Evolutionary mechanisms that facilitate the maintenance of microbial diversity

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctorate of Philosophy
with a
Major in Bioinformatics and Computational Biology
in the
College of Graduate Studies
University of Idaho
by
Michael T. France

Major Professor: Larry J. Forney, Ph.D.
Committee Members: Eva M. Top, Ph.D.; Steve Krone, Ph.D.; Luke Harmon, Ph.D.
Department Chair: Eva M. Top, Ph.D.

August 2018

Authorization to Submit Dissertation

This dissertation of Michael T. France, submitted for the degree of Doctorate of Philosophy with a Major in Bioinformatics and Computational Biology and titled "Evolutionary mechanisms that facilitate the maintenance of microbial diversity," has been reviewed in final form. Permission, as indicated by the signatures and dates below is now granted to submit final copies for the College of Graduate Studies for approval.

Advisor:	Larry J. Forney, Ph.D.	Date
Committee Members:	Eva M. Top, Ph.D.	Date
	Steve Krone, Ph.D.	Date
	Luke Harmon, Ph.D.	Date
Department Chair:	Eva M. Top, Ph.D.	Date

Abstract

Bacterial populations in the environment have been shown to harbor tremendous within species diversity. This diversity is important because it drives the evolution and ecology of these populations. It is the toolbox from which solutions to novel problems can be drawn. While several mechanisms have been proposed previously, it is still not clear how diversity is created and maintained in environmental bacterial populations. Bacteria typically live and evolve in environments that have physical structure. Their rate of growth is constrained by the abundance of routinely scarce essential nutrients, leading to generation times measured in days or weeks. I hypothesized that prolonged generation times and spatial structure would protract selective sweeps and limit competition to a local scale, allowing for the accumulation of genetic diversity via spontaneous mutations. In Chapter 2 I review our current knowledge of how spatial structure and slow growth rates influence the evolution of bacterial populations. I put my hypothesis to the test in Chapters 3 and 4, wherein I detail the effects of spatial structure and slow growth rates, respectively, on the accumulation and maintenance of diversity in bacterial populations. In Chapter 5 I employ the concepts detailed above in order to investigate the evolution of antibiotic resistance in bacterial biofilms. Biofilms are bacterial assemblages held together by an extracellular matrix. They have spatial structure and contain large subpopulations that experience slow growth rates. Finally, in Chapter 6 I used genomic comparisons to characterize and compare the genetic diversity present within and between two closely related bacterial species. Together, these studies give new insight into how diversity is created and maintained in environmental bacterial populations.

Acknowledgments

I would be remiss if I didn't acknowledge the many people who made this PhD possible. First and foremost I thank my advisor, Larry Forney. Over the past 5+ years he has mentored me in science, writing, career choices, and life in general. Without his efforts, I would not be the scientist I am today and for that I am eternally grateful.

Second, I thank Susanna Remold, I began my career in science as a undergraduate in her laboratory at the University of Louisville. Her teaching and mentorship provided me with a strong foundation upon which to build. Without the opportunities she gave me, none of this would have been possible.

I also thank my committee members: Eva Top, Steve Krone, Luke Harmon, and Paul Joyce. Eva has always been willing to give a much needed second opinion on my writing/thinking. Steve's mathematical biology course gave life to the mathematics presented in Chapters 2 and 4 and Luke's comments on my projects helped increase the scope of my work. The work of the late Paul Joyce laid the foundation for many of my studies. Without his input and insight these studies would never have been accomplished.

For the spatial structure and slow growth work, I could always turn to Thubault Stalder and Ben Ridenhour for help thinking through the experiments, data analysis, and writing about them. I also thank Ana Cornea and Hanna Khelet-Delgado, who both helped on the biofilm project. Ana in particular taught me how to conduct evolution experiments in biofilms. For help with the *Lactobacillus* genomics study I thank Kenetta Nunn, Roxana Hickey, and Helena-Mendes Soares. All three of them were willing and eager to share what they knew about the vaginal microbiome.

I also thank Maria Schneider and Lisha Abendroth. Maria managed the Forney lab and directly facilitated all the work presented in my dissertation. She is always willing to help any and everyone with their work whenever they happen to need it, regardless of how busy she is at the time. Through her tireless work, and sometimes strict direction, the Forney lab space is always well-stocked, clean, and organized. I am convinced nothing would get done in the lab without her. Similarly, Lisha Abendroth runs the BCB graduate program and is always willing to help a confused graduate student navigate the university administration. Without Lisha's help, my experience in the PhD program would have undoubtedly been rocky and unpleasant.

Finally, I thank IBEST and all of its faculty members, students, postdocs, and administrators. IBEST provided the perfect setting for me to learn and develop. The influence of this vibrant community is spread throughout this work, from intellectual discussions that prompted new ideas to taking breaks to rejuvenate my intellectual powers, and cannot sufficiently detailed here.

Dedication

To my parents, Tim and Gayle, who have always supported
my ambitions and pushed me to succeed,
To my brother Cory, who I have always admired and respected,
To my in-laws, Ann and Kirk, whose encouragement and
love for education is inspiring,

and most important of all,

To my wife Shannon,

who is my best friend,
whose enduring curiosity in biology renews my own interest,
and whose love and support made this all possible.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Populations of bacteria are not, as many originally thought, largely clonal with a single dominant genotype (Ørskov and Ørskov, 1983; Smith et al., 1993). Many studies have shown that tremendous diversity can be found within bacterial species (Rainey et al., 2000; Konstantinidis et al., 2006; Cordero and Polz, 2014). This within-species diversity is important because it necessary for adaptation to occur and is the source of solutions to novel challenges (Fisher, 1930). Some of the first studies to demonstrate within species diversity in bacteria used electrophoresis banding patterns of allozymes to type environmentally isolated bacterial strains (Milkman, 1973; Selander and Leven, 1980; Ochman and Selander, 1984; McArthur et al., 1988). For example, Selander and Leven (1980) surveyed the variation in 20 enzymes for a set of *Escherichia coli* isolates and found that their set of 109 isolates included 98 unique banding patterns. A similar survey of *Burkholderia cepacia* (then named *Pseudomonas cepacia*) strains isolated along an environmental gradient revealed even more extensive genetic diversity (McArthur et al., 1988). However, these early estimates of within species bacterial diversity were probably conservative as they were derived from structural variation in essential enzymes, likely under strong purifying selection (Jordan et al., 2002; Luo et al., 2015). Advances in sequencing technologies have allowed researchers to directly assess the genetic diversity within bacterial species by sequencing whole genomes of multiple strains. These studies have shown that, in addition to substantial allelic diversity (Thompson et al., 2005; Konstantinidis et al., 2006), members of a bacterial species do not even necessarily have the same repertoire of genes in their genome (Tettelin et al., 2005; Cordero and Polz, 2014). A comparison of 61 *E. coli* strains revealed that they only had around 1000 genes in common, roughly 6% of the total number of genes encoded by the set of genomes (Lukjancenko et al., 2010).

Genome comparison studies have provided insight into the diversity present within a bacterial species but they typically lack the context necessary to address diversity within individual bacterial populations. Here we make a distinction between bacterial species, which include all strains that share 95% average nucleotide identify (Konstantinidis et al., 2006), and bacterial populations, which are individuals belonging to the same bacterial species that are found in the same physical location. Initially diversity within bacterial populations was examined by isolating multiple strains from a habitat of interest and measuring variation in a phenotype or genotype within the resulting strain collection (Cho and Tiedje, 2000; Oda et al., 2002; Picard and Bosco, 2008; France and Remold, 2016). More recently, shotgun metagenomics has been applied to directly characterize the genetic diversity in microbial populations (Brito and Alm, 2016; Tett et al., 2017; Truong et al., 2017). By sequencing DNA extracted directly from an environmental sample, metagenomic analysis can characterize the diversity present within the abundant species that make up a microbial community. Such an analysis of a human gut bacterial community identified 178 species but an excess of 200,000 genomic variants that resulted in over 5,000 distinct haplotypes (Kuleshov et al., 2016). These studies have highlighted the diversity present in many natural microbial populations and provided further evidence that microbial populations are far from clonal.

Diversity within bacterial populations is largely determined through the action of evolutionary processes like mutation, gene flow, selection, genetic drift, and gene flow. Of these four evolutionary processes,

selection has historically garnered the most attention both within the scientific community and in the general public. Selection enacts changes in genotype frequencies based on the differential reproductive success of individuals (Darwin, 1859). However, the other three processes also play critical roles in shaping the evolution of bacterial populations. Mutation creates genetic diversity by producing new genotypes via changes in the genetic code of individuals, while gene flow is the process by which new genotypes are introduced into a population either via the movement of individuals or the transfer of genetic material (Slatkin, 1985). Genetic drift, perhaps the most enigmatic of the four, accounts for changes in genotype frequencies that happen due to chance alone and results from sampling bias in finite population sizes (Masel, 2011). Besides these four processes, bacteria can also evolve through the lateral transfer of genetic material from one individual to another, outside the paradigm of vertical inheritance in clonal populations (Thomas and Nielsen, 2005).

Decades of research has characterized how these processes drive evolution in well-mixed laboratory populations of bacteria (Kawecki et al., 2012; Kassen, 2014). However, the populations used in these studies differ from the overwhelming majority of environmental bacterial populations in at least two crucial aspects. First, bacterial populations usually live and evolve in environments that have spatial structure (Kim et al., 2014; Nadell et al., 2016). This contrasts sharply with the common practice of cultivating bacteria in shaken test tubes. Second, bacteria in the environment typically grow much slower than those used in laboratory studies. Evolutionary biologists, in particular, have exploited the rapid growth of some bacterial strains to fit hundreds of thousands of generations in a relatively short time frame (Lenski et al., 1991; Elena and Lenski, 2003; Kawecki et al., 2012). Yet the majority of bacteria in the environment experience generation times on the order of days or weeks (Jannasch, 1969; Harris and Paul, 1994; Debellis et al., 1998; Rousk and Bååth, 2011; Kirchman, 2016). I hypothesized that these two factors, spatial structure and slowed growth rates, facilitate the emergence and maintenance of diversity in bacterial populations.

In Chapter 2, I review our current understanding of how spatial structure and reduced growth rates can impact bacterial evolution, while Chapters 3 and 4 are devoted to empirical tests of my hypothesis. In Chapter 3, I used experimental studies backed by model simulations in order to systematically characterize the relationship between spatial structure and the maintenance of diversity in bacterial populations. Results from this study give insight into the maximum amount of migration that can occur between bacterial subpopulations before the effects of spatial structure are masked. Chapter 4 is devoted to characterizing the effect that growing slowly has on bacterial evolution. To do so, I compared the evolution of *E. coli* populations in chemostats that operated with dilution rates that provided generation time of 2 (fast) and 24 hours (slow). The results produced from this study indicate that slowly growing populations experience higher per generation mutation rates and maintain more genetic diversity than rapidly growing populations. I also show that, over the course of the 72 generations, the slowly growing populations adapted faster than the rapidly growing populations. Results from both of these studies are consistent with my hypothesis and highlight the role that spatial structure and reduced growth rates might play in the maintenance of diversity in environmental bacterial populations.

In Chapter 5 I apply the knowledge gleaned in Chapters 2, 3 and 4 in order to understand how antibiotic resistance evolves in bacterial biofilms. Biofilms are microbial assemblages held together by an extracellular matrix and are both common in the environment and medically relevant (Hall-Stoodley et al., 2004; Høiby et al., 2011). The bacterial populations within biofilms are spatially structured (Tolker-Nielsen and Molin, 2000) and contain large subpopulations that are growing slowly (Wentland et al., 1996). I therefore hypothesized that growth within a biofilm would facilitate the accumulation and persistence of spontaneous antibiotic resistant mutants. My results indicate that antibiotic resistant variants accumulate prior to treatment, sweep to high frequency during treatment and can persist at that high frequency for extended periods of time after antibiotic treatment is halted. This study underlines the importance of heritable resistance mutations in the recalcitrance of biofilm infections to antibiotic therapy.

Finally, in Chapter 6, I analyze the genetic diversity present within and between two common bacterial inhabitants of the human vagina (*Lactobacillus crispatus* and *L. iners*). These species are closely related and have the capability to perform many of the same functions in the vaginal environment (e.g. production of lactic acid). Studies have shown that they rarely cooccur at high frequency and that swings in the dominant species of *Lactobacillus* can be common (Ravel et al., 2011; Gajer et al., 2012). These data suggest that *L. crispatus* and *L. iners* compete for shared niche space in the vagina. I show that the genomes of *L. iners* and *L. crispatus* differ in many respects, several of which may impact competition between these two species. My results provide insight into the factors that might drive the complicated temporal dynamics often exhibited by vaginal communities.

Chapter 2: Effects of spatial structure and reduced growth rates on evolution in bacterial populations

"Effects of spatial structure and reduced growth rates on evolution in bacterial populations *forthcoming as a chapter in the book Molecular Mechanisms of Microbial Evolution*

2.1 Abstract

The evolutionary forces that create and maintain the awesome diversity observed in microbial communities and populations are not well understood. For the most part, previous studies on microbial evolution have been done using model species that are grown in well mixed homogenous environments in which cells experience continuous or episodic periods of exponential growth. The relevance of these experimental systems to the evolution of naturally occurring populations is limited because bacterial populations in most environments reside in spatially structured heterogeneous habitats in which cell growth is slowed by nutrient limitation and cells often experience prolonged periods of stasis. Here we review and discuss how spatial structure and slow growth influence the evolution of microbial populations. We focus our discussion on microbial populations contained within biofilms, which are complex assemblages of microbial cells enclosed in self-made extracellular matrices. Biofilm bound populations have spatial structure and contain subpopulations that are growing slowly. Studies have shown that spatial structure limits competition to a local scale thereby protracting selective sweeps. Additionally, reduced growth rates directly impact the rate at which selection can alter genotype frequencies in populations. Combined, these characteristics lead to the emergence and maintenance of genetic diversity within biofilms. We contend that the findings of studies on evolutionary processes in bacterial biofilms can readily be extrapolated to other spatially structured microbial habitats, such as soils and sediments, in which nutrient limitation causes slow growth.

2.2 Introduction

The biological diversity among and within microbial species is important because it is the pool from which solutions to novel challenges are selected. The diversity of species present in many microbial communities and populations far exceeds that found in their macro-organism counterparts (Dykhuizen, 1998; Whitman et al., 1998), with a single gram of soil estimated to contain as many as 50,000 bacterial species (Schloss and Handelsman, 2006; Roesch et al., 2007). While some of this diversity can be explained by functional differences between the constituent species, much of it exists within ecotypes, which are lineages of genetically and ecologically distinct strains within a named species (Kopac et al., 2014; Shapiro and Polz, 2014; Cohan, 2016). These ecotypes coexist in small spaces seemingly in defiance of Gause's

law of competitive exclusion (Gause, 1934). Thus, there is a need to understand evolutionary processes as they play out in naturally occurring microbial habitats so that we can better comprehend the emergence and maintenance of genetic diversity in the microbial world.

Evolution is defined as changes in the frequency of genotypes in a population over time (Barton et al., 2007). The processes that drive evolution can be divided into four fundamental forces: mutation, selection, gene flow and genetic drift. Decades of prior research on microbial evolution has characterized the action of these evolution forces in well-mixed laboratory populations (Kawecki et al., 2012). Yet the vast majority of microbial populations found outside the laboratory differ dramatically in at least two key ways. First, they typically reside in structured environments such as soils, sediments and biofilms (Costerton et al., 1978; Whitman et al., 1998). Biofilms in particular can be found in nature (Costerton et al., 1978; Besemer et al., 2009; Burmolle et al., 2012), the built environment (Rogers et al., 1994; Wang et al., 2013), and even on/within the human body (Palestrant et al., 2004; Macfarlane and Dillon, 2007; Verstraelen and Swidsinski, 2013). Second, the majority of environmental populations are growing markedly slower than their laboratory counterparts. The growth of microbial populations is limited by the abundance of routinely scarce essential nutrients leading to generation times for populations in soils and the oceans frequently on the order of days and weeks (Jannasch, 1969; Harris and Paul, 1994; Rousk and Bååth, 2011; Kirchman, 2016). If we want to understand how diversity is created and maintained in natural microbial populations, we must endeavor to characterize the effects of spatial structure and reduced growth rates on microbial evolution.

While there are many examples of microbial populations that either exhibit spatial structure, or are growing slowly, in this chapter we chose to focus our discussion on populations of bacteria growing within biofilms. Biofilms are complex assemblages of microbial cells that are enclosed in a self-made extracellular matrix (Flemming and Wingender, 2010). They are often attached to surfaces but can also be found as free floating pellicles or globules of cells. This choice was made for several reasons, the first of which is that populations within biofilms both exhibit inherent spatial structure (Nadell et al., 2016) and often contain large subpopulations that are growing slowly (Sternberg et al., 1999; Stewart et al., 2016). However, they are also fairly ubiquitous in the natural world and, among other things, play critical roles in human disease (Parsek and Singh, 2003; James et al., 2007; Hall-Stoodley and Stoodley, 2009), biofouling of industrial equipment (Brooks and Flint, 2008; Flemming, 2011), and the cycling of nutrients in the environment (Paul et al., 1991; Battin et al., 2003; Meckenstock et al., 2015). Despite their prevalence and importance, surprisingly few studies have explored evolution within biofilm populations (Adams and Rosenzweig, 2014; Steenackers et al., 2016), although this subject has gained more attention over the past 5 years (Poltak and Cooper, 2011; Traverse et al., 2013; Ellis et al., 2015). This is perhaps because biofilms are inherently more complex than their planktonic counterparts (Stoodley et al., 2002; Stewart and Franklin, 2008; Flemming et al., 2016), which complicates both the technical execution of experiments and the analysis and interpretation of the resulting data. In the following chapter, we review how growth within a biofilm influences microbial evolution. We break our discussion down into three overarching topics: spatial structure, the effects of growing slowly, and horizontal gene transfer (HGT)

in biofilms. While the discussion within focuses on biofilms, our observations may apply equally well to other, non-biofilm populations, that are either spatially structured or growing slowly.

2.2.1 Biofilms and the evolution of antibiotic resistance

The evolution of antibiotic resistance is a notable example of why it is imperative that we advance our understanding of how evolutionary processes operate in biofilms populations. Biofilms have been shown to play a critical role in the pathogenesis of many bacterial species (Lynch and Robertson, 2008; Percival et al., 2012; Mulcahy et al., 2014). Estimates from the Centers for Disease Control and Prevention indicate that over 65% of all bacterial infections are caused by pathogens residing with biofilms (Klevens et al., 2007). Prominent examples of biofilm infections include those caused by *Pseudomonas aeruginosa* in the lungs (Mulcahy et al., 2014), *Acinetobacter baumannii* in wounds (Qi et al., 2016), and *Staphylococcus aureus* on indwelling devices (Otto, 2008). Infections caused by pathogens residing within biofilms have been repeatedly shown to be extremely difficult to eradicate, regardless of the method used, including antibiotics (Høiby et al., 2010), disinfectants (Kumar and Anand, 1998), and mechanical removal (Simões et al., 2005). Furthermore, residing in biofilms has been shown to provide protection against the human immune system (Jensen et al., 2010; Domenech et al., 2013). Consequently these infections are associated with higher rates of morbidity and mortality (Lynch and Robertson, 2008; Høiby et al., 2011) and are more likely to progress from an acute to a chronic state (Wolcott et al., 2010). We speculate that because biofilm infections are both pervasive and persistent, they may be a hot spot for the evolution of antibiotic resistance.

Perhaps the most alarming characteristic of populations within biofilms is their inherent recalcitrance to antibiotic treatment. Initial investigations demonstrated that this trait was not heritable, suggesting that it did not result from genetic change. This led to the proposition of several phenotypic mechanisms have been proposed to explain the characteristic recalcitrance of populations within biofilms. Early studies suggested the biofilm matrix might slow the diffusion of antibiotics, providing for zones of sub-inhibitory antibiotic concentrations (Hoyle et al., 1992; Stewart, 1996). While others focused on the wide range of physiological states present in biofilm populations, some of which are more tolerant to certain antibiotics (Xu et al., 1998; Sternberg et al., 1999). For example, β-lactam antibiotics that inhibit bacterial cell wall synthesis may not be effective against non-growing or slowly growing cells found in the interior of a biofilm. More recent studies have shown that biofilm populations harbor antibiotic tolerant individuals that emerge from the differential expression of toxin-antitoxin system genes (e.g. hipAB) (Lewis, 2007, 2010; Maisonneuve and Gerdes, 2014). The tolerant phenotype exhibited by these individuals allows them to persist, but not grow, during antibiotic treatment. This phenotype is readily reversed (Amato et al., 2014; Schumacher et al., 2015) and is not the result of heritable genetic mutations that confer antibiotic resistance (Olsen, 2015; Brauner et al., 2016). While these phenotypic mechanisms certainly play a role in the recalcitrance of biofilm populations to antibiotic treatment, they do not preclude a significant contribution of genetic change. In fact, by allowing portions of the population to survive during treatment, tolerance mechanisms may actually facilitate the evolution of antibiotic resistance. This was demonstrated in a recent study that showed antibiotic tolerant *Escherichia coli* subpopulations acquired

mutations that provided resistance to ampicillin (Levin-reisman et al., 2017). More work is needed to detail the potential interplay between phenotypic tolerance and the evolution of heritable antibiotic resistance in biofilm infections.

There is also a growing body of research that suggests that the accumulation of genetic diversity in biofilm populations may also directly contribute to their recalcitrance (Boles and Singh, 2008; Eastman et al., 2011; Tyerman et al., 2013). This was perhaps missed by the initial investigations into biofilm recalcitrance because these studies were largely conducted by exposing relatively young biofilms to antibiotics for relatively short periods of time and then examining survivors (Williams et al., 1997). By not providing enough time for evolution to enact change, this experimental design may have precluded any observation of the evolution of heritable antibiotic resistance. This is especially important given the often-prolonged nature of biofilm infections. In their 2004 paper Boles *et al.* 2004 demonstrated that *Pseudomonas aeruginosa* biofilm populations accumulate genetic diversity and that this diversity allows populations to better resist environmental stress. They connected these observations to the "insurance hypothesis" from ecological theory, which states that genetic diversity fortifies populations against future challenges (Yachi and Loreau, 1999). Other studies have shown that the diversity accumulated in biofilm populations includes spontaneous genetic mutants that are resistant to antibiotics (Boles and Singh, 2008; Driffield et al., 2008; Ponciano et al., 2009; Tyerman et al., 2013). Likewise, our own unpublished data from studies on *Escherichia coli* biofilms suggests this observation may be true for a wide array of antibiotics (Figure 2.1). Mutants resistant to kanamycin, chloramphenicol, cycloserine, and tetracycline spontaneously increased in frequency by more than two orders of magnitude over the course of 15 days of growth in the absence of antibiotics, as compared to the frequency in the inoculum. Future efforts to study microbial evolution in biofilms are likely to reveal the mechanisms behind the accumulation of these antibiotic resistant variants.

2.3 Effect of spatial structure on microbial evolution

Populations of bacteria housed within biofilms exhibit inherent spatial structure (Nadell et al., 2016). The extracellular matrix that encloses a biofilm provides the population with a three-dimensional configuration and restricts the movement of individual cells (Flemming and Wingender, 2010). Studies on the evolution and ecology of macroorganisms have demonstrated the profound effects of spatial structure on adaptation and diversity (Tillman, 1994; Pannell and Fields, 2014). The key findings from these studies are: (1) spatial structure promotes diversity; (2) interactions among and between species occur within localized niches; (3) organisms adapt at a local scale; and (4) forces such as migration, gene flow, and genetic drift can either promote or retard local adaptation, depending on specific circumstances. The limited information available on the evolution of microbial populations in spatially structured environments suggests that many of these findings will apply equally well to populations of microorganisms.

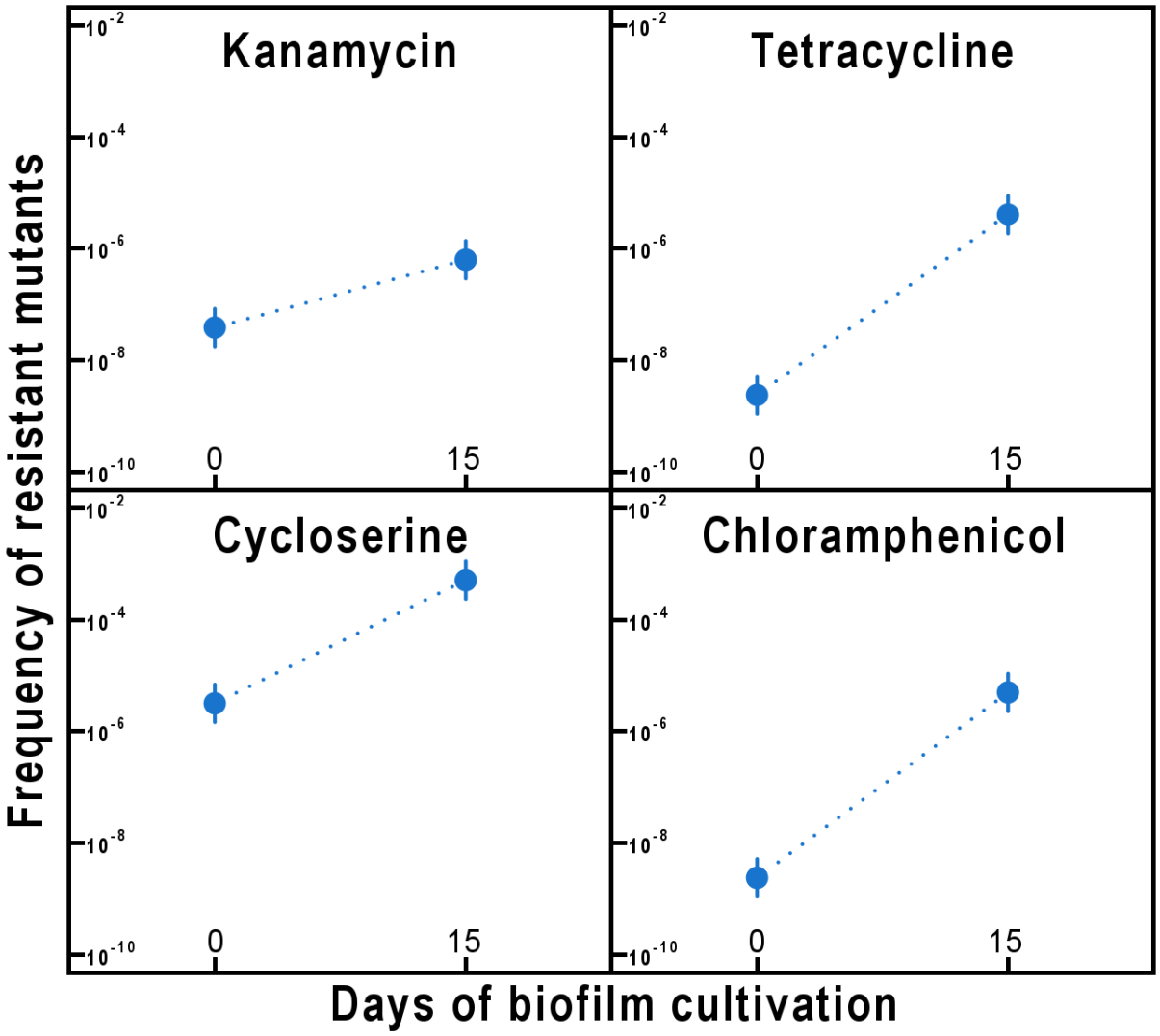

Figure 2.1: Increase in the frequency of antibiotic resistant mutants in a single biofilm population of *Escherichia coli* MG1655 during 15 days of antibiotic-free cultivation. Biofilms were grown in flow cells as described in Ponciano et al. (2009). Resistance was determined by plating on selective media containing one of the following antibiotic concentrations (kanamycin 25μg/mL, tetracycline 10μg/mL, cycloserine 40μM, chloramphenicol 25μg/mL).

2.3.1 Spatial structure and maintenance of diversity

Previous studies have shown that spatial structure has a profound impact on the evolution of microbial populations (Rainey and Travisano, 1998; Perfeito et al., 2008; Kryazhimskiy et al., 2012; Nahum et al., 2015). In structured populations the position of individuals within the spatial landscape matters because competitive interactions do not occur on a global scale like they do in well-mixed populations (Kerr et al., 2002; Kim et al., 2014). Instead, individual cells only compete for resources with a small subset of the population that is in close physical proximity (Figure 2.2). This means that the relative fitness of a given genotype is dependent not on the average fitness of the entire population, but rather just a subset of cells that are in the immediate neighborhood. Thus, spatial structure helps maintain diversity in microbial populations by reducing the scale of competitive interactions. This was elegantly demonstrated by Kerr et al. (2002) who showed that when toxin-producing, toxin-resistant and toxin-susceptible strains of *Escherichia coli* compete on the surface of an agar plate (a spatially structured environment), then all three strains persisted. In contrast, when environmental spatial structure was removed by growing the populations in well-mixed cultures, then only a single genotype—the toxin resistant strain—persisted (Kerr et al., 2002). The fate of only three genotypes were tracked in this study. Yet mutational processes in naturally occurring biofilm populations that have very large population sizes are likely to create a plethora of genotypes, many of which might coexist due to spatial structure. We postulate that spatial structure is likely to play a critical role in the maintenance of this diversity through time. This is consistent with the results from several studies that have already shown that biofilms tend to accumulate extraordinary genetic diversity over relatively short periods of time (Figure 2.3; (Boles et al., 2004; Ponciano et al., 2009; Tyerman et al., 2013)).

2.3.2 Spatial structure and clonal interference

Previous studies have shown that spatial structure can increase the amount of time required for beneficial mutations to fix (Whitlock, 2003; Habets et al., 2007; Perfeito et al., 2008). Instead of rapidly sweeping through the entire population all at once, beneficial mutations in biofilms must sweep through the physical space as well, thereby slowing an overall increase in biofilm population fitness (Nahum et al., 2015). This was clearly shown in the work of Perfeito et al. (2008). They demonstrated that the increase in fitness after 275 generations of evolution was significantly lower in *Escherichia coli* populations that were grown on the surface of an agar medium as compared to populations grown in a liquid medium (Perfeito et al., 2008). This stands in contrast to the idealized process of evolution in asexual populations wherein adaptation is driven by a series of beneficial mutations that occur and sweep to high frequency in the population one at a time (Elena et al., 1996), with each new mutation building on the fitness gains provided by prior mutations. In reality, the events seem to play out differently in large populations of microorganisms, and there is increasing evidence that multiple beneficial mutations arise simultaneously and compete against each other (Lang et al., 2013). This effect, known as clonal interference (Sniegowski and Gerrish, 2010), has been theorized to slow the rate of adaption by creating "wasted genetic potential" (Muller, 1932). We posit that spatial structure is likely to enhance and amplify the effects of clonal

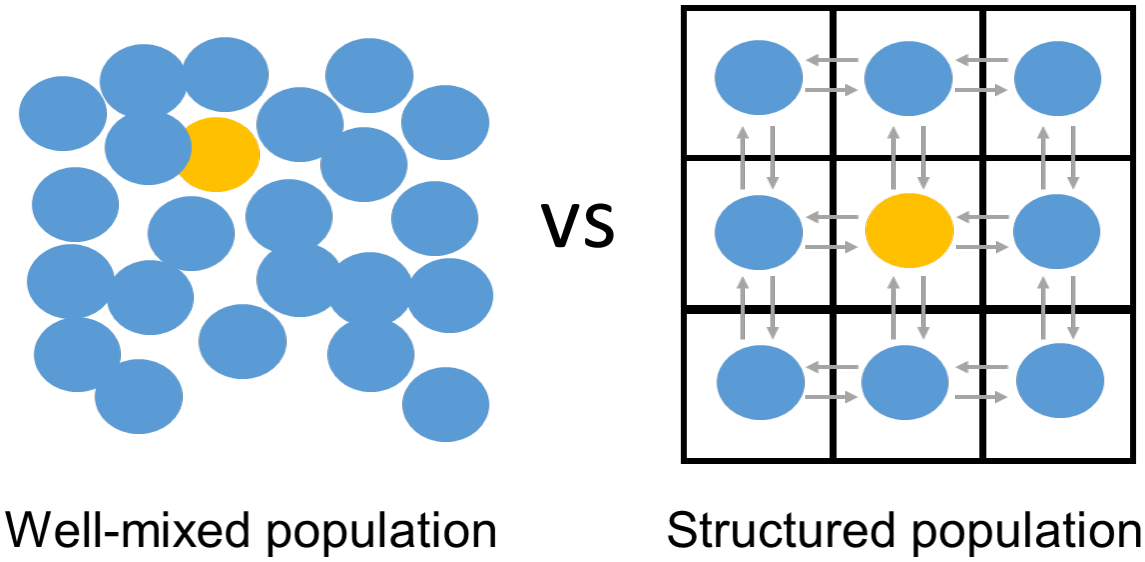

Figure 2.2: Simplified example of competition in well-mixed versus structured populations. In well-mixed populations competition occurs at a global scale and the focal (yellow) individual must compete against the entire population. In structured populations competition is limited to a local scale, and the focal individual (yellow) only compete against the subset of the population that is physically near it.

interference in biofilm populations. This can be attributed to the increased amount of time required for a beneficial mutation to sweep through a biofilm population, which provides an opportunity for new beneficial mutations to emerge. As a consequence, there are likely countless beneficial mutations present in a biofilm at any given time (Figure 2.4).

While the clonal interference that results from evolution in structured environments is generally predicted to slow adaptation in the short-term, it may ultimately allow members of the population to attain higher fitness values in the long-run. This "tortoise and hare" effect (Nahum et al., 2015) is best visualized using Sewall Wright's fitness landscape, which relates genotypic space to fitness (Wright, 1932). When epistasis is rare and the fitness effects of mutations are largely additive, the landscape is a smooth surface (Figure 2.5A). Conversely, when epistasis is common the landscape becomes "rugged" with sharp transitions in fitness over few mutational steps (Figure 2.5B). In the absence of spatial structure, adaptation is expected to drive the entire population up the nearest fitness peak (Figure 2.5A; (Gillespie, 1983)). In contrast, spatial structure effectively fragments populations into metapopulations that simultaneously adapt along multiple paths in the fitness landscape (Wakeley, 1998; Cherry and Wakeley, 2003). If the fitness landscape is smooth, no benefit is derived from increased exploration as the metapopulations are all expected to climb the same fitness peak. However, if the fitness landscape is rugged, the largely independent exploration of the landscape can prevent the population from becoming stuck at local fitness optima and instead allow certain metapopulations to reach fitness peaks that might otherwise be unattainable Figure 2.5B; (Burch and Chao, 2000; Rozen et al., 2008; Nahum et al., 2015)).

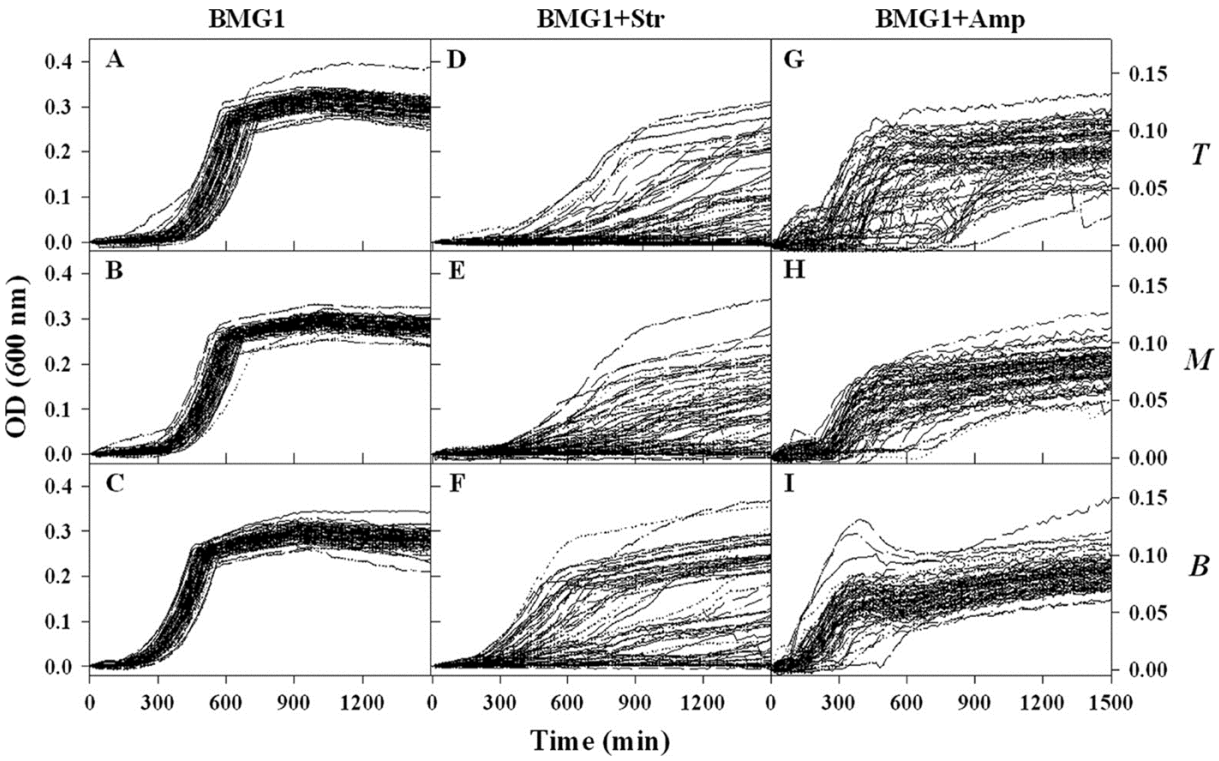

Figure 2.3: Diversity in growth kinetics displayed by 192 clones isolated from a single antibiotic naive *Escherichia coli* k12 MG1655 biofilm. Sixty-four clones were isolated from the top (2,250 to 2,400 μm; A, D, and G), middle (1,050 to 1,200 μm; B, E, and H), and bottom (0 to 100 μm; C, F, and I) of a biofilm. The growth of each clone was tested in three different media, two of which contain sub-lethal antibiotic concentrations: BMG1 (panels A, B, and C), BMG1+streptomycin (5μg/mL; panels D, E, and F), and BMG1+ampicillin (5μg/mL; panels G, H, and I). OD, optical density. Figure reproduced from: Ponciano, J. M., La, H., Joyce, P., Forney, L. J. (2009) Evolution of diversity in spatially structured *Escherichia coli* populations. Appl. Environ. Microbiol. 75(19): 6047-6054

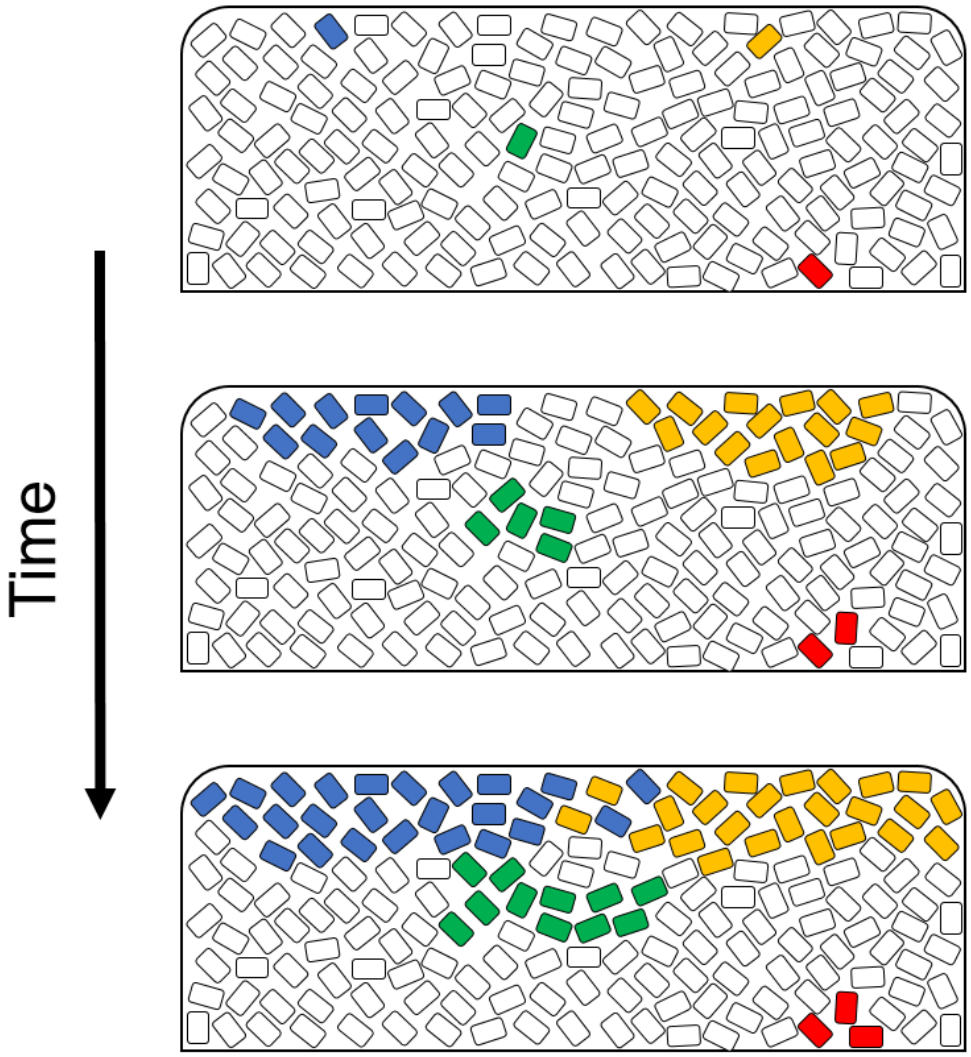

Figure 2.4: Spatial structure causes the beneficial mutations colored in blue, yellow, red, and green to sweep locally through the biofilm. In a well-mixed population, the beneficial mutation with the highest fitness would rapidly sweep through the population, thereby collapsing genetic diversity.

