composition and diversity (Broeckling et al., 2008; Sweeney et al., 2021). These factors can, in combination, result in a strong spatial pattern to fungal communities. RC_{bray} calculations between all possible pairwise comparisons within a site do show that diversifying forces can explain most of the compositional differences between samples. Comparisons of phylogenetic structure between samples of different depths do not show a consistent trend as observed for bacterial samples. These results suggest that the community assembly processes of soil fungi and bacterial communities can differ (Schmidt et al., 2014).

To understand how potential environmental filters affect community assembly processes, we correlated pH, moisture content, and enzymatic activity of sample sites with βNTI. In general, greater the differences in the magnitude of the filters between samples, corresponded to greater phylogenetic differences between the communities. Previous research has shown moisture content to be important in shaping bacterial communities (Lupatini et al., 2019). However, we did not see any consistent trend with moisture content based on the pairwise βNTI values between samples within a site. For some sites, we instead see a weak negative correlation. These associations may reflect statistical artifacts reflecting the low sample numbers at greater differences in moisture values.

Another potentially important factor is pH, and it has been shown to be a major driver of bacterial community assembly (Lauber et al., 2009; Tripathi et al., 2018). Highly acidic and basic pH can impose selection pressure and act as environmental filter (Tripathi et al., 2018; Jiao and Lu, 2020). As expected, βNTI values based on ASVs show consistent positive correlation for bacterial samples (Figure 4.7). However, the strength

of correlation is season dependent - spring season showing stronger correlation compared to fall season. Fungal samples show weak positive to moderate correlation depending on site and season (Figure S4.27). Some previous research has shown pH to be influential in determining fungal community at fine scales (Glassman et al., 2017) whereas, in some systems only alpha diversity is shown to be correlated with pH (Wang et al., 2015). Thus, though bacterial communities are influenced by pH at broad scales, the impact on fungi seems to be highly context specific. Other environmental factors (e.g., plant types, land-use, soil, and root chemistry) could be driving this context dependency.

Microbes actively modify the environment they live on. This can influence colonization process of late arriving microbes thereby shaping community composition in a deterministic way through priority effects (Cline and Zak, 2015). One mechanism of habitat modification is through secretion of exoenzymes (Cline and Zak, 2015; Pierre-Emmanuel et al., 2016). Hence, we measured whether the difference in enzymatic activities correlated with differences in phylogenetic community structures (βNTI). We used the overall activity of N-acetylaminoglucosidase, β-glucosidase, and phosphatase (Figures S4.32-S4.37). Results show that the strength of correlation is site and season dependent. Overall, the correlation is stronger in spring compared to the fall season, possibly because organisms are more active during the spring season. Likewise, the strength of correlation is stronger at the site 3 for both fungal and bacterial samples. Priority effects can also act through practices of land management through modification of soil characteristics. Hence, we assessed whether different assembly processes are dominant in samples derived from different land-use history (Figure 4.5 and Figure 4.6, right panels). For both fungal and bacterial communities, comparison between samples

from native and restored land show that stochastic processes are more dominant in these samples whereas diversifying forces are dominant in samples derived from native vs agricultural and agricultural vs restored lands comparisons. This is further supported by RC_{bray} calculations (Figures S4.19-S4.25). Thus, deterministic processes that actively modify soil characteristics can shape both bacterial and fungal community assembly processes.

There has been a recent push in the field of microbial ecology to move away from the use of OTUs, and ASVs are advocated instead (Callahan et al., 2017). However, most of the previous studies that use neutral and null models use OTUs instead of ASVs (Stegen et al., 2012; Sieber et al., 2019). Hence, to test whether the use of ASVs and OTUs has any effect on the analysis outcomes, we performed analysis using both taxonomic units. For neutral models, we see a better fit on both bacterial and fungal samples based on the OTUs rather than the ASVs for the same rarefication depths. But the pattern between sites is largely consistent regardless of the choice. NTI calculations do not show much difference between OTUs and ASVs. However, the inferences of processes based on both βNTI and RC_{bray} change to some extent depending on the choice of taxonomic unit. For example, the seasonal pattern of stochastic process is reversed in OTUs and ASVs (compare Figure 4.8 and S4.38). We also see a change in the strength of correlation with pH, moisture, and enzymatic activity based on whether we use OTUs or ASVs. These results suggest that the choice of a taxonomic unit can change the inference of processes that are dominant during the community assembly.