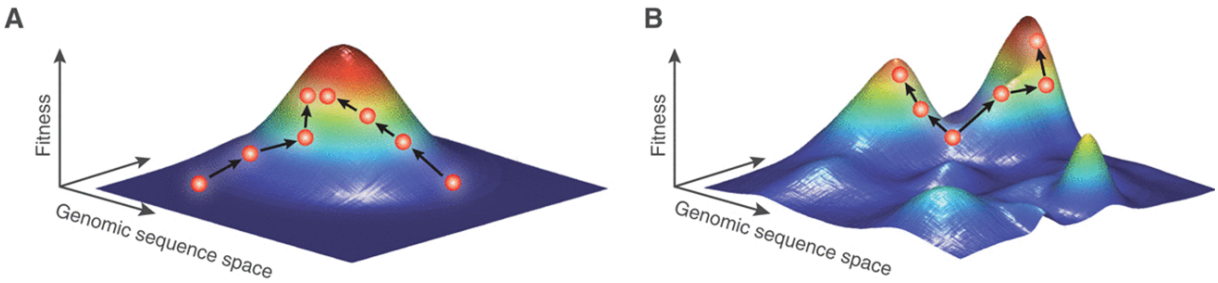

Figure 2.5: Examples of smooth (A) and rugged (B) fitness landscapes. For a smooth landscape with a single fitness peak (A), sampling more of the landscape is not beneficial because all paths lead to the same location. Whereas, for a rugged landscape with multiple peaks (B), spatial structure allows a population to simultaneously climb multiple fitness peaks, increasing the likelihood that the population reaches higher peaks. Figure reproduced from: Conrad. T.M., Lewis, N.E., Palsson, B. Ø. (2011) Microbial laboratory evolution in the era of genome-scale science. Molecular Systems Biology 7:509.

2.3.3 Emergence of antagonistic pleiotropy in biofilms

Spatial structure creates environmental heterogeneity in biofilm populations (Xu et al., 1998; Stewart and Franklin, 2008; Stewart et al., 2016). Instead of being uniformly distributed as they are in well-mixed populations, the availability of nutrients and terminal electron acceptors in biofilms is altered by reaction-diffusion processes (Erban and Chapman, 2007)). As molecules diffuse into the biofilm structure they are consumed by bacterial cells creating environmental gradients wherein resource concentrations are highest at the interface with the bulk media (exterior) and lowest in interior regions of biofilms Figure 2.6; (Stewart, 1998)). This heterogeneity is further compounded by the complex three dimensional structure of biofilms and the fact that biofilms are often comprised of multiple species with distinctive physiologies and a patchy distribution (Tolker-Nielsen and Molin, 2000; Webster et al., 2006; Elias and Banin, 2012; de los Rios et al., 2015). This environmental heterogeneity drives adaptation in response to selective pressures that are localized and specific to individual microenvironments (Kraemer and Boynton, 2017). For example, mutations that improve aerobic metabolism may be selected for in the upper reaches of a biofilm where oxygen is readily available, but selected against in the lower, anaerobic parts of a biofilm (de Beer et al., 1994; Stewart and Franklin, 2008).

Not surprisingly, mutations that are beneficial in one environment might be detrimental in another. Such instances of antagonistic pleiotropy have been shown to be relatively common in nature (Anderson et al., 2013; Kassen, 2014; Schenk et al., 2015; Ferenci, 2016). Should antagonistic pleiotropy be common in biofilms, the complex array of microenvironments would be expected to fragment the population into multiple subpopulations (metapopulations). Antagonistic pleiotropy would drive these metapopulations to diverge and evolve towards specialization in their particular environment (Futuyma and Moreno, 1988; Devictor et al., 2010). The effect of local adaption to specific microenvironments is also likely to be amplified by the spatial structure of biofilm populations, which can prevent "entire biofilm" selective sweeps during reasonable time periods. Indeed, population-wide selective sweeps in biofilms may only result

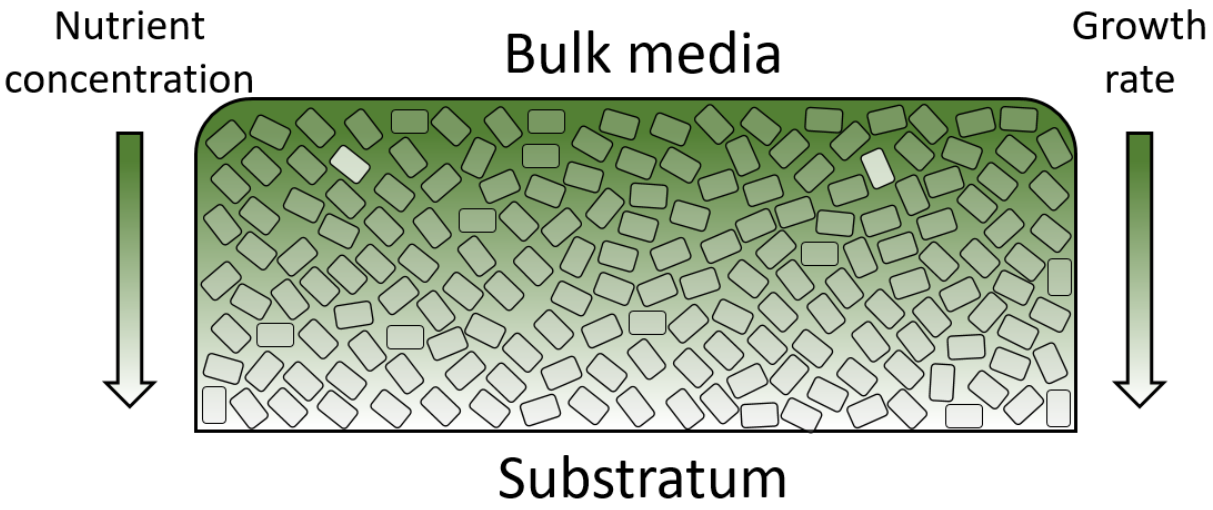

Figure 2.6: Simple schematic of the distribution of nutrient concentrations and growth rates in a theoretical biofilm population.

from strong directional selective pressures like that provided by the application of antibiotics.

2.4 Effect of growth rate on mutation and selection

Variation in cellular growth rates is an important consequence of the complex array of environments found in biofilms (Sternberg et al., 1999; Stewart and Franklin, 2008; Stewart et al., 2016). Those cells located at or near the source of incoming nutrients are likely to exhibit higher growth rates than those located in more nutrient-depleted zones. Large swaths of the interior of biofilms may not be growing (dividing) at all and exist in a quiescent state. Although it has been chronically overlooked, we suggest that a microbial population's growth rate is likely to strongly influence its evolution. While mutations can occur independent of chromosomal replication and cellular division, selection requires some combination of cellular growth and death to affect evolutionary changes in microbial populations. The relative influence of these two processes, mutation and selection, may therefore vary with the rate at which the population is growing.

2.4.1 Replication dependent and independent mutational processes

Mutation, unlike selection, can occur independent of cell growth and death. The mutations that underpin evolutionary change in biofilm populations fall into two broad categories: replication-dependent and replication-independent mutations. The former largely consist of point mutations introduced during the replication of bacterial chromosomes in actively growing cells, while the latter arise through several different mechanisms (Table 2.1), including error prone repair of damaged DNA (Goodman, 2002;

Table 2.1: Processes that cause mutations independent of genome duplication

Mutational Mechanism	References
Repair of DNA damage by:	
Uv radiation	(Bridges, 1992; Truglio et al., 2006)
Gamma radiation	(Wijker et al., 1998; Rodgers and Mcvey, 2016)
Cytosine deamination	(Duncan, 1980; Frederico et al., 1993)
Depurination	(Schaaper and Loeb, 1980; Suzuki et al., 1994)
DNA oxidation	(Imlay, 2003; Bjelland, 2003)
Non-enzymatic methylation	(Rydberg and Lindahl, 1982; Mazin et al., 1985)
Movement of mobile elements	(Kidwell and Lisch, 2001; Frost et al., 2005)
Recombination	(Anderson and Roth, 1981; Bull et al., 2001)
Slipped-strand mispairing	(Levinson and Gutman, 1987; Torres-Cruz and van der Woude, 2003)

Rodgers and Mcvey, 2016), recombination (Anderson and Roth, 1981; Bull et al., 2001), movement of mobile genetic elements (Frost et al., 2005; Foster, 2007) and slipped-strand mispairing (Levinson and Gutman, 1987; Torres-Cruz and van der Woude, 2003). These replication-independent mutations occur in all cells regardless of their growth rate, probably even in those that are quiescent (Bull et al., 2001; Kivisaar, 2003, 2010). This phenomenon can be seen in Figure 2.7, which demonstrates a gradual linear increase in rifampicin resistant mutants in stationary phase populations of *E. coli*.

The failure to adequately consider replication independent mutations is reflected in the fact that mutation rates are expressed in terms of "mutations per generation" which by logical extension implies that if there are no generations (i.e., cells are not replicating their chromosome and dividing), then the expected number of new mutants is zero. Additionally, the vast majority of experiments done to measure mutation rates are done in studies in which mutation rates are determined using exponentially growing cells (Rosche and Foster, 2000; Foster, 2006). This could be misleading since microbes in many natural habitats grow slowly or episodically and exponential growth is uncommon (Debellis et al., 1998; Dixon and Turley, 2001; Rousk and Bååth, 2011; Kirchman, 2016). Harris and Paul (1994) estimated the generation time of bacteria in agricultural soils to be 160 days and 107 days in grassland soils. Other investigators have estimated the generation times of bacteria in pelagic marine environments to be 8-9 days (Carlucci and Williams, 1978), while Lomstein et al. (2012) have projected biomass turnover times of hundreds to thousands of years in deep sub-seafloor sediments.

Given the low bacterial growth rates in most habitats it could be that replication-independent mutations are the principle means by which genetic diversification occurs. Thus, in studies of evolution in biofilms it may be important to consider both replication-dependent and replication-independent mutations (Eastman et al., 2011) because growth rates may vary depending on the location of cells in the

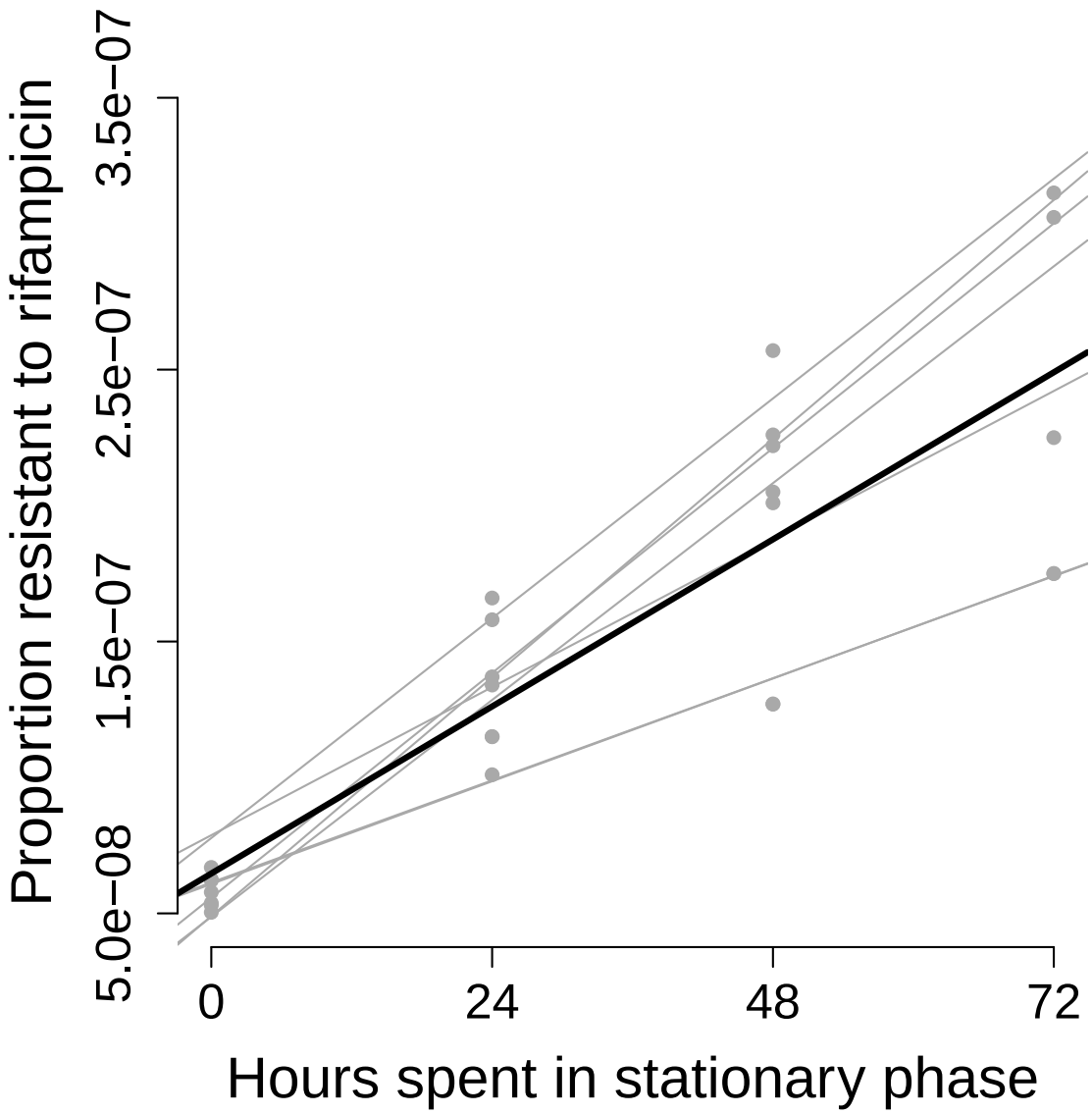

Figure 2.7: Increase in the frequency of rifampicin resistant mutants in seven planktonic stationary phase cultures of *E. coli*. Grey points are estimates of the frequency from each individual culture at each time point. Grey and black lines resulted from a simple linear regression of the mutant frequency versus the time spent in stationary phase for the individual cultures (grey) and the overall average of all cultures (black).

matrix. Lastly, because of the potentially important role of mutation in slowly growing or quiescent cells, we suggest that mutation rates be expressed as mutations per unit time instead of mutations per generation.

2.4.2 Strength of selection when growing slowly

On the other hand, cell growth or death must occur for selection to affect changes in allele frequencies in the population. This implies that the rate at which a population is growing (or dying) has a direct influence on the rate at which selection can alter allele frequencies. Consider the following simplified example: If a beneficial mutation requires 250 generations to reach fixation then this would require 25 days in a population that has a growth rate of 10 generations of per day. In contrast, fixation would require 1750 days in a population that experiences only 1 generation per week. When applied to biofilms this implies that the times required for beneficial mutations to sweep may well depend on their location in the matrix, with fixation occurring more quickly near the interface with the bulk media and much more slowly in deeper regions near the substratum. Moreover, as previously discussed, many of the molecular mechanisms that cause mutations are not bound to chromosomal replication or cellular division. Perhaps mutation rates might best be expressed as a continuous time process rather than instead of being indexed to the production of offspring—an event that is required for natural selection to operate. By extension of this reasoning the effect of selection is undetectable in populations that are in stasis (i.e. cells are neither replicating or dying).

The connection between growth rate and selection in bacterial populations can be made clear by examining mathematical models that relate mutation and selection in well-mixed populations. If we ignore gene flow and genetic drift, a deleterious mutant accumulates at a rate defined in terms of mutation, which provides new variants, and selection, which purges lower-fitness mutants from the population. Traditional models (Equation 2.1) define the change in mutant abundance (Δx) as being approximately equal to the mutation rate μ minus the selection coefficient s multiplied by the current proportion of mutants x (Fisher, 1928), i.e.

$$\begin{aligned} \Delta x &\approx \mu - sx \\ x_t &= (1-s)x_{t-1} + \mu \end{aligned} \tag{2.1}$$

The selection coefficient s in this model is defined as the covariance of x and its fitness (measured at time t). If we assume the time-step between $t-1$ and t is 1 (note that the units here are not specified, the usual choice would be 1 generation), we can break the mutation rate up into two portions: (1) k is the portion of the time-step during which chromosomal replication is not occurring. And, (2) $(1-k)$ is the proportion of the time-step where chromosomal replication occurs. Now assume that replication dependent and independent mutations occur at rate η and v, respectively. Then

$$\begin{aligned} \mu &\equiv vk + \eta(1-k) \\ \therefore x_t &= (1-s)x_{t-1} + vk + \eta(1-k). \end{aligned} \tag{2.2}$$

Equation 2.2 highlights the balance between the two types of mutation. If k is large (as it might be in slow growing populations), then v has increased importance in evolution. Conversely, if most of the time the bacteria are undergoing chromosomal replication (i.e. k is small), then η has increased importance. Note that, because the selection coefficient is estimated at unit length, the change due to natural selection is of fixed size in comparison.

This analysis suggests that it is important to consider the possible effect of growth rate on selection when thinking about evolution in biofilm populations. This matter might be important for at least two reasons. First, there is large variation in growth rates throughout a biofilm (Wentland et al., 1996; Sternberg et al., 1999) so it is likely that the relative contribution of replication dependent and independent mutations varies through the population (Eastman et al., 2011). Specifically, cells near the bulk fluid interface are likely to experience faster growth rates and may therefore incur more growth dependent mutations (Sternberg et al., 1999), while nutrient starved cells may incur a higher proportion of replication independent mutations (Bull et al., 2001; Saumaa et al., 2002; Kivisaar, 2003). Second, cells in the basal layers of biofilms experience conditions that severely limit or preclude growth (shown in Figure 2.6). The influence of selection on these slowly growing or quiescent populations is likely minimal ($s \approx 0$). Yet mutations are expected to accumulate unabated through replication independent processes.

2.5 Horizontal gene transfer in biofilms

Genetic diversity can also be created by the introduction of genetic elements into cells (Thomas and Nielsen, 2005). This process is termed horizontal gene transfer (HGT) and it stands in contrast to the vertical inheritance of genes through common descent (Thomas and Nielsen, 2005). HGT is an important component of microbial evolution because it allows for the rapid dissemination of genetic information. Comparative genomics studies have shown that the genomes of most microorganisms contain considerable amounts of genetic material has been acquired through HGT and not common descent (Ochman et al., 2000). Considerable attention has been given to HGT because it has played a critical role in the acquisition of virulence determinants and the spread of antibiotic resistance (Davies and Davies, 2010). Below we describe how residing within a biofilm might influence HGT. We focus our attention on conjugation and transformation because relatively few studies have examined transduction in biofilms.

Conjugation is probably the HGT process most studied in biofilms. Many have argued that conjugation rates are likely to be elevated in densely populated biofilms because cell-cell contact is required for plasmid transfer and immobilization in the extracellular matrix may facilitate successful mating by decreasing the chances that the donor and recipient conjugation machinery is not disrupted by sheer forces (Hausner and Wuertz, 1999). On the other hand, spatial structure may slow and limit plasmid spread

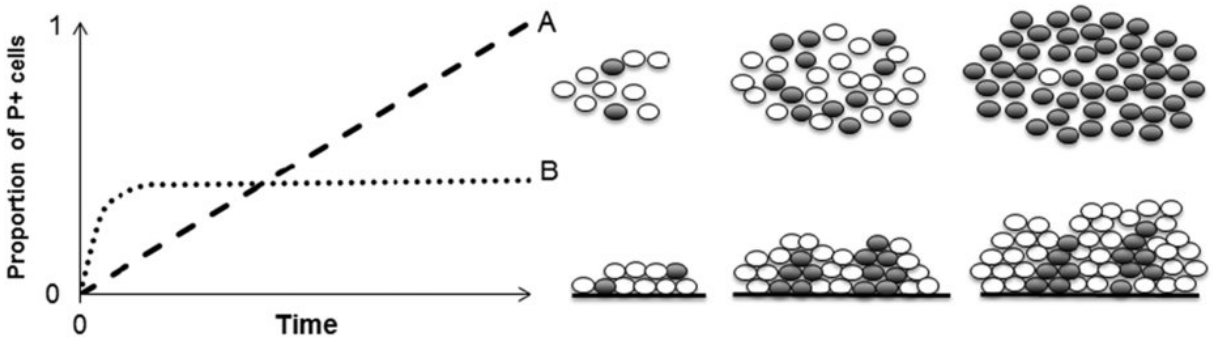

Figure 2.8: Theoretical spread of a self-transmissible plasmid in a well-mixed (A) and structured population (B). This is demonstrated schematically in the drawing on the right. Plasmid free bacteria (white) and plasmid containing bacteria (dark gray) in well-mixed (top) and structured (bottom) populations. Figure is reproduced from: Stalder, T., Top, E. (2016) Plasmid transfer in biofilms: a perspective on limitations and opportunities. npj Biofilms and Microbiomes 2: 16022.

in biofilms through the same mechanisms that slow selective sweeps (Figure 2.8). Correspondingly, some studies have indicated that plasmid transfer rates are elevated in biofilms (Hausner and Wuertz, 1999; Kajiura et al., 2006; Meervenne et al., 2014) while others have suggested the exact opposite (Christensen et al., 1996, 1998). Arguments against elevated transfer rates in biofilms are based on either the need for cell-cell contact that may limit plasmid transfer to nearest neighbors in the biofilm matrix (Tolker-Nielsen and Molin, 2000; Seoane et al., 2011) or the depletion of energy sources (nutrients) in the deeper regions of a biofilm (Fox et al., 2008). Furthermore, Król et al. (2011) demonstrated that conjugal transfer of IncP-1 in *E. coli* biofilms was dependent on oxygen levels and population densities (Król et al., 2011). This is was supported by the findings of Stalder and Top (2016) who observed limited plasmid transfer and spread in mixed species biofilms of *E. coli* and *P. putida* (Figure 2.9; (Stalder and Top, 2016)). Additional studies are needed to define the conditions under which conjugation is favored or limited in biofilm populations.

Transformation occurs when extra-cellular DNA (eDNA) is taken up by another bacterium and integrated into the chromosome. This process may be a prominent mechanism of horizontal gene transfer in microbial biofilms because eDNA has been shown to be common component of the extracellular matrix (Molin and Tolker-Nielsen, 2003; Chiang and Tolker-Nielsen, 2010). Extracellular DNA is thought to play a structural role in biofilms but may also be taken up by the bacteria embedded in the matrix (Whitchurch et al., 2002; Hobley et al., 2015). Furthermore, studies have identified transformation in biofilms formed by *Streptococcus mutans* (Li et al., 2001), *Acinetobacter* spp. (Hendrickx et al., 2003; Merod and Wuertz, 2014), *Gonococcus* sp. (Kouzel et al., 2015) and multispecies oral biofilms (Hannan et al., 2010). The study on *Gonococcus* biofilms by Kouzel et al. (2015) further demonstrated that transformation rates were higher in early stage biofilms than in planktonic populations. However, they also showed that transformed rates were diminished in the biofilm populations after 24 hours, a result that was also observed in the studies on *Acinetobacter* sp. and *Streptococcus mutans* studies (Li et al., 2001; Hendrickx et al., 2003). The mechanism responsible for this trend has not been identified, although

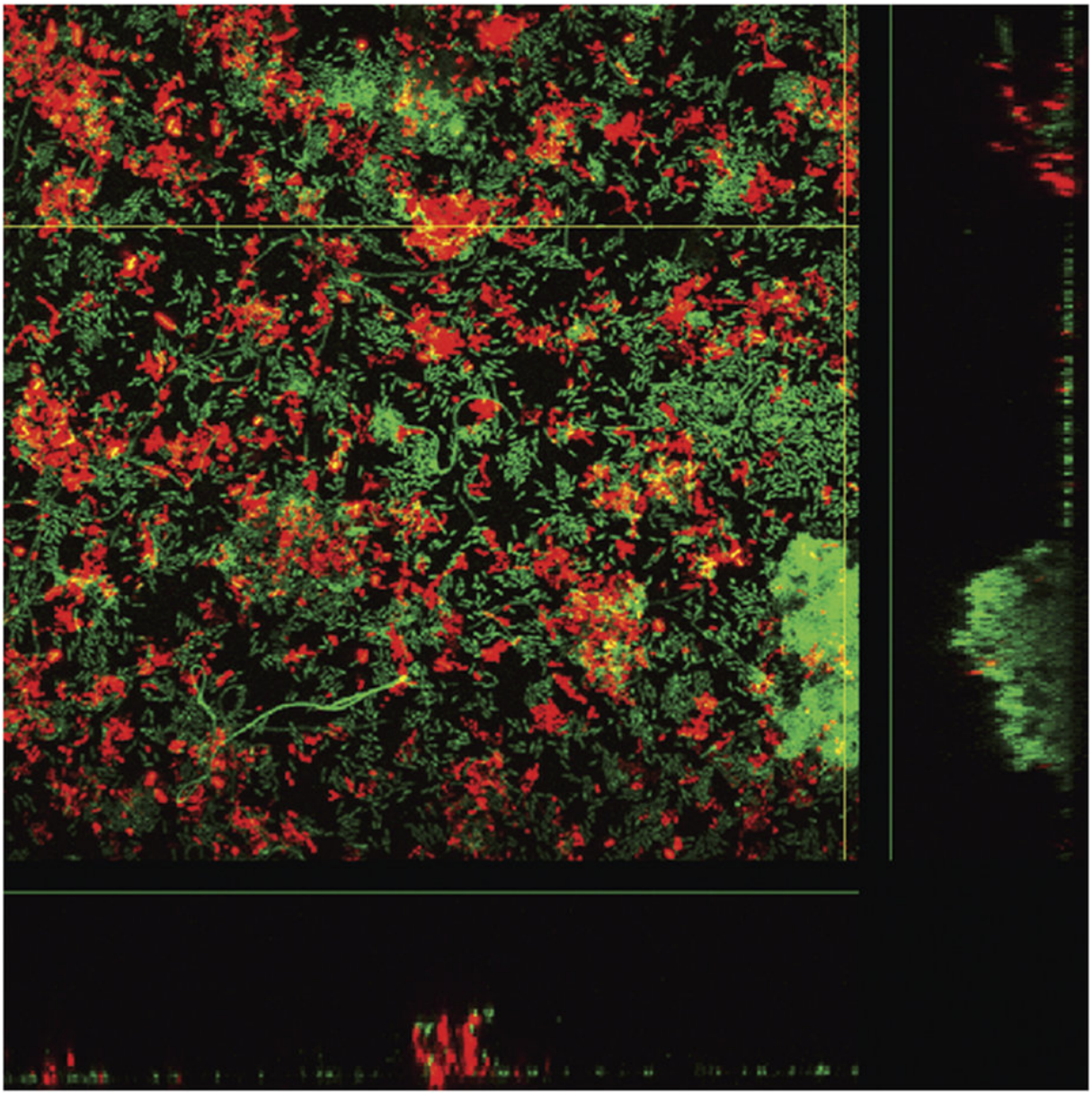

Figure 2.9: Confocal laser scanning microscopy photographs of plasmid transfer in a dual-species biofilm. Plasmid donor cells (red) are *E. coli* K12 MG1655 carrying plasmid pB10 marked with dsRed and recipient cells (green) are *P. putida* KT2244 marked with gfp. Transfer of the dsRed marked plasmid from *E. coli* into the gfp marked *P. putida*, generates the yellow/orange transconjugants seen in the image. Figure reproduced from: Stalder, T., Top, E. (2016) Plasmid transfer in biofilms: a perspective on limitations and opportunities. npj Biofilms and Microbiomes 2: 16022.

Kouzel et al. (2015) suggest that it could result from either oxygen deprivation or the increased matrix density of older biofilms (Kouzel et al., 2015). While it is still not clear whether transformation rates are generally elevated or reduced in biofilms, we speculate that this may be an important and underappreciated means of HGT in biofilms. We base our speculation on the abundance of eDNA in the matrix, and the fact that many environmental biofilms are complex multi-species conglomerates.

2.6 Conclusions

Despite the fact that few studies have been done on the evolution of populations that reside in biofilms, we have synthesized the available findings along with principles of macroecology and evolutionary theory to speculate about the factors that influence adaptive radiation and evolution in biofilms. In our view, four fundamental characteristics of biofilms must be taken into account. The first is that the extracellular matrix restricts the movement of individual bacteria within biofilms, limiting the scale of competitive interactions. Second, reaction-diffusion processes within biofilms create microenviroments that exert a spectrum of selective pressures. Third, mutagenic processes that are independent of chromosome replication and cell division undoubtedly occur, which leads to genetic diversification of cells; even those that are growing slowly or not at all. Finally, the growth rates of certain metapopulations within the biofilm are greatly reduced, which can protract selective sweeps thus allowing clonal interference and the persistence of genetic variants. These characteristics lead to the emergence and maintenance of genetic diversity within biofilms and create metapopulations that, by chance, may survive better if environmental conditions change. Our observations should apply equally well to other spatially structured, non-biofilm environments such as soils and sediments.

It has been proposed that adaptive evolution in bacterial (haploid) populations is driven by strong selection and weak mutation (SSWM; (Orr, 2002; Joyce et al., 2008). In this scenario, there is a finite number of rare mutations that can only be beneficial or deleterious. Selection is strong enough that the entire population is essentially composed of only one genotype and adaption occurs through the stepwise fixation of single mutations. In many ways, this scenario captures our understanding of how evolution operates in microbial populations residing in homogeneous, unstructured environments. Here we propose that evolution in biofilms and other spatially structured environments is more akin to weak-selection and strong mutation; a conceptual model that we refer to as "anti-SWWM." Perhaps the ideas put forward here can be used in the future to guide the development of new models of bacterial evolution in spatially structured environments and a better understanding of the extraordinary diversity found in the microbial world.

Chapter 3: The relationship between spatial structure and the maintenance of diversity in microbial populations

3.1 Abstract

Spatial structure is pervasive in the microbial world, yet we know relatively little about how it influences the evolution of microbial populations. It is thought that spatial structure limits the scale at which competitive interactions can occur thereby protracting selective sweeps. The degree of evolutionary independence provided by the protraction of selective sweeps can allow the population to simultaneously explore multiple evolutionary paths. Sewall Wright thought this was an important means by which asexual microbial populations might overcome sign epistasis. But how structured must a microbial population be before these effects are realized? In this study we used a combination of experimental and simulation studies to explore the relationship between spatial structure and the maintenance of diversity in microbial populations. The degree of spatial structure experienced by evolving *Escherichia coli* metapopulations was manipulated by varying the rate of migration between sets of 12 independent subpopulations. Each of the 12 subpopulations were inoculated with equal proportions of two competing genotypes that initially had equal fitness. It was expected that, in the absence of structure, one of the two genotypes would acquire a beneficial mutation and sweep to higher frequency (major genotype) while also purging the other (minor genotype) from the metapopulation. After 150 generations of evolution, the frequencies of these two genotypes were again determined for each of the 12 subpopulations. Our results showed that structured metapopulations maintained higher frequencies of the minor genotype and that the frequency of this genotype decreased exponentially as the rate migration between the subpopulations was increased. Beyond a certain rate of migration—here when 2,000 individuals were migrated between the subpopulations per day—the minor' genotype was rapidly expunged from the metapopulation. However, at migration rates less than this threshold, spatial structure was able to protract selective sweeps, allowing for the persistence of both genotypes. These results demonstrate that structured microbial populations can evolve along multiple evolutionary trajectories even when migration rates between the subpopulations are relatively high.

3.2 Introduction

The majority of microbial populations live and evolve in environments that have some degree of spatial structure (Franklin and Mills, 2007). By that we mean the individuals that make up a population are distributed across physical space with each having a distinct location. This includes the bacteria that inhabit porous environments likes soils and sediments (Becker et al., 2006; Deschesne et al., 2007) as well as those that dwell in biofilms and microbial mats (Battin et al., 2007). Even populations that simply reside on surfaces like countertops (Flores et al., 2013) and epithelial tissues (Poulsen et al., 1994; Schreiber et al., 2004; Monier and Lindow, 2005) experience some degree of spatial structure. Despite the

pervasiveness of structured populations in nature, most studies and models of microbial evolution focus only on unstructured populations, like those found in well-mixed liquid cultures (Kawecki et al., 2012; Kassen, 2014). This disparity between the microbial population found in nature and those commonly used in laboratory studies is important because spatial structure has been shown to have a profound impact on the evolution of microbial populations (Rainey and Travisano, 1998; Kerr et al., 2002; Habets et al., 2006, 2007; Perfeito et al., 2008; Ponciano et al., 2009). A simple demonstration of this can be seen in Rainey and Travisano (1998). Here, the authors showed that simply growing *Pseudomonas flourescens* in liquid media without shaking could lead to the rapid evolution of multiple ecotypes—an outcome that was not observed when the strain was grown in an unstructured environment (Rainey and Travisano, 1998). If we are to comprehend the complexities of microbial evolution in the natural environment, we must develop an understanding of the impact of spatial structure on evolutionary processes.

Theory predicts that spatial structure has a direct impact on natural selection. In the absence of structure, competition between individuals occurs on a global scale, allowing selection to enact rapid, population-wide changes in the frequencies of genotypes (Dykhuizen and Hartl, 1983). In unstructured populations, the reproductive success of an individual is determined by its fitness relative to that of the entire population average. Spatial structure, on the other hand, limits competition to a local scale (Tillman, 1994; Kerr et al., 2002; Kim et al., 2014; Nadell et al., 2016), wherein each individual competes only against the subset of the population that is in close proximity to it. Beneficial mutations that occur in a structured population must sweep through physical space in addition to sweeping in frequency through the population. This can protract selective sweeps (Habets et al., 2007), thereby facilitating the maintenance of diversity (Amarasekare, 2003; Wang et al., 2005). Although this effect is often cited as fact in the literature (Steenackers et al., 2016), there are relatively few demonstrations of it. Those that do exist either rely only on comparisons between structured and unstructured populations (Habets et al., 2006; Nahum et al., 2015), or do not address how spatial structure impacts diversity (Kryazhimskiy et al., 2012). Here, we directly characterized the relationship between the degree of spatial structure a microbial population experiences as it evolves and the maintenance of diversity.

Adaption by microbial populations has canonically been viewed as occurring through the sequential fixation of beneficial mutations, each providing some incremental increase fitness (i.e. under Strong Selection Weak Mutation, (Gillespie, 1983)). In the absence of sex and recombination, each fixation event is associated with a collapse of genetic diversity, as all loci on the chromosome are linked (Koch, 1974; Levin, 1981; Dykhuizen and Hartl, 1983). Under this scenario, the range of accessible genotypes is largely restricted to only those that are a single mutational step away from the dominant genotype, and adaptation proceeds along a single evolutionary trajectory. This mode of adaptation is efficient if sign epistasis is rare and the fitness landscape is correspondingly smooth. However, if sign epistasis is common, as many empirical studies have demonstrated (Carvalho et al., 2011; Kvitek and Sherlock, 2011; Chou et al., 2014), and the fitness landscape correspondingly rough, asexual populations may become trapped at suboptimal fitness peaks surrounded by regions of lower fitness (Weinreich et al., 2005). Sewall Wright proposed, in his shifting balance theory (SBT), that evolution in a structured environment may

provide a mechanism by which populations might overcome epistasis (Wright, 1932). By splitting a single population in into several subpopulations connected by gene flow, spatial structure may allow the population to simultaneously explore several regions of the fitness landscape, increasing the range of easily accessible genotypes. This would allow structured populations to eventually reach higher adaptive peaks in a rough fitness landscape (Nahum et al., 2015). Although spatial structure is an integral component of SBT, it is not clear what level of gene flow is needed to allow subpopulations to evolve semi-independently.

Here we used empirical studies and model simulations to characterize the effect of spatial structure on the maintenance of diversity in microbial populations. On the empirical side, we evolved replicate *Escherichia coli* metapopulations under varying degrees of spatial structure. Each metapopulation consisted of a set of 12 subpopulations and structure was introduced by manipulating the migration rate between them (Kryazhimskiy et al., 2012). The subpopulations were founded with equal proportions of two differentiable genotypes. One was the wild type *E. coli* K12 MG1655 strain and the other was a spontaneous cycloserine resistant mutant derived from the wild type *E. coli* strain. It was expected that one of these two genotypes would acquire a beneficial mutation and, in the unstructured metapopulations, outcompete the other genotype, leading to its eventual loss from the metapopulation. However, we hypothesized that spatial structure would protract selective sweeps, allowing for the persistence of both genotypes. Additionally, we predicted that, in the structured metapopulations, there would be some variation in the frequencies of these two genotypes among the 12 subpopulations. To further detail the relationship between spatial structure and the maintenance of genetic diversity, we constructed a branching process model for microbial evolution that allowed us to test our hypothesis under a broader parameter set.

3.3 Methods

3.3.1 Microbial strains and conditions

For all experiments, *Escherichia coli* K12 MG1655 was grown at 37°C in a glucose limited (12.2 mM) M9 salts minimal media, supplemented with Wolfe's vitamins and trace elements (Wolin et al., 1963). Agar plates were prepared by adding 1.5% agar to the media. To generate two genotypes of *E. coli* K12 MG1655 that could be distinguished by plating on selective media we isolated a cycloserine resistant mutant from the ancestral strain. Resistance to this antibiotic is caused by mutations in the cycA alanine transporter gene and do not generally affect fitness (data not shown, (Cseh and Umenhoffer, 2006)). Approximately 10^7 colony forming units (CFUs) of the cycloserine sensitive ancestor were spread on agar plates containing 40μM cycloserine. After 48 hours of incubation, a single cycloserine resistant colony was randomly selected, restreaked on selective media, and then inoculated into 5mL of fresh media. The resulting culture was incubated overnight at 37°C and then archived at -80°C in 20% glycerol.

3.3.2 Experimental evolution

Each *E. coli* metapopulation consisted of 12 subpopulations that were grown in 200μL of the glucose limited media each in a separate well of a 96 well plate. We tested four different degrees of spatial structure: no structure, low structure, high structure, and complete structure. In the no structure metapopulations, the twelve subpopulations are mixed together before being subcultured, while in the low and high structure metapopulations, only genotypes with a frequency greater than 10% (low structure), or equal to 100% (high structure), had a high probability of migrating and surviving the subsequent bottleneck. The entire experiment was replicated three times (see Appendix A for a more in depth protocol).

To start the evolution experiment, the cycloserine sensitive and cycloserine resistant *E. coli* genotypes were separately inoculated into 5mL of media and incubated overnight at 37°C, with shaking at 185 rpm. The two resultant cultures were then serially diluted in 0.85% saline solution and spread plate on either antibiotic free media (sensitive genotype) or media containing 40muM cycloserine (resistant genotype). The plates were incubated at 37°C for 24 hours, after which twelve colonies of each genotype were randomly selected and resuspended in media. The resulting cell suspensions were then standardized to the same optical density. Pairs of resistant and sensitive genotypes were combined together creating twelve new cell suspensions, each of which contained approximately equal proportions of the sensitive and resistant genotypes. These cell suspension mixtures were then diluted and plated on antibiotic free and antibiotic containing media to determine the starting frequencies of the two genotypes in each subpopulation. The twelve cell suspension mixtures where then diluted 1:10 in fresh media and then 15μL was inoculated into 185μL of fresh media in four wells of a 96 well plate, one each for the four metapopulation types (complete structure, high structure, low structure, and no structure).

The degree of spatial structure experienced by the evolving *E. coli* metapopulations was manipulated by altering the migration rates between the 12 distinct subpopulations. This is similar to an approach used previously (Kryazhimskiy et al., 2012). All metapopulations were subcultured into fresh media daily (1:128 dilution, providing 7 generations per day), however the protocol differed depending on the degree of spatial structure in question. For the completely structured metapopulations no migration occurred between the subpopulations. In the case of the no structure metapopulations, the 12 subpopulations were combined together in equal volumes, and this mixture was used as the inoculum for each of the 12 subpopulations. In the low and high structure metapopulations, the subpopulations were first subcultured into fresh media. Then the 12 subpopulations were combined together in equal volumes and the resulting mixture was diluted in fresh media by a factor of 10^{-3} (low structure) or 10^{-5} (high structure) and then 10μL of the mixture was added to each subpopulation. This provides a migration rate of 2,000 individuals per day between any two subpopulations in the low structure and 20 individuals per day in the high structure.

The evolution experiment was carried out for 21 days (150 generations) with daily serial transfers and migration events. At seven day intervals, the optical densities (600nm) of the individual subpopulations were tracked at 15-minute intervals over 24 hours using a BioTek PowerWave HT (BioTek Labs, Suwa-

nee GA). Growth kinetic parameters were estimated from the resulting data using the R package grofit (Kahm et al., 2010). At the end of the experiment, the 12 subpopulations were again plated on antibiotic free and antibiotic containing media to determine the frequency of the two genotypes. The frequency of the genotype that was least abundant in each metapopulation was determined and is hereafter referred to as the minor genotype frequency of MGF.

3.3.3 Simulation

To further characterize the effect of spatial structure on the maintenance of diversity in microbial populations, we used a simple branching process simulation (python code available at: `github.com/michaelfrance/simulatedmetapopulations/`). The experimental setup was recreated computationally. Individual genotypes were tracked through evolutionary time with the following associated metadata: a census count, a fitness, which of the 12 subpopulations they are currently located in, which subpopulation they originated from, which genetic background they have, and an ancestry barcode. Mutation, reproduction with selection, migration, and serial transfer were all modeled as Poisson processes with lambda equal to the product of the lineage census size and the beneficial mutation rate, the fitness relative to the local subpopulation average, the migration rate, or the transfer probability, respectively. We modeled the beneficial mutation rate (U_b), and the selective effects of new mutations were assigned as a random draw from an exponential distribution with β=0.1 (Orr, 2003; Kassen and Bataillon, 2006). We tested 6 values for the degree of spatial structure (migration rates of 0, 2, 20, 200, 2,000, and 250,000 individuals migrated between subpopulations per day) and 4 values for U_b (10^{-7},10^{-8},10^{-9},10^{-10} beneficial mutations per generation). Subpopulations were initialized with equal proportions of the two genotypes. The simulated metapopulations were then allowed to evolve for 1000 generations with regular tracking of average fitness of the metapopulation, the frequency of the minor genotype, and the variance in these two traits across the 12 subpopulations. Simulations were conducted with 1000 replicates for each of the 24 parameters sets. We found our results to be generally insensitive to U_b with the only difference being that the dynamics were faster in metapopulations that evolved with a larger U_b. For this reason, we only report results from simulation with U_b=10^{-7} beneficial mutations per generation.

3.3.4 Statistical analysis

Statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC). The specific growth rate estimates and the minor genotype frequencies for each the individual subpopulations were modeled using similar approaches. In the first, a linear mixed model was used with the growth rate estimates as the response variable and the timepoint, the degree of spatial structure, and their interactions as predictor variables, as well as the experimental replicate as a random effect. To compare the variance around the metapopulation average growth rate estimates we analyzed the squared residuals output from the previous model, excluding the data from day zero. A general linear model was employed again with the squared residuals as the response variable and time, the degree of spatial structure, and their

interaction as predictor variables, with the replicate included as a random effect. All pairwise post hoc comparisons were tested, resulting p-values were corrected using the Benjamini and Hochberg approach (Benjamini and Hochberg, 1995). Similar models were used to analyze the minor genotype frequency data, except that time was dropped from the predictor variables because only the final timepoint was analyzed. Again, post hoc comparisons were constructed and corrected for multiple comparisons using the same approach. Simulation data was analyzed using three statistical models. In the first, relative fitness was the response variable with the degree of spatial structure and time as categorical predictor variables. While the second and third models included the minor genotype frequency at the 1000th generation and the variance in the minor genotype frequency across the subpopulations as response variables, respectively, with the degree of spatial structure as a categorical predictor variable.

3.4 Results

To characterize the relationship between spatial structure and the maintenance of diversity in microbial populations, we evolved replicate *E. coli* metapopulations under varying degrees of spatial structure (Figure 3.1). We tested four degrees of spatial structure: complete structure (no migration), high structure (20 individuals migrated between each subpopulation per day), low structure (2,000 individuals migrated per day), and no structure (250,000 individuals migrated per day). During the evolution experiment, we monitored changes in the growth kinetics of the 12 subpopulations. We found that the average growth rate of the metapopulations increased by approximately 30% over the 150 generations, with the bulk of the change occurring between generations 50 and 100 (Figure 3.2). There were no substantial differences in the overall change in growth rate across the tested degrees of spatial structure. Concordantly, the most influential predictor variable in the analysis of these data was time($F_{3,552=1319.4}$), while the degree of spatial structure had only a modest impact on the observed change in specific growth rate ($F_{3,552}$=16.56). This result is consistent with at least one selective sweep having occurred in all metapopulations over this timescale.

We hypothesized that metapopulations that evolved in environments with more spatial structure would have more variation among their subpopulations. One approach we used to test this hypothesis was to compare the squared residuals from our model of the relative growth rate estimates. These values represent the departure of the individual subpopulation estimates from their metapopulation average. Convergent evolution is likely to obscure some of the variation among the subpopulations in this analysis because all subpopulations evolved an increased growth rate during the evolution experiment. Nevertheless, we find that metapopulations that evolved in environments with a higher degree of spatial structure had more variance in their relative growth rate ($F_{3,552}$=7.05, $p<0.0001$;Figure 3.3A). The variance in the complete and high structure metapopulations were similar and greater than that in the low and no structure metapopulations. We also identified a time dependence in our variance estimates ($F_{3,552}$=7.97, $p<0.0001$; Figure 3.3B), with the squared residuals peaking at the 100th generation time point. It is at this timepoint that the increase in the average growth rates of the metapopulations was first observed.

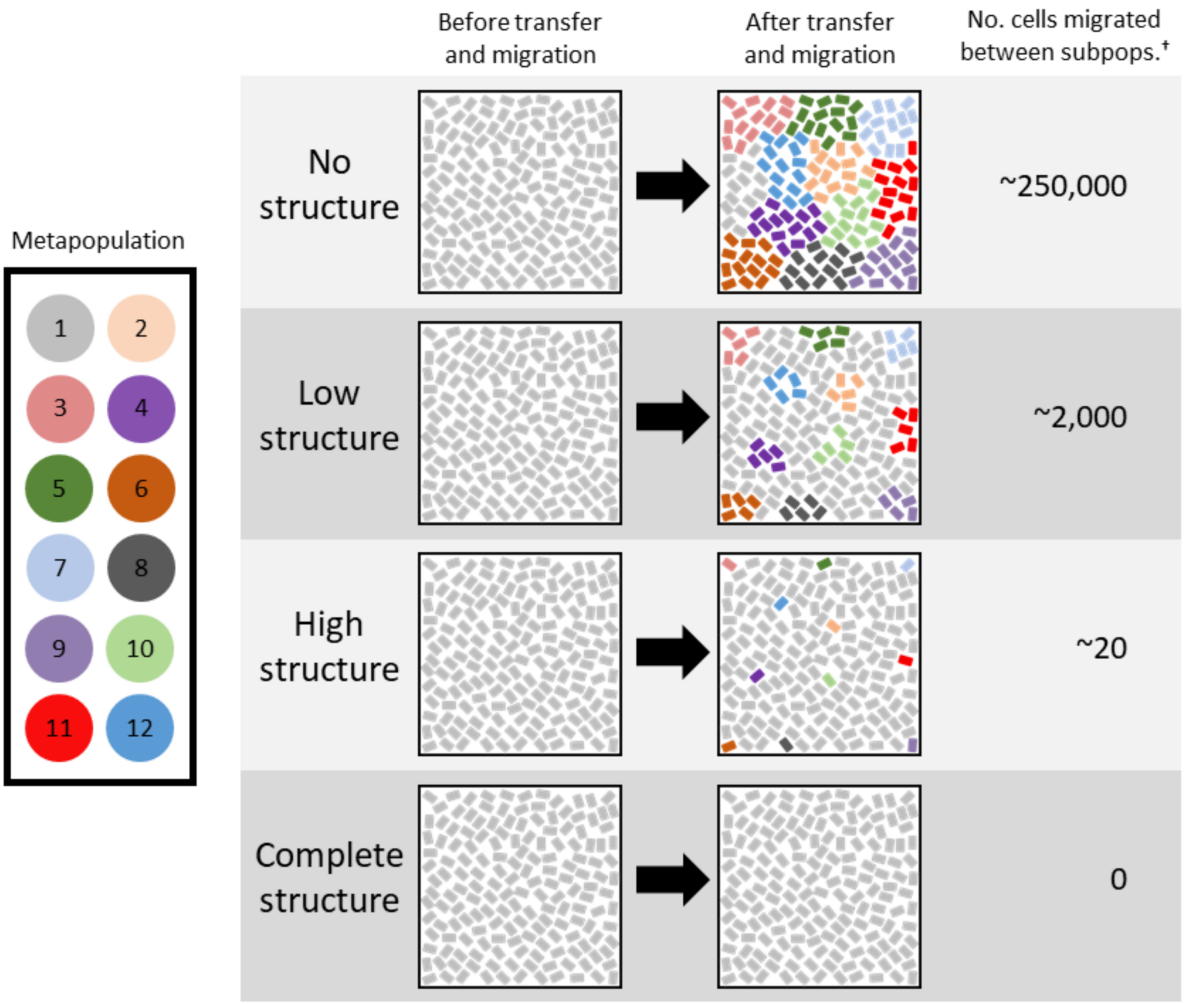

Figure 3.1: We manipulated the degree of spatial structure experienced by evolving *E. coli* metapopulations by altering the migration rate among a set of 12 distinct subpopulations. In all but the completely structured metapopulations, migration occurred between each of the 12 subpopulations daily (every 7 generations). The rate of migration was controlled by combining equal proportions of the 12 subpopulations and then diluting the resulting mixture to provide the desired number of migrants per subpopulation. The appropriately diluted migrant mixture was then added to each subpopulation. † Numbers give an approximation for the number of CFUs transferred between any two subpopulations per day.

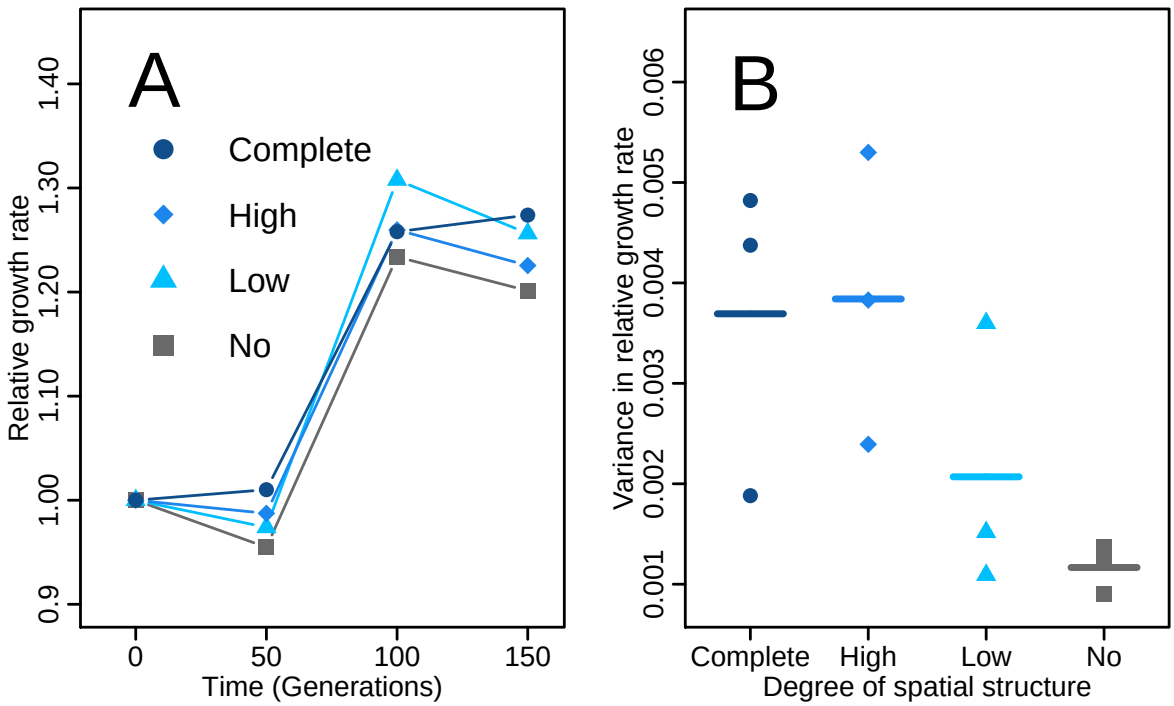

Figure 3.2: Change in the average growth rate over 150 generations of evolution for metapopulations that experienced: no structure (squares, grey), low structure (diamonds, light blue), high structure (diamonds, blue), and complete structure (circles, dark blue). Growth rate values are reported relative to estimates for the metapopulation growth rate at the start of the evolution experiment.

To provide a marker for genetic diversity we inoculated each subpopulation with equal proportions of two differentiable genotypes that initially had the same fitness: one that was cycloserine sensitive and one that was cycloserine resistant. In the absence of structure, it was expected that a beneficial mutation would occur in one of the two genetic backgrounds and that genetic background would sweep to fixation, leading to the loss of the other. Conversely, we hypothesized that spatial structure might facilitate the maintenance of both genotypes in the metapopulations, at least over the time period tested. We defined the minor genotype frequency (MGF) as the average relative abundance of the lesser of the two genotypes across the entire metapopulation. We found that metapopulations that evolved with higher degrees of spatial structure had corresponding greater minor genotype frequencies after the 150 generations of evolution ($F_{3,132}$=8.66, p<0.0001;Figure 3.4A). The complete and high structure metapopulations had minor genotype frequencies of 0.40 and 0.27, while the low and no structure metapopulations had frequencies of 0.14 and 0.10. Post hoc comparisons revealed that both the complete and high structure metapopulation MGF estimates were significantly greater than the no structure estimate (Figure 3.4A). This result is consistent with our hypothesis that spatial structure promotes the maintenance of diversity.

We next asked whether metapopulations that evolved in environments that had a higher spatial structure, also had more variation in MGF among their twelve subpopulations. If the set of subpopulations are evolving semi-independently, it was expcted that they would have some variation in their MGF. To test this hypothesis, we analyzed the squared residuals from the statistical model that was used to analyze the average MGF in each metapopulation. The squared residuals measured the difference between the frequency of the minor genotype in each subpopulation and their corresponding metapopulation average MGF. We found that metapopulations that evolved in environments that had more structure, also had a greater variance in the frequency of the minor genotype across the 12 subpopulations (Figure 3.4B). Variance in the MGF was greatest for the complete structure metapopulations, followed by the high structure, with the low and no structure being equivalent. These results are consistent with there being some degree of evolutionary independence among the subpopulations that make up the high structure metapopulations but not those that evolved in the low structure environment.

It was not clear if genetic diversity would vary with spatial structure in a discrete manner, i.e. populations either look like they evolved in environments that did or did not have spatial structure, or if it would instead vary continuously as a function of the degree of spatial structure. To answer this question, we employed a branching processes simulation for microbial evolution. The simulation was constructed to match the experimental setup: each simulated metapopulation included 12 subpopulations and the degree of spatial structure was manipulated by altering the rate of migration between the subpopulations. Six different degrees of spatial structure were tested with migrations rates between subpopulations of: 0, 2, 20, 200, 2,000, and 250,000 individuals per day. We simulated 1000 generations of evolution, with 1000 replicates for each parameter set. During this time period, at least one selective sweep occurred in all of the simulated metapopulations. However, we did find a negative correlation between the degree of spatial structure and that rate at which these metapopulations adapt (Figure 3.5), which is likely caused

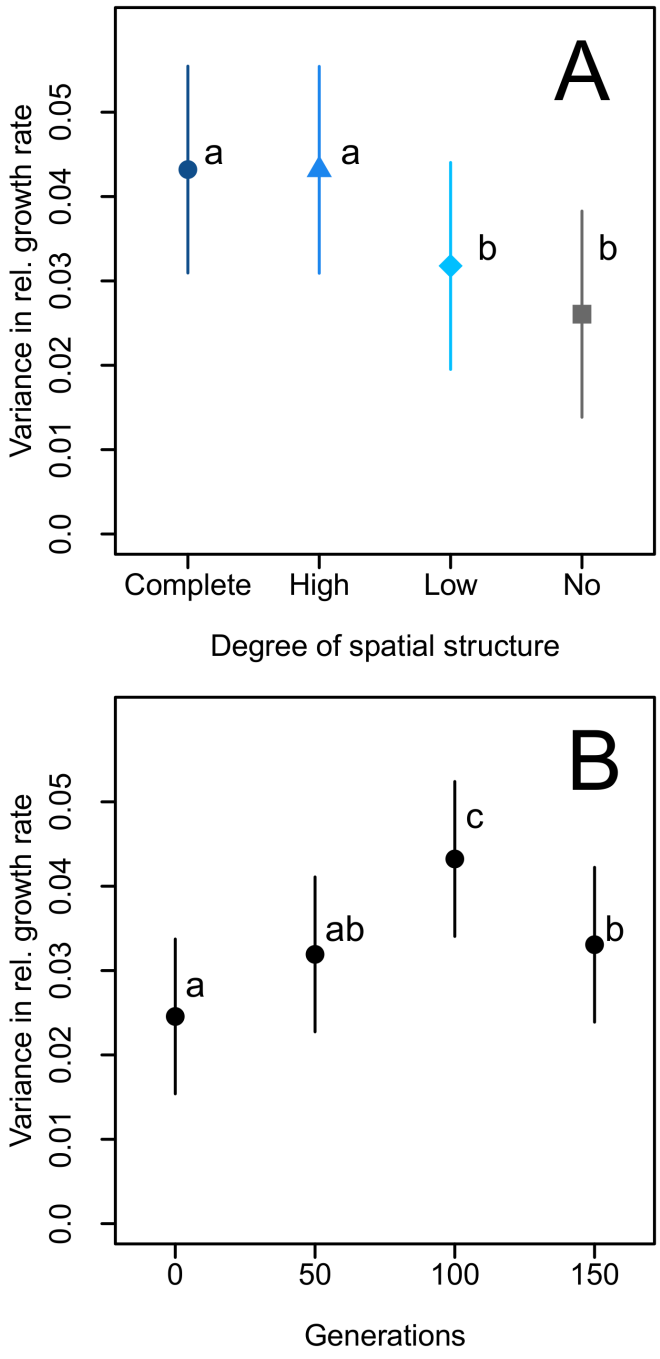

Figure 3.3: Differences in the variation in relative growth rates among the 12 subpopulations. Metapopulations that experienced higher degrees of spatial structure had corresponding greater variation in growth rates among their constituent subpopulations (A). Independent of the degree of spatial structure, there were also differences in variation dependent on time (B). Variation in growth rates peaked at 100 generations, corresponding to the observed change in the metapopulation average growth rate. Points that are labeled with the same letter(s) do not differ significantly ($p<0.05$), after adjusting for multiple comparisons using the Benjimini and Hochberg procedure.

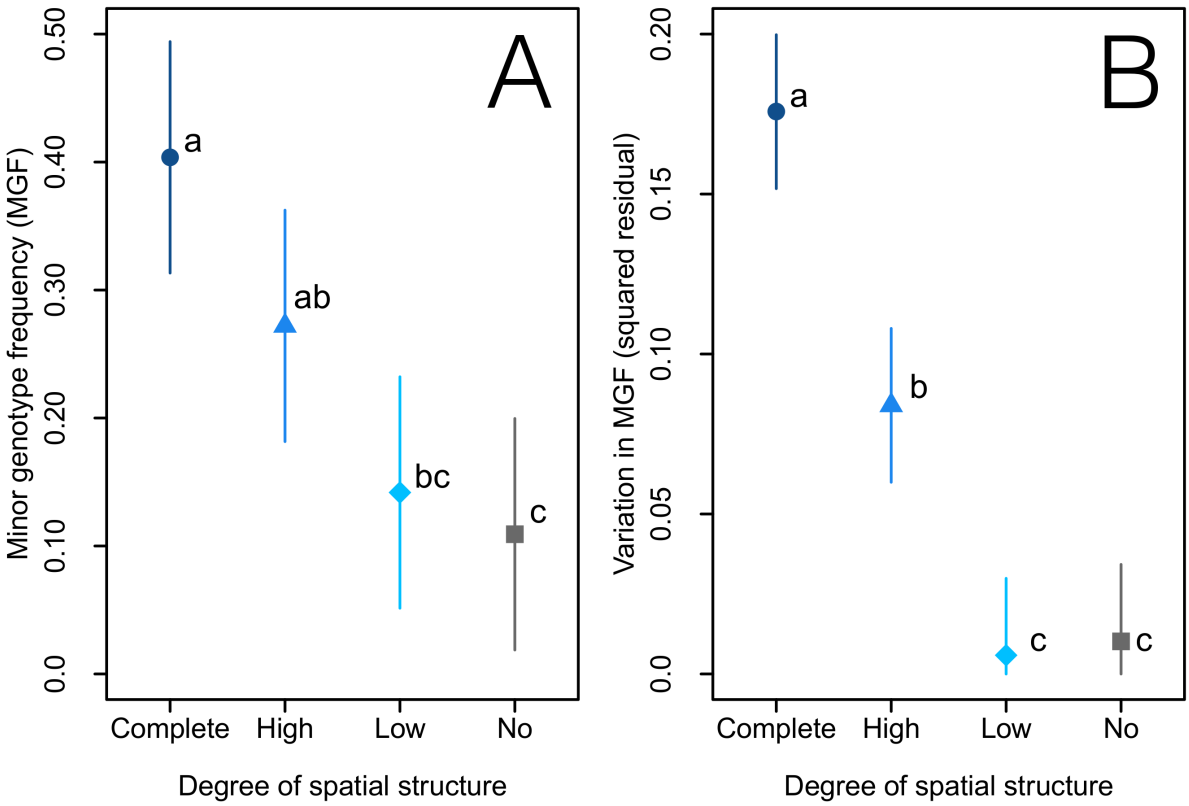

Figure 3.4: (A) Minor genotype frequency at generation 150 in metapopulations experiencing different degrees of spatial structure. (B) Variance (squared residuals) in minor genotype frequency across the 12 subpopulations that make up each individual metapopulation. Each point is the average of three replicates. Points that are labeled with the same letter(s) do not differ significantly ($p<0.05$), after adjusting for multiple comparisons using the Benjimini and Hochberg procedure.

by the reduction in effective population size that results from spatial structure.

To investigate the relationship between the degree of spatial structure and the maintenance of genetic diversity, we compared the frequency of the minor genotype at the 1000th generation across the six different degrees of spatial structure. The distribution of MGFs for the 1000 replicate simulation runs under varying degrees of spatial structure can be found in Figure 3.6. We recovered the result that structured populations have higher MGFs, consistent with spatial structure being correlated with genetic diversity ($F_{5,5995}$=978.83, p<0.0001; Figure 3.7A). Based on these data, it appears that the relationship between spatial structure and genetic diversity resembles an exponential decay function. . Between migration rates of 0 and 2,000 individuals per day, there is an exponential decline in the average minor genotype frequency at the cessation of the simulation. However, there is relatively little differences between metapopulations that experienced migration rates at or above 2,000 individuals per day, with all of them closely resembling metapopulations that evolved in the absence of structure. This result suggests a "threshold" relationship between spatial structure and the maintenance of diversity. In highly structured populations, small changes in migration rates between the subpopulations led to large differences in the observed MGF at the end of the evolution. However, beyond a certain point of migration, here 2,000 individuals per day, the metapopulations evolved as if the environment did have spatial structure.

Our experimental data also indicated that spatial structure can allow for the semi-independent evolution of the subpopulations. This is critical component of Sewall Wright's shifting balance theory. Here we use data from our simulation to determine how much migration can occur between structured subpopulations before they lose their independence. The experimental data suggested that structured metapopulations experiencing migration rates of 20 but not 2,000 individuals per day, evolved with some degree of independence. Our analysis of the simulation data mirrors this result and provides more insight into the maximum rate of migration that a metapopulation can experience while still maintaining some degree of evolutionary independence ($F_{5,5995}$=1766.14, p<0.0001; Figure 3.7B). These data indicate that there is an exponential decay relationship between the two and suggest that metapopulations experiencing a migration rate above 2,000 individuals per day, largely act as a single, unstructured population.

3.5 Discussion

Our results demonstrate that spatial structure can drive the maintenance of genetic diversity and evolutionary independence in asexual microbial populations. They further place a limit on how much migration can occur before the impact of spatial structure on microbial evolution is degraded. The majority of bacterial populations in nature live and evolve in environments that have physical structure (Franklin and Mills, 2007). These environmental bacterial populations have been shown to possess a wealth of diversity, with several strains of the same species existing in the same population (McArthur et al., 1988; Lebret et al., 2012; Lieberman et al., 2013). We argue that spatial structure plays a central role in the maintenance of this diversity.

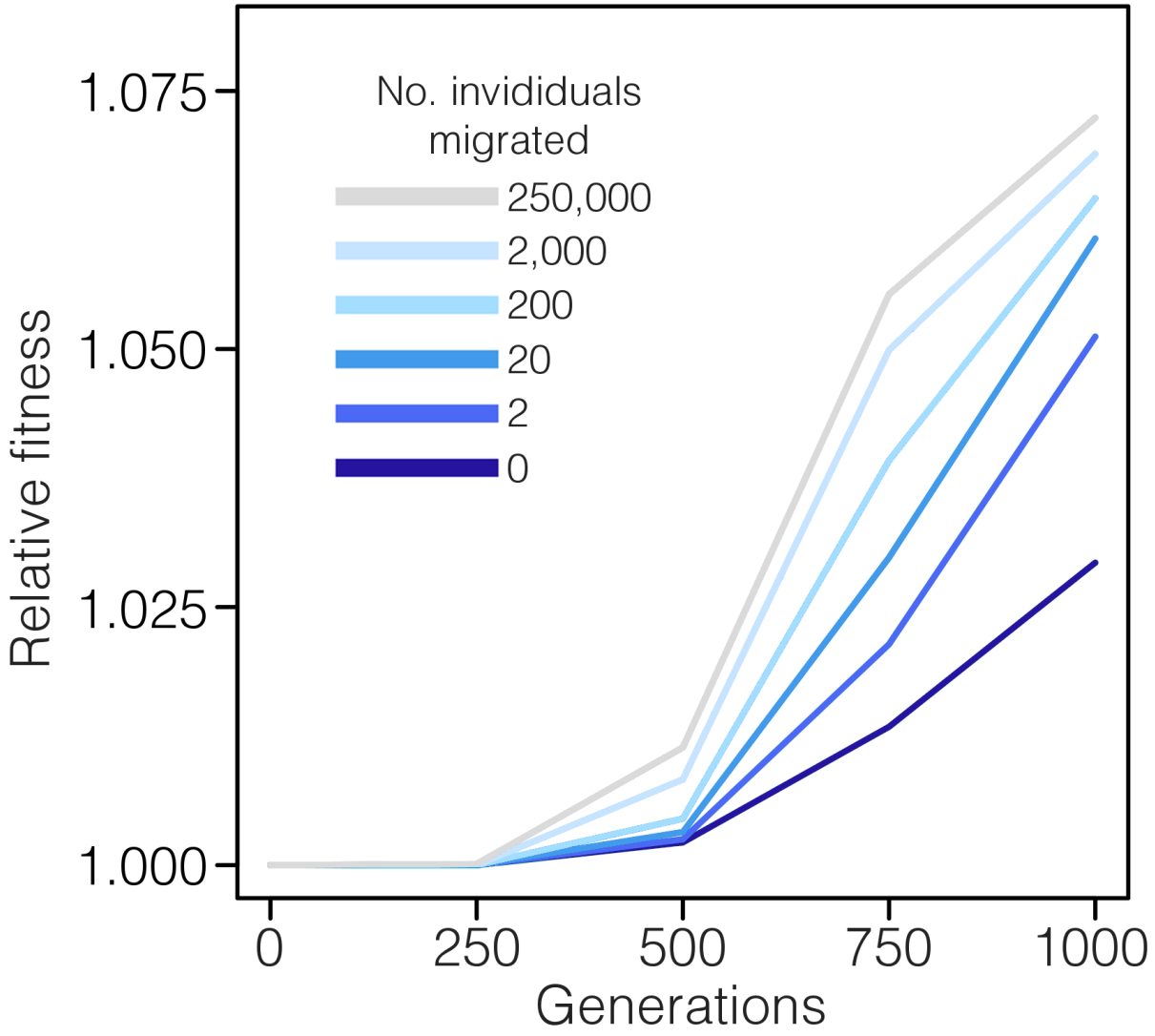

Figure 3.5: Average relative fitness of simulated metapopulations that evolved while experiencing varying degrees of spatial structure. Each line is the average result across 1000 replicate simulation runs with $U_b = 10^{-7}$ per individual per generation. This trend was recovered across all tested beneficial mutation rates ($U_b = 10^{-7}$, 10^{-8}, 10^{-9}, 10^{-10}; data not shown).

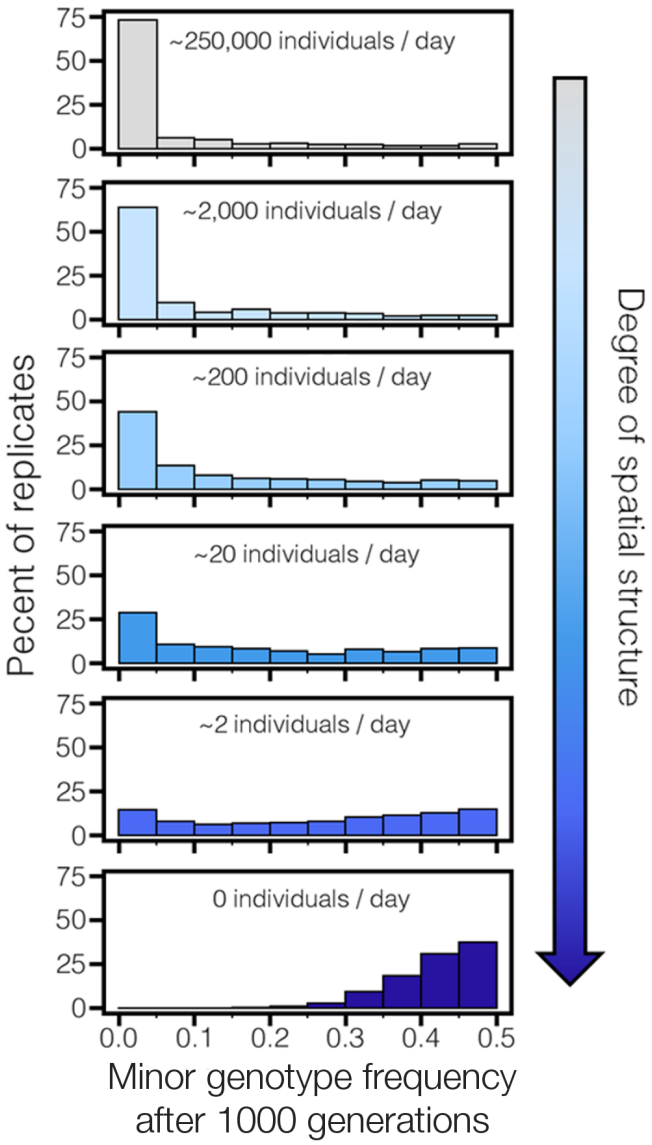

Figure 3.6: Frequency of the minor genotype after 1000 generations of evolution in simulated populations experiencing varying degrees of spatial structure. Each plot contains a histogram of the results for 1000 replicate simulation runs under each parameter set. The degree of spatial structure varies over 6 orders of magnitude from top to bottom. Results displayed are derived from simulations where $U_b = 10^{-7}$ per individual per generation. Similar patterns were observed when (U_b equaled 10^{-8}, 10^{-9}, 10^{-10} (data not shown).

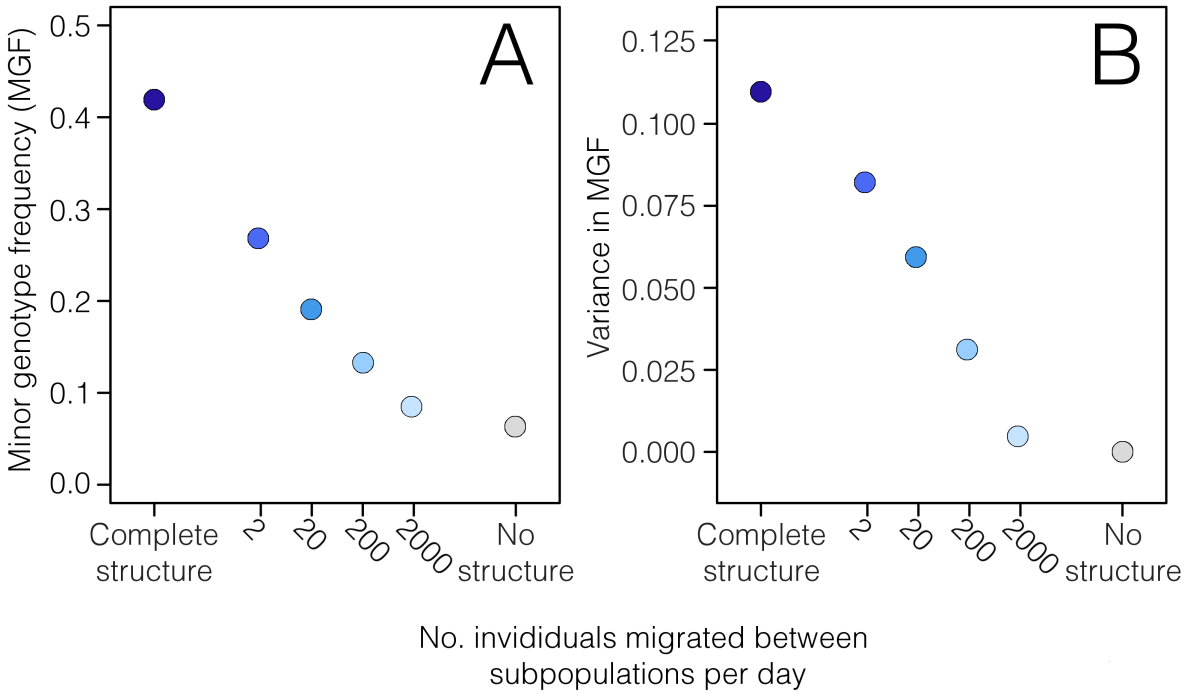

Figure 3.7: Average minor genotype frequency in metapopulations that evolved for 1000 generations while experiencing varying degrees of spatial structure (A). Variance in the frequency of the minor genotype across the 12 subpopulations (B). The results shown are derived from 1000 replicate simulation runs for each of the six tested degrees of spatial structure with $U_b = 10^{-7}$ per individual per generation. The observed trends were not sensitive to the rate of beneficial mutations.

One limitation of our study is that our implementation of spatial structure differs from that experienced by environmental microbial populations in at least two ways. The first is that the individual subpopulations that made up each metapopulation were well-mixed and structure was introduced by manipulating the migration rate between the sets of subpopulations. Selection likely operated within each subpopulation without any impact from spatial structure. In comparison, spatial structure is experienced by many environmental populations at the level of the individual. Each individual cell that makes up a biofilm has its movement restricted by the extracellular matrix and therefore only competes against the subset of cells that are physically near it (Flemming et al., 2016). It is likely the effects of spatial structure are amplified in populations that evolve under this finer degree of structure. The second difference is that migration between each of our twelve subpopulations occurred at the same rate, as if they were all equidistant from one another. In environmental populations, it is instead more likely that the rate of migration decays with the distance between the source and sink subpopulations (Hanson et al., 2012). Placing restrictions on migration between the subpopulations should slow the speed at which selective sweeps spread through the metapopulation, further enhancing the effect of spatial structure (Perfeito et al., 2008).

In his shifting balance theory, Sewall Wright proposed that spatial structure might provide a means by which asexual populations could overcome epistasis (Wright, 1932). By splitting a single population into several subpopulations connected by migration, spatial structure can allow populations to explore multiple evolutionary trajectories. Our results suggest that migration rates between subpopulations of up to 2,000 individuals per day (0.06% of the bottleneck subpopulation size) affords some degree of evolutionary independence. It is not clear how much migration microbial populations are subjected to in the environment and it certainly varies depending on the population in question. However, we speculate that most microbial populations do experience migration at a lower rate than this threshold and certainly with more irregularity than that which was experienced in this study. It may be best to view these populations not as a single genotype evolving along a single trajectory, but rather as a cloud of genotypes that are evolving along disparate trajectories. This would allow these populations to explore more of the fitness landscape and potentially reach higher adaptive peaks over the long term.

In this study, we've focused on perhaps the simplest form of diversity—two competing genotypes that start with the same fitness and at the same abundance. In the absence of structure, we've shown that one genotype rapidly subverts the other, likely due to the acquisition of a beneficial mutation. Intuitively, you might expect that any amount of migration between subpopulations would allow the genotype that acquires the beneficial mutation to spread to the surrounding subpopulations with ease. Yet, we've also shown that both genotypes are maintained at relatively high frequency in our highly structured metapopulations. We speculate that this is likely due to clonal interference (Gerrish and Lenski, 1998; Sniegowski and Gerrish, 2010). Because only 20 individuals were exchanged per day, a beneficial mutation must sweep to a high frequency before it is guaranteed to migrate between the subpopulations. This provides more time for a beneficial mutation to occur in the other genetic background, elsewhere

in the metapopulation. A fit genotype that migrates into a subpopulation that already has a resident segregating beneficial mutation must face stiffer competition. By sheltering subpopulations from beneficial mutations that occur elsewhere in the population, spatial structure can amplify the effects of clonal interference. Which, in turn, can further facilitate the maintenance of diversity in structured populations.

While the results from our experimental and simulation studies generally agree, there are some minor differences between the two. Adaptation to the glucose limited environment occurred rapidly in our experimental metapopulations—a dramatic change in the average growth rate of population was observed after 100 generations of evolution. It took our simulated metapopulations much longer to achieve this outcome, even when $U_b = 10^{-7}$ per individual per generation. Estimates for the rate of beneficial mutations in *E. coli* range from 10^{-10} to 10^{-8} per individual per generation (Imhof and Schlotterer, 2001; Rozen et al., 2002). There are at least two factors that might contribute to this disparity. The first is that, in the experiment, each subpopulation was inoculated with a mixture of two single colony suspensions. These inocula likely included some amount of genetic diversity that originated during the growth of the two colonies. Whereas the simulated metapopulations were initialized with equal proportions of two single genotypes. This initial diversity may have sped up adaptation in the experimental metapopulations (Fisher, 1930). Another possible explanation for this difference could be that the initial stage of adaption to this glucose limited environment might not be driven by point mutations, but instead by gene duplication. In general, gene duplications occur with a higher frequency than point mutations (Anderson and Roth, 1981; Elliott et al., 2013) and have been shown to be adaptive in carbon limited environments (Sonti and Roth, 1989; Brown et al., 1998; Maharjan et al., 2013). This would explain both the speed of adaptation observed in the evolved *E. coli* metapopulations, and the repeatability of the phenotypic outcome. In the three biological replicates, across the four degrees of spatial structure, all subpopulations increased their average growth rate by approximately 30% over the 150 generations. This is unlikely to have resulted from shared ancestry because each subpopulation was inoculated with an independent mixture of two single colony isolations.