In this paper, we explored the potential roles of stochastic and deterministic processes in community assembly of bacterial and fungal communities. However, these

inferences are critically dependent on the underlying model assumptions. Unless the validity of the assumptions is independently examined, we need to be careful with the inferences we make from these models (Dayton, 1973). In addition, the multicausality nature of ecological patterns warrant that we do not ignore alternate explanations just because a pattern matches model predictions (Quinn and Dunham, 1983). In our results, the bacterial communities show strong fit to neutral models. On the other hand, null models show homogenizing and diversifying forces to be dominant in a significant portion of pairwise community comparisons. Even though these models seem to make different predictions, a good proportion of sample comparisons done using null models can be explained by stochastic processes. If these different models, despite their widely different underlying assumptions, converge to similar predictions, we can make robust inferences of the processes (Levins, 1966). Thus, we can confidently infer that the stochastic processes largely explain the patterns observed in bacterial communities.

Another potential issue lies with the choice of models itself. There are several available null models, but it is not always clear which of these models are more ‘realistic’. For example, we can perform randomization of phylogeny based null models in different ways instead of simply randomizing the tip labels in phylogeny (Kembel, 2009), and each of these may lead to different predictions. Hence the choice of a particular model also has the potential to change result interpretations. In addition, several recent studies have combined phylogeny and composition based null models to tease apart different assembly mechanisms such as drift, dispersal limitation, homogenizing dispersal, homogeneous and variable selection (Stegen et al., 2012, 2015; Tripathi et al., 2018). In this framework, phylogeny based null model (βNTI) is first used to break processes into

3 different types – homogeneous selection, variable section, and stochastic processes. Composition based null model (RC_{bray}) is then used to further breakdown stochastic processes into dispersal limitation, homogenizing dispersal, and drift. Though we have used both phylogeny and composition based null models, we do not exactly follow this framework for several reasons. First, this framework relies on the assumption that phylogenetic structure reflects microbial function and processes. Though a strong phylogenetic signal can support the claim that phylogenetic structure is correlated with function, it does not invalidate potential alternative hypothesis that composition better represents function of a microbiome. Inference based on phylogenetic structure alone may not be sufficient to tease apart assembly processes as many of the assumptions within this framework are only weakly supported (Gerhold et al., 2015). In addition, βNTI and RC_{bray} calculations are based on different assumptions and are complimentary but not interdependent. Thus, we use both phylogeny and composition based null models as complimentary approaches to infer similar processes. Although by doing this, the number of processes that we can infer is limited (i.e., homogenizing forces, diversifying forces, and stochastic process), we can make more robust inferences of processes compared to the generally used framework.

References

Apprill, A., Mcnally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* 75, 129–137.

Arif, I., Batool, M., and Schenk, P. M. (2020). Plant microbiome engineering: expected benefits for improved crop growth and resilience. *Trends in Biotechnology* 38, 1385–1396.

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., et al. (2018). Structure and function of the global topsoil microbiome. *Nature 2018 560:7717* 560, 233–237.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology 2019 37:8* 37, 852–857.

Broeckling, C. D., Broz, A. K., Bergelson, J., Manter, D. K., and Vivanco, J. M. (2008). Root exudates regulate soil fungal community composition and diversity. *Applied and Environmental Microbiology* 74, 738–744.

Cadotte, M. W., and Tucker, C. M. (2017). Should environmental filtering be abandoned? *Trends in Ecology & Evolution* 32, 429–437.

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

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

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., et al. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology 2019 17:9* 17, 569–586.

Chase, J. M. (2007). Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17430–17434.

Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M., and Inouye, B. D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* 2, art24.

Chase, J. M., and Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, 2351–2363.

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters* 7, 241–253.

Chen, J., Wang, P., Wang, C., Wang, X., Miao, L., Liu, S., et al. (2020). Fungal community demonstrates stronger dispersal limitation and less network connectivity than bacterial community in sediments along a large river. *Environmental Microbiology* 22, 832–849.

Cline, L. C., and Zak, D. R. (2015). Initial colonization, community assembly and ecosystem function: fungal colonist traits and litter biochemistry mediate decay rate. *Molecular Ecology* 24, 5045–5058.

Dayton, P. K. (1973). Two cases of resource partitioning in an intertidal community: Making the right prediction for the wrong reason. *The American Naturalist* 107, 662–670.

D'Haens, G. R., and Jobin, C. (2019). Fecal microbial transplantation for diseases beyond recurrent *Clostridium difficile* infection. *Gastroenterology* 157, 624–636.

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America* 112, E1326–E1332.

Dodds, W. K., Zeglin, L. H., Ramos, R. J., Platt, T. G., Pandey, A., Michaels, T., et al. (2020). Connections and feedback: aquatic, plant, and soil microbiomes in heterogeneous and changing environments. *BioScience* 70, 548–562.