To conclude, we've shown that spatial structure can facilitate the maintenance of diversity in asexual microbial populations. We've further demonstrated that the effects of spatial structure were not eroded even with modest amounts of migration. Spatial structure can also impact the speed and outcome of adaptation, if the fitness landscape is rough. We argue that spatial structure is a critical determinant of evolutionary dynamics in environmental microbial populations. that should be kept in mind as we begin to characterize the diversity within such populations.

3.6 Acknowledgements

We thank Steve Krone, Eva Top, Ben Ridenhour, Celeste Brown and Thibault Stalder for insightful discussions.

Chapter 4: Evolution in slowly growing bacterial populations

4.1 Abstract

The majority of natural microbial populations are growing markedly slower than their laboratory counterparts. In the environment, the growth of microbial populations is limited by the abundance of routinely scarce essential nutrients. We hypothesized that the rate at which a population is growing may impact its evolution. While the action of natural selection is inherently linked to the growth and death of cells, mutations can occur through mechanisms other than errors made during the replication of the chromosome. These replication-independent mutations occur through a variety of different mechanisms and should increase the per generation mutation rate experienced by slowly growing bacterial populations. We therefore hypothesized that selective sweeps would be protracted in slowly growing populations, facilitating the accumulation of diversity via replication-independent mutations. To test this hypothesis, we evolved replicate *Escherichia coli* populations in chemostats with growth rates of 2 hours (fast) and 24 hours (slow). During the evolution, we tracked the frequency of a selectively neutral spontaneous mutation in the populations as a marker for genetic diversity. Our results indicate that, across 72 generations of evolution, the slowly growing populations accumulated and maintained higher frequencies of this mutant. We further demonstrate that the slow growth populations evolved higher fitness and experienced more phenotypic change than the fast growth populations. These results indicated that a bacterial population's growth rate can influence their evolution. Because the vast majority of populations in the environment are not experiencing rapid growth, we speculate that growing slowly might play a critical role in the emergence and maintenance of bacterial genetic diversity.

4.2 Introduction

The frequency of mutants in a bacterial population is primarily influenced through the action of two evolutionary processes: mutation and natural selection. Spontaneous mutations increase the frequency of mutants in a population while natural selection generally acts to purge deleterious mutations from the population and increase the abundance of beneficial mutations (Fisher, 1930). Although both processes are often expressed on a per generation basis, they actually do not both occur on a per generation basis in asexually reproducing organisms. The ability of selection to alter genotype frequencies is inherently linked to the growth and death of individuals in the population (Darwin, 1859). The time it takes for beneficial mutations to fix in a population or for deleterious mutations to be purged from a population, scales with the generation time of the population in question. However, mutation is not similarly constrained—new mutants can be introduced into a population even in the absence of cellular division (Ryan, 1955; Kivisaar, 2010). The relative impact of selection and mutation on the frequency of mutants in a population may be different in bacterial populations that are growing slowly. We therefore hypothesized that slow growth would protract selective sweeps allowing for the accumulation and maintenance

of genetic diversity through spontaneous mutations.

The mutational processes that introduce genetic diversity into bacterial populations can be divided into two broad categories: those that result from mistakes made during the replication of the chromosome (replication-dependent) and those that result from other means (replication-independent) (Ryan, 1955; Eastman et al., 2011). Replication-independent mutations can be caused by a variety of different mechanisms including, but not limited to: the repair of damaged DNA (Kivisaar, 2010), the movement of transposable elements (Kidwell and Lisch, 2000; Frost et al., 2005), and recombination (Anderson and Roth, 1981; Bull et al., 2001). The overall rate of mutation experienced by a bacterial population is expected to be some linear combination of the replication-dependent and replication-independent rates (Eastman et al., 2011). While the number of replication-dependent mutations that occur per unit time varies with the generation time of the population, the occurrence of replication-independent mutations does not. Replication-independent mutations can instead be thought of as a "metronome" of genetic change; one that constantly introduces new mutations through time.

Under this paradigm of replication-dependent and replication-independent mutational processes, two expectations can be derived. The first is that slowly growing bacterial populations should experience a higher mutation rate when measured per generation of growth. This is because, for each generation of growth, a slowly growing population will experience more replication-independent mutations than a population which is experiencing rapid growth. The second is that rapidly growing populations will experience a higher mutation rate when it is measured per unit time (hours, days, etc.). This is because, for each unit of time, a rapidly growing population will undergo more generations of growth and therefore incur more replication-dependent mutations. However, without empirical estimates for the replication-dependent and independent mutation rates, it impossible to predict the magnitude of the difference in mutation rate experienced by slowly versus rapidly growing bacterial populations.

Growing slowly may also influence the adaptation of bacterial populations. Under the strong selection weak mutation paradigm outlined by (Gillespie, 1983, SSWM), asexual bacterial populations are expected to adapt via the sequential fixation of single beneficial mutations. However, when population sizes are large, as in the case for many bacterial populations, multiple beneficial mutations are predicted to occur simultaneously giving rise to a process termed "clonal interference" (Gerrish and Lenski, 1998; Sniegowski and Gerrish, 2010). Eventually the genotype with the greatest fitness advantage outcompetes the others, leading to the loss of all other genotypes from the population. By sampling more mutations per generation via the replication-independent processes, slowly growing populations may have greater access to beneficial mutations, leading to an exaggerated form of clonal interference. Instead of being driven by the sweep of one or a few beneficial mutations, adaptive evolution in slowly growing populations may result from the simultaneous sweeps of many beneficial mutations. Because the summed starting frequencies of such a set of beneficial mutations is higher, together they require fewer generations of evolution before their effect on mean population fitness is realized. We therefore hypothesized that the mean fitness of slowly growing populations might actually increase more rapidly, at least when measured

on a per generation basis.

The majority of bacterial populations in the environment do not routinely experience sustained rapid growth (Roszak and Colwell, 1987). Instead, such populations more typically grow slowly because their rate of growth is limited by the availability of at least one scarce essential nutrient (Church et al., 2000; Demoling et al., 2007). Many estimates for the generation times of bacteria in the environment have ranged between days and weeks (Harris and Paul, 1994; Bååth, 1998; Rousk and Bååth, 2011; Kirchman, 2016). Despite its prevalence in nature, the effect of growing slowly on the evolution of bacterial populations remains under studied. To test whether slowly growing populations harbor more genetic diversity and adapt faster than populations experiencing rapid growth, we compared the evolution of *E. coli* populations that experienced two different growth rates: fast, with a generation time of two hours and slow with a generation time of 24 hours. The populations were evolved for 72 generations which took 144 hours for the fast growth populations and 1728 hours for the slow growth populations. Our findings indicate that slowly growing populations do experience higher mutation rates when measured per generation. We further show that the slow growth populations accumulated and maintained higher frequencies of cycloserine resistant mutants. Finally, we show that our slow growth populations evolved a higher fitness than the fast growth populations. These results are consistent with our hypotheses and provide a much needed first look at how evolution proceeds when bacterial populations grow slowly.

4.3 Materials and methods

4.3.1 Bacterial strains and media

All experiments were conducted using *Escherichia coli* K12 MG1655. The culture media consisted of 1X M9 salts and 27.35 mM glycerol supplemented with Wolfe's vitamins and trace elements (Wolin et al., 1963). Agar plates were created by adding 1.5% agar to the recipe. The three ancestral strains used in the evolution experiment were independent clones derived from a freezer stock *E. coli* K12 MG1655. Spontaneous cycloserine resistant mutants (hereafter referred to as the cyc^{r} ancestor) were also derived from each of the three ancestor strains for use in the fitness assay. Briefly, 10^7 colony forming units (CFUs) of each ancestor were spread on separate agar plates containing 40 μM cycloserine, a single cycloserine resistant colony was randomly selected, restreaked on selective agar again, and then archived at -80°C in 20% glycerol.

4.3.2 Chemostat construction and operation

Chemostats were constructed similar to what has been described previously (see `http://eclf.net` for the general design). Each chemostat vessel consisted of a single 50 mL conical tube with three needles inserted through the top of the cap: one for the media inlet (25 mM, 18 G), the air inlet (150 mM, 22 G), and the air/media effluent (90 mM, 16 G). The holes in the cap created by the insertion of the needles

were sealed using aquarium silicone (Aqueon, Franklin, WI). Prior to use, silicone was also applied to the threads of the conical tube in order to create an air tight seal around the lid. All connections were made using standard luer locks and Tygon silicone tubing with a 1/32" interior diameter (Tygon, Valley Forge, Pennsylvania). The chemostat was operated as follows. Sterile media was drawn from a bottle using a peristaltic pump (Watson Marlow, Gloucestershire, UK) and pumped into the chemostat vessel via the media inlet needle. An aquarium pump (Rena Air, Charlotte, NC) was used to drive air across a 0.22 μM filter, thereby sterilizing it, and then into the vessel via the air input needle. Because the chemostat vessel was sealed, the air was forced to exit through the outlet needle, taking any excess media along with it into a waste bottle. The steady-state volume in the chemostat vessel was determined by the height of outlet needle, and in this case, was 10mL. Differential flow rates were achieved by varying the diameter of the tubing that spanned the peristaltic pump (Watson Marlow, 0.88 mM, fast growth; 0.25 mM, slow growth).

4.3.3 Experimental evolution

A total of six chemostat populations were used in this study: three that grew with a generation time of 2 hours (fast) and three that grew with a generation time of 24 hours (slow). Each of the three ancestral *E. coli* strains were used to found one population that experience fast growth and one population that experienced slow growth. Each ancestor was inoculated from the archived freezer stocks into 5 mL of media and then grown overnight at 37°C with shaking at 185 rpm. In order to return the resulting cultures to exponential phase, they were diluted 1:4 in fresh media and then incubated for an additional 4 hrs. Then 5 mL aliquots were drawn from each culture with a syringe and the injected into the appropriate chemostat via the outlet port. The flow of media was commenced at dilution rates of 0.345 hr^{-1} for the fast growth populations and 0.028 hr^{-1} for the slow growth populations. The evolution experiment was continued for a total of 72 generations (6 days of cultivation for the fast growth populations and 72 days for the slow growth populations). To minimize the effect of wall growth, approximately every other week the slowly growing populations were temporarily removed from their chemostats, the entire apparatus was autoclaved, and then the populations were reintroduced into the chemostats via the outlet port. The chemostats were sampled regularly (multiple times per day from the fast growth populations, every other day from the slow growth populations) via the outlet port and spread plated on selective and nonselective media to determine both the total population size and the frequency of cycloserine resistant mutants. At 24 generation intervals larger samples were taken ($\sim$1 mL) from the populations and archived at -80°C in 20% glycerol.

4.3.4 Determination of mutation rates

The replication-dependent rate of mutation towards cycloserine resistance experienced by actively growing cells was determined using fluctuation assays as described previously (Foster, 2006). The replication-independent rate of mutation due to mechanisms independent of the replication of the chromosome was determined by tracking the linear increase in the frequency of cycloserine resistant mutants

in stationary phase cultures. Cultures of *E. coli* K12 MG1655 were inoculated from freezer stocks into 5mL of the minimal glycerol media and then incubated until they had reached stationary phase (24 hrs at 37°C and shaking at 185 rpm). The proportion of the population resistant to cycloserine was determined at 0, 24 and 48 hours in stationary phase by spread plating on selective and nonselective media. A separate linear trendline was fit to the resulting data from each culture and the slopes of these lines were recorded as an estimate for the replication-independent mutation rate towards cycloserine resistance. This measure was expressed as mutations per colony forming unit per hour.

The mutation rate towards cycloserine resistance experienced by the fast and slow growth chemostat populations was also determined over the 72 generations of evolution. During the course of the evolution experiment, samples were taken regularly from the populations and spread plated on selective and non-selective media to determine the proportion of the population resistant to cycloserine. Trendlines were fit to sections of the resulting data that demonstrated a linear increase in the proportion of cycloserine resistant mutants and the slopes of these lines were used as estimates for the *in vivo* mutation rate for the fast and slow growth populations.

4.3.5 Characterization of Monod growth kinetics

To determine whether evolution by the fast and slow growth populations resulted in similar phenotypic changes, we characterized the Monod parameters on growth on the minimal glycerol media. These parameters (μmax and K_s) were determined for the three ancestral strains and the evolved populations at 24, 48, and 72 generations. Growth kinetics were determined using media with eight different glycerol concentrations (0 mM, 1.3 mM, 2.7 mM, 4.8 mM, 6.8 mM, 13.7 mM, 27.3 mM, and 41.1 mM) by measuring the change in optical density (600nm) at 15-minute intervals over 24 hours using a BioTek PowerWave HT (BioTek Labs, Suwanee GA). Growth kinetic parameters were then estimated by fitting the data to a logistic growth function using the nls function in R. Estimates for μmax and K_s were then derived by fitting the resulting specific growth rate estimates and their associated glycerol concentrations to the Monod equation, also with the nls function.

4.3.6 Fitness determinations

The fitness of the evolved chemostat populations were determined at 24, 48 and 72 generations as described previously (Dykhuizen and Hartl, 1983; Ziv et al., 2013). Each competition experience was conducted between an evolved population and the cyc^r ancestor. Additional competition experiments were conducted between the ancestors and cyc^r ancestors to determine if the cycloserine resistance mutations influence fitness. For each fitness assay, overnight cultures of the two strains were diluted back 1:4 in fresh media and then incubated for an additional 4 hrs at 37°C to return them to exponential phase. Each of the two resulting exponential phase cultures were then inoculated into separate chemostats as previously described. Flow was commenced at the appropriate dilution rate for 3 days to allow the populations

time to reach steady state prior to starting the competition. Then, a 5mL aliquot was drawn from each chemostat, combined together, and the mixture was inoculated into a separate chemostat. The ratio of the two populations (the evolved and the cyc^r ancestor) was then tracked over 12 generations for the fast growth populations and 3 generations for the slow growth populations. The difference in growth rate between the two populations was calculated as the linear relationship between ln(evolved/cyc^r ancestor) and time, in generations. These estimates were then standardized by subtracting the measured difference in growth rates between the associated ancestor and its cyc^r derivative. Values are reported as relative fitness as calculated by the difference between the evolved and ancestral divided by ancestral. All competition experiments were performed three times.

4.3.7 Statistical analysis

The following statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC). To compare the difference in the accumulation of cycloserine resistant mutants between the slow and fast growth populations, we first log transformed the data. Doing so minimized the effect of the large swings in the abundance of cycloserine resistant mutants that were observed in the slow growth chemostats towards the end of the evolution experiment. The log transformed proportions of cycloserine resistant mutants were analyzed using a repeated measures model with growth rate, time (in generations), and their interaction as predictor variables and an autoregressive covariance structure for observations from the same population. To determine whether the fast and slow growth populations differed in rate of adaptation we modeled our estimates of relative fitness using a linear mixed model with growth rate, time (in generations), and their interaction as predictor variables and the ancestral genotype as a random variable. The Monod parameter estimates were analyzed using two similar models except with our estimates for μmax and K_s as response variables.

4.4 Results

We hypothesized that (a) slowly growing bacterial populations would accumulate more genetic diversity and (b) adapt faster than rapidly growing populations. To test these two hypotheses, we contrasted the evolution of *E. coli* populations in chemostats that experienced either fast (generation time of 2 hours) or slow (generation time of 24 hours) growth for 72 generations. Throughout the text we express our results in terms of generations of evolution rather than hours because this is the timescale on which selection operates.

4.4.1 Slow growth populations accumulate higher proportions of cycloserine resistant mutants

We asked whether slowly growing populations accumulated and maintain more genetic diversity than rapidly growing populations. To answer this question, we tracked the proportion of each population that

was resistant to the antibiotic cycloserine. The mutations responsible for cycloserine resistance do not affect fitness (data not shown, (Cseh and Umenhoffer, 2006)) and the frequency of these mutants was used as a selectable marker for diversity. Previous studies have shown that, in chemostats, the frequency of spontaneously generated, neutral mutations is determined through the action of two processes: ongoing spontaneous mutations and selection (Kubitschek, 1974; Dykhuizen and Hartl, 1983). Ongoing spontaneous mutations were expected to drive linear increases in the frequency of such mutants while selective sweeps of beneficial mutations in other genetic backgrounds were expected to purge even neutral diversity from the populations. The timing and magnitude of the effect of these selective sweeps are both inherently stochastic, and this introduce substantial variation among replicate populations (Figure 4.1). Nevertheless we observed the expected linear increase in the frequency of cycloserine resistant mutants in both the fast and slow growth populations. Decreases in the frequency of these mutants were also observed in both population types, although this effect was more pronounced in the slow growth populations (Figure 4.1).

We showed that during the evolution, the slow growth populations accumulated higher proportions of cycloserine resistant mutants than the fast growth populations (Figure 4.1). While the cycloserine resistant mutants never surpassed a frequency of approximately 1 in 10^{-5} in the fast growth populations, the slow growth populations reached this frequency after only 24 generations of evolution. The slow growth populations further maintained higher frequencies of these mutants throughout the majority of the 72 generations of evolution (Figure 4.1). We log transformed the frequencies of cycloserine resistant mutants and then determined whether the linear relationship between the transformed values and the generations of evolution, differed between the slow and fast growth populations. Our analysis showed that, on average, the slopes of these trajectories were positive for the slow growth populations and slightly negative for the fast growth populations (0.014 vs. -0.005, $F_{1,24.6}<0.001$). This indicates that the slow growth populations exhibited increasing proportions of cycloserine resistant variant across the 72 generations of evolution, while the fast growth populations did not. This result is consistent with our hypothesis that slowly growing populations would accumulate and maintain more genetic diversity than rapidly growing populations.

4.4.2 Estimated rates of mutation experienced by the populations match the model that accounts for replication-independent mutations

We next asked if the observed difference in the frequency of cycloserine resistant mutants between our fast and slow growth populations could be explained by the replication-independent mutation rate. To answer this question, we measured the replication-dependent and independent rates of mutations towards cycloserine resistance and compared them to our estimates for the mutation rates experienced by the slow and fast growth populations during the evolution experiment. The replication-dependent mutation rate was estimated using a standard fluctuation test (Foster, 2006) and was found to be $1.3\text{x}10^{-7}$ per CFU per generation. The replication-independent mutation rate was estimated by measuring the linear increase in cycloserine resistant variants in stationary phase batch cultures over 48 hrs and was found to be $3.3\text{x}10^{-8}$

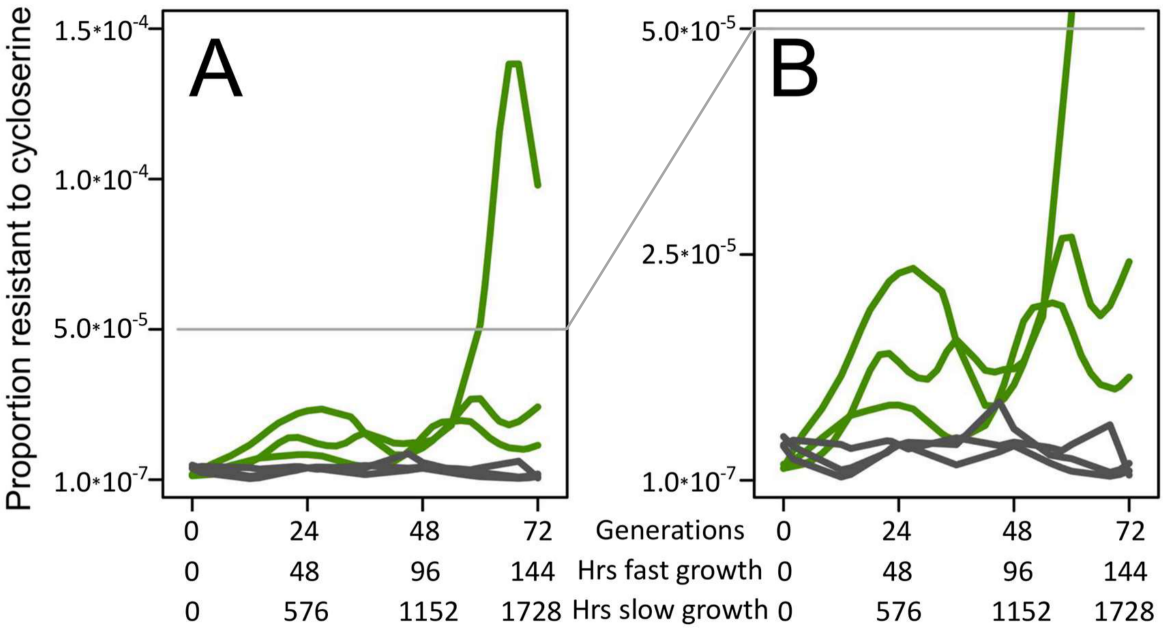

Figure 4.1: Proportion of the population that is resistant to the antibiotic cycloserine in slow (green) and fast (grey) growth populations. Spread plating on selective and nonselective media was used to track the frequency of these mutants in the populations over the 72 generations of evolution. Each line summarizes the data from a wholly independent chemostat with a spline fit. Time is displayed on the x-axis in terms of generations and the associated number of hours slow and fast growth populations required to reach the specified number of generations. (A) presents the full graph while (B) has been zoomed in to better illustrate the dynamics observed in the fast growth populations.

per CFU per hour (Figure 4.2A).

If the replication-independent mutation rates are not accounted for, then the rate of mutation, when measured on a per generation basis, should not differ between the fast and slow growth populations (Equation 1, Figure 4.2B). However, when the replication-independent mutation rate is accounted for (Equation 2, Figure 4.2B), two expectations can be derived: (1) when the rate of mutation is measured on a per generation basis, it should be higher in the slow growth populations and (2) when the rate of mutation is measured on a per unit time basis, it should be higher in the fast growth populations. Estimates for rate of mutation experienced by the fast and slow growth populations were derived from linear increases in the proportion of cycloserine resistant mutants observed in each population during the evolution experiment (Figure 4.1). Such linear increases in the proportion of cycloserine resistant mutants were found during the following intervals in each population: ancestor 1 (fast, generations 12-26, r^2=0.86; slow generations 0-24 r^2=0.71), ancestor 2 (fast, generations 12-24, r^2=0.83; slow, generations 0-24, r^2=0.98) and ancestor 3 (fast, generations 0-12, r^2=0.84; slow generations 33-52, r^2=0.61). The slopes of these six lines were recorded as estimates for the per generation mutation rate towards cycloserine resistance in the fast and slow growth populations. In keeping with expectation (1), we estimated the per generation rate of mutation towards cycloserine resistance to be 8.32×10^{-7} new mutants per CFU per generation for the slow growth populations and 2.17×10^{-7} new mutants per CFU per generation for the fast growth populations (Figure 4.2C; paired t test, t=-4.6, p=0.045). Conversely, when these rates were instead expressed on a per unit time basis, the estimated mutation rate for the fast growth was found to be 1.7×10^{-7} new mutations per CFU per hour, higher than our estimated rate of mutation for the slow growth population which was 3.0×10^{-8} new mutants per CFU per hour. This is in agreement with expectation (2). These findings are consistent with the expected rates of mutations derived from the equation that accounts for the replication-independent mutation rate (Equation 2, Figure 4.2C)

4.4.3 Adaption to slow and fast growth environments involved similar changes in phenotype

The initial steps of adaptation to the chemostat environment has been shown previously to involve changes in the Monod growth parameters of μmax and K_s (Dykhuizen and Hartl, 1983). However, these studies were largely conducted at high flow rates that support corresponding high growth rates. Here, we asked whether our fast and slow growth populations both evolved via changes in these two parameters. The Monod growth parameters were μmax and K_s were estimated after 0, 24, 48 and 72 generations of evolution. Our analysis of these data revealed that, while the two population types appear to have adapted along similar trajectories, the slow growth populations had progressed further in phenotypic space. Both fast and slow chemostat populations evolved a higher μmax and, in some replicates, a lower K_s (Figure 4.3). However, while the slow growth populations demonstrated positive change in these parameters after only 24 generations of evolution, two of three fast growth populations regressed slightly and only one evolved a slightly higher μmax during this time.

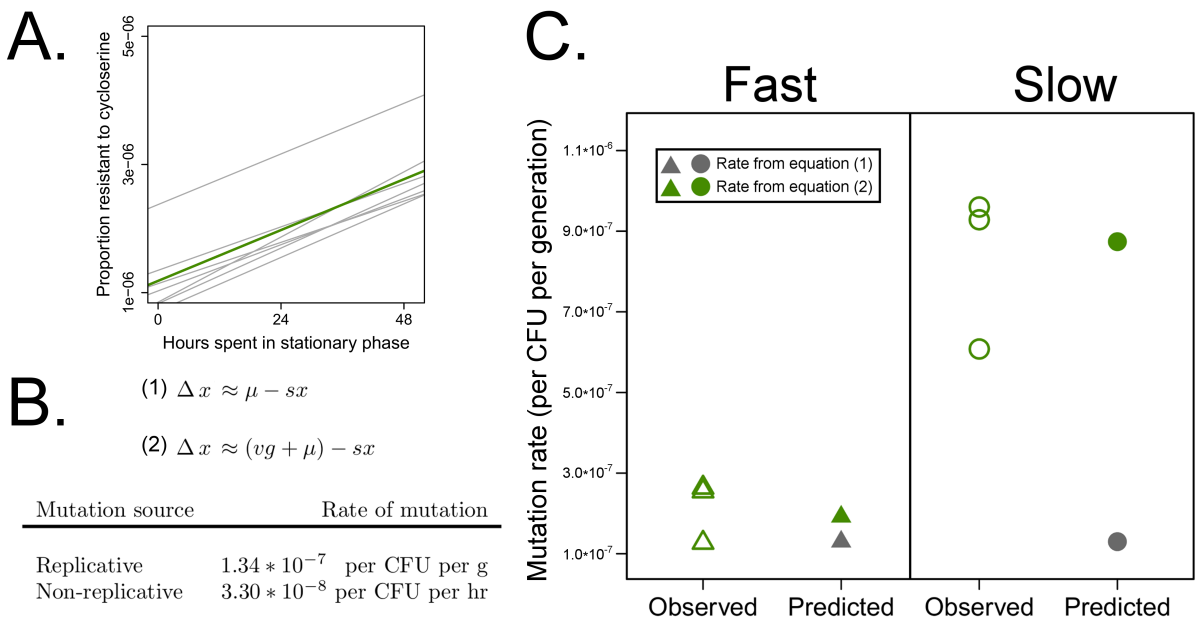

Figure 4.2: (A) Characterization of the rate of mutation towards cycloserine resistance during stationary phase of batch cultures. Each of the seven grey lines is a linear fit of the increase in the proportion of a single stationary phase population that was resistant to cylcoserine over 48 hours. The green line is a composite average of these seven independent measures of the replication-independent mutation rate towards resistance to cycloserine. (B) Equation (1) is a representation of the change in the frequency of mutant x (Δx) as a function of the mutation rate (μ) and the selection coefficient (s), without accounting for the growth independent mutations. Equation (2) is an expansion of (1) with the inclusion of the replication independent mutations (v) scaled to the generation time (g). (B) also includes our point estimates for the replicative (derived from standard fluctuation assays) and non-replicative mutation rates (derived from part A of this figure). (C) Comparison between the rate of mutation towards cycloserine observed in the chemostats (fast growth, open triangles; slow growth, open circles) and that predicted by equations (1) and (2).

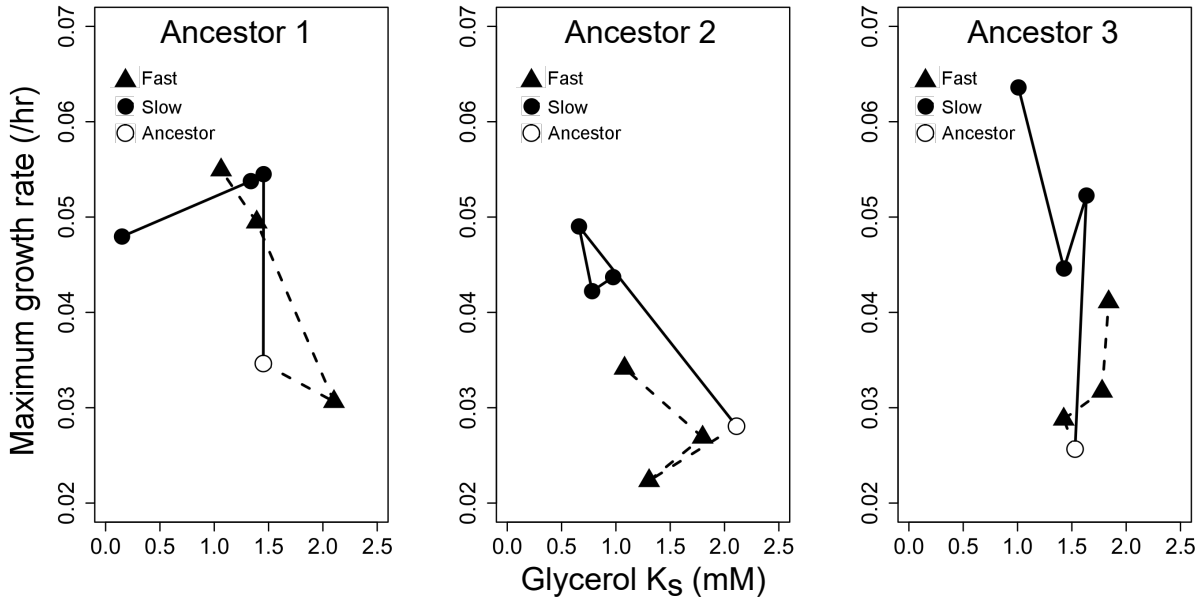

Figure 4.3: Monod growth parameters (K_s, x-axis; μmax, y-axis) in the minimal glycerol media for the three ancestors (open circles) and the slow (filled circles, solid lines) and fast (filled triangles, dashed lines) growth populations after 24, 48 and 72 generations of evolution (Left, ancestor 1; Middle, ancestor 2; Right, ancestor 3). Each point represents the average of three independent determinations of these parameters and the lines connect the estimates sequentially through time for the populations.

4.4.4 Slow growth populations reached higher fitness over the 72 generations than fast growth populations

We hypothesized that the additional mutations supplied, per generation, by replication independent processes might allow the slow growth populations to adapt at a faster per generation rate. To test this hypothesis, we measured the fitness of our fast and slow growth populations after 0, 24, 48, and 72 generations of evolution. Fitness was estimated in competition experiments wherein the evolved populations were competed against their corresponding cyc^r ancestral clone. Fitness was defined as the slope of the line relating the ln(evolved/ancestor) abundances over time. Our analysis showed that the slow growth populations averaged higher fitness values at all timepoints tested (Figure 4.4; $F_{1,61}$=37.27, P<0.001). The average relative fitness of the slow growth populations at the end of the evolution experiment was 1.71, while fast growth populations only reached an average fitness of 1.18 (Tukey adjusted post hoc comparison; t=-5.5, p<0.0001). This corresponds to average rates of adaptation of 0.0099 per generation in the slow growth populations and 0.0025 per generation for the fast growth populations. Conversely, when these rates are considered on a per day basis, the fast growth populations adapted more quickly (0.0099 per day in the slow growth populations, and 0.03 per day in the fast growth populations). This is not surprising because the fast growth populations required fewer days to reach the 72 generations of evolution.

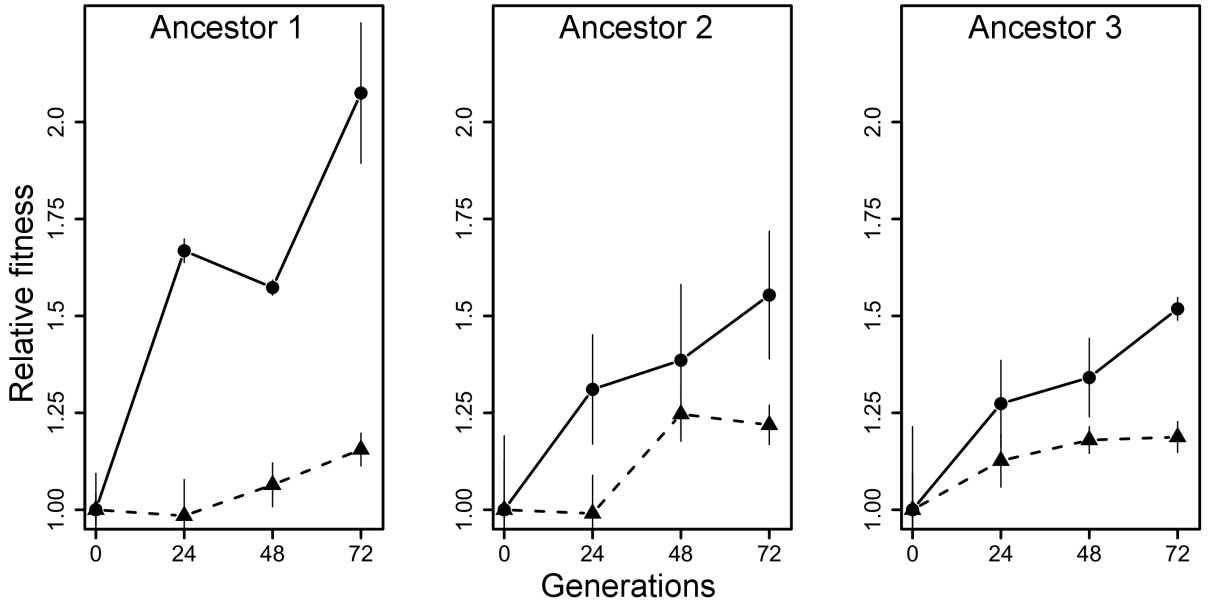

Figure 4.4: Fitness of the chemostat populations, relative to the ancestor, after 24, 48 and 72 hours of evolution (Left, ancestor 1; Middle, ancestor 2; Right, ancestor 3). In all three cases, the slowly growing populations (circles, solid lines) achieved a greater relative fitness than the rapidly growing populations (triangles, dashes). Each point is the average of three independent fitness assays, error bars represent the standard error of the mean.

The size of a population has been previously shown to influence the rate at which it adapts (Lanfear et al., 2014). For this reason, we compared the sizes of our slow and fast growth populations to determine if this might be a confounding factor in our experiments. However, it was expected that the size of the slow and fast growth populations would be similar because the density of cells in a chemostat is determined by the concentration of the growth limiting nutrient and not the flow rate of the chemostat. Throughout the evolution experiment, samples were regularly drawn from the populations and spread plate on nonselective media to estimate population size. Some variation in population size was observed in each of the six populations with estimates typically ranging between $1x10^9$ and $5.0x10^{10}$ colony forming units (CFUs). The effective population size of each population was then calculated as the harmonic mean of the population size estimates for that population over time. We found that average effective population size of the fast and slow growth population were $7.2x10^9$ and $8.1x10^9$ CFUs, respectively, and did not significantly differ from one another (paired t-test, t=-0.24, p=0.82). This indicates that variation in effective population size can not explain the observed differences in the rate of adaptation.

4.5 Discussion

We compared the evolution of *E. coli* populations that had experienced either fast (generation time of 2 hours) or slow (generation time of 48 hours) growth to determine the effect a bacterial population's

growth rate might have on its evolution. Although often overlooked, mutations can be caused by mechanisms other than errors made during the replication of the chromosome (Ryan, 1955; Kivisaar, 2010). Conversely, selection's ability to alter genotype frequencies in a population hinges on the growth and/or death of individuals (Darwin, 1859). We therefore hypothesized that slow growth would protract selective sweeps, allowing for the accumulation of diversity via replication-independent mutations.

To test our hypothesis, we compared the frequency of cycloserine resistant mutants in our fast and slow populations over the 72 generations of evolution. We have shown that slowly growing *E. coli* populations accumulate and maintain higher proportions of cycloserine resistant mutants than rapidly growing bacterial populations. This result is consistent with our hypothesis that slowly growing populations harbor more genetic diversity than rapidly growing populations. However, we note that our estimates for genetic diversity are derived from mutations in a single locus: the *cycA* gene. The relevance of these estimates to overall genetic diversity in the populations is not clear, although neutral markers have been widely used to estimate genetic diversity in the past (Dykhuizen, 1990; Kirk and Freeland, 2011). One advantage of using this particular marker is that resistance to cycloserine is caused by point mutations, insertions, and deletions in the *cycA* gene (Russell, 1972; Cseh and Umenhoffer, 2006), allowing it to track the genetic diversity that results from a wide range of mutational mechanisms. This is an important characteristic because replication-independent mutations are expected to be caused by a wide variety of mechanisms (Kivisaar, 2010).

Previous studies have described the population dynamics of other spontaneous neutral mutations for bacterial populations grown in chemostats (Atwood et al., 1951; Dykhuizen and Hartl, 1983). These studies have shown that such mutants accumulate in the populations over time due to ongoing spontaneous mutations (Novick and Szilard, 1950; Kubitschek, 1974). However, as beneficial mutations occur elsewhere in the population and sweep to high frequency, they raise the average fitness of the population. Because the prior accumulation of the neutral mutants are derived from the previous genetic background, their fitness is lower than the new average and they are purged from the population (Novick and Szilard, 1950; Atwood et al., 1951; Dykhuizen and Hartl, 1983). The process then repeats itself as the more fit dominant genotypes begin to generate new mutants via spontaneous mutation (reviewed in Dykhuizen and Hartl, 1983).

The tracked proportions of cycloserine resistant mutants were observed to have an oscillatory behavior with substantial variation between replicate populations experiencing the same growth conditions 4.1. There are at least two possible explanations for this variation that are not mutually exclusive. First, each pair of fast and slow growth populations were founded using different clones of *E. coli* K12 MG1655. It is possible that some of the variation in the frequency of cycloserine resistant mutants may have resulted from differences in the genetic makeup of these initial inocula. Second, the timing and magnitude of fitness effects are inherently stochastic. Differences in the timings of selective sweeps are expected to result in variation in when the accumulated cycloserine resistant mutants are purged. The magnitude of the fitness effects of the sweeping beneficial mutants could impact the speed at which the cycloserine

resistant mutants are purged from the populations. In one of three slow growth populations we also observed a rapid increase in the frequency of cycloserine resistant mutants towards the end of the evolution experiment. We speculate that this could have been driven by positive selection on a cycloserine resistant genotype. This could have either resulted from a beneficial mutation occurring in a cycloserine resistant background or from the sweep of a beneficial mutation generating a cycloserine resistant mutant when at low frequency.

We expected that the frequency of cycloserine resistant mutants would be higher in the slow growth populations due to a protraction of selective sweeps and an accumulation of replication-independent mutations. To examine whether this was a reasonable explanation for the observed accumulation of these mutants, we empirically determined the replication-dependent and replication-independent rates of mutation towards cycloserine resistance and compared them to the rate of mutation observed in the fast and slow growth populations. We showed that the rate of mutation in observed in the slow growth populations matched the model which accounts for replication-independent mutation rate better than the model which did not. However, this comparison relies on the validity of our estimates of the replication-dependent and replication-independent mutation rates. The replication-dependent mutation rate was determined using a standard fluctuation test which involved growing bacterial populations from a low cell density to a high cell density over a period of 24 hours (Foster, 2006). The populations used during this test certainly also experienced replication-independent mutations over the 24 hour time period. However, the expected number of replication-independent mutations that occur per hour is the product of the replication-independent mutation rate and the size of the population. For the majority of the fluctuation test, the population sizes were small. This means that the number cycloserine resistant mutants generated via replication-independent mechanisms was minimal and likely did not confound our estimate for the replication-dependent mutation rate. The replication-independent mutation rate was estimated by measuring the linear increase in the frequency of cycloserine resistant mutants during stationary phase. However, it could be that could be that observed increase during the 48 hours of stationary phase resulted from replication-dependent mutations that occurred when a fraction of the stationary phase population died and then was replaced by new growth. Using our estimate for the replication-dependent mutation rate (1.34×10^{-7} per CFU per generation), we calculate that at least 20% of the stationary phase populations would need to die and regrow every hour to account for the observed linear increase—a thermodynamically unreasonable expectation. We therefore conclude that our estimate for the replication-independent mutation rate likely was not driven by population turnover during stationary phase.