Dong, M., Kowalchuk, G. A., Liu, H., Xiong, W., Deng, X., Zhang, N., et al. (2021). Microbial community assembly in soil aggregates: A dynamic interplay of stochastic and deterministic processes. *Applied Soil Ecology* 163, 103911.

Dove, N. C., Barnes, M. E., Moreland, K., Graham, R. C., Berhe, A. A., and Hart, S. C. (2021). Depth dependence of climatic controls on soil microbial community activity and composition. *ISME Communications 2021 1:1* 1, 1–11.

Duncker, K. E., Holmes, Z. A., and You, L. (2021). Engineered microbial consortia: strategies and applications. *Microbial Cell Factories 2021 20:1* 20, 1–13.

Elzhov, T. v, Mullen, K. M., Spiess, A.-N., and Bolker, B. (2016). minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds version 1.2-1 from CRAN.

Hannula, S. E., Kielak, A. M., Steinauer, K., Huberty, M., Jongen, R., de Long, J. R., et al. (2019). Time after time: temporal variation in the effects of grass and forb species on soil bacterial and fungal communities. *mBio* 10, e02635-19.

Estrela, S., Sánchez, Á., and Rebolleda-Gómez, M. (2021). Multi-replicated enrichment communities as a model system in microbial ecology. *Frontiers in Microbiology* 12, 760.

Fine, P. V. A., and Kembel, S. W. (2011). Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* 34, 552–565.

Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. v., and Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology* 29, 600–614.

Glassman, S. I., Wang, I. J., and Bruns, T. D. (2017). Environmental filtering by pH and soil nutrients drives community assembly in fungi at fine spatial scales. *Molecular Ecology* 26, 6960–6973.

Hao, J., Chai, Y. N., Lopes, L. D., Ordóñez, R. A., Wright, E. E., Archontoulis, S., et al. (2021). The effects of soil depth on the structure of microbial communities in agricultural soils in Iowa (United States). *Applied and Environmental Microbiology* 87, e02673-20.

Hubbell, S. P. (2006). Neutral theory and the evolution of ecological equivalence. *Ecology* 87, 1387–98.

Ihrmark, K., Bödeker, I. T. M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al. (2012). New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology* 82, 666–677.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q., and Wei, G. (2018). Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome* 6, 1–13.

Jiao, S., and Lu, Y. (2020). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environmental Microbiology* 22, 1052–1065.

Jiao, S., Yang, Y., Xu, Y., Zhang, J., and Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME Journal* 14, 202–216.

Kang, S., and Mills, A. L. (2006). The effect of sample size in studies of soil microbial community structure. *Journal of Microbiological Methods* 66, 242–250.

Kassambara, A. (2020). “ggplot2” Based Publication Ready Plots [R package ggpubr version 0.4.0]. Available at: https://CRAN.R-project.org/package=ggpubr [Accessed August 16, 2021].

Kembel, S. W. (2009). Disentangling niche and neutral influences on community assembly: Assessing the performance of community phylogenetic structure tests. *Ecology Letters* 12, 949–960.

Köberl, M., Wagner, P., Müller, H., Matzer, R., Unterfrauner, H., Cernava, T., et al. (2020). Unraveling the complexity of soil microbiomes in a large-scale study subjected to different agricultural management in Styria. *Frontiers in Microbiology* 11, 1052.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., and Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* 29, 592–599.

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology* 75, 5111–5120.

Levins, R. (1966). The strategy of model building in population biology. *American Scientist* 54, 421–431.

Lupatini, M., Suleiman, A. K. A., Jacques, R. J. S., Lemos, L. N., Pylro, V. S., van Veen, J. A., et al. (2019). Moisture is more important than temperature for assembly of both potentially active and whole prokaryotic communities in subtropical grassland. *Microbial Ecology* 77, 460–470.

McKenney, E. A., Koelle, K., Dunn, R. R., and Yoder, A. D. (2018). The ecosystem services of animal microbiomes. *Molecular Ecology* 27, 2164–2172.

Mundra, S., Kjønaas, O. J., Morgado, L. N., Krabberød, A. K., Ransedokken, Y., and Kauserud, H. (2021). Soil depth matters: shift in composition and inter-kingdom co-occurrence patterns of microorganisms in forest soils. *FEMS Microbiology Ecology* 97, fiab022.

Naylor, D., Sadler, N., Bhattacharjee, A., Graham, E. B., Anderton, C. R., McClure, R., et al. (2020). Soil microbiomes under climate change and implications for carbon cycling. *Annual Review of Environment and Resources* 45, 29–59.

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews* 77, 342–356.

Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research* 47, D259–D264.

Ofiţeru, I. D., Lunn, M., Curtis, T. P., Wells, G. F., Criddle, C. S., Francis, C. A., et al. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences of the United States of America* 107, 15345–15350.

Orrock, J. L., and Watling, J. I. (2010). Local community size mediates ecological drift and competition in metacommunities. *Proceedings of the Royal Society B: Biological Sciences* 277, 2185–2191.

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* 18, 1403–1414.

Pierre-Emmanuel, C., François, M., Marc-André, S., Myriam, D., Stéven, C., Fabio, Z., et al. (2016). Into the functional ecology of ectomycorrhizal communities: environmental filtering of enzymatic activities. *Journal of Ecology* 104, 1585–1598.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41, D590.

Quinn, J. F., and Dunham, A. E. (1983). On hypothesis testing in ecology and evolution. *The American Naturalist* 122, 602–617.

R Development Core Team R A Language and Environment for Statistical Computing. Vienna, Austria R Foundation for Statistical Computing. - References - Scientific Research Publishing. Available at: https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=2600003 [Accessed December 9, 2021].

Saleem, M., Hu, J., and Jousset, A. (2019). More than the sum of its parts: microbiome biodiversity as a driver of plant growth and soil health. *Annual Review of Ecology, Evolution, and Systematics* 50, 145–168.

Schmidt, S. K., Nemergut, D. R., Darcy, J. L., and Lynch, R. (2014). Do bacterial and fungal communities assemble differently during primary succession? *Molecular Ecology* 23, 254–258.

Sieber, M., Pita, L., Weiland-Bräuer, N., Dirksen, P., Wang, J., Mortzfeld, B., et al. (2019). Neutrality in the metaorganism. *PLOS Biology* 17, e3000298.

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* 8, 732–740.

Song, Z., Schlatter, D., Kennedy, P., Kinkel, L. L., Kistler, H. C., Nguyen, N., et al. (2015). Effort versus reward: preparing samples for fungal community characterization in high-throughput sequencing surveys of soils. *PLOS ONE* 10, e0127234.

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME Journal* 7, 2069–2079.

Stegen, J. C., Lin, X., Fredrickson, J. K., and Konopka, A. E. (2015). Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in Microbiology* 6.

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME Journal* 6, 1653–1664.

Sweeney, C. J., de Vries, F. T., van Dongen, B. E., and Bardgett, R. D. (2021). Root traits explain rhizosphere fungal community composition among temperate grassland plant species. *New Phytologist* 229, 1492–1507.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551, 457–463.

Tripathi, B. M., Stegen, J. C., Kim, M., Dong, K., Adams, J. M., and Lee, Y. K. (2018). Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *The ISME Journal* 12, 1072–1083.

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

Vellend, M., Srivastava, D. S., Anderson, K. M., Brown, C. D., Jankowski, J. E., Kleynhans, E. J., et al. (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* 123, 1420–1430.

Wang, C., Michalet, R., Liu, Z., Jiang, X., Wang, X., Zhang, G., et al. (2020). Disentangling large- and small-scale abiotic and biotic factors shaping soil microbial communities in an Alpine cushion plant system. *Frontiers in Microbiology* 11, 925.

Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J. C., et al. (2013). Phylogenetic beta diversity in bacterial assemblages across ecosystems: Deterministic versus stochastic processes. *ISME Journal* 7, 1310–1321.

Wang, J. T., Zheng, Y. M., Hu, H. W., Zhang, L. M., Li, J., and He, J. Z. (2015). Soil pH determines the alpha diversity but not beta diversity of soil fungal community along altitude in a typical Tibetan forest ecosystem. *Journal of Soils and Sediments* 15, 1224–1232.

Webb, C. O., Ackerly, D. D., McPeek, M. A., and Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecological System* 33, 475–505.

Woodcock, S., van der Gast, C. J., Bell, T., Lunn, M., Curtis, T. P., Head, I. M., et al. (2007). Neutral assembly of bacterial communities. *FEMS Microbiology Ecology* 62, 171–180.

Yan, Z.-Z., Chen, Q.-L., Li, C.-Y., Thi Nguyen, B.-A., Zhu, Y.-G., He, J.-Z., et al. (2021). Biotic and abiotic factors distinctly drive contrasting biogeographic patterns between phyllosphere and soil resistomes in natural ecosystems. *ISME Communications 2021 1:1* 1, 1–9.

Zhang, K., Shi, Y., Cui, X., Yue, P., Li, K., Liu, X., et al. (2019). Salinity is a key determinant for soil microbial communities in a desert ecosystem. *mSystems* 4, e00225-18.

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews* 81, e00002-17.