We observed that slowly growing populations accumulated cycloserine resistant mutants at a higher, per generation rate, than rapidly growing populations. Disparity in generation times between populations has been suggested as a factor that might invalidate the molecular clock hypothesis (MCH). First theorized by Zuckerkandl and Pauling (1962), the MCH holds that the rate of change in a proteins sequence is constant through time and across lineages. A population's generation time is thought to introduce variation in the rate at which proteins evolve by changing the number of mutations introduced per unit time,

via replication-dependent mutations (Laird et al., 1969; Thomas et al., 2010). Thus, populations that grow slowly reproduce less often and therefore experience a lower, per unit time, rate of mutation. Weller and Wu (2015) utilized comparisons between spore forming and non-spore forming Firmicute species in order to test this hypothesis in bacteria. They showed that the spore forming bacteria, who presumably experience longer generation-times, had corresponding lower rates of molecular evolution. The results from our study are not in conflict with this finding. If expressed on a per day basis, the fast growth populations did indeed experience higher mutation rates than the slow growth populations.

We also examined the effect a population's growth rate might have on the rate at which it adapts. We showed that our slowly growing populations evolved a higher average fitness over the 72 generations than the rapidly growing populations. We further demonstrated that the fast and slow growth populations had the same effective population sizes and were evolving via similar phenotypic change, indicating that neither played a role in the observed difference in the rate of adaption. Given previous estimates for the beneficial mutation rate in bacteria (Orr, 2003; Kassen and Bataillon, 2006), it was expected that many such mutations should occur in both the slow and fast growth populations, even without accounting for the replication-independent mutations. However, these replication-independent mutations should supply the slowly growing populations with more beneficial mutations for each generation of evolution and may even give slowly growing population access to rare large effect beneficial mutations. A greater number of beneficial mutants would effectively increase the summed frequencies of more fit individuals in the populations, allowing them to increase the average fitness of the populations more rapidly. Another possible explanation is that if the slowly growing populations experienced a higher beneficial mutation rate, it may give them increased access to rare large effect beneficial mutations.

If indeed adaptation in the slow growth populations is driven by multiple beneficial mutations, as opposed to just one, then there should be also be a diversity of high frequency genotypes in the populations as well. Well-mixed populations, like those in chemostats, are typically thought to evolve along a single trajectory via successive selective sweeps of single genotypes (Gillespie, 1983). Clonal interference allows a bacterial population to instead explore multiple evolutionary trajectories simultaneously. This mode of adaptation is predicted to be beneficial when sign epistasis is common, and the adaptive landscape correspondingly rough (Weinreich et al., 2005). Evolution along multiple trajectories can help populations that are evolving on such fitness landscapes by allowing them to circumvent suboptimal fitness peaks surrounded by fitness valleys (Wright, 1941; Nahum et al., 2015). Therefore, we speculate that growing slowly might help alleviate the barriers to adaptation that are caused by epistasis and asexual reproduction.

Although they were originally thought to be largely clonal (Ørskov and Ørskov, 1983), many studies have since demonstrated that bacterial populations in the environment contain a wealth of diversity (Selander and Leven, 1980; Smith et al., 1993; Wilder et al., 2009; Kuleshov et al., 2016). A central goal of evolutionary study has been to identify and characterize the factors that shape this within-species diversity (Rainey et al., 2000; Kassen, 2014). In this study, we demonstrated that a bacterial population's

growth rate can influence the emergence and maintenance of diversity and the rate of adaptation. Most microbial populations in the environment are growing slowly—they experience generation times on the order of days or even weeks. We speculate that these slowly growing environmental populations likely accumulate genetic diversity via the ongoing "metronome" of genetic change that results from replication-independent mutational processes. Extended further, this concept implies that as a bacterial population's growth rate approaches zero, selection's ability to purge this genetic diversity is minimized.

4.6 Acknowledgements

We thank John Roth, Thibault Stalder, and Eva Top for insightful discussions as well as Bruce Levin and Omar Cornejo for providing a descriptive manual on the construction and use of chemostats.

Chapter 5: Spatial structure facilitates the accumulation and persistence of antibiotic resistant mutants in biofilms

submitted as "Spatial structure facilitates the accumulation and persistence of antibiotic resistant mutants in biofilms"

5.1 Abstract

The emergence and spread of antibiotic resistance in bacterial pathogens is a global crisis. Because 65% of all bacterial infections are caused by pathogens that reside in biofilms we sought to investigate how biofilms influence the evolution of antibiotic resistance. We hypothesized that the inherent spatial structure of biofilms facilitates the accumulation and persistence of spontaneously evolved antibiotic resistant mutants. To test this, we tracked the frequency of mutants resistant to kanamycin and rifampicin in biofilm populations of *Escherichia coli* before, during, and following an antibiotic treatment regimen. Our results show that biofilms accumulate resistant mutants even before the start of treatment. Upon exposure to an antibiotic the resistant mutants swept to high frequency. Following the conclusion of treatment these resistant mutants remained at unexpectedly high frequencies in the biofilms for over 45 days. In contrast, if samples from antibiotic treated biofilms were used to found well mixed liquid cultures and propagated by serial transfer, the frequency of resistant cells dramatically decreased as they were out competed by sensitive clones. These observations suggest that evolution in spatially structured biofilms may account for the emergence and persistence of mutants that are resistant to antibiotics in bacterial infections.

5.2 Introduction

The US Centers for Disease Control and Prevention estimates that the majority of bacterial infections are caused by pathogens residing in biofilms (Costerton et al., 1999; Poterra, 1999), which are assemblages of microbial cells held together by an extra-cellular matrix (Stoodley et al., 2002). Yet studies on the evolution of antibiotic resistance have largely used serially passaged planktonic populations as their experimental system. This distinction is important because the environment of bacterial populations in biofilms differs dramatically from their planktonic counterparts. Individual cells within a biofilm are restricted in movement, providing the population with spatial structure. Biofilm populations also experience complex diffusional gradients of nutrients and waste products that result in a wide array of physiological states and growth rates (Stewart and Franklin, 2008). Infections caused by pathogens that form biofilms have been shown to be more difficult to eradicate with antibiotics than planktonic popu-

lations (Xu et al., 2000; Wolcott et al., 2010), often requiring longer treatment regimens (Römling and Balsalobre, 2012). We speculate that these differences may significantly affect the evolution of antibiotic resistance, making it all the more important that this process be better understood.

We posit that the spatial structure of microbial biofilms strongly influences the evolution of antibiotic resistance. The fitness of antibiotic resistant clones is often reduced relative to their antibiotic sensitive ancestors (Andersson and Levin, 1999; Andersson and Hughes, 2010). In unstructured populations, this fitness cost increases the probability that resistant clones will be lost from the population in the absence of positive selective pressures exerted by antibiotics. This is because unstructured bacterial populations experience global competition (Hibbing et al., 2010) wherein each individual competes against the entire population for resources and their reproductive success depends on their fitness relative to that of their competitors. In contrast, the situation in biofilms is strikingly different because individual cells are fixed in space by the extracellular polymeric matrix that they themselves produce (Sutherland, 2001). Thus, each individual only competes against a small subset of the population that is in close physical proximity. This greatly limits the spatial scale at which natural selection can operate and protracts selective sweeps thus allowing less fit variants (including those resistant to antibiotics) to persist and even accumulate (Gordo and Campos, 2006; Habets et al., 2007; Perfeito et al., 2008)). On this basis, we supposed that growth within a biofilm might facilitate the evolution and persistence of antibiotic resistant mutants.

Researchers have long known that resistance to most antibiotics, including aminoglycosides (Shakil et al., 2008), rifampins (Jin, 1988), polymixins (Port et al., 2014), flouroquionolones (Wolfson and Hooper, 1989), and β-lactams (Sun et al., 2014), can be achieved through mutations in target genes. These mutations occur in the absence of antibiotic selection at rates that are dependent on the number and mutation rate of the responsible genes, as well as the bacterial strain in question (Courvalin, 2008). However, these heritable mechanisms of resistance have seldom been considered in the context of the recalcitrance of biofilms to antibiotic therapy. This is perhaps because it is commonly held that mutations are rare and of little consequence over short periods of time. However, this is misleading because although per cell mutation rates may be low, the sizes of bacterial populations in biofilms can be quite large. As a result, mutants resistant to any particular antibiotic may be common within any given biofilm population. Furthermore, these mutations are likely to be clinically relevant since that they have been identified in whole genome sequencing data from infections caused by several biofilm forming pathogens including: *Staphylococcus aureus* (Mwangi et al., 2007; Howden et al., 2011), *Mycobacterium tuberculosis* (Eldholm et al., 2014), *Enterococcus faecalis* (Arias et al., 2011), *Klebsiella pneumoniae* (López-Camacho et al., 2014), *Pseudomonas aeruginosa* (Tsukayama et al., 2004; Marvig et al., 2015; Sommer et al., 2016; Haidar et al., 2017), and *Acinetobacter baumannii* (Liu et al., 2016).

Resistance is not the only means by which bacteria within biofilms can survive antibiotic exposure. Previous studies have shown that biofilm populations harbor antibiotic tolerant subpopulations that emerge from the differential expression of toxin-antitoxin system genes (e.g. hipAB) (Lewis, 2007, 2010). These tolerant cells persist during antibiotic treatment in a reversible state of dormancy (Schumacher

et al., 2015). The tolerant phenotype does not result from genetic change and is therefore not passed on to any of the surviving individuals' offspring. Another example is the biofilm matrix itself, which can slow diffusion of the antibiotic thereby creating zones of lowered antibiotic concentrations where sensitive cells can survive (Hoyle et al., 1992; Stewart, 1996). Additionally, diffusional gradients of nutrients and resources in a biofilm allow for cells in a wide range of physiological states some of which may be more tolerant to certain antibiotics (Xu et al., 1998; Sternberg et al., 1999). For example, β-lactam antibiotics such as penicillins, cephalosporins and carbapenems that inhibit bacterial cell wall synthesis may not be effective against the non-growing or slowly growing cells found in the interior of a biofilm. None of these mechanisms account for the emergence of heritable drug resistance that is the bane of physicians attempting to cure chronic infections caused by biofilms.

The goal of this study was to better understand the role of mutation and selection in the development of biofilm recalcitrance to antibiotics by tracking the frequency of antibiotic resistant mutants in biofilm populations of *Escherichia coli* K12 MG1655 prior to, during, and after treatment with either rifampicin or the aminoglycoside antibiotic kanamycin. We expected that (1) prior to antibiotic treatment, biofilm populations would accumulate antibiotic resistant mutants; (2) when these populations were treated with antibiotics, the resistant mutants would increase in frequency; and (3) refugia within a biofilm would facilitate the persistence of resistant mutants following the cessation of treatment. Our results for both kanamycin and rifampicin treated biofilms are consistent with these hypotheses and provide insight into how antibiotic resistance evolves and persists in biofilm populations during typical antibiotic treatment regimens.

5.3 Materials and Methods

5.3.1 Strains, media, and biofilm cultivation

Escherichia coli K12 MG1655 was grown in an M9 salts based minimal glucose (12.2mM) media supplemented with Wolfe's vitamins and trace elements (Wolin et al., 1963). Agar plates were prepared by supplementing the medium with 1.5% agar and, when appropriate, 20 μg/mL of either kanamycin or rifampicin. All planktonic cultures were grown at 37°C with shaking at 185 rpm. The minimum inhibitory concentration (MIC) of the ancestor was determined by plating 5μL droplets (containing approximately 10^4 CFUs) onto agar plates containing: 0 μg/mL, 1.5 μg/mL, 3.0 μg/mL, 4.5 μg/mL, 6 μg/mL, 10 μg/mL, 15 μg/mL or 20 μg/mL of either rifampicin or kanamycin. Plates were incubated for 24 hours and then examined for the presence of growth. We found the MIC of *E. coli* K12 MG1655 to be <1.5μg/mL for kanamycin and 4.5μg/mL for rifampicin.

Biofilms were cultivated at 25°C in custom acrylic flow cells with a glass substratum (total volume 10.4 mL) as described in Ponciano et al. (2009). Media was supplied to the flow cells via a peristaltic pump fitted with syringe flow breakers to prevent upstream contamination and bubble traps to prevent

the accumulation of bubbles (Figure 5.1). Flow cells were inoculated with 200μL of an overnight culture using a needle and syringe. Following inoculation, a 24h incubation period without flow was used to allow the bacteria time to adhere to the glass substrate. The flow of media was then commenced with a hydraulic retention time of two hours (5.4 mL/h).

Biofilms were destructively sampled using a calcium/alginate entrapment technique. A liquid solution of 3% alginate was added to the flow cells over the course of two hours at a rate of 20 mL/h, Followed by a 1h incubation period. Next, a 61.1mM calcium chloride solution was added over the course of 2h at 20mL/h followed by another 1hr incubation. The calcium alginate mixture solidified providing a gel encased biofilm. The flow cells were then disassembled and scalpel blades were used to remove three equally spaced, 1cm by 2cm horizontal sections from each biofilm. Biofilm sections were dissolved by placing them in a 0.85% saline (5mL) solution and incubated at 37oC and shaking at 185rpm for two hours.

The frequency of antibiotic resistant mutants in each biofilm section was then determined by diluting the sample in 0.85% saline and plating on selective and non-selective media in triplicate. All of the resistant genotypes identified in this study were capable of forming a colony on plates that contained 20μg/mL of the respective drug, significantly higher than the ancestral MIC. This ensures that non-heritable persister strategies were not included in our experiments, as they, by definition, are not capable of growth in the presence of the antibiotic.

5.3.2 Treatment regimen

Biofilms were cultivated for 15 days in the absence of antibiotics and then treated with either kanamycin or rifampicin (30 μg/mL) for 15 days (5.1). After the treatment regimen, the biofilms were cultivated for a further 45 days in the absence of antibiotics. Three independent biofilms were destructively sampled using the calcium alginate technique at the start of treatment (day 15), at several time points during treatment (kanamycin: days 18, 21, and 25; rifampicin: days 20, and 25), at the cessation of treatment (day 30), and at several time points following treatment (days 45, 60, and 75). A total of 45 biofilms were included in this study.

The triplicate biofilms harvested on the final day of treatment with either kanamycin or rifampicin (day 30) were used to inoculate planktonic cultures (5μL of biofilm into 5mL of media). These planktonic populations were then subcultured daily into fresh media (5μL of culture into 5mL media). This passaging regime provided 10 generations of growth per day and was carried out for a total of 25 days (250 generations).

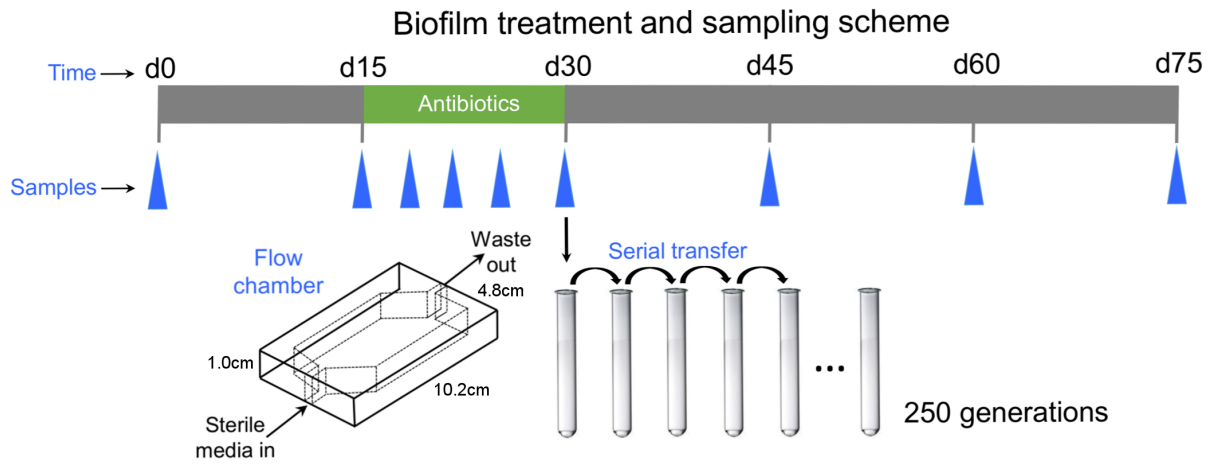

Figure 5.1: Biofilm cultivation apparatus including flow breakers, bubble traps and flow cell devices. Biofilms were grown for 75 days and treated with antibiotics from day 15 to day 30. Triplicate biofilms were destructively sampled at the time points denoted with stars (18, 21, 25 and 30 days for kanamycin and 20, 25 and 30 days for rifampicin). Day 30 biofilm samples were used to inoculate a planktonic population that was serially passaged in the absence of antibiotics for 250 generations.

5.3.3 Determination of mutation rates and relative fitness

Mutation rates towards resistance to kanamycin and rifampicin were determined in triplicate for exponentially growing cultures as described in Rosche and Foster (2000) and analyzed using a maximum likelihood method (Hall et al., 2009). For each replicate of the assay 24 cultures were inoculated with $<$300 cells and grown until reaching stationary phase. Three of the 24 cultures were diluted and plated on non-selective plates to determine the average population size and the remaining cultures were plated undiluted on selective media.

Mutation rates in stationary phase populations were determined by monitoring the increase in number of resistant mutants present during 5 days of stationary phase incubation by plating on selective and non-selective media every 24h. The rate of mutation was calculated as the slope of the increase in the abundance of antibiotic resistant mutants. There was no observed change in total population size during the 5 days.

The fitness of randomly selected antibiotic resistant mutants relative to the sensitive ancestor was determined as previously described (Lenski et al., 1991). Cultures of antibiotic resistant mutants and the sensitive ancestor were grown separately overnight and then 2.5μL of each was added to 5mL of media. Mixtures of these clones were grown for 24 hours resulting in approximately 10 generations of growth. Final and initial densities of the resistant clone and the sensitive ancestor were determined by plating on selective and non-selective media. The relative fitness of each resistant clone was calculated as the ratio of the natural logarithm of the final over the initial population sizes.

5.3.4 Statistical Analysis

All statistical analyses were performed using SAS 9.3 (SAS Institute, Cary, NC). The fraction of antibiotic resistant clones in the biofilm populations before, during and after the treatment regime was analyzed using a factorial ANOVA with the log transformed fraction of resistant clones as the response variable and the antibiotic (either rifampicin or kanamycin), time point of the experiment, and their interaction as response variables. Post-hoc comparisons of the fraction of resistance clones at various timepoints were used to test four hypotheses: that antibiotic resistant mutants accumulate in biofilms in the absence of antibiotics (day 0 versus day 15), that these mutants sweep to high frequency during treatment (day 15 versus day 30), that the evolved high proportion of resistant mutants can persist in the absence of selection in spatially structured environments (day 30 versus days 45,60, and 75) and that growth within a biofilm was necessary for the persistence of the evolved resistance (day 75 versus final abundance in planktonic). To correct for multiple comparisons, the Benjamini and Hochberg procedure was applied to the twelve comparisons using a false discovery rate of 0.05. Differences in the rate of mutation towards resistance to kanamycin and rifampicin and the relative fitness of resistant clones were determined using an unpaired two-sample t test. Figures were prepared using R 2.1.4 (Team 2013) and the plotrix package (Lemon, 2006). All of the data, SAS and R code for this study are available at https://github.com/michaelfrance/msphere2017_biofilms.

5.4 Results

Replicate biofilms were cultivated in custom flow cells as previously described (Ponciano et al., 2009), and after 15 days of growth in the absence of antibiotics, they were subjected to a 15 day antibiotic treatment regimen that was then followed by an additional 45 days of cultivation in the absence of antibiotics. After the initial 15 days of growth in the absence of antibiotics, *E. coli* K12 MG1655 was observed to have formed a robust biofilm that was attached to the glass substrate. The biofilms housed approximately 10^{10} colony forming units (CFUs) from this point on, even during the 15-day antibiotic treatment regime. At various times during the course of the experiment triplicate biofilms were destructively sampled and the average frequencies of antibiotic resistance in the populations were determined (5.1).

5.4.1 Populations within biofilms accumulated antibiotic resistant mutants in the absence of antibiotics

We first tested whether antibiotic resistant mutants accumulated prior to treatment by comparing the frequency of rifampicin and kanamycin resistant mutants in the initial inoculum to that of 15-day old antibiotic-naïve biofilm populations. The results showed that both kanamycin and rifampicin resistant mutants were present at low frequency in the inoculum and increased in frequency during a 15-day growth period in the absence of antibiotics (Figures 5.2A and 5.2B). Kanamycin resistant mutants increased from 2.31×10^{-8} in the inoculum to 1.07×10^{-6} in the 15 day old biofilms, which was a 46-fold increase (post

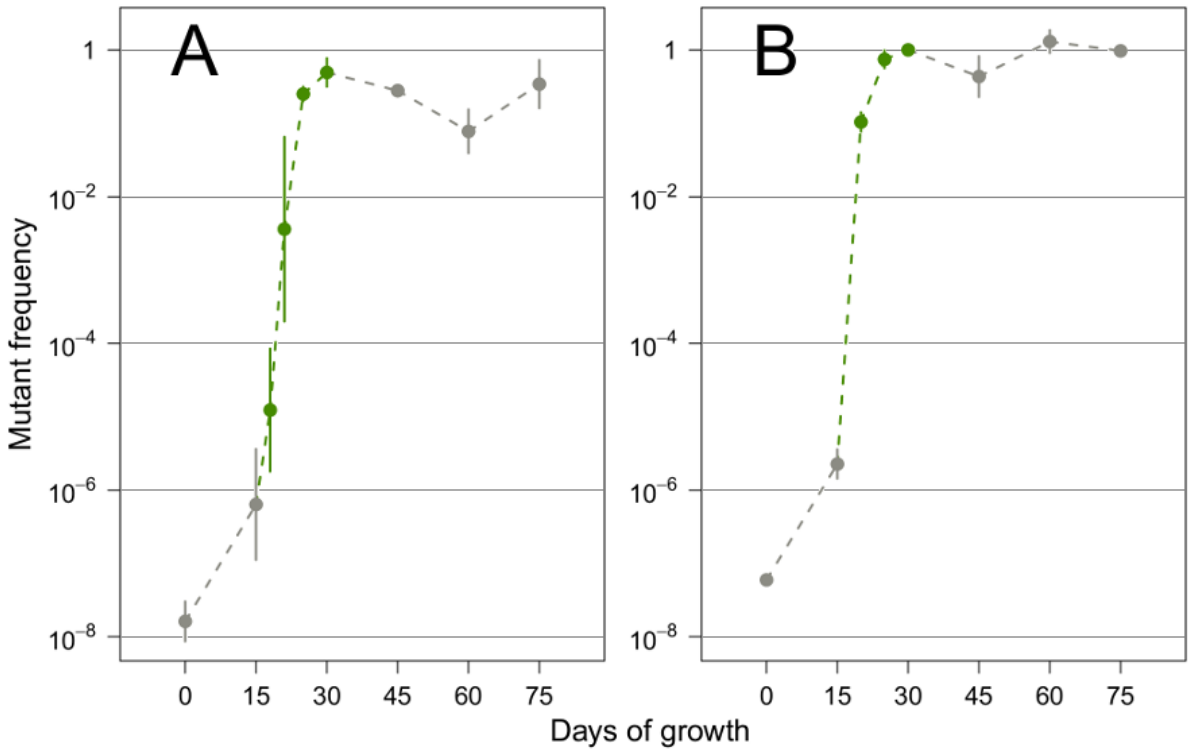

Figure 5.2: Log base ten transformed frequency of mutants resistant to either (A) kanamycin or (B) rifampicin before (grey), during (green) and after (grey) a 15 day antibiotic treatment regime. Each point represents the average of three independent biofilms which were destructively sampled. For the experiments using kanamycin a total of 24 biofilms were used, while 21 were used in the rifampicin study. Errors bars represent 95% confidence intervals.

hoc comparison; t = 3.6, p<0.001). Similarly, rifampicin resistant mutants increased from 5.81 x 10^{-8} in the inoculum to 2.46 x 10^{-6} during the 15 days of biofilm growth, which was a 42-fold increase (post hoc comparison; t = 21.7, p<0.001).

The observed increase in frequency of resistant mutants prior to treatment (Figures 5.2A and 5.2B) could be driven by either: 1) ongoing spontaneous mutations that generate new resistant genotypes, or 2) selection that favors resistant genotypes in the absence of the antibiotics. To somehow clarify which of these processes was responsible would require that we directly measure the mutation rates towards resistance as well as the selective effects of these mutations, in the biofilm habitats. However, the complex nature of these populations makes obtaining these estimates in vivo prohibitively difficult. Instead we used estimates from planktonic populations to give us some insight into which evolutionary process is more likely to be responsible.

If mutation alone was responsible, then the rate of mutation (expressed per day) would need to match

or exceed the rate of increase observed in the biofilms ($6.98*10^{-8}$, per day for kanamycin resistant mutants and at least $1.60*10^{-7}$ per day for rifampicin resistant mutants). Because biofilms contain large subpopulations that are growing very slowly or not at all, we determined the rate of mutation towards resistance in both exponentially growing and stationary phase populations. We found that, in exponentially growing cultures, mutations that confer resistance to kanamycin and rifampicin occurred at rates of $9.48x10^{-8}$ and $3.54x10^{-8}$ mutations per CFU per generation, respectively (Figure 5.3A; t=4.83, p=.0085). Less than one generation of growth per day would be required to explain the increased frequency of kanamycin resistant mutants while four and a half generations of growth per day would be needed to explain the increased frequency of rifampicin resistant mutants. In stationary phase cultures, kanamycin and rifampicin resistant mutants occurred at rates of $6.38x10^{-7}$ mutations per CFU per day and $6.58x10^{-8}$ mutations per CFU per day, respectively (Figure 5.3B; t=8.97, p<.0001). These estimates for the per day mutation rate during stationary phase exceeded, in the case of kanamycin resistance, or approached, in the case of rifampicin resistance, the rate of increase in mutant frequencies observed in the biofilm populations prior to treatment.

Conversely, if selection was primarily responsible for the observed increase in the frequency of antibiotic resistant mutants prior to treatment, then the resistance mutations should not have large fitness costs in the absence of an antibiotic. To characterize the spectrum of fitness costs associated with the mutations that cause resistance to kanamycin and to rifampicin, we randomly selected 18 spontaneously resistant mutants and determined their fitness, relative to that of the sensitive ancestor, in the absence of the antibiotic. The cost of resistance was calculated as the fitness of the resistant mutants relative to the ancestor minus one. We found that mutations causing resistance to kanamycin were usually associated with a large fitness cost in this strain (Figure 5.4, average s=-0.72). Only one of the eighteen kanamycin resistant mutants exhibited a fitness similar to that of the ancestor and six had relative fitness values below the detection limit of the assay. In contrast, we found that the mutations that cause resistance to rifampicin were, on average, associated with only a small fitness cost and several resistant mutants were more fit than the ancestor even in the absence of rifampicin (Figure 5.4; average s=-0.056).

5.4.2 Antibiotic resistant mutants sweep during treatment but do not always reach fixation

We expected that upon exposure to antibiotics, the preexisting resistant subpopulations would markedly increase in frequency but might not necessarily reach fixation in the population. To test this, we harvested kanamycin treated biofilms at 18, 21, 25 and 30 days and rifampicin treated biofilms at 20, 25 and 30 days (Figure 5.1) and determined the frequency of resistant mutants by plating on selective and non-selective media. As expected, the frequency of resistant mutants increased exponentially during the 15-day treatment regime. However, the results obtained with the two antibiotics differed slightly. In the kanamycin treated biofilms, the frequency of resistant mutants increased from 1.07×10^{-6} at the start of treatment to 0.523 after 15d of treatment (Figure 5.2A, post hoc comparison, t = 11.0, p<0.001). Note that even after 15 days of treatment, kanamycin resistant mutants only accounted for roughly half of the

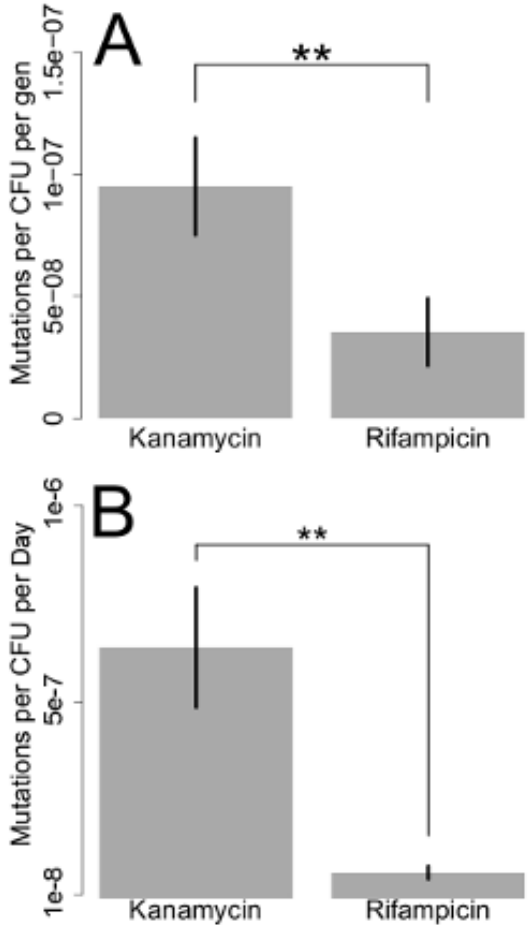

Figure 5.3: Comparison of the rate of mutation towards kanamycin and rifampicin resistance during exponential (A, t=4.83, p=0.0085) and stationary phases (B, t=8.97, p<0.001). * p<0.05, ** p<0.01, *** p<0.001, error bars represent 95% confidence intervals.

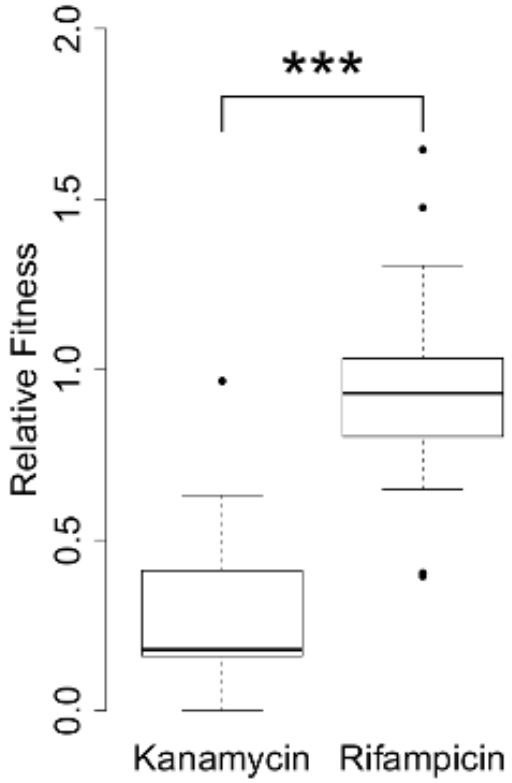

Figure 5.4: Comparison of the fitness of randomly selected resistant clones when grown in the absence of antibiotics relative to the sensitive ancestor, with kanamycin n=12 and with rifampicin n=18 (t=5.79, p<0.0001). * p<0.05, ** p<0.01, *** p<0.001, error bars represent 95% confidence intervals.

total population. By comparison the frequency of rifampicin resistant mutants increased from 1.07 x 10^{-6} resistant mutants per CFU in the naive biofilms to almost fixation at the end of treatment (Figure 5.2B, post hoc comparison, t = 60.9, p<0.001).

5.4.3 Antibiotic resistant mutants can persist following the cessation of antibiotic treatment

We next asked whether these resistant mutants could persist following the cessation of treatment. To answer this, we monitored the frequency of the resistant mutants in the post-treatment biofilms for an additional 45 days in the absence of antibiotics. In both cases the high abundance of resistant mutants did not substantially change despite the removal of antibiotic selection. For kanamycin, the resistant clones comprised 52% of the population at the end of antibiotic treatment and still comprised 39% of the population after 45 days of antibiotic free cultivation (Figure 5.2A, t=-0.30, p=0.77). Likewise, the rifampicin resistant mutants, which had nearly fixed in the population by the end of antibiotic treatment, persisted in the biofilms after the removal of antibiotic selection (Figure 5.2B, t=-0.14, p=0.89). These results demonstrate that the antibiotic resistance which developed during the treatment of biofilms can persist for extended periods of times even in the absence of antibiotics.

Finally, we determined whether growth within a biofilm was required for the persistence of the evolved high frequency of resistant mutants. These resistance mutations, including those studied here, are often associated with a fitness cost (Figure 5.4). Because well mixed populations experience global competition, we postulated that the frequency of the resistant mutants would decline when the populations were evolved in planktonic cultures. To test this, we founded planktonic populations using inocula from biofilm samples taken on the last day of treatment, and serially passaged them for 250 generations in the absence of antibiotics. In the case of cultures founded from kanamycin treated biofilms, the high frequency of antibiotic resistant mutants steadily declined over time (Figure 5.5A) and after 250 generations the frequencies of resistant mutants in planktonic populations had returned to roughly that found in the inocula used to found the biofilms. This is in stark contrast to persistence of kanamycin resistant mutants that was observed in the corresponding biofilm populations (Figure 5.2A, t=9.77, P<0.01). In comparison, rifampicin resistance persisted in both the biofilm and planktonic populations (Figure 5.5B, t=0.02, p=1.00). We think that the persistence of rifampicin resistant clones in planktonic cultures may be due to the low fitness costs of rifampicin resistance, which would be expected to slow the sweep of rifampicin sensitive cells. These results indicate that growth within a spatially structured biofilm facilitates the persistence of antibiotic resistant cells, even when resistance exacts a high cost on fitness.

5.5 Discussion

The evolution of antibiotic resistance by bacterial pathogens is a major challenge in the treatment of infectious diseases. In this study we characterized the evolution of antibiotic resistance in biofilm popu-

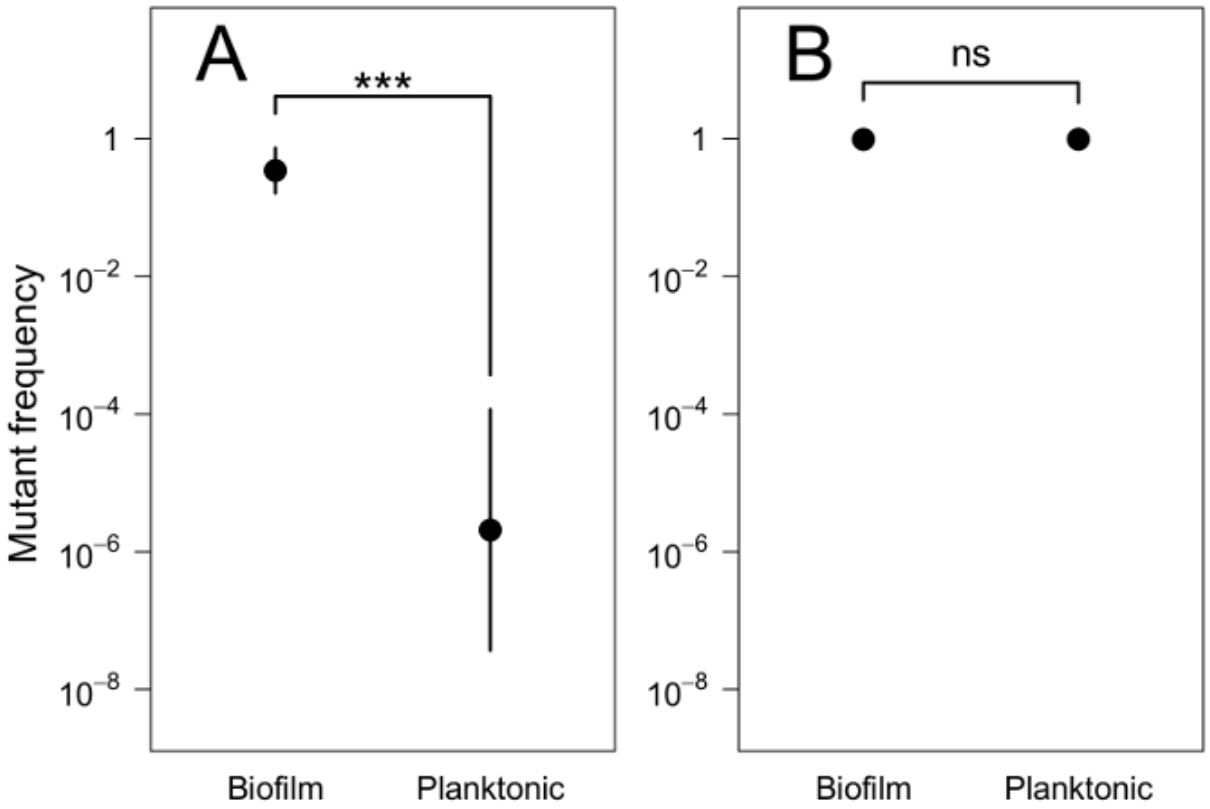

Figure 5.5: Comparison of the persistence of mutants resistant to either kanamycin (A, t=-9.77, p=0.0003) or rifampicin (B, t=0.02, p=0.982) in the biofilm and planktonic populations after 45 days (biofilm) or 250 generations (planktonic) days of growth in the absence of the associated antibiotic. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, error bars represent 95% confidence intervals.

lations of *E. coli* before, during, and following a 15 day antibiotic treatment regime. We have shown that while the mutations that confer antibiotic resistance occur infrequently, they accumulate and are orders of magnitude more common in spatially structured biofilms than in well-mixed cell suspensions, even in the absence of antibiotics. The resistant clones in these biofilms are pre-adapted and akin to the "seed banks" familiar to plant scientists in which pre-existing seeds germinate when conditions are favorable (Leck et al. 1989).

The accumulation of resistant mutants prior to treatment could be driven by continual spontaneous mutations that generate new resistant genotypes, by selection that favors resistant mutants in the absence of antibiotics, or by some combination of these two. This distinction has important implications for the generality of our findings. If selection is largely responsible, then the observed increase in resistant mutants would hinge on whether the resistant mutants were able to outcompete the sensitive ancestor in the absence of the antibiotic. However, if the increase in the frequency of resistant mutants was mainly driven by spontaneous mutations, it would only depend on whether resistance is obtainable via single step mutations. Boles and Singh (2008) observed the accumulation of gentamycin resistant mutants in *Pseudomonas aeruginosa* biofilms and suggested that selection was responsible. Our data suggests the opposite. A neutral process of spontaneous mutation is more likely to be the primary driver of the initial accumulation of antibiotic resistant variants in biofilm populations. To be clear, we are not suggesting that the accumulation of antibiotic resistant variants prior to treatment is driven by any foresight on the part of the bacteria, but instead we think that the population accumulates genetic diversity through a neutral and continual process of spontaneous mutations. Below we provide six lines of evidence supporting this supposition.

First, we observed the accumulation of mutants resistant to two antibiotics that differ in their mechanism of action as well as the spectrum of mutation that provide the phenotype. We further show that resistance to these two drugs emerges at different rates in this *E. coli* strain and that the resulting mutants differ in the magnitude of their effect on fitness, at least in planktonic populations. The observed accumulation of mutants resistant to these disparate antibiotics supports a neutral mechanism as the causative agent (see Figure 2.1 for data from other antibiotics). Second, the scale of the observed increase in the frequency of antibiotic resistant mutants prior to treatment was small. Both rifampicin and kanamycin resistant mutants increased in frequency by two orders of magnitude, from approximately 1 in 10^8 CFUs to 1 in 10^6 CFUs, over the 15 days of cultivation. Expressed on a per day basis, this amounts to a rate of accumulation that does not differ substantially from our estimates of the mutation rate in either exponential or stationary phase cultures (see Appendix B for details). While we could not directly measure the rate of mutation in the biofilm populations, it is unlikely that it is much lower than our planktonic estimates as several studies have hinted that mutation rates may, in fact, be elevated in biofilms (Boles and Singh, 2008; Driffield et al., 2008; Ryder et al., 2012). Third, we have shown that resistance to one of the antibiotics, kanamycin, was associated with a large fitness cost, at least when the bacterium was grown in planktonic cultures. Although the relevance of this estimate to growth in the biofilms is not straightforward, we think it unlikely that the sign of this fitness effect

would change given that aminoglycoside resistance mutations have been show to slow ribosome function (Galas and Branscomb, 1976). Fourth, if selection did favor kanamycin resistant mutants in the absence of the antibiotic, these mutants would be expected to further increase in frequency during the 45 days of post-treatment cultivation. Instead the resistant mutants remained at roughly the same frequency in the populations throughout this time period. Fifth, biofilms are spatially structured and contain a plethora of largely independent subpopulations that are growing slowly (Sternberg et al., 1999). These characteristics impede changes in allele frequency; a prerequisite for selective sweeps (Gordo and Campos, 2006; Habets et al., 2007; Perfeito et al., 2008). The overall impact of selection on biofilm populations was probably diminished, thereby facilitating the accumulation of diversity through spontaneous mutations. Finally, in order to invoke positive selection on the resistant mutants in the absence of the associated antibiotic, the observed change in the frequency of the antibiotic resistant mutants would need to exceed the expected accumulation due to spontaneous mutation. This is not the case for our observations of the accumulation of either rifampicin or kanamycin resistant variants.

We also demonstrated that resistant mutants increased exponentially in frequency once the biofilms were treated with either kanamycin or rifampicin. This result was expected given that antibiotics select for any resistant genotypes in the population. However, somewhat unexpectedly, kanamycin resistant mutants did not fix in the population and after 15 days of treatment roughly half of the population was still sensitive to the antibiotic. These sensitive genotypes may have survived due to some combination of the previously mentioned phenotypic mechanisms of biofilm recalcitrance (antibiotic diffusion (Stewart, 1996), antibiotic action antagonism (Brown et al., 1988), or the occurrence of persister cells (Lewis, 2007)). In particular, aminoglycosides have been shown to not be as effective at killing under anaerobic conditions like that found in the interior of biofilm populations (Shakil et al., 2008). This problem is likely further compounded by the relatively low fitness of spontaneous kanamycin resistant mutants (Figure 5.4). In comparison, anaerobic conditions do not have as large of an effect on the activity of rifampicin and spontaneous rifampicin resistant mutants do not suffer the same fitness cost (Maggi et al., 1966; Reynolds, 2000). Accordingly, rifampicin resistant mutants almost fixed in biofilm populations. Whether or not resistant genotypes fix in the population has important implications for their persistence since, as the ancestral sensitive genotype declines in frequency, so does its chances of reemerging following the cessation of treatment.

Perhaps the most important result from this study is our demonstration that antibiotic resistant mutants persist at high frequency in biofilms even in the absence of antibiotics. We observed that after 45 days the frequency of resistant mutants had not changed much from that found at the cessation of treatment. We suggest that this outcome probably derives from characteristics of the biofilm mode of growth. Biofilms are spatially structured and contained large subpopulations of slowly growing cells (Sternberg et al., 1999). Spatial structure and reduced growth rates likely limit the effectiveness of selection and facilitate that persistence of less fit variants like kanamycin resistant mutants. This explanation is supported by our finding that the same kanamycin resistant mutants were rapidly expunged from well mixed planktonic populations, where selection should act efficiently. The persistence of large numbers of

resistant mutants in the absence of antibiotics might also allow for the acquisition of additional compensatory mutations (Andersson and Hughes, 2010) that diminish the cost of resistance. We speculate that once the cost of resistance is ameliorated, the resistance mutants are more likely to persist even in well mixed populations, further facilitating the spread of antibiotic resistance (Normak and Normak, 2002; Andersson, 2003; Andersson and Hughes, 2010).

The results from this study demonstrate that biofilms play a critical role in the evolution of antibiotic resistance. Novel strategies must be developed for the treatment of persistent or recurrent biofilm infections that account for the inevitable rise and persistence of resistance in these populations. Our results are derived from experiments using two distinct antibiotics: one whose resistance mutations are associated with a large fitness cost, and one whose resistance mutations are not. Previous studies have indicated the cost associated antibiotic resistance mutations can vary widely, although they are rarely as costly as the kanamycin resistance mutations characterized in this study (Melnyk et al., 2015). We speculate that our observation of a high fitness cost associated with kanamycin resistance mutations may have resulted from the growth conditions used in this study—a previous study indicated that, for *E. coli*, the fitness cost of some resistance mutations are exacerbated when the strain is grown in minimal media, like that used in this work (Petersen et al., 2009). Despite the stark differences in the fitness cost of rifampicin and kanamycin resistance mutations, our results for the two antibiotics are fairly similar. The primary differences were that rifampicin resistant variants nearly fixed during treatment while the frequency of kanamycin resistant variants did not, and that kanamycin resistance declined in the unstructured populations. Furthermore, while our study focused on a single strain of *E. coli*, we expect that our results are likely applicable to clinically relevant biofilms formed by other species. This is because all bacterial populations experience the evolutionary processes of mutation and natural selection and spatial structure is an inherent characteristic of biofilm populations. We anticipate that spatial structure will allow for the accumulation and persistence of spontaneously generated mutants that are resistant to any given antibiotic.

Chapter 6: Genomic comparisons of *Lactobacillus crispatus* and *Lactobacillus iners* reveal potential ecological drivers of community composition in the vagina

"Genomic comparisons of *Lactobacillus crispatus* and *Lactobacillus iners* reveal potential ecological drivers of community composition in the vagina" *Applied and Environmental Microbiology, vol. 82, 2016, pp. 7063-7073*

6.1 Abstract

Lactobacillus crispatus and *Lactobacillus iners* are common inhabitants of the healthy human vagina. These two species are closely related and are thought to perform similar ecological functions in the vaginal environment. Temporal data on the vaginal microbiome have shown that nontransient instances of cooccurrence are uncommon, while transitions from an *L. iners*-dominated community to one dominated by *L. crispatus*, and vice versa, occur often. This suggests that there is substantial overlap in the fundamental niches of these species. Given this apparent niche overlap, it is unclear how they have been maintained as common inhabitants of the human vagina. In this study, we characterized and compared the genomes of *L. iners* and *L. crispatus* to gain insight into possible mechanisms driving the maintenance of this species diversity. Our results highlight differences in the genomes of these two species that may facilitate the partitioning of their shared niche space. Many of the identified differences may impact the protective benefits provided to the host by these two species.

6.2 Importance

The microbial communities that inhabit the human vagina play a critical role in the maintenance of vaginal health through the production of lactic acid and lowering the environmental pH. This precludes the growth of nonindigenous organisms and protects against infectious disease. The two most common types of vaginal communities are dominated by either *Lactobacillus iners* or *Lactobacillus crispatus*, while some communities alternate between the two over time. We combined ecological theory with state-of-the-art genome analyses to characterize how they these two species might partition their shared niche space in the vagina. We show that the genomes of *L. iners* and *L. crispatus* differ in many respects, several of which may drive differences in their competitive abilities in the vagina. Our results provide insight into factors that drive the complicated temporal dynamics of the vaginal microbiome and demonstrate how closely related microbial species partition shared fundamental niche space.

6.3 Introduction

The microbial communities that inhabit the vaginas of healthy reproductive age women commonly contain high proportions of *Lactobacillus* species (Vásquez et al., 2002). Studies have shown that these communities can be divided into five different types, four of which are dominated by either *Lactobacillus iners*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, or *Lactobacillus jensenii* (Ravel et al., 2011). These four species are closely related and are thought to perform similar ecological functions in the vaginal environment (namely, the production of lactic acid) (Linhares et al., 2011; O'Hanlon et al., 2011). Instances of cooccurrence among these species are rare, and temporal data have demonstrated that shifts in the dominant *Lactobacillus* are common (Gajer et al., 2012), suggesting that the species compete for shared niche space in the vagina. Ecological theory predicts that multiple species cannot occupy the same niche indefinitely, as one will eventually outcompete the others (Gause, 1934). Therefore, it is unclear how the four *Lactobacillus* species have been maintained as common inhabitants of the vaginal niche.

Previous studies have shown that competing species can partition their shared niche space through a variety of mechanisms. One such mechanism, termed resource partitioning, occurs when competing species specialize in the use of different subsets of resources, thereby dividing the niche into multiple niches and allowing them to cooccur (Schoener, 1974; Hunt et al., 2008; Mendes-Soares and Rychlik, 2009). However, we argue that it is unlikely that the vaginal lactobacilli are dividing their shared niche space in this way because they rarely cooccur (Ravel et al., 2011). Species can also partition shared niche space temporally through a mechanism termed conditional differentiation. This occurs when the species differ in competitive ability across the niche's range of environmental conditions (Chesson, 2000; Amarasekare, 2003; Mammola and Isaia, 2014). The abundance of the species is then determined by the abiotic and biotic factors that influence their competitive interactions. For example, in their 2014 work, Mammola and Isaia showed that variation in the temperature and humidity levels in caves allowed two competing spider species to partition their shared niche space (Mammola and Isaia, 2014). We argue that given the complex temporal fluctuations exhibited by vaginal *Lactobacillus* species, these species likely partition their shared niche via this mechanism. In the present study, we characterized and compared the genomes of two of the four prominent vaginal *Lactobacillus* species, *L. crispatus* and *L. iners*, to identify possible ecological factors that might drive these temporal fluctuations in the dominant *Lactobacillus* species.

Little is known about the abiotic and biotic factors that might be relevant to competitive interactions between vaginal *Lactobacillus* species. We speculate that the host's physiology plays a critical role in shaping the vaginal environment through at least two different mechanisms. First, the host is the exclusive source of nutrients available in the environment. These nutrients originate both from the mucus produced by the cervix, which contains a rich mixture of carbohydrates, fatty acids, and trace elements (Wolf et al., 1980), and from vaginal epithelial cells, which in reproductive-age women are loaded with glycogen (Cruickshank and Sharman, 1934). The amount and composition of cervical mucus, as well as the amount of glycogen available, vary among women and through time in a single woman. Temporal

variation in these characteristics occurs both on the scale of the menstrual cycle as well as through the lifetime of an individual (Milwidsky et al., 1980; Bigelow et al., 2004; Farage and Maibach, 2006; Grande et al., 2015; Mirmonsef et al., 2015). Additionally, the host can effect change in the vaginal communities via the immune system. The vagina contains various components of the innate and adaptive immune systems that protect this important interface from infection (Wira et al., 2005). The activity of the host's immune system also exhibits variation between women and through time within women (Wira et al., 2002), and some of this variation results from interactions with the microbes inhabiting the vagina (Doerflinger et al., 2014). While it is clear that host physiology must play some role in shaping the microbial communities inhabiting the vagina, the magnitude and nature of its effect are not known.

Although there are four dominant vaginal *Lactobacillus* species, we chose to focus our efforts on genomic comparisons of *L. crispatus* and *L. iners* for two reasons. First, of the five community types, the two most commonly identified in reproductive-age women (>75% relative abundance) are dominated by one of these two species (Ravel et al., 2011; Gajer et al., 2012). Second, although both of these species produce lactic acid as the end product of fermentation, they differ in many other respects, including their relationship to host health. While the presence of *L. crispatus* in the vagina has always been associated with good health, some studies have hinted that communities dominated by *L. iners* may provide the host with fewer protective benefits (Tamrakar et al., 2007; Petricevic et al., 2014; Wertz et al., 2008). For example, *L. iners*, but not *L. crispatus*, commonly cooccurs with many of the bacterial species that colonize the vagina during incidences of bacterial vaginosis (Tamrakar et al., 2007; Fredricks et al., 2005; Ferris et al., 2007). Communities dominated by *L. iners* have also been associated with a higher vaginal pH than that in communities dominated by *L. crispatus* (Ravel et al., 2011). In addition, these two species also differ in their specificity for the vaginal habitat. While *L. iners* has almost exclusively been isolated from human vaginal secretions, *L. crispatus* has also been identified in other habitats, like the vertebrate gastrointestinal tract (Ojala et al., 2010), although it is unclear whether *L. crispatus* is a frequent colonizer of these other habitats. We argue that by focusing our efforts on these two common but markedly different vaginal colonizers, we are likely to identify both ecologically and medically relevant factors that govern competitive interactions between the vaginal *Lactobacillus* species.

We employed comparative genomics techniques to investigate niche partitioning by *L. crispatus* and *L. iners* and to identify putative ecological factors that govern competitive interactions between these species. The two previous comparative genomic studies of the vaginal *Lactobacillus* species have either focused on *L. crispatus* only (Ojala et al., 2014), or on broad functional differences between vaginal versus nonvaginal *Lactobacillus* (Mendes-Soares et al., 2014). In this study, we characterized and compared the genomes of 15 *L. iners* and 15 *L. crispatus* strains. We report differences in the size, functional makeup, and evolution of the genomes of these two species, consistent with niche partitioning via conditional differentiation. Our analysis indicates that *L. iners* has a genome whose size is drastically reduced, likely making the species dependent on exogenous sources (e.g., cervical mucus or other vaginal species) of vital nutrients. In comparison, *L. crispatus* has a larger genome that contains a broader array of metabolic machinery, likely allowing it to function under a more diverse subset of environmental conditions. These

genotypic differences may provide the species with differential competitive abilities across the range of conditions common to the vaginal environment, facilitating the partitioning of their shared niche space.

6.4 Materials and Methods

6.4.1 Sequences

All available complete and draft genome sequences for *L. iners* and *L. crispatus* were obtained from the National Center for Biotechnology Information (NCBI) FTP site (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/) in March 2016. Of the 30 *L. crispatus* and *L. iners* strains used in this study, only two were not isolated from the human vagina (*L. iners* DSM 13335 was isolated from urine, and *L. crispatus* ST1 was isolated from the crop of a chicken). A complete list of the strains with accompanying genome quality statistics can be found in Table 6.1. The number of contigs for each genome ranged from 7 to 97 contigs for *L. iners* and 1 to 295 contigs for *L. crispatus*. Additionally, sequence data for *Lactobacillus acidophilus* NCFM, *Lactobacillus casei* BD-II, *Lactobacillus delbrueckii* ATCC 11842, *Lactobacillus helveticus* CNRZ32, *Lactobacillus johnsonii* NCC 533, *Lactobacillus plantarum* WCFS1, *Lactobacillus rhamnosus* GG, *Bacillus subtilis* 168, *Enterococcus faecalis* ATCC 29212, and *Streptococcus pneumoniae* R6 were obtained from NCBI for use as outgroups in the construction of the phylogenetic tree.

6.4.2 Gene prediction and functional annotation

Open reading frames (ORFs) were predicted in all genomes using Glimmer version 3.0, with a maximum overlap of 50 bp between ORFs and a minimum ORF length of 110 bp (Salzberg et al., 1998; Delcher et al., 1999, 2007). Contigs from draft genomes were concatenated with 25-bp spacers of ambiguous DNA sequences. Any ORFs containing spacer sequences were later removed from the analysis. The identified ORFs were translated to protein sequences, and orthologous groups were identified among them using OrthoMCL version 2.0, with a percent identity threshold of 50% (Li et al., 2003). Pangenome, core-genome, and accessory-genome accumulation curves were generated from the OrthoMCL output using a python script. This python script and all others used in the analysis are available at `github.com/michaelfrance/lactobacillus_genomics`.

Core- and accessory-gene sets were assigned functional annotations using Kyoto Encyclopedia of Genes and Genomes (KEGG) BlastKOALA searches against the prokaryotic species database (Kanehisa et al., 2016). However, this method proved incapable of assigning functions to the genes that are found in only a single strain. We therefore assigned this gene set to functions based on results from Blast searches against the RefSeq nonredundant protein database (Altschul et al., 1997). The top 20 matches for each Blast search with >50% identity (calculated as the number of matches divided by the total length of the two aligned sequences) were examined and, when in agreement, used to assign a broad functional category.

Table 6.1: Bacterial genome data

Strain	Source	Genome size (Mbp)	No. of contigs	No. of ORFs
L. crispatus				
JV-V01	Human vagina	2.22	86	2,151
MV-1A-US	Human vagina	2.25	7	2,383
125-2-CHN	Human vagina	2.30	30	2,196
MV-3A-US	Human vagina	2.44	76	2,458
CTV-05	Human vagina	2.36	25	2,425
FB049-03	Human vagina	2.46	5	2,474
FB077-07	Human vagina	2.70	10	2,688
SJ-3C-US	Human vagina	2.09	201	2,199
214-1	Human vagina	2.07	187	2,100
2029	Human vagina	2.19	295	2,545
EM-LC1	Human vagina	1.83	63	1,839
ST1	Chicken crop	2.04	1	2,060
VCM6	Human vagina	2.34	253	2,397
VCM7	Human vagina	2.10	247	2,108
VCM8	Human vagina	2.33	255	2,427
L. iners				
UPII 60-B	Human vagina	1.32	31	1,288
UPII 143-D	Human vagina	1.26	21	1,213
LactinV 01V1-a	Human vagina	1.29	92	1,506
LactinV 03V1-b	Human vagina	1.30	67	1,454
LactinV 09V1-c	Human vagina	1.31	35	1,384
LactinV 11V1-d	Human vagina	1.31	27	1,362
LEAF 2052A-d	Human vagina	1.32	28	1,268
LEAF 2053A-b	Human vagina	1.37	37	1,288
LEAF 2062A-h1	Human vagina	1.30	24	1,278
LEAF 3008A-a	Human vagina	1.27	25	1,216
SPIN 1401G	Human vagina	1.28	52	1,224
SPIN 2503 V10-D	Human vagina	1.28	31	1,293
ATCC 55195	Human vagina	1.24	7	1,152
AB-1	Human vagina	1.29	7	1,230
DSM 13335	Human urine	1.28	12	1,212

6.4.3 Phylogenetic analysis

OrthoMCL was also used to identify the genes that *L. crispatus*, *L. iners*, and the outgroup strains have in common (50% identify, 90% coverage thresholds). The 39 strains were found to have 242 genes in common spanning more than 250 kb. The orthologous gene clusters were aligned with ClustalW version 2 (Larkin et al., 2007) and concatenated into a single alignment using Phyutility version 2.2.6 (Smith and Dunn, 2008). The rCluster hierarchical clustering algorithm implemented by PartitionFinder version 1.1.1 (Lanfear et al., 2012) was used to identify the optimal partitioning scheme for the concatenated alignment based on codon positions; a total of 301 partitions were used. RAxML version 8 was used to construct a maximum likelihood phylogenetic tree from the partitioned concatenated multiplesequence alignment (Stamatakis, 2014). Each partition was modeled separately under the general time-reversible gamma substitution model (Tavare, 1986). One thousand bootstrap replicates were performed, and convergence occurred after 700 bootstrap replicates. A separate phylogeny was constructed in the same manner from an alignment of the *L. iners* cytolysin gene and matching sequences from the NCBI nonredundant nucleotide database.

6.4.4 Evolutionary analysis

We used the program TimeZone version 1.0 to characterize the evolution of each species' core genome separately (Chattopadhyay et al., 2013). This program executes a pipeline designed to identify genes exhibiting signatures of positive selection and recombination. TimeZone relies on the use of a reference genome and is therefore only suitable for the analysis of genes shared among all strains (i.e., core genes). In both cases, the most complete genome was used as the reference sequence (*L. crispatus* ST1 and *L. iners* AB-1).A percent identity and coverage cutoff of 75% were used for both within-species analyses. We used TimeZone to calculate the ratio of nonsynonymous to synonymous nucleotide changes (dN/dS) for the core genome of each species. In general, dN/dS ratios below 1 are indicative of purifying selection, while dN/dS ratios above 1 are indicative of positive selection.

Evolution can also occur through the horizontal acquisition of foreign DNA. To identify these horizontally acquired sequences, we analyzed two different sets of genes. The first included the genes uniquely conserved in the core genomes of either *L. crispatus* or *L. iners*. The second set included those genes present in only a single strain of *L. crispatus* or *L. iners*. To determine whether these sequences were more likely to have been inherited or horizontally acquired, we used Blast searches against the RefSeq nonredundant protein database (as described in the gene prediction and functional annotation methodology section) to identify the genus of the closest matching homologous sequence in the database. If the closest matching sequences in the database were identified in other *Lactobacillus* species, we inferred that the gene is more likely to have been inherited, whereas if the gene matched only non-*Lactobacillus* species, we considered it to have been horizontally acquired.

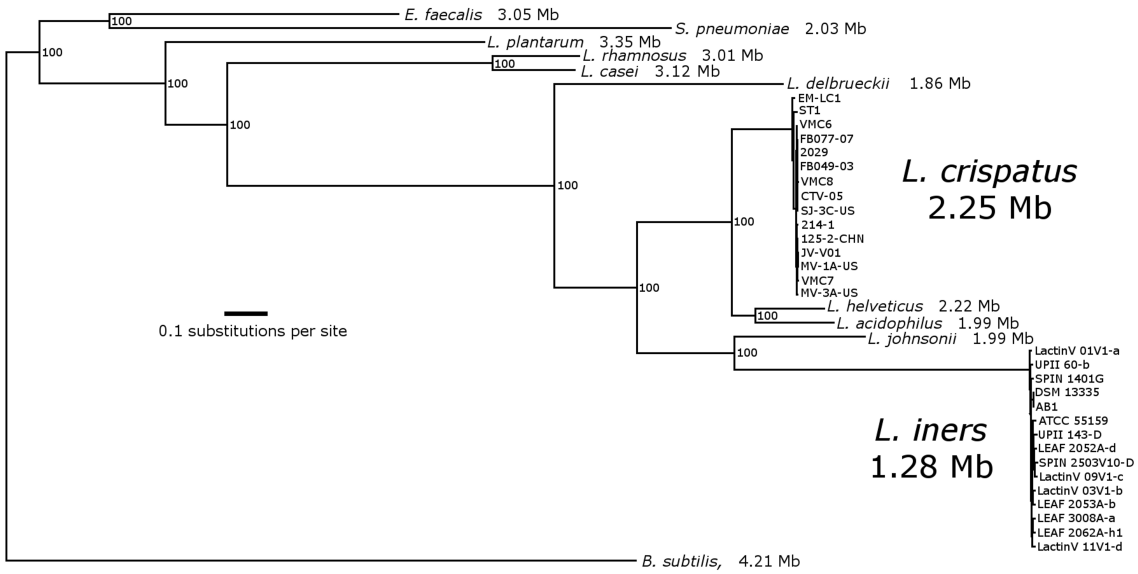

Figure 6.1: Maximum likelihood tree of the phylogenetic relationships between the strains of *L. iners* and *L. crispatus* used in this study. The phylogeny was constructed from a partitioned concatenated alignment of the 242 genes shared between the included *L. crispatus* and *L. iners* strains, as well as several outgroup species. Genome size in megabase pairs is mapped onto the tips of the tree to give an idea of how this trait has evolved along the phylogeny.

6.5 Results

6.5.1 Phylogenetic analysis

To understand the phylogenetic relationships among the strains used in our study, we constructed a maximum likelihood phylogenetic tree using a partitioned concatenated multiple-sequence alignment of the 242 genes that were present in all of the *L. crispatus* and *L. iners* strains, as well as the outgroup species. In the resultant phylogenetic tree, the 15 strains of *L. crispatus* and 15 strains of *L. iners* clustered together into single clades (Figure 6.1). Branching patterns within the *L. crispatus* clade, but not the *L. iners* clade, are generally well supported. Branch lengths within the two species are orders of magnitude shorter than those between species, indicating that the majority of the observed diversity in these 242 genes occurred between rather than within species. Additionally, our analysis indicates that these two vaginal species are not sisters of one another; rather, *Lactobacillus johnsonii* is sister to *L. iners*, and both *Lactobacillus helveticus* and *Lactobacillus acidophilus* are sisters to *L. crispatus* (Fig. 1). In the following analyses, we used this tree to determine whether specific traits of *L. crispatus* and *L. iners* are more likely to be derived characteristics of the species or are ancestral.

6.5.2 Differences in genome size

Perhaps the most obvious difference between the genomes of *L. iners* and *L. crispatus* are a matter of scale. While the average size of the *L. crispatus* genome was 2.25 Mbp, the *L. iners* genome was

only 1.28 Mbp on average (Table 6.1; Welch's t test, t=17.8, P <0.001). *L. crispatus* was also found to have roughly twice as many open reading frames (ORFs) as *L. iners* (Table 6.1; Welch's t test, t=15.9, P<0.001). Based on the phylogeny presented in Figure 6.1, the reduced genome of *L. iners* appears to be a derived characteristic of the species. In comparison, *L. crispatus* has maintained a larger genome, more similar to that of other vagina-associated *Lactobacillus* species (*L. delbrueckii*, *L. acidophilus*, and *L. johnsonii*).

Next, we sorted the predicted open reading frames into orthologous gene sets using OrthoMCL. These orthologous gene sets were then categorized as either core genes, meaning those present in all strains of a species, or accessory genes, whose presence varies across the strains (Medini et al., 2005). Additionally, the union of core and accessory genes is defined as the pangenome and contains all orthologous gene sets identified in these species. Consistent with the genome size data, we found that the pangenome of *L. crispatus* had almost twice as many genes as *L. iners* (4,300 versus 2,300 genes; Figure 6.2A). Furthermore, the pangenome accumulation curves indicate that this difference is likely to be maintained as more strains are sequenced for each species (Figure 6.2A). Similarly, the core genome of *L. crispatus* was larger than that of *L. iners* (1,442 genes versus 993 genes; Figure 6.2C). Finally, the accessory genome of *L. crispatus* contained 2,884 genes, of which 45% were present in only one strain (singletons), while the accessory genome of *L. iners* contained only 1,233 genes, of which 56% were singletons (Figure 6.2B). These differences in genome size are consistent with *L. crispatus* having access to a broader array of metabolic pathways.

6.5.3 Functional differences and similarities

Niche partitioning by *L. crispatus* and *L. iners* through conditional differentiation could be driven by differences in the functional makeup of the two species genomes. To investigate this possibility, we used the BlastKOALA function from the Kyoto Encyclopedia of Genes and Genomes (KEGG) to assign the core genes of both species to metabolic pathways and functions (Table 6.2). One might expect that given the larger genome size of *L. crispatus*, this species might have access to a broader array of metabolic functions. In may respects, our functional analysis confirms this expectation. While both *L. crispatus* and *L. iners* rely heavily on fermentation to generate energy, we found that they may differ in respects to the carbon sources they are capable of fermenting. In total, *L. crispatus* has 85 enzymes related to carbohydrate metabolism, whereas *L.iners* has only 59 enzymes (Table 6.2). Both species have the genetic capability to metabolize glucose, mannose, maltose, and trehalose. However, only *L. crispatus* has the genetic capability to ferment lactose, galactose, sucrose, and fructose (Figure 6.3). Our analysis also indicates that the two species differ in regard to the isomers of lactic acid that they can produce as end products of fermentation: *L. iners* can only produce L-lactic acid, while *L. crispatus* can produce L- and D-lactic acid. Furthermore, we found that the core genome of *L. crispatus* also contains the gene pyruvate oxidase which converts pyruvate into acetate, generating hydrogen peroxide in the process. These differences in the genetic potential for carbon metabolism may influence competitive interactions between these two species.

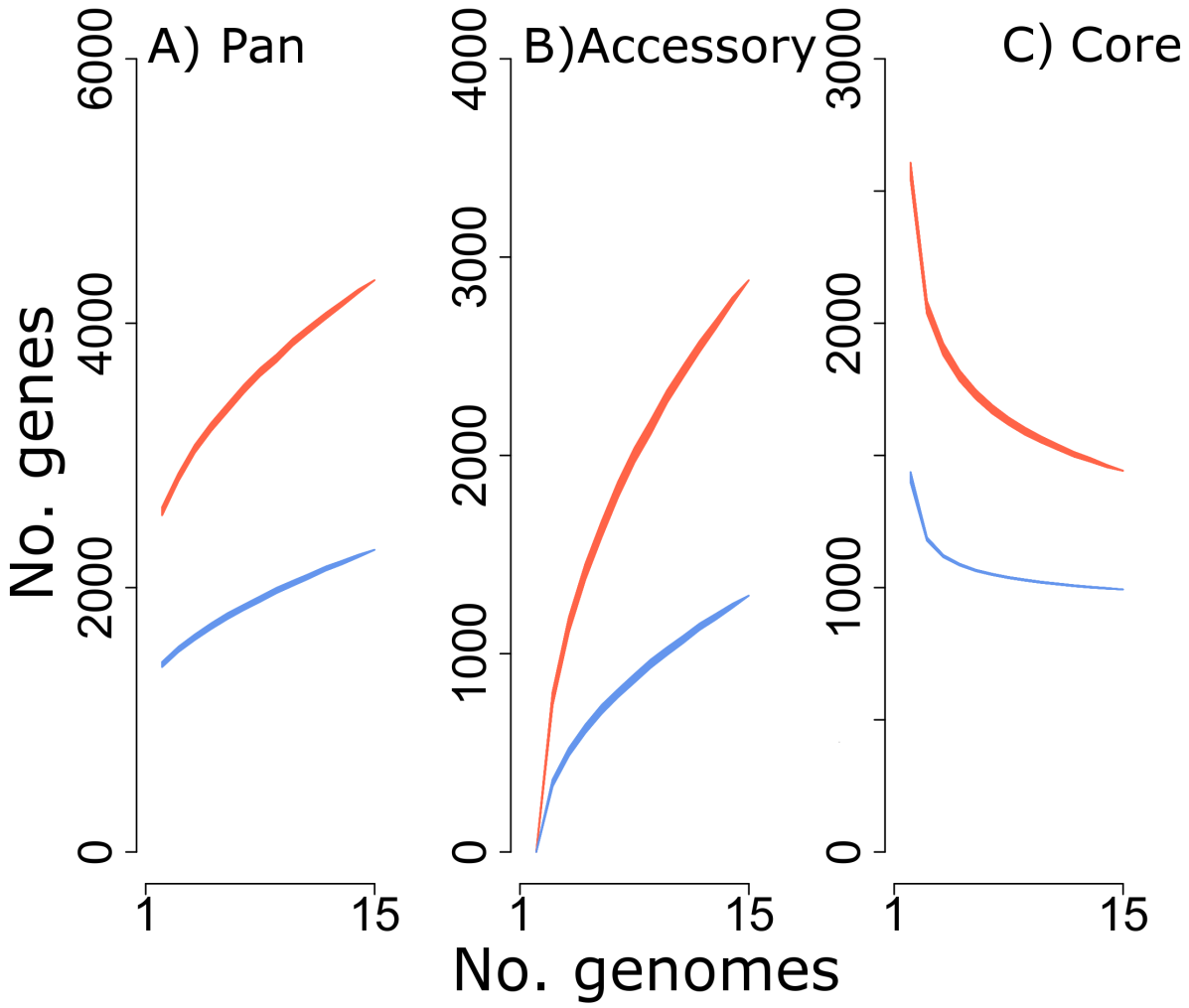

Figure 6.2: Pangenome, accessory-genome, and core-genome accumulation curves for *Lactobacillus crispatus* (red) and *Lactobacillus iners* (blue). Line thickness represents the 95% confidence interval around the mean.

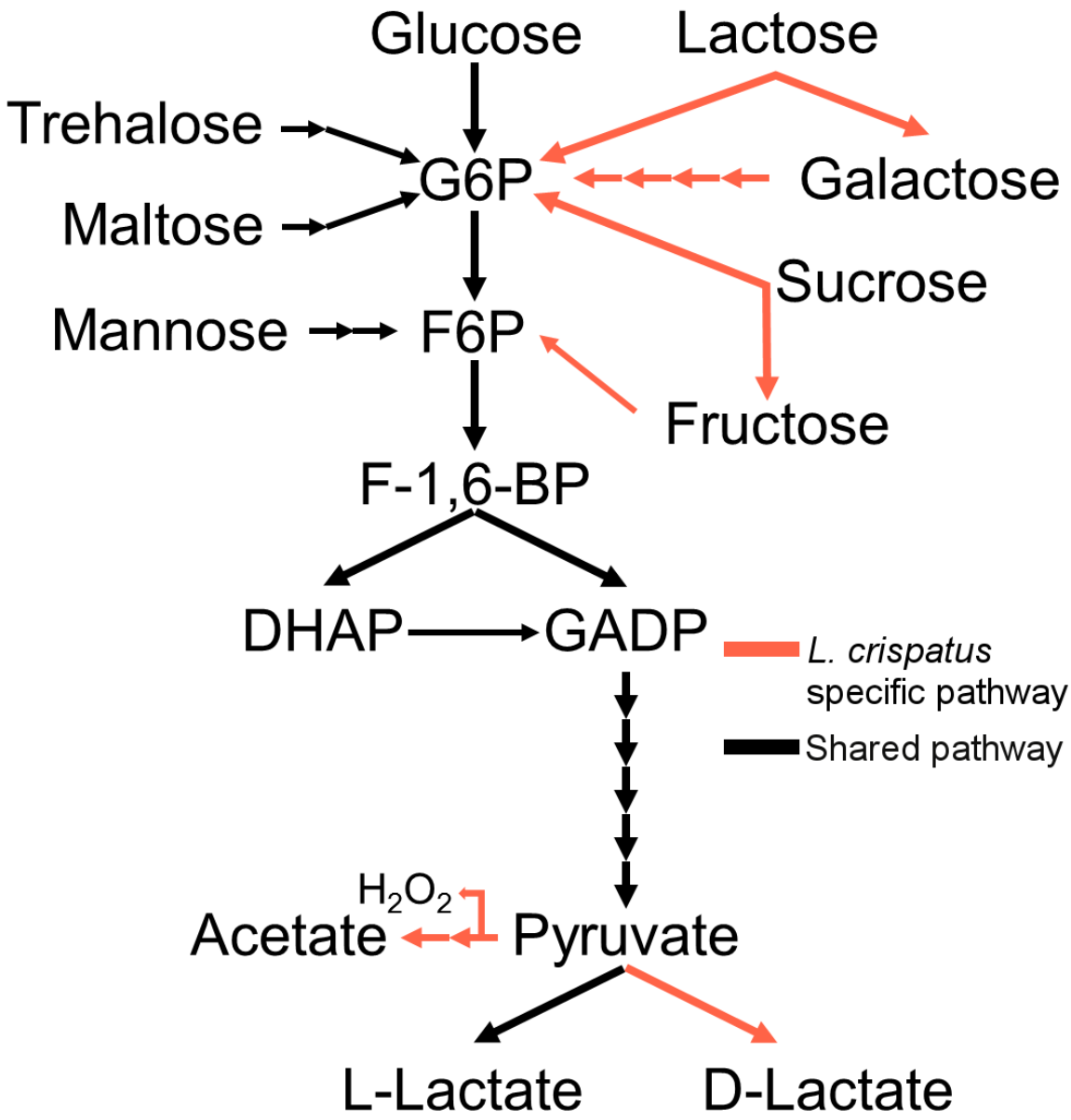

Figure 6.3: Graphical representation of the fermentation pathways encoded by the core genome of *L. crispatus* and *L. iners*. Included are those shared by *L. crispatus* and *L. iners* (black) and those unique to *L. crispatus* (red). G6P, glucose-6-phosphate; F6P, fructose 6-phosphate; DHAP, dihydroxyacetone phosphate; GADP, glyceraldehyde 3-phosphate.

Table 6.2: Functional category and metabolic pathways encoded by the core genome

Functional category/pathway[a]	No. found in core genes of: *L. crispatus*	*L. iners*
Carbohydrate metabolism	**85**	**59**
Glycolysis	17	14
Citric acid cycle	3	1
Pentose phosphate pathway	14	12
Fructose and mannose metabolism	18	14
Galactose metabolism	11	8
Starch and sucrose metabolism	16	10
Amino acid metabolism	**54**	**43**
Ala, Asp, and Glu metabolism	11	10
Gly, Ser, and Thr metabolism	9	3
Cys and Met metabolism	8	5
Lysine biosynthesis	12	4
Arginine biosynthesis	3	1
Lipid metabolism	**21**	**17**
Nucleic acid metabolism	**51**	**56**
Metabolism of cofactors and vitamins	**33**	**27**
Thiamine metabolism	5	3
Riboflavin metabolism	5	1
Vitamin B6 metabolism	2	1
Nicotinate metabolism	5	5
CoA biosynthesis	5	5
Folate biosynthesis	2	5
Membrane transporter	**70**	**54**
ABC transporter	39	31
Phosphate transport system	23	15
Bacterial secretion system	8	8
Replication and repair	**41**	**36**
DNA replication	14	14
Base excision repair	9	7
Nucleotide excision repair	7	7
Mismatch repair	16	15
Homologous recombination	19	19
Transcription	**4**	**5**
Translation	**78**	**79**
Peptidoglycan biosynthesis	**14**	**14**

[a] Entries in bold font represent functional categories while indented entries are specific metabolic pathways within each category. Enzymes can appear in multiple pathways but are only counted once in the functional category total. CoA, coenzyme A.

We found that *L. crispatus* and *L. iners* also differ in their repertoire of enzymes related to the biosynthesis and metabolism of amino acids. The core genome of *L. crispatus* encodes 54 different amino acid-related enzymes, while that of *L. iners* encodes only 43 enzymes (Table 6.2). More specifically, the core genome of *L. crispatus* has a complete pathway for the biosynthesis of lysine, while the *L. iners* core genome is almost completely devoid of these genes. *L. crispatus* also has more genes related to cysteine and methionine biosynthesis and glycine, serine, and threonine biosynthesis. However, we also found that the two species have similar numbers of genes related to alanine, aspartate, and glutamine metabolism (Table 6.2). In addition to the genes related to the biosynthesis of the essential amino acids, *L. crispatus*, but not *L. iners*, also has the genetic capability to transport and break down putrescine, a product of ornithine catabolism. These differences are consistent with *L. iners* being more reliant on exogenous sources of amino acids than *L. crispatus.*

Bacterial cells rely on transport proteins to import extracellular supplies of carbohydrates, amino acids, nucleic acids, and inorganic ions. If *L. iners* truly relies more heavily on exogenous sources of nutrients, one might expect that its core genome would contain more transport proteins. However, our analysis indicates that the reverse is true. The core genome of *L. crispatus* contains eight more ABC transporter genes and eight more phosphate transport system genes (Table 6.2). These additional ABC transport genes include a complete phosphonate transporter, an oligopeptide transporter, and an iron transporter. The unique presence of this iron transport system in *L. crispatus* may allow the species to more effectively sequester iron. Interestingly, we found that the core genome of *L. iners* has an ABC-type zinc transporter and an osmoprotectant transporter that are not found in *L. crispatus.* Additionally, the core genome of these two species also contain a number of phosphotransferase system (PTS) genes. *L. crispatus* and *L. iners* share fructose, galactose, glucose, maltose, mannose, and trehalose PTS genes, while only *L. crispatus* has the sucrose PTS gene. This result is consistent with the unique capability of *L. crispatus* to ferment sucrose. Our functional analysis also indicates that *L. crispatus* and *L. iners* are functionally similar in many respects (Table 6.2). This is not surprising given that the two species are closely related (Figure 6.1) and that many metabolic pathways are necessary to life. The two species have similar numbers of genes related to the metabolism of lipids, nucleic acids, and cofactors. However, the two species vary somewhat in cofactor pathways: *L. crispatus* has more genes related to the metabolism of riboflavin, while *L. iners* has more genes related to the metabolism of folate. The two species also have similar numbers of genes related to peptidoglycan biosynthesis, transcription, and translation, as well as replication and repair of their chromosome. Included in these replication and repair proteins are complete sets of base excision repair, nucleotide repair excision, mismatch repair, and homologous recombination proteins.

In addition to a core genome, bacterial species also have an accessory genome that contains genes that are present in some but not all strains of a given species. We first split the accessory genome into unique genes, which are those present in only one strain, and variable genes, which are those present in multiple strains. We then characterized the function of the variable genes using the same BlastKOALA approach described above. Unfortunately, this approach only annotated about 16% of this gene set, likely due to

its focus on well-annotated metabolic pathways. However, our analysis indicated that the two species differ in the functional makeup of their variable genome (Figure 6.4). We found that *L. crispatus* has more variable genes in every functional category except translation, where *L. iners* has slightly more, and replication and repair, where the two species have similar numbers. Specifically, we found that two *L. crispatus* strains contain a complete pathway for the metabolism of L-rhamnose, a deoxy-sugar commonly found in the outer membrane of some bacterial species. Another seven *L. crispatus* strains have the genetic capability to break down raffinose, an oligosaccharide common in plant matter. Additionally, four of the *L. crispatus* strains contained both genes needed for a functional multidrug transport protein. In comparison, our analysis of the *L. iners* variable genes revealed lactose and galactose transporters in only four and six strains, respectively. However, we did not detect the additional enzymes required to route these sugars into central metabolism.

We also functionally characterized the genes that are unique to individual *L. crispatus* and *L. iners* strains, although we found that the BlastKOALA approach we previously used to annotate the core and variable genes was not well suited for this task. Instead, we used Blast searches against NCBI's nonredundant protein database to predict the function of these unique genes. Our analysis indicated that the set of unique *L. crispatus* genes is enriched for transposable elements (Figure 6.5A, $P<0.001$). On average, each *L. crispatus* strain contained approximately 12 unique transposons or integrases in their genome compared to <1 per *L. iners* genome. In contrast, we found that the set of unique genes for *L. iners* was enriched for restriction-modification enzymes (Figure 6.5B, $P<0.001$). While each *L. iners* genome contained approximately three unique restriction-modification enzymes, each *L. crispatus* genome contained only one. Other notable differences between the unique proteins of *L. crispatus* and *L. iners* include that *L. crispatus* has twice as many amino acid metabolism genes and three times as many carbohydrate metabolism genes, while *L. iners* has twice as many replication and repair genes.

6.5.4 Evolutionary differences

Thus far, we have focused on describing differences between these two species based on the presence or absence of specific genes and metabolic pathways. However, further analysis of the sequence diversity within a gene can be used to gain insight into its evolution. These evolutionary analyses can predict the strength of selection acting on the gene and can identify whether a gene has been horizontally transferred into the genome. We characterized the genetic diversity present in the core genes of *L. crispatus* and *L. iners* using the program TimeZone. Our analysis identified 35,992 single nucleotide polymorphisms (SNPs) across 890 *L. crispatus* core genes and 32,089 SNPs across 791 *L. iners* core genes. Using these SNPs, we estimated the nucleotide diversity (the average number of polymorphisms per base pair of a gene between any two randomly selected sequences) for each gene. Comparisons of the mean nucleotide diversity for each species indicated that the core genome of *L. crispatus* is more diverse than the core genome of *L. iners* (Figure 6.6, 0.0173 versus 0.0114, respectively; Welch's t test; $t=7.0$; $P<0.001$). However, closer analysis of these data revealed that this result was primarily driven by a series of highly diverse *L. crispatus* genes, including several uncharacterized proteins and transposable elements. A non-

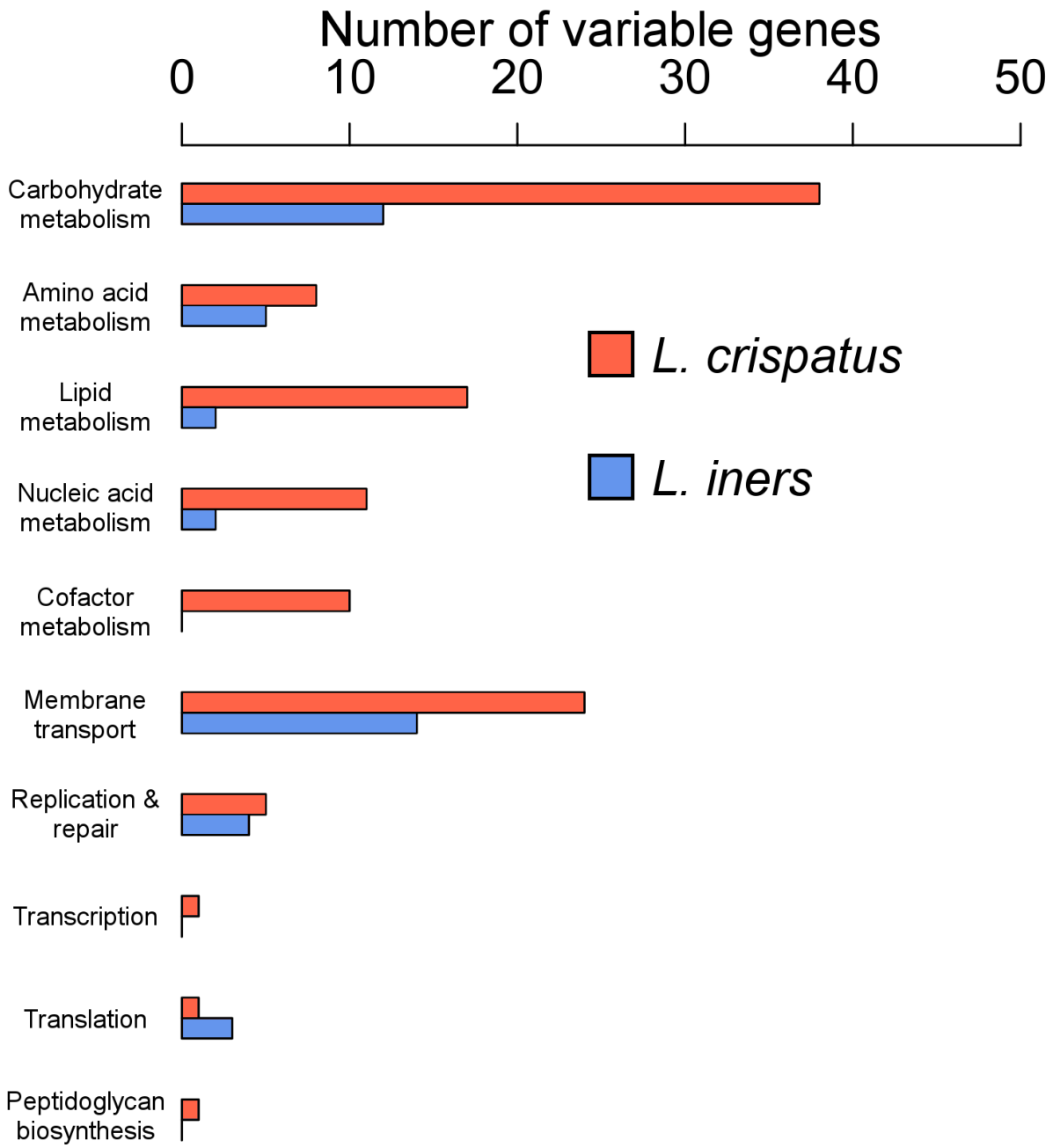

Figure 6.4: Broad functional characteristics of the variable accessory genes found in *L. crispatus* (red) and *L. iners* (blue).

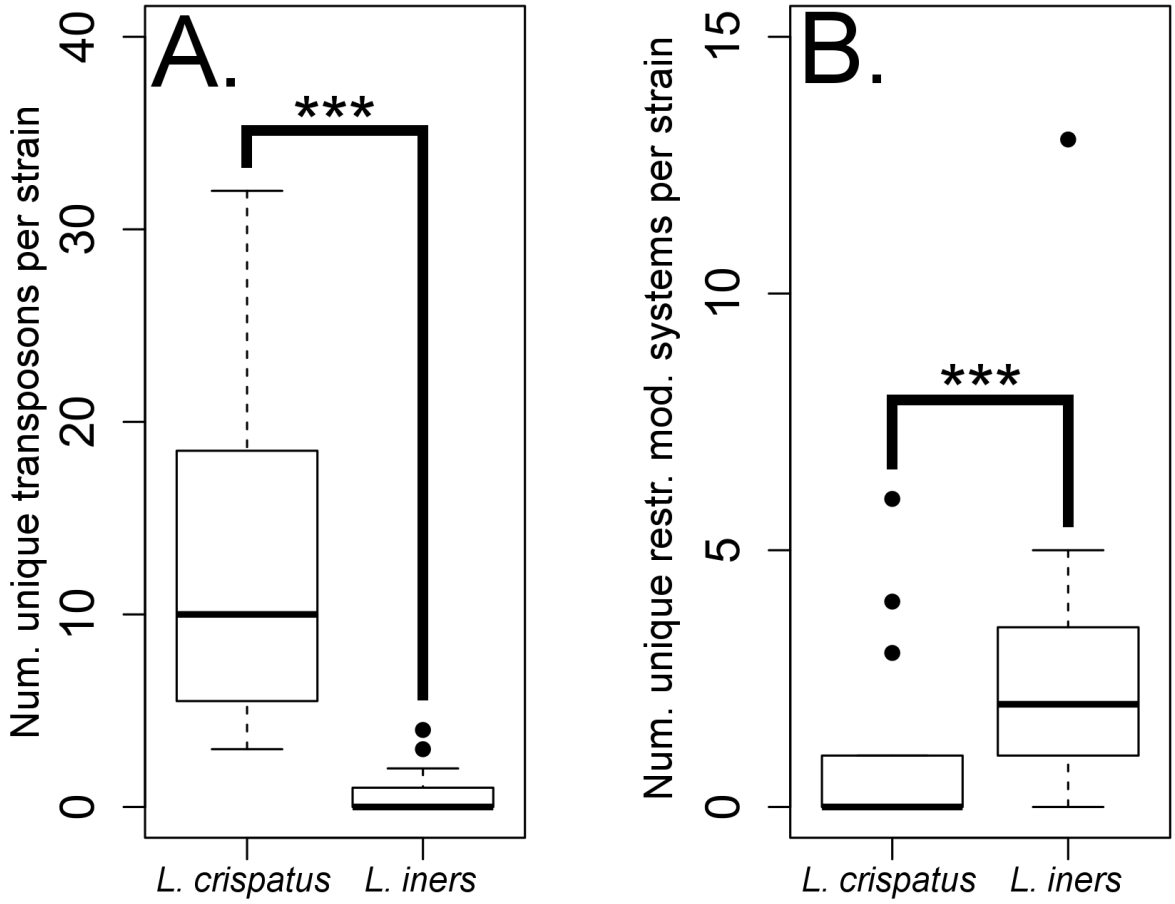

Figure 6.5: Boxplot comparisons of the number of transposable elements (A) and restriction-modification (restr. mod.) enzymes (B) per strain in the set of unique *L. crispatus* and *L. iners* genes. Boxes span the first and third quartiles, with the inner line representing the median value. Whiskers represent 1.5X the length of the inner quartile range, and points are outliers. Significance was assessed using a two-sided Poisson comparison (***, P<0.001).

parametric Wilcoxon rank sum test demonstrated that the median nucleotide diversity is greater for *L. iners* than *L. crispatus* (0.010 versus 0.006, respectively; w=26,218; P<0.001).

The strength of selection acting on a gene can be inferred based on the ratio of nonsynonymous to synonymous nucleotide changes (dN/dS). Under neutral evolution conditions, the rate of nonsynonymous changes is expected to be equal to the rate of synonymous changes, and their ratio is expected to be one. Values above one indicate an overabundance of nonsynonymous changes, which is usually interpreted as a signature of positive selection, while those under 1 indicate a lack of such changes and usually interpreted as a signature of purifying selection. Of the identified SNPs in the core genes, only 8,946 SNPs (24.8%) and 7,646 SNPs (23.8%) were nonsynonymous mutations in *L. crispatus* and *L. iners*, respectively. The vast majority of the core genes for both species have dN/dS ratios below the neutral expectation of one, indicating that these genes are likely under strong purifying selection (Figure 6.6). This result is not unexpected given that the core genes typically carry out essential cellular processes (e.g., transcription, translation, and central metabolism). Our analysis indicates that the median dN/dS ratio for the *L. crispatus* core genome was 0.126, roughly twice the observed median for *L. iners* (Figure 6.6; 0.069; Wilcoxon rank sum test, w=265,218; P<0.001). This result is consistent with purifying selection being stronger on the core genome of *L. iners* than on the core genome of *L. crispatus*.

Although the majority of the core genes from both species bear the signature of strong purifying selection (low dN/dS), there is a number which shows signatures of positive selection (dN/dS, >1; Figure 6.6). We argue that insight into the ecology of the two species in the vaginal environment can be drawn from the products these candidate positively selected genes encode (Table 6.3). For example, our analysis indicates that the riboflavin synthesis protein RibF may be subject to positive selection in *L. crispatus*. This is interesting because *L. crispatus*, but not *L. iners*, has the riboflavin biosynthesis pathway encoded in its core genome. Another *L. crispatus* core gene with a high dN/dS ratio encodes an enterocin A immunity protein that protects its carrier against the activity of a bacteriocin (O'Keeffe et al., 1999). Our analysis also indicated that the gene encoding the lactocepin S-layer protein (LCRIS_RS05305) had nine nonsynonymous mutations, with only one synonymous (Table 6.3, dN/dS=2.5). This protein has been shown to degrade human proinflammatory chemokines in other *Lactobacillus* species (von Schillde et al., 2012). Other candidate positively selected genes in *L. crispatus* can be found in Table 6.3. In comparison, we found that *L. iners* had only two genes, with dN/dS values of 2 (Table 6.3). Unfortunately, we were unable to annotate either gene beyond "uncharacterized protein."

Bacterial species can also evolve through the acquisition of foreign DNA through horizontal gene transfer events. The acquisition of foreign DNA can have a dramatic impact on the ecology of bacterial strains and species (Gyles and Boerlin, 2014). We used Blast searches against the nonredundant protein database to identify candidate horizontally acquired genes in the core and accessory genomes of *L. iners* and *L. crispatus*. If the Blast searches demonstrated that the gene is present in other *Lactobacillus* species, we inferred that it is likely to have been inherited, whereas if the gene is present only in non-*Lactobacillus* species, we considered it to have been horizontally acquired. Our analysis indicates that only one of the

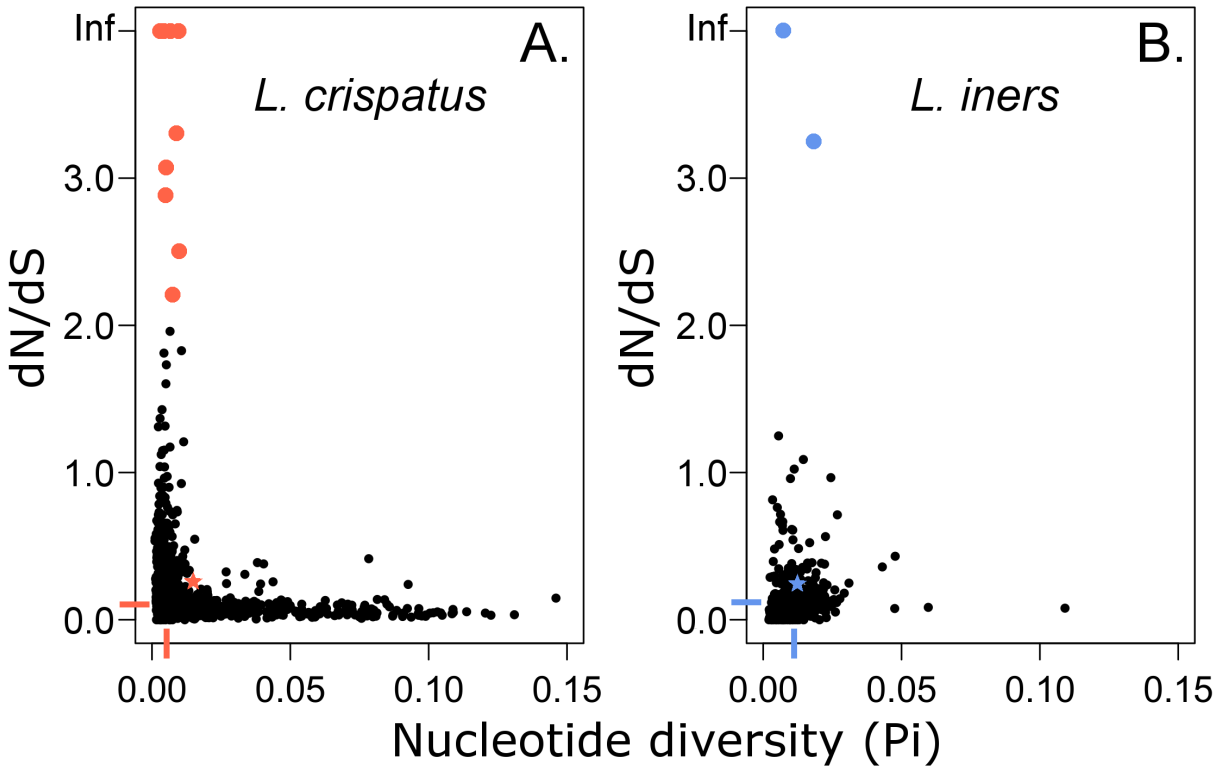

Figure 6.6: Scatterplot of nucleotide diversity (x axis) versus dN/dS ratio (y axis) for the core genes of *L. crispatus* (A) and *L. iners* (B). Nucleotide diversity was estimated as the average number of pairwise nucleotide differences per site (Pi). Stars in the plot represent the average values for these two parameters, while the colored lines on the axes represent the median values.

Table 6.3: Core genes exhibiting dN/dS>2

Species	Gene name	Annotation	Length (nt)	No. of synonymous changes	No. of nonsynonymous changes	dN/dS[a]
L. crispatus	*comEB*	dCMP deaminase	159	0	4	inf
	LCRIS_RS01075	β-Propeller of dehydrogenase	194	0	3	inf
	purR	PUR operon repressor	276	0	9	inf
	psiE	Phosphate starvation	135	0	7	inf
	ribF	Riboflavin biosynthesis	316	0	6	inf
	LCRIS_RS09970	Haloacid dehydrogenase	257	0	5	inf
	LCRIS_RS09110	Enterocin A immunity	107	0	4	inf
	LCRIS_RS05415	Homocysteine methyltransferase	76	0	3	inf
	LCRIS_RS09270	tet(R) transcription regulator	174	1	9	3.307
	LCRIS_RS03325	α/β-Superfamily hydrolase	306	1	10	3.072
	LCRIS_RS06755	α/β-Superfamily hydrolase	291	1	8	2.871
	LCRIS_RS05305	Lactocepin S-layer protein	166	1	9	2.504
	LCRIS_00991	D-Gluconic acid reductase	278	1	11	2.209
L. iners	RS0100505	Uncharacterized protein	94	0	4	inf
	RS06445	Uncharacterized protein	75	1	9	3.25

[a]inf, infinity

unique core genes of *L. crispatus* is likely to have been horizontally acquired, while we found *L. iners* to have 14 such genes. The single *L. crispatus* unique core gene matched glycosyltransferase found in the *Chlamydia trachomatis* genome. The 14 core genes in *L. iners* that were potentially acquired by horizontal gene transfer matched genes in *Gardnerella vaginalis* (n=4), *Chlamydia trachomatis* (n=2), *Aerococcus christensenii* (n=2), *Parvimonas micra*, *Facklamia hominis*, *Finegoldia magna*, *Streptococcus sp.*, and *Enterococcus faecium*. Most of these species are commonly identified in the human vagina, further reinforcing the notion that they may have been horizontally acquired. These 14 genes include several toxin-antitoxin proteins, a zinc and a phosphate transporter, two DNA repair proteins, and several uncharacterized proteins. Furthermore, our analysis indicated that the cytolysin gene of *L. iners* is also likely to have been horizontally acquired. We found that the *L. iners* sequence for this gene most closely matches cytolysins identified in *G. vaginalis* and various *Streptococcus* species. We extracted these matching sequences from the database and constructed a maximum likelihood tree to identify their phylogenetic relationships (Figure 6.7). Our analysis indicated that the *L. iners* cytolysin is most closely related to the *G. vaginalis* cytolysin but has diverged substantially in sequence since being acquired by *L. iners*.

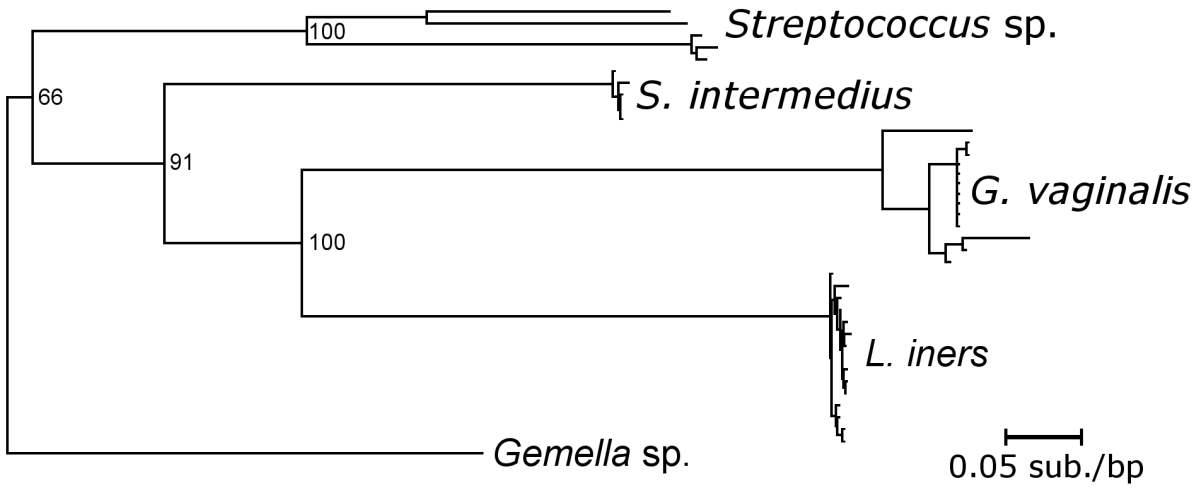

Figure 6.7: Maximum likelihood tree of the phylogenetic relationships between the *L. iners* cytolysin gene and the closest matching sequences from the NCBI nonredundant nucleotide database. Phylogeny is rooted with the single *Gemella sp.* sequence as the outgroup. sub., substitutions.

Genes that appear in only a single strain of a species are also likely to have been horizontally acquired. We used the same Blast approach to identify these genes and to determine the closest matching sequence in the database. Our analysis indicated that the majority of the unique genes for both species mostly closely matched genes from other *Lactobacillus* species (*L. crispatus*, 68.7%; *L. iners*, 50.1%). However, we also found that some of these unique accessory genes, both from *L. crispatus* and *L. iners*, did not match genes of other *Lactobacillus* species but instead matched genes from other genera. Notably, we found *L. iners* to have twice as many such genes as *L. crispatus* (n=156 versus n=78, respectively), as well as a more diverse pool of genera to which these sequences matched. These 78 *L. crispatus* genes

matched sequences from three different genera: *Chlamydia* (n=59), *Mycoplasma* (n=6), and *Streptococcus* (n=3). The other 10 *L. crispatus* genes matched sequences identified in genomes of multiple genera. In comparison, the 156 *L. iners* genes matched sequences from six different genera: *Streptococcus* (n=25), *Chlamydia* (n=20), *Anaerococcus* (n=14), *Gardnerella* (n=13), *Peptoniphilus* (n=13), and *Atopobium* (n=7). Another 84 of these *L. iners* singleton genes matched sequences identified in genomes of multiple genera. Species from all of the identified genera are routinely found in vaginal samples, supporting the notion that these genes have been acquired through horizontal transfer events.

6.6 Discusion

Lactobacillus crispatus and *L. iners* are both common inhabitants of the healthy human vagina. These two species are closely related and perform similar ecological functions, namely, the production of lactic acid. They are rarely found to coexist for extended periods of time, and transitions between an *L. crispatus*-dominated community and one dominated by *L. iners* are common, making it likely that there is substantial overlap in their fundamental niches. Ecological theory predicts that two species cannot occupy the same niche indefinitely (Gause, 1934), making it unclear how these two species have been maintained as common inhabitants of the human vagina. In this study, we characterized and compared the genomes of *L. crispatus* and *L. iners* to identify possible ecological factors that drive niche partitioning by these species. Our results highlight several key differences in the genomes of these two species that we believe may influence their ecology in the vaginal environment.

The typical *L. iners* genome is almost half the size of the *L. crispatus* genome. We have shown that this is likely a derived trait unique to *L. iners* (Figure 6.1). The reduced genome size of *L. iners* limits the number of proteins encoded by both its core and accessory genomes, which corresponds to reduced metabolic capabilities. Our analysis demonstrated that *L. iners* is likely capable of fermenting fewer carbon sources than *L. crispatus* and lacks more of the machinery necessary to synthesize essential amino acids. With fewer metabolic pathways available, *L. iners* likely relies more heavily on exogenous sources for essential resources than *L. crispatus*. Their dependence on nutrients derived from the host or other community members is likely facilitated by the species' ability to bind to human fibronectin, which allows it to maintain close contact with host tissues (McMillan et al., 2012). The limited genetic repertoire of *L. iners* likely makes the species more sensitive to environmental fluctuations. Accordingly, temporal studies on the vaginal microbiome have indicated that communities dominated by this species may be more unstable than those dominated by *L. crispatus* (Gajer et al., 2012; Jakobsson and Forsum, 2007).

The process of genome reduction is widespread among host-associated bacterial species and can lead to genomes that are less than 250 kb in size (Andersson and Kurland, 1998; Moran and Wernegreen, 2000; McCutcheon and Moran, 2012). We do not completely understand the driving force behind this process, although several mechanisms have been proposed. While we cannot definitively determine what drove the reduction of the *L. iners* genome, we can discuss the consistency of our data with the various proposed mechanisms. First, Muller's Ratchet posits that frequent bottlenecks and a lack of recombination allow

slightly deleterious mutations to accumulate in genes, leading to their deterioration and subsequent loss (McCutcheon and Moran, 2012; Moran, 1996). However, we argue that our results are not consistent with this hypothesis, as almost all of the remaining core genes of *L. iners* exhibit signatures consistent with strong purifying selection. Furthermore, we showed that *L. iners* has retained the genes required for several DNA repair pathways, including homologous recombination. Instead, we argue that our results are more consistent with the Black Queen hypothesis, which argues that selection for streamlined genomes and a loss of functional redundancy may be common for host-associated and free-living bacteria that experience relatively constant environments (Bliven and Maurelli, 2012; Morris et al., 2012). Additional support for this hypothesis is provided by the species' limited accessory genomes, which are enriched for genes encoding mechanisms by which bacteria can resist the integration of foreign DNA into their genome (e.g., restriction-modification enzymes) (Bickle and Krüger, 1993; Snyder, 1995).

While the current *L. crispatus* genome is larger than that of *L. iners*, it is smaller than those of other *Lactobacillus* species, particularly those that are not host associated. We argue that both species have reduced their genome size and differ only in the magnitude of this effect. Some of our results even suggest that *L. crispatus* may still be undergoing this process of genome reduction. We have shown that the accessory genome of this species is bloated with transposable elements. Other studies have demonstrated that the accumulation of mobile genetic elements often precedes genome reduction (McCutcheon and Moran, 2012; Toh et al., 2006; Burke and Moran, 2011). This effect is thought to be driven by the reduction in the effective population size that often accompanies the transition from being free living to becoming host associated. A smaller effective population size can allow for the fixation of mildly deleterious mutations, like the incorporation of selfish DNA into the genome (Moran, 1996). Later, when these selfish genes are purged from the genome, they may take with them adjacent stretches of the genome. The abundance of transposable elements in the genomes of *L. crispatus* suggest that strains of this species might be currently undergoing genome reduction. This hypothesis is further supported by the genome of *L. crispatus* EM-C1, which is already 400 kb smaller than the others included in this study. In contrast with this evidence, our results also indicate that most of the genes in the *L. crispatus* core genome are experiencing strong purifying selection, suggesting that selection is still acting efficiently on this species. Only time and further observation will reveal whether *L. crispatus* is evolving along this trajectory.

Lactobacillus crispatus and *L. iners* rely on the fermentation of carbon-containing substrates to produce energy for the cell (Carr et al., 2002). The primary source of carbon and energy in the vaginal ecosystem is thought to be glycogen produced by host epithelial tissues (Cruickshank and Sharman, 1934; Milwidsky et al., 1980). However, the core genome of neither *L. crispatus* nor *L. iners* contains the enzymes necessary to degrade this polysaccharide. The two species must therefore rely on either the host (Nasioudis et al., 2015; Spear et al., 2014) or other bacterial species for the initial breakdown of glycogen. Our analysis indicated that the two species do have the genetic capability to ferment a number of other sugars, including several glycogen breakdown products. The two species share the ability to ferment glucose, trehalose, maltose, and mannose, of which glucose and maltose are common glycogen breakdown products. In addition to these shared functions, the core genome of *L. crispatus* also includes

the enzymes necessary to ferment lactose, galactose, fructose, and sucrose. A previous study showed that the application of a sucrose gel can select for *L. crispatus* in the *Rhesus macaque* vagina (Hu et al., 2015), lending credence to this result. Furthermore, fructose can be found in high abundance in male ejaculate (Owed and Katz, 2005). We speculate that the abundance of these sugars in the vaginal environment may influence competitive interactions between *L. crispatus* and *L. iners*.

One unique trait that may influence the ecology of *L. iners* is its ability to produce inerolysin, a pore-forming cytolysin. Previous studies have demonstrated that this cytolysin is similar in sequence to those produced by *Gardnerella vaginalis* and *Streptococcus intermedius* (Rampersaud et al., 2011). Our analysis confirms this result and is consistent with this gene having been horizontally acquired by *L. iners*. The gene was not identified in any other *Lactobacillus* species included in this study, and our Blast searches against the nonredundant database did not reveal any matching *Lactobacillus* sequences. Our phylogeny shows that the inerolysin gene is most closely related to that found in *Gardnerella vaginalis*, although its sequence is heavily divergent. This derived trait of *L. iners* may allow it to liberate resources from host cells (Rampersaud et al., 2011). We speculate that this may give *L. iners* a competitive advantage in the vaginal environment when nutrients are scarce and the ability to liberate them directly from host tissue is favored. Indeed, microbial surveys have suggested that *L. iners* is capable of persisting under other potentially adverse conditions in the vagina (Ferris et al., 2007; Zozaya-hinchliffe et al., 2010; Mayer et al., 2015). Additionally, because the glycogen content of the vaginal epithelium is linked to circulating estrogen levels (Cruickshank and Sharman, 1934), the abundance of nutrients in the vagina may vary across the female reproductive cycle as well as through a woman's lifetime. If *L. iners* does indeed have a competitive advantage in times of low nutrient abundance, it may also be selected for during times of low circulating estrogen.

We, along with others (Rampersaud et al., 2011), have demonstrated that the cytolysin of *L. iners* is likely to have been horizontally acquired from a species outside *Lactobacillus*. Our analysis also identified several other core and accessory *L. iners* genes that are likely to have been horizontally acquired. These candidate horizontally acquired genes were found to most closely match those of species in other genera, including *Chlamydia*, *Streptococcus*, *Parvimonas*, *Gardnerella*, and *Atopobium*. In particular, several of these genes, including the cytolysin, match most closely to genes found in *Gardnerella vaginalis*. Historically, this species has been associated with bacterial vaginosis, a condition that increases the risk of women to preterm birth, sexually transmitted infections, and other adverse sequelae. However, several recent studies have highlighted that *G. vaginalis* is frequently found in healthy asymptomatic vaginal samples from women of all ages (Ravel et al., 2011; Hillier and Lau, 1997; Hickey et al., 2015). In comparison, we identified fewer candidate horizontally acquired genes in *L. crispatus*, and those that we did find matched sequences in a smaller subset of genera. We argue that while there is ongoing gene flow between both *L. iners* and *L. crispatus* and species from these other genera, the rate at which this occurs may be higher for *L. iners*. Furthermore, more of the genes acquired by *L. iners* are conserved across all of the strains analyzed, indicating that they may have been selectively favored in the evolution of the species. We speculate that the acquisition of these genes (e.g., cytolysin, zinc and phosphate trans-

porters, and toxin antitoxin system genes) have shaped the ecology of *L. iners* in the vaginal environment.

Clinicians have long considered the dominance of *L. crispatus* in the vagina to be associated with good health, while only recently have researchers also demonstrated that healthy women can be colonized by *L. iners* (McMillan et al., 2012). Our analysis pointed out several differences in the genomes of *L. crispatus* and *L. iners* that may influence how the species influence host health. Recent studies have demonstrated that the isomers of lactic acid have differential effects on the host immune system (Witkin et al., 2013). We have shown that while *L. crispatus* can produce D- and L-lactic acid, *L. iners* has the capability to produce L-lactic acid only. *L. iners* also does not have the genetic capability to produce hydrogen peroxide via oxidation of pyruvate, a pathway which we and others (Ojala et al., 2014) have demonstrated is available to *L. crispatus*. The production of hydrogen peroxide is thought to be one mechanism by which *Lactobacillus* species can prevent anaerobes from colonizing the vagina (Felten et al., 1999). *L. crispatus* also has the capability to breakdown putrescine, a malodorous amino acid commonly found in vaginal secretions during episodes of bacterial vaginosis (Wolrath et al., 2001). We found the core genome of *L. crispatus* to contain an iron transport system that is absent in the core genome of *L. iners*. This transport system may allow *L. crispatus* to sequester the iron released by the host during menses (Hallberg et al., 1966), thereby preventing other species, including vaginal pathogens, from acquiring this vital resource. We also found that the core genome of *L. crispatus* contains a gene encoding lactocepin, a serine protease that has been shown to degrade the proinflammatory chemokine interferon-gamma-inducible protein 10 (IP-10) (von Schillde et al., 2012). Our analysis of this gene indicated that it may be experiencing positive selection in *L. crispatus*, which could reflect adaptation via changes in this function. In vitro tests have demonstrated that colonization of vaginal epithelial cells with *L. iners* resulted in a more-proinflammatory signaling response from the host tissue than colonization by *L. crispatus* (Doerflinger et al., 2014). However, it has yet to be determined if the lactocepin produced by *L. crispatus* is capable of preventing or reducing inflammation in the vagina. Based on these two results,we conclude that the presence of *L. crispatus* in the vagina may offer more protective benefits to the host than *L. iners*.

As closely related frequent colonizers of the human vagina, *L. crispatus* and *L. iners* are likely to compete for shared niche space. We have shown that the two species share the key fermentation pathways needed to metabolize glycogen breakdown products (glucose and maltose). However, their genomes differ in many other respects, likely providing them with differential competitive abilities across the range of environmental conditions common to the human vagina. For example, should sucrose or fructose be introduced into the vagina, *L. crispatus* may be selected for given its unique capability to ferment these sugars, whereas *L. iners* may be favored when nutrients are rare and its cytolysin is needed to liberate them from host tissues. More study is needed to detail the biotic and abiotic factors (e.g., pH, nutrient availability, and coinhabiting microbial species) under which *L. iners* and *L. crispatus* are favored. Additionally, this study is limited in scope given that it only focused on two of the four common vaginal lactobacilli. Analysis of *L. gasseri* and *L. jensenii* is needed to gain a more complete picture of niche partitioning and competitive interactions in the human vagina. Identifying the conditions under which particular vaginal *Lactobacillus* species flourish will enable a deeper understanding of the complicated

temporal dynamics of the vaginal microbiome.

6.7 Acknowledgments

We thank three anonymous reviewers for their insightful suggestions.
This research was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under grant U19AI084044 and by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant P30 GM103324.

6.8 Funding Information

This work, including the efforts of LARRY J. FORNEY, was funded by HHS|NIH| National Institute of Allergy and Infectious Diseases (NIAID) (U19AI084044). This work, including the efforts of LARRY J. FORNEY, was funded by HHS|NIH|National Institute of General Medical Sciences (NIGMS) (P30 GM103324).

References

Adams, J., Rosenzweig, F., 2014. Experimental microbial evolution: History and conceptual underpinnings. Genomics 104, 393–398.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., Lipman, D. J., 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research 25 (17), 3389–3402.

Amarasekare, P., 2003. Competitive coexistence in spatially structured environments: A synthesis. Ecology Letters 6, 1109–1122.

Amato, S. M., Fazen, C. H., Henry, T. C., Mok, W. W. K., Orman, M. A., Sandvik, E. L., Volzing, K. G., Brynildsen, M. P., 2014. The role of metabolism in bacterial persistence. Frontiers in Microbiology 5 (70).

Anderson, J. T., Lee, C.-r., Rushworth, C. A., Colautti, R. I., Mitchell-Olds, T., 2013. Genetic trade-offs and conditional neutrality contribute to local adaptation. Molecular Ecology 22, 699–708.

Anderson, P., Roth, J., 1981. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rrna (rrn) cistrons. Proceedings of the National Academy of Sciences 78 (5), 3113–3117.

Andersson, D. I., oct 2003. Persistence of antibiotic resistant bacteria. Current Opinion in Microbiology 6 (5), 452–456.

Andersson, D. I., Hughes, D., 2010. Antibiotic resistance and its cost: Is it possible to reverse resistance? Nature Reviews Microbiology 8 (4), 260–71.

Andersson, D. I., Levin, B. R., 1999. The biological cost of antibiotic resistance. Current Opinion in Microbiology 2 (5), 489–93.

Andersson, S. G., Kurland, C. G., 1998. Reductive evolution of resident genomes. Trends in Microbiology 6 (7), 263–268.

Arias, C. A., Panesso, D., McGrath, D. M., Qin, X., Mojica, M. F., Miller, C., Diaz, L., Tran, T. T., Rincon, S., Barbu, E. M., Reyes, J., Roh, J. H., Lobos, E., Sodergren, E., Pasqualini, R., Arap, W., Quinn, J. P., Shamoo, Y., Murray, B. E., Weinstock, G. M., 2011. Genetic basis for in vivo daptomycin resistance in enterococci. New England Journal of Medicine 365 (10), 892–900.

Atwood, K. C., Schneider, L. K., Ryan, F. J., 1951. Periodic selection in *Escherichia coli*. Proceedings of the National Academy of Sciences 37 (3), 146–155.

Bååth, E., 1998. Growth rates of bacterial communities in soils at varying pH: A comparison of the thymidine and leucine incorporation techniques. Microbial Ecology 36 (3), 316–327.

Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B., Patel, N. H., 2007. Evolution, 1st Edition. Cold Spring Harbor Laboratory Press, New York.

Battin, T. J., Kaplan, L. a., Denis Newbold, J., Hansen, C. M. E., 2003. Contributions of microbial biofilms to ecosystem processes in stream mesocosms. Nature 426 (6965), 439–442.

Battin, T. J., Sloan, W. T., Kjelleberg, S., Daims, H., Head, I. M., Curtis, T. P., Eberl, L., 2007. Microbial landscapes: New paths to biofilm research. Nature Reviews Microbiology 5 (1), 76–81.

Becker, J. M., Parkin, T., Nakatsu, C. H., Wilbur, J. D., Konopka, A., 2006. Bacterial activity, community structure, and centimeter-scale spatial heterogeneity in contaminated soil. Microbial Ecology 51 (2), 220–231.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57 (1), 289–300.

Besemer, K., Singer, G., Hödl, I., Battin, T. J., 2009. Bacterial community composition of stream biofilms in spatially variable-flow environments. Applied and Environmental Microbiology 75 (22), 7189–7195.

Bickle, T. A., Krüger, D. H., 1993. Biology of DNA restriction. Microbiological Reviews 57 (2), 434–450.

Bigelow, J. L., Dunson, D. B., Stanford, J. B., Ecochard, R., Gnoth, C., Colombo, B., 2004. Mucus observations in the fertile window: A better predictor of conception than timing of intercourse. Human Reproduction 19 (4), 889–892.

Bjelland, S., oct 2003. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. Mutation Research 531 (1-2), 37–80.

Bliven, K. A., Maurelli, A. T., 2012. Antivirulence genes: Insights into pathogen evolution through gene loss. Infection and Immunity 80 (12), 4061–4070.

Boles, B. R., Singh, P. K., 2008. Endogenous oxidative stress produces diversity and adaptability in biofilm communities. Proceedings of the National Academy of Sciences 105 (34), 12503–12508.

Boles, B. R., Thoendel, M., Singh, P. K., 2004. Self-generated diversity produces "insurance effects" in biofilm communities. Proceedings of the National Academy of Sciences 101 (47), 16630–16635.

Brauner, A., Fridman, O., Gefen, O., Balaban, N. Q., 2016. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nature Reviews Microbiology 14 (5), 320–30.

Bridges, B., 1992. Mutagenesis after exposure of bacteria to ultraviolet-light and delayed photoreversal. Molecular and General Genetics 233 (3), 331–336.

Brito, I. L., Alm, E. J., 2016. Tracking strains in the microbiome: Insights from metagenomics and models. Frontiers in Microbiology 7, 712.

Brooks, J. D., Flint, S. H., 2008. Biofilms in the food industry: Problems and potential solutions. International Journal of Food Science and Technology 43 (12), 2163–2176.

Brown, C. J., Todd, K. M., Rosenzweig, R. F., 1998. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. Molecular Biology and Evolution 15 (8), 931–42.

Brown, M., Allison, D. G., Gilbert, P., 1988. Resistance of bacterial biofilms to antibiotics: a growth-rate related effect. Journal of Antimicrobial Chemotherapy 22 (6), 777–780.

Bull, H. J., Lombardo, M. J., Rosenberg, S. M., 2001. Stationary-phase mutation in the bacterial chromosome: Recombination protein and DNA polymerase IV dependence. Proceedings of the National Academy of Sciences 98 (15), 8334–8341.

Burch, C. L., Chao, L., 2000. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature 406 (6796), 625–628.

Burke, G. R., Moran, N. A., 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. Genome Biology and Evolution 3 (1), 195–208.

Burmolle, M., Kjoller, A., Sorensen, S. J., 2012. An invisible workforce: Biofilms in the soil. In: Lear, G., Lewis, G. D. (Eds.), Microbial Biofilms: Current Research and Applications. Horizon Scientific Press, Poole, UK, pp. 61–71.

Carlucci, A. F., Williams, P. M., 1978. Simulated in situ growth rates of pelagic marine bacteria. Naturwissenschaften 65 (10), 541–542.

Carr, F. J., Chill, D., Maida, N., 2002. The lactic acid bacteria: A literature survey. Critical Reviews in Microbiology 28 (4), 281–370.

Carvalho, M., Reis, A. M., Gordo, I., Silva, R. F., Dionisio, F., 2011. Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations. PLoS Genetics 7 (7), e1002181.

Chattopadhyay, S., Paul, S., Dykhuizen, D. E., Sokurenko, E. V., 2013. Tracking recent adaptive evolution in microbial species using TimeZone. Nature Protocols 8 (4), 652–665.

Cherry, J. L., Wakeley, J., 2003. A diffusion approximation for selection and drift in a subdivided population. Genetics 163 (1), 421–428.

Chesson, P., 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31 (2000), 343–366.

Chiang, W., Tolker-Nielsen, T., 2010. Extracellular DNA as a matrix component in microbial biofilms. In: Kikuchi, Y., Rykova, E. (Eds.), Extracellular Nucleic Acids. Springer-Verlag, Berlin, Germany.

Cho, J. C., Tiedje, J. M., 2000. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. Applied and Environmental Microbiology 66 (12), 5448–5456.

Chou, H. H., Delaney, N. F., Draghi, J. A., Marx, C. J., 2014. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. PLoS Genetics 10 (2), e1004149.

Christensen, B. B., Sternberg, C., Andersen, J. B., Eberl, L., Møller, S., Givskov, M., Molin, S., 1998. Establishment of new genetic trait in a microbial biofilm community. Applied and Environmental Microbiology 64 (6), 2247–2255.

Christensen, B. B., Sternberg, C., Molin, S., 1996. Bacterial plasmid conjugation on semi-solid surfaces monitored with the green fluorescent protein (GFP) from *Aequorea victoria* as a marker. Gene 173 (1), 59–65.

Church, M. J., Hutchins, D. A., Ducklow, H. W., 2000. Limitation of bacterial growth by dissolved organic matter and iron in the southern ocean. Applied and Environmental Microbiology 66 (2), 455–466.

Cohan, F. M., 2016. Bacterial speciation: Genetic sweeps in bacterial species. Current Biology 26 (3), R112–R115.

Cordero, O. X., Polz, M. F., 2014. Explaining microbial genomic diversity in light of evolutionary ecology. Nature Reviews Microbiology 12 (4), 263–273.

Costerton, J. W., Geesey, G. G., Cheng, K.-J., 1978. How bacteria stick. Scientific American 238, 86–95.

Costerton, J. W., Stewart, P. S., Greenberg, E. P., 1999. Bacterial biofilms: A common cause of persistent infections. Science 284 (5418), 1318–1322.

Courvalin, P., 2008. Predictable and unpredictable evolution of antibiotic resistance. Journal of Internal Medicine 264 (1), 4–16.

Cruickshank, R., Sharman, A., 1934. The biology of the vagina in the human subject. II. The bacterial flora and secretion of the vagina at various age-periods and their relation to glycogen in the vagina epithelium. BJOG: An International Journal of Obstetrics & Gynaecology 41 (2), 208–226.

Cseh, B., Umenhoffer, K., 2006. An assay for measuring in vivo mutation rates. Mutation Research 595, 184–190.

Darwin, C., 1859. On the origin of species. John Murray, London.

Davies, J., Davies, D., 2010. Origins and evolution of antibiotic resistance. Microbiology and Molecular Biology Reviews 74 (3), 417–433.

de Beer, D., Stoodley, P., Roe, F., Lewandowski, Z., 1994. Effects of biofilm structures on oxygen distribution and mass-transport. Biotechnology and Bioengineering 43 (11), 1131–1138.

de los Rios, A., Ascaso, C., Wierzchos, J., Vincent, W. F., Quesada, A., 2015. Microstructure and cyanobacterial composition of microbial mats from the High Arctic. Biodiversity and Conservation 24, 841–863.

Debellis, T., Kernaghan, G., Bradley, R., Widden, P., 1998. Growth rate of bacterial communities in soils at varying pH : A comparison of the thymidine and leucine incorporation techniques. Microbial Ecology 36 (3), 316–327.

Delcher, A. L., Bratke, K. a., Powers, E. C., Salzberg, S. L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23 (6), 673–679.

Delcher, a. L., Harmon, D., Kasif, S., White, O., Salzberg, S. L., 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Research 27 (23), 4636–4641.

Demoling, F., Figueroa, D., Bååth, E., 2007. Comparison of factors limiting bacterial growth in different soils. Soil Biology and Biochemistry 39 (10), 2485–2495.

Deschesne, A., Pallud, C., Grundmann, G. L., 2007. Spatial distribution of bacteria at the microscale in soil. In: Franklin, R., Mills, A. (Eds.), The spatial distribution of microbes in the environment. Springer, Dordrecht.

Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., Venail, P., Villéger, S., Mouquet, N., 2010. Defining and measuring ecological specialization. Journal of Applied Ecology 47 (1), 15–25.

Dixon, J. L., Turley, C. M., 2001. Measuring bacterial production in deep-sea sediments using 3H-thymidine incorporation: Ecological significance. Microbial Ecology 42 (4), 549–561.

Doerflinger, S. Y., Throop, A. L., Herbst-Kralovetz, M. M., 2014. Bacteria in the vaginal microbiome alter the innate immune response and barrier properties of the human vaginal epithelia in a species-specific manner. Journal of Infectious Diseases 209 (12), 1989–1999.

Domenech, M., Ramos-Sevillano, E., García, E., Moscoso, M., Yuste, J., 2013. Biofilm formation avoids complement immunity and phagocytosis of *Streptococcus pneumoniae*. Infection and Immunity 81 (7), 2606–2615.

Driffield, K., Miller, K., Bostock, J. M., Neill, A. J. O., Chopra, I., 2008. Increased mutability of *Pseudomonas aeruginosa* in biofilms. Journal of Antimicrobial Agents and Chemotherapy 61 (5), 1053–1056.

Duncan, B., 1980. Mutagenic deamination of cytosine residues in DNA. Nature 287, 560–561.

Dykhuizen, D. E., 1990. Experimental studies of natural selection in bacteria. Annual Review of Ecology and Systematics 21, 373–398.

Dykhuizen, D. E., 1998. Santa Rosalia revisited: Why are there so many species of bacteria? Antonie van Leeuwenhoek 73 (1), 25–33.

Dykhuizen, D. E., Hartl, D. L., 1983. Selection in chemostats. Microbiology Reviews 47 (2), 150–168.

Eastman, J. M., Harmon, L. J., La, H. J., Joyce, P., Forney, L. J., 2011. The onion model, a simple neutral model for the evolution of diversity in bacterial biofilms. Journal of Evolutionary Biology 24 (11), 2496–2504.

Eldholm, V., Norheim, G., von der Lippe, B., Kinander, W., Dahle, U. R., Caugant, D. A., Mannsåker, T., Mengshoel, A. T., Dyrhol-Riise, A. M., Balloux, F., 2014. Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. Genome Biology 15 (11), 490.

Elena, S. F., Cooper, V. S., Lenski, R. E., 1996. Punctuated evolution caused by selection of rare beneficial mutations. Science 272 (5269), 1802–4.

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

Elias, S., Banin, E., jan 2012. Multi-species biofilms: living with friendly neighbors. FEMS Microbiology Reviews.

Elliott, K. T., Cuff, L. E., Neidle, E. L., 2013. Copy number change: Evolving views on gene amplification. Future Microbiology 8 (7), 887–99.

Ellis, C. N., Traverse, C. C., Mayo-Smith, L., Buskirk, S. W., Cooper, V. S., 2015. Character displacement and the evolution of niche complementarity in a model biofilm community. Evolution 69 (2), 283–293.

Erban, R., Chapman, S. J., 2007. Reactive boundary conditions for stochastic simulations of reaction-diffusion processes. Physical biology 4 (1), 16–28.

Farage, M., Maibach, H., 2006. Lifetime changes in the vulva and vagina. Archives of Gynecology and Obstetrics 273 (4), 195–202.

Felten, A., Barreau, C., Bizet, C., Lagrange, P. H., Philippon, A., 1999. *Lactobacillus* species identification, H_2O_2 production, and antibiotic resistance and correlation with human clinical status. Journal of Clinical Microbiology 37 (3), 729–733.

Ferenci, T., 2016. Trade-off mechanisms shaping the diversity of bacteria. Trends in Microbiology 24 (3), 209–223.

Ferris, M. J., Norori, J., Zozaya-hinchliffe, M., Martin, D. H., 2007. Cultivation-independent analysis of changes in bacterial vaginosis flora following metronidazole treatment. Journal of Clinical Microbiology 45 (3), 1016–1018.

Fisher, R. A., 1928. The possible modification of the response of the wild type to recurrent mutations. American Naturalist 62, 115–116.

Fisher, R. A., 1930. The genetical theory of natural selection. Clarendon, Oxford, UK.

Flemming, H.-C., 2011. Microbial biofouling: Unsolved problems, insufficient approaches, and possible solutions. In: Flemming, H.-C., Wingender, J., Szewzyk, U. (Eds.), Biofilm Highlights. Springer, Berlin, Heidelberg.

Flemming, H. C., Wingender, J., 2010. The biofilm matrix. Nature Reviews Microbiology 8 (9), 623–633.

Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S. A., Kjelleberg, S., 2016. Biofilms: an emergent form of bacterial life. Nature Reviews Microbiology 14 (9), 563–575.

Flores, G. E., Bates, S. T., Caporaso, J. G., Lauber, C. L., Leff, J. W., Knight, R., Fierer, N., 2013. Diversity, distribution and sources of bacteria in resedential kitchens. Environmental Microbiology 5 (2), 588–96.

Foster, P. L., 2006. Methods for determining spontaneous mutation rates. Methods in Enzymology 409, 195–213.

Foster, P. L., 2007. Stress-induced mutagenesis in bacteria. Critical Reviews in Biochemistry and Molecular Biology 42 (5), 373–397.

Fox, R. E., Zhong, X., Krone, S. M., Top, E. M., 2008. Spatial structure and nutrients promote invasion of IncP-1 plasmids in bacterial populations. The ISME Journal 2 (10), 1024–1039.

France, M., Remold, S., 2016. Interference competition among household strains of *Pseudomonas*. Microbial Ecology 72 (4), 821–830.

Franklin, R., Mills, A. (Eds.), 2007. The spatial distribution of microbes in the environment. Springer, Dordrecht.

Frederico, L. A., Kunkel, T. A., Shaw, B. R., 1993. Cytosine deamination in mismatched base pairs. Biochemistry 32, 6523–6530.

Fredricks, D. N., Fiedler, T. L., Marrazzo, J. M., 2005. Molecular identification of bacteria associated with bacterial vaginosis. The New England Journal of Medicine 353 (18), 1899–1911.

Frost, L. S., Leplae, R., Summers, A. O., Toussaint, A., sep 2005. Mobile genetic elements: the agents of open source evolution. Nature Reviews Microbiology 3 (9), 722–32.

Futuyma, D. J., Moreno, G., 1988. The evolution of ecological specialization. Annual Review of Ecology and Systematics 19, 207–233.

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schutte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z., Zhou, X., Abdo, Z., Forney, L. J., Ravel, J., 2012. Temporal dynamics of the human vaginal microbiota. Science Translational Medicine 4 (132), 132ra52.

Galas, D. J., Branscomb, E. W., 1976. Ribosome slowed by mutation to streptomycin resistance. Nature 262, 617–619.

Gause, G., 1934. The struggle for existence. Williams & Wilkins, Baltimore, MD.

Gerrish, P. J., Lenski, R. E., 1998. The fate of competing beneficial mutations in an asexual population. Genetica 102/103 (1-6), 127–144.

Gillespie, J. H., 1983. Some properties of finite populations experiencing strong selection and weak mutation. American Naturalist 121 (5), 691–708.

Goodman, M. F., 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annual Review of Biochemistry 71 (1), 17–50.

Gordo, I., Campos, P. R. A., 2006. Adaptive evolution in a spatially structured asexual population. Genetica 127 (1-3), 217–229.

Grande, G., Milardi, D., Vincenzoni, F., Pompa, G., Biscione, A., Astorri, a. L., Fruscella, E., De Luca, A., Messana, I., Castagnola, M., Marana, R., 2015. Proteomic characterization of the qualitative and quantitative differences in cervical mucus composition during the menstrual cycle. Molecular BioSystems 11 (6), 1717–1725.

Gyles, C., Boerlin, P., 2014. Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. Veterinary Pathology 51 (2), 328–40.

Habets, M. G. J. L., Czárán, T., Hoekstra, R. F., de Visser, J. A. G. M., 2007. Spatial structure inhibits the rate of invasion of beneficial mutations in asexual populations. Proceedings of the Royal Society B-Biological Sciences 274 (1622), 2139–2143.

Habets, M. G. J. L., Rozen, D. E., Hoekstra, R. F., De Visser, J. A. G. M., 2006. The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. Ecology Letters 9 (9), 1041–1048.

Haidar, G., Philips, N. J., Shields, R. K., Snyder, D., Cheng, S., Potoski, B. A., Doi, Y., Hao, B., Press, E. G., Cooper, V. S., Clancy, C. J., Nguyen, M. H., 2017. Ceftolozane-Tazobactam for the treatment of multidrug-resistant *Pseudomonas aeruginosa* infections: Clinical effectiveness and evolution of resistance. Clinical Infectious Diseases 65 (1), 110–120.

Hall, B. M., Ma, C.-X., Liang, P., Singh, K. K., 2009. Fluctuation AnaLysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. Bioinformatics 25 (12), 1564–1565.

Hall-Stoodley, L., Costerton, J. W., Stoodley, P., 2004. Bacterial biofilms: From the natural environment to infectious diseases. Nature Reviews Microbiology 2, 95–108.

Hall-Stoodley, L., Stoodley, P., jul 2009. Evolving concepts in biofilm infections. Cellular Microbiology 11 (7), 1034–43.

Hallberg, L., Hogdahl, A.-M., Nilsson, L., Rybo, G., 1966. Menstrual blood loss and iron deficiency. Acta Medica Scandinavica 180, 639–650.

Hannan, S., Ready, D., Jasni, A. S., Rogers, M., Pratten, J., Roberts, A. P., 2010. Transfer of antibiotic resistance by transformation with eDNA within oral biofilms. FEMS Immunology and Medical Microbiology 59 (3), 345–349.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., Martiny, J. B. H., 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. Nature Reviews Microbiology 10 (7), 497–507.

Harris, D., Paul, E. A., 1994. Measurement of bacterial growth rates in soil. Applied Soil Ecology 1 (4), 277–290.

Hausner, M., Wuertz, S., 1999. High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. Applied and Environmental Microbiology 65 (8), 3710–3713.

Hendrickx, L., Hausner, M., Wuertz, S., 2003. Natural genetic transformation in monoculture *Acinetobacter* sp. Applied and Environmental Microbiology 69 (3), 1721–1727.

Hibbing, M. E., Fuqua, C., Parsek, M. R., Peterson, S. B., 2010. Bacterial competition: Surviving and thriving in the microbial jungle. Nature Reviews Microbiology 8 (1), 15–25.

Hickey, R. J., Zhou, X., Settles, M. L., Erb, J., Malone, K., Hansmann, M. A., Shew, M. L., Pol, B. V. D., Fortenberry, J. D., Forney, L. J., Forney, J., 2015. Vaginal microbiota of adolescent girls prior to the onset of menarche resemble those of reproductive-age women. mBio 6 (2), e00097–15.

Hillier, S. L., Lau, R. J., 1997. Vaginal microflora in postmenopausal women who have not received estrogen replacement therapy. Clinical Infectious Diseases 25 (Suppl 2), 123–126.

Hobley, L., Harkins, C., MacPhee, C. E., Stanley-Wall, N. R., 2015. Giving structure to the biofilm matrix: An overview of individual strategies and emerging common themes. FEMS Microbiology Reviews 39 (5), 649–669.

Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S., Ciofu, O., apr 2010. Antibiotic resistance of bacterial biofilms. International Journal of Antimicrobial Agents 35 (4), 322–32.

Høiby, N., Ciofu, O., Johansen, H. K., Song, Z.-j., Moser, C., Jensen, P. Ø., Molin, S., Givskov, M., Tolker-Nielsen, T., Bjarnsholt, T., 2011. The clinical impact of bacterial biofilms. International Journal of Oral Science 3 (2), 55–65.

Howden, B. P., Mcevoy, C. R. E., Allen, D. L., Chua, K., Gao, W., Paul, F., Bell, J., Coombs, G., Bennett-wood, V., Porter, J. L., Robins-browne, R., Davies, J. K., Seemann, T., Stinear, T. P., 2011. Evolution of multidrug resistance during *Staphylococcus aureus* infection involves mutation of the essential two component regulator walKR. PLoS pathogens 7 (11), e1002359.

Hoyle, B. D., Alcantara, J., Costerton, J. W., 1992. *Pseudomonas aeruginosa* biofilm as a diffusion barrier to piperacillin. Antimicrobial Agents and Chemotherapy 36 (9), 2054–2056.

Hu, K.-t., Zheng, J.-x., Yu, Z.-j., Chen, Z., Cheng, H., Pan, W.-g., Yang, W.-z., Wang, H.-y., Deng, Q.-w., Zeng, Z.-m., 2015. Directed shift of vaginal microbiota induced by vaginal application of sucrose gel in *Rhesus macaques*. International Journal of Infectious Diseases 33, 32–36.

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., Polz, M. F., 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science 320 (5879), 1081–1085.

Imhof, M., Schlotterer, C., 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proceedings of the National Academy of Sciences 98 (3), 1113–1117.

Imlay, J. a., jan 2003. Pathways of oxidative damage. Annual Review of Microbiology 57, 395–418.

Jakobsson, T., Forsum, U., 2007. *Lactobacillus iners*: A marker of changes in the vaginal flora? Journal of Clinical Microbiology 45 (9), 3145.

James, G. A., Swogger, E., Wolcott, R., deLancey Pulcini, E., Secor, P., Sestrich, J., Costerton, J. W., Stewart, P. S., 2007. Biofilms in chronic wounds. Wound Repair and Regeneration 16 (1), 37–44.

Jannasch, H. W., 1969. Estimations of Bacterial Growth Rates in Natural Waters. Journal of Bacteriology 99 (1), 156–160.

Jensen, P. O., Givskov, M., Bjarnsholt, T., Moser, C., 2010. The immune system vs. *Pseudomonas aeruginosa* biofilms. FEMS Immunology and Medical Microbiology 59 (3), 292–305.

Jin, D., 1988. Mapping and sequencing of mutations in *Escherichia coli rpoB* gene that lead to rifampicin resistance. Journal of Molecular Biology 202 (1), 45–58.

Jordan, I. K., Rogozin, I. B., Wolf, Y. I., Koonin, E. V., 2002. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. Genome Research 12 (6), 962–968.

Joyce, P., Rokyta, D. R., Beisel, C. J., Orr, H. A., 2008. A general extreme value theory model for the adaptation of DNA sequences under strong selection and weak mutation. Genetics 180 (3), 1627–1643.

Kahm, M., Hasenbrink, G., Ludwig, J., 2010. grofit : Fitting biological growth curves with R. Journal of Statistical Software 33 (7), 1–21.

Kajiura, T., Wada, H., Ito, K., Anzai, Y., Kato, F., 2006. Conjugative plasmid transfer in the biofilm formed by *Enterococcus faecalis*. Journal of Health Science 52 (4), 358–367.

Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. Journal of Molecular Biology 428 (4), 726–731.

Kassen, R., 2014. Experimental evolution and the nature of biodiversity. Roberts, Greenwood Village, CO.

Kassen, R., Bataillon, T., 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nature Genetics 38 (4), 484–489.

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., Whitlock, M. C., 2012. Experimental evolution. Trends in Ecology and Evolution 27 (10), 547–560.

Kerr, B., Riley, M. A., Feldman, M. W., Bohannan, B. J. M., 2002. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. Nature 418 (6894), 171–174.

Kidwell, M., Lisch, D., 2001. Perspective: Transposable elements parasitic DNA and genome evolution. International Journal of Organic Evolution 55 (1), 1–24.

Kidwell, M. G., Lisch, D. R., 2000. Transposable elements and host genome evolution. Trends in Ecology and Evolution 15 (3), 95–99.

Kim, W., Racimo, F., Schluter, J., Levy, S. B., Foster, K. R., 2014. Importance of positioning for microbial evolution. Proceedings of the National Academy of Sciences 111 (16), E1639–47.

Kirchman, D. L., 2016. Growth rates of microbes in the oceans. The Annual Review of Marine Science 8, 285–309.

Kirk, H., Freeland, J. R., 2011. Applications and implications of neutral versus non-neutral markers in molecular ecology. International Journal of Molecular Sciences 12 (6), 3966–3988.

Kivisaar, M., 2003. Stationary phase mutagenesis : Mechanisms that accelerate adaptation of microbial populations under environmental stress. Environmental Microbiology 5 (10), 814–827.

Kivisaar, M., 2010. Mechanisms of stationary-phase mutagenesis in bacteria: Mutational processes in pseudomonads. FEMS Microbiology Letters 312 (1), 1–14.

Klevens, R. M., Edwards, J. R., Richards Jr., C. L., Horan, T. C., Gaynes, R. P., Pollock, D. A., Cardo, D. M., 2007. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. Public Health Reports 122 (2), 160–166.

Koch, A. L., 1974. The pertinence of the periodic selection phenomenon to prokaryote evolution. Genetics 77, 127–142.

Konstantinidis, K. T., Ramette, A., Tiedje, J. M., 2006. The bacterial species definition in the genomic era. Philosophical Transactions of the Royal Society B: Biological Sciences 361 (1475), 1929–1940.

Kopac, S., Wang, Z., Wiedenbeck, J., Sherry, J., Wu, M., Cohan, F. M., 2014. Genomic heterogeneity and ecological speciation within one subspecies of *Bacillus subtilis*. Applied and Environmental Microbiology 80 (16), 4842–4853.

Kouzel, N., Oldewurtel, E. R., Maier, B., 2015. Gene transfer efficiency in gonococcal biofilms: Role of biofilm age, architecture, and pilin antigenic variation. Journal of Bacteriology 197 (14), 2422–31.

Kraemer, S., Boynton, P., 2017. Evidence for microbial local adaptation in nature. Molecular Ecology 26 (7), 1860–1876.

Król, J. E., Nguyen, H. D., Rogers, L. M., Beyenal, H., Krone, S. M., Top, E. M., 2011. Increased transfer of a multidrug resistance plasmid in *Escherichia coli* biofilms at the air-liquid interface. Applied and Environmental Microbiology, 77 (15), 5079–5088.

Kryazhimskiy, S., Rice, D. P., Desai, M. M., 2012. Population subdivision and adaptation in asexual populations of *Saccharomyces cerevisiae*. Evolution 66 (6), 1931–1941.

Kubitschek, H. E., 1974. Operation of selection pressure on microbial populations. Symposium of the Society for General Microbiology 24, 105–130.

Kuleshov, V., Jiang, C., Zhou, W., Jahanbani, F., Batzoglou, S., Snyder, M., 2016. Synthetic long-read sequencing reveals intraspecies diversity in the human microbiome. Nature Biotechnology 34 (1), 64–69.

Kumar, C. G., Anand, S. K., 1998. Significance of microbial biofilms in food industry: A review. International Journal of Food Microbiology 42 (1-2), 9–27.

Kvitek, D. J., Sherlock, G., 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genetics 7 (4), e1002056.

Laird, C. D., McConaughy, B. L., McCarthy, B. J., 1969. Rate of fixation of nucleotide substitutions in evolution. Nature 224, 149–154.

Lanfear, R., Calcott, B., Ho, S. Y. W., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29 (6), 1695–1701.

Lanfear, R., Kokko, H., Eyre-Walker, A., 2014. Population size and the rate of evolution. Trends in Ecology & Evolution 29 (1), 33–41.

Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., Desai, M. M., 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500 (7464), 571–4.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., Higgins, D. G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23 (21), 2947–2948.

Lebret, K., Kritzberg, E. S., Figueroa, R., Rengefors, K., 2012. Genetic diversity within and genetic differentiation between blooms of a microalgal species. Environmental Microbiology 14 (9), 2395–2404.

Lemon, J., 2006. Plotrix: a package in the red light district of R. R-News 6, 8–12.

Lenski, R. E., Rose, M. R., Simpson, S. C., Tadler, S. C., 1991. Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. American Naturalist 138 (6), 1315–1341.

Levin, B. R., 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics 99 (1), 1–23.

Levin-reisman, I., Ronin, I., Gefen, O., Braniss, I., 2017. Antibiotic tolerance facilitates the evolution of resistance. Science, eaaj2191.

Levinson, G., Gutman, G. A., 1987. Slipped-strand mispairing : A major mechanism for DNA sequence evolution. Molecular Biology and Evolution 4 (3), 203–221.

Lewis, K., 2007. Persister cells, dormancy and infectious disease. Nature Reviews Microbiology 5 (1), 48–56.

Lewis, K., 2010. Persister cells. Annual Review of Microbiology 64, 357–72.

Li, L., Stoeckert, C. J., Roos, D. S., sep 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome research 13 (9), 2178–89.

Li, Y.-h., Lau, P. C. Y., Lee, J. H., Ellen, R. P., Cvitkovitch, D. G., 2001. Natural genetic transformation of *Streptococcus mutans* growing in biofilms. Journal of Bacteriology 183 (3), 897–908.

Lieberman, T. D., Flett, K. B., Yelin, I., Martin, T. R., McAdam, A. J., Priebe, G. P., Kishony, R., 2013. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. Nature Genetics 46 (1), 82–87.

Linhares, I. M., Summers, P. R., Larsen, B., Giraldo, P. C., Witkin, S. S., 2011. Contemporary perspectives on vaginal pH and lactobacilli. American Journal of Obstetrics and Gynecology 204 (2), 120–e1.

Liu, X., Zheng, H., Zhang, W., Shen, Z., Zhao, M., 2016. Tracking cefoperazone/sulbactam resistance development in vivo in *A . baumannii* isolated from a patient with hospital-acquired pneumonia by whole-genome sequencing. Frontiers in Microbiology 7, 1268.

Lomstein, B., Langerhuus, A., D'Hondt, S., Jorgensen, B., Spivack, A., 2012. Endospore abundance, microbial growth, and necromass turnover in deep sub-seafloor sediment. Nature 484 (7392), 101–104.

López-Camacho, E., Gómez-Gil, R., Tobes, R., Manrique, M., Lorenzo, M., Galván, B., Salvarelli, E., Moatassim, Y., Salanueva, I. J., Pareja, E., Codoñer, F. M., Alvarez-Tejado, M., Garcillán-Barcia, M. P., De la Cruz, F., Mingorance, J., 2014. Genomic analysis of the emergence and evolution of multidrug resistance during a *Klebsiella pneumoniae* outbreak including carbapenem and colistin resistance. Journal of Antimicrobial Chemotherapy 69 (3), 632–636.

Lukjancenko, O., Wassenaar, T. M., Ussery, D. W., 2010. Comparison of 61 sequenced *Escherichia coli* genomes. Microbial Ecology 60 (4), 708–720.

Luo, H., Gao, F., Lin, Y., 2015. Evolutionary conservation analysis between the essential and nonessential genes in bacterial genomes. Scientific Reports 5, 13210.

Lynch, A. S., Robertson, G. T., 2008. Bacterial and fungal biofilm infections. Annual Review of Medicine 59, 415–428.

Macfarlane, S., Dillon, J. F., may 2007. Microbial biofilms in the human gastrointestinal tract. Journal of Applied Microbiology 102 (5), 1187–96.

Maggi, N., Pasqualucci, C. R., Ballotta, R., Sensi, P., 1966. Rifampicin: A new orally active rifamycin. Chemotherapy 11 (5), 285–292.

Maharjan, R. P., Gaffé, J., Plucain, J., Schliep, M., Wang, L., Feng, L., Tenaillon, O., Ferenci, T., Schneider, D., 2013. A case of adaptation through a mutation in a tandem duplication during experimental evolution in *Escherichia coli*. BMC Genomics 14 (1), 441.

Maisonneuve, E., Gerdes, K., 2014. Molecular mechanisms underlying bacterial persisters. Cell 157 (3), 539–548.

Mammola, S., Isaia, M., 2014. Niche differentiation in *Meta bourneti* and *M. menardi* (Araneae, Tetragnathidae) with notes on the life history. International Journal of Speleology 43 (3), 343–353.

Marvig, R. L., Dolce, D., Sommer, L. M., Petersen, B., Ciofu, O., Campana, S., Molin, S., Taccetti, G., Johansen, H. K., 2015. Within-host microevolution of *Pseudomonas aeruginosa* in Italian cystic fibrosis patients. BMC Microbiology 15 (1), 218.

Masel, J., 2011. Genetic drift. Current Biology 21 (20), R837–R838.

Mayer, B. T., Srinivasan, S., Fiedler, T. L., Marrazzo, J. M., Fredricks, D. N., Schiffer, J. T., 2015. Rapid and profound shifts in the vaginal microbiota following antibiotic treatment for bacterial vaginosis. Journal of Infectious Diseases 212, 793–802.

Mazin, A., Gimadutdinov, O., Turkin, S., Burtseva, N., Vaniushin, B., 1985. Non-enzymatic DNA methylation by S-adenosylmethionine results in the formation of minor thymine residues and 5-methylcytosine from cytosine. Molecular Biology 19 (4), 903–914.

McArthur, J. V., Kovacic, D. A., Smith, M. H., 1988. Genetic diversity in natural populations of a soil bacterium across a landscape gradient. Proceedings of the National Academy of Sciences 85 (24), 9621–9624.

McCutcheon, J. P., Moran, N. A., 2012. Extreme genome reduction in symbiotic bacteria. Nature Reviews Microbiology 10, 13–26.

McMillan, A., Macklaim, J. M., Burton, J. P., Reid, G., 2012. Adhesion of *Lactobacillus iners* AB-1 to Human fibronectin: A key mediator for persistence in the vagina? Reproductive Sciences 20 (7), 791–6.

Meckenstock, R. U., Elsner, M., Griebler, C., Lueders, T., Stumpp, C., Aamand, J., Agathos, S. N., Albrechtsen, H. J., Bastiaens, L., Bjerg, P. L., Boon, N., Dejonghe, W., Huang, W. E., Schmidt, S. I., Smolders, E., SØrensen, S. R., Springael, D., van Breukelen, B. M., 2015. Biodegradation: Updating the concepts of control for microbial cleanup in contaminated aquifers. Environmental Science and Technology 49 (12), 7073–7081.

Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005. The microbial pan-genome. Current Opinion in Genetics and Development 15 (6), 589–594.

Meervenne, E. V., Weirdt, R. D., Coillie, E. V., Devlieghere, F., Herman, L., Boon, N., 2014. Biofilm models for the food industry: Hot spots for plasmid transfer? Pathogens and Disease 70 (3), 332–338.

Melnyk, A. H., Wong, A., Kassen, R., 2015. The fitness costs of antibiotic resistance mutations. Evolutionary Applications 8 (3), 273–283.

Mendes-Soares, H., Rychlik, L., 2009. Differences in swimming and diving abilities between two sympatric species of water shrews: *Neomys anomalus* and *Neomys fodiens* (Soricidae). Journal of Ethology 27 (3), 317–325.

Mendes-Soares, H., Suzuki, H., Hickey, R. J., Forney, L. J., 2014. Comparative functional genomics of *Lactobacillus* spp. reveals possible mechanisms for specialization of vaginal lactobacilli to their environment. Journal of Bacteriology 196 (7), 1458–1470.

Merod, R. T., Wuertz, S., 2014. Extracellular polymeric substance architecture influences natural genetic transformation of *Acinetobacter baylyi* in biofilms. Applied and Environmental Microbiology 80 (24), 7752–7757.

Milkman, R., 1973. Electrophoretic variation in *Escherichia coli* from natural sources. Science 182 (4116), 1024–1026.

Milwidsky, A., Palti, Z., Gutman, A., 1980. Glycogen metabolism of the human endometrium. Journal of Clinical Endrochrinology and Metabolism 51 (4), 765–770.

Mirmonsef, P., Modur, S., Burgad, D., Gilbert, D., Golub, E. T., French, A. L., McCotter, K., Landay, A. L., Spear, G. T., 2015. Exploratory comparison of vaginal glycogen and *Lactobacillus* levels in premenopausal and postmenopausal women. Menopause 22 (7), 702–709.

Molin, S., Tolker-Nielsen, T., 2003. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. Current Opinion in Biotechnology 14 (3), 255–261.

Monier, J., Lindow, S. E., 2005. Spatial organization of dual-species bacterial aggregates on leaf surfaces. Applied Environmental Microbiology 71 (9), 5484–5493.

Moran, N. A., 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proceedings of the National Academy of Sciences 93 (7), 2873–2878.

Moran, N. A., Wernegreen, J. J., 2000. Lifestyle evolution in symbiotic bacteria: Insights from genomics. Trends in Ecology and Evolution 15 (8), 321–326.

Morris, J. J., Lenski, R. E., Zinser, E. R., 2012. The Black Queen Hypothesis : Evolution of dependencies through adaptive gene loss. mBio 3 (2), e00036–12.

Mulcahy, L. R., Isabella, V. M., Lewis, K., jul 2014. *Pseudomonas aeruginosa* biofilms in disease. Microbial Ecology 68 (1), 1–12.

Muller, C. J., 1932. Some genetic aspects of sex. American Naturalist 66 (703), 118–138.

Mwangi, M. M., Wu, S. W., Zhou, Y., Sieradzki, K., de Lencastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E. D., Tomasz, A., 2007. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. Proceedings of the National Academy of Sciences 104 (22), 9451–9456.

Nadell, C. D., Drescher, K., Foster, K. R., 2016. Spatial structure, cooperation and competition in biofilms. Nature Reviews Microbiology 14 (9), 589–600.

Nahum, J. R., Godfrey-Smith, P., Harding, B. N., Marcus, J. H., Carlson-Stevermer, J., Kerr, B., 2015. A tortoise-hare pattern seen in adapting structured and unstructured populations suggests a rugged fitness landscape in bacteria. Proceedings of the National Academy of Sciences 112 (24), 7530–7535.

Nasioudis, D., Beghini, J., Bongiovanni, A., Giraldo, P., Linhares, I., Witkin, S. S., 2015. Alpha-amylase in vaginal fluid: Association with conditions favorable to dominance of *Lactobacillus*. Reproductive Sciences 22 (11), 1393–1398.

Normak, B. H., Normak, S., aug 2002. Evolution and spread of antibiotic resistance. Journal of Internal Medicine 252 (2), 91–106.

Novick, A., Szilard, L., 1950. Experiments with the chemostat on spontaneous mutations of bacteria. Proceedings of the National Academy of Sciences 36 (12), 708–719.

Ochman, H., Lawrence, J., Groisman, E., 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405 (6784), 299–304.

Ochman, H., Selander, R. K., 1984. Standard reference strains of *Escherichia coli* from natural populations. Journal of Bacteriology 157 (2), 690–693.

Oda, Y., Wanders, W., Huisman, L. a., Meijer, W. G., Gottschal, J. C., Forney, L. J., 2002. Genotypic and phenotypic diversity within species of purple nonsulfur bacteria isolated from aquatic sediments. Applied and Environmental Microbiology, 68 (7), 3467–3477.

O'Hanlon, D. E., Moench, T. R., Cone, R. A., 2011. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. BMC Infectious Diseases 11 (1), 200.

Ojala, T., Kankainen, M., Castro, J., Cerca, N., Edelman, S., Westerlund-wikström, B., Paulin, L., Holm, L., Auvinen, P., 2014. Comparative genomics of *Lactobacillus crispatus* suggests novel mechanisms for the competitive exclusion of *Gardnerella vaginalis*. BMC Genomics 15, 1070.

Ojala, T., Kuparinen, V., Koskinen, J. P., Alatalo, E., Holm, L., Auvinen, P., Edelman, S., Westerlund-Wikström, B., Korhonen, T. K., Paulin, L., Kankainen, M., 2010. Genome sequence of *Lactobacillus crispatus* st1. Journal of Bacteriology 192 (13), 3547–3548.

O'Keeffe, T., Hill, C., Ross, R. P., 1999. Characterization and heterologous expression of the genes encoding enterocin A production , immunity , and regulation in *Enterococcus faecium* DPC1146. Applied and Environmental Microbiology 65 (4), 1506–1515.

Olsen, I., 2015. Biofilm-specific antibiotic tolerance and resistance. European Journal of Clinical Microbiology & Infectious Diseases 34 (5), 877–886.

Orr, H. A., 2002. The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56 (7), 1317–1330.

Orr, H. A., 2003. The distribution of fitness effects among beneficial mutations. Genetics 163, 1519–1526.

Ørskov, F., Ørskov, I., 1983. Summary of a workshop on the clone concept in the epidemiology, taxonomy, and evolution of the Enterobacteriaceae and other bacteria. The Journal of Infectious Diseases 148 (2), 346–357.

Otto, M., 2008. Staphylococcal biofilms. Current Topics in Microbiology and Immunology 322, 207–228.

Owed, D., Katz, D., 2005. A review of the physical and chemical properties of human semen and the formulation of a semen simulant. Journal of Andrology 26 (4), 459–469.

Palestrant, D., Holzknecht, Z. E., Collins, B. H., Parker, W., Miller, S. E., Bollinger, R. R., 2004. Microbial biofilms in the gut: Visualization by electron microscopy and by acridine orange staining. Ultrastructural Pathology 28 (1), 23–27.

Pannell, J. R., Fields, P. D., 2014. Evolution in subdivided plant populations: Concepts, recent advances and future directions. New Phytologist 201 (2), 417–432.

Parsek, M. R., Singh, P. K., jan 2003. Bacterial biofilms: an emerging link to disease pathogenesis. Annual Review of Microbiology 57, 677–701.

Paul, B. J., Duthie, H. C., Taylor, W. D., 1991. Nutrient cycling by biofilms in running waters of differing nutrient status. Journal of the North American Benthological Society 10 (1), 31–41.

Percival, S. L., Hill, K. E., Williams, D. W., Hooper, S. J., Thomas, D. W., Costerton, J. W., 2012. A review of the scientific evidence for biofilms in wounds. Wound Repair and Regeneration 20 (5), 647–657.

Perfeito, L., Pereira, M. I., Campos, P. R. a., Gordo, I., 2008. The effect of spatial structure on adaptation in *Escherichia coli*. Biology Letters 4 (1), 57–59.

Petersen, A., Aarestrup, F. M., Olsen, J. E., 2009. The in vitro fitness cost of antimicrobial resistance in *Escherichia coli* varies with the growth conditions. FEMS Microbiology Letters 299 (1), 53–59.

Petricevic, L., Domig, K. J., Nierscher, F. J., Sandhofer, M. J., Fidesser, M., Krondorfer, I., Husslein, P., Kneifel, W., Kiss, H., 2014. Characterisation of the vaginal *Lactobacillus* microbiota associated with preterm delivery. Scientific Reports 4, 5136.

Picard, C., Bosco, M., 2008. Genotypic and phenotypic diversity in populations of plant-probiotic *Pseudomonas* spp. colonizing roots. Naturwissenschaften 95 (1), 1–16.

Poltak, S. R., Cooper, V. S., mar 2011. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. The ISME Journal 5 (3), 369–78.

Ponciano, J. M., La, H. J., Joyce, P., Forney, L. J., 2009. Evolution of diversity in spatially structured *Escherichia coli* populations. Applied and environmental microbiology 75 (19), 6047–6054.

Port, G. C., Vega, L. a., Nylander, A. B., Caparon, M. G., 2014. *Streptococcus pyogenes