Chapter 5 - Conclusion and Implications

"We shall end by establishing a new science. But first let you and me unlock the door and then anybody can go in who likes." -Ronald Ross's letter to McKendrik (1911) cited in Heesterbeek (2002).

This dissertation contributes to understanding of how ecology and complex life history determines virulence in facultative pathogens. For a long time, facultative pathogens were considered as an intermediate in transition from true saprophytes to the fully adapted avirulent obligatory pathogens (Theobald Smith, 1887). The facultative pathogens of old times still exist, so does their virulence. The avirulence hypothesis is now fully substituted by the 'trade-off hypothesis' (Alizon et al., 2009). However, the trade-off hypothesis mostly focuses on the transmission-virulence trade-off, sometimes with little justification (Bull & Lauring, 2014; Ebert & Bull, 2003). Owing to its unique life-history, other trade-offs or trait correlations can influence virulence in facultative pathogens. Using mathematical modeling and experimental evolution, I show how different potential trade-offs and ecology drive virulence evolution in facultative pathogens.

In the first chapter, I reviewed different mechanisms by which heterogeneous selection and trade-offs arise. Many evolutionary analyses treat the trade-off mechanisms as 'black box' to analyze the consequences (Mauro & Ghalambor, 2020), and the field of virulence evolution is no exception. However, if we are to devise more refined and predictive models of virulence evolution, we must consider the molecular and physiological mechanisms by which the trade-offs arise and evolve.

In the second chapter, I provided a theoretical analysis of how heterogeneous selection pressure and multiple trade-offs influence virulence evolution. Historically, the fitness of pathogen is approximated by the basic reproduction number R_0. Since it omits pathogen life history outside of the host, it cannot be a sufficient fitness metric for evolutionary analysis of facultative pathogens. In addition, it turns out that deriving the R_0 is not straightforward either. Thus, I first provided a framework to derive R_0 for pathogens whose lifecycle alter between host and non-host environments. For evolutionary invasion analysis, I used the adaptive dynamics framework. The model is general so that it can be applied to many facultative pathogens. But the application to a specific system will depend on the trade-off shape and life-history details of that system. Therefore, system specific empirical estimations of trait correlations and trade-offs are necessary for meaningful predictions.

In the third chapter, I focused on virulence in *S. maltophilia* and how it can evolve in different environmental conditions. This is an important facultative human pathogen that also infects *C. elegans* in the wild. Using experimental evolution and genetic manipulation, I found out that the type IV secretion system and its effector proteins are involved in virulence against *C. elegans*. Since the same effectors also influence interference competition, I argue that coincidental selection in environment for interference competition can help maintain high virulence. My work shows that selection in environment can influence within-host consequences in facultative pathogens. While I have identified some effector proteins that contribute to virulence, it is still unclear what determines the variation in virulence in different strains. The field will benefit from further

research on other potential virulence mechanisms (e.g., type II secretion system) and effector proteins.

The theme of the fourth chapter is slightly different from the previous three chapters. In it, I look at the community level of microorganisms to figure out how stochastic and niche-based processes shape microbiome communities. With the help of neutral and null models, I show that stochastic processes are dominant in shaping bacterial communities whereas niche-based processes dominate fungal communities in soil. While I focus on understanding the community assembly processes of soil microbiomes, the framework can also be extended to pathogens. Since facultative pathogen such as *S. maltophilia* can reside and potentially reproduce in soil reservoirs, its prevalence can be shaped by both stochastic and deterministic factors. With the help of neutral and null models, we can determine whether the pathogens prevalence fall within the neutral and null expectations. If they fall outside these expectations, it can provide a motivation to look at the factors that shape their prevalence in the environment. These factors can potentially determine epidemic potential of these pathogens. Though I know of no studies that exploit this framework in understanding pathogen dynamics, I believe it can be a good addition to the available tools.

References

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

Bull, J. J., & Lauring, A. S. (2014). Theory and empiricism in virulence evolution. *PLoS Pathogens*, *10*(10), 1–3.

Ebert, D., & Bull, J. J. (2003). Challenging the trade-off model for the evolution of virulence. *Trends in Microbiology*, *11*(1), 15–20.

Heesterbeek, J. (2002). A brief history of R_0 and a recipe for its calculation. *Acta Biotheoretica*, 50(3), 189-204.

Mauro, A. A., & Ghalambor, C. K. (2020). Trade-offs, pleiotropy, and shared molecular pathways: A unified view of constraints on adaptation. *Integrative and Comparative Biology*, *60*(2), 332–347.

Theobald Smith, B. (1887). Parasitic bacteria and their relation to saprophytes. *The American Naturalist*, *XXI*(I), 1–9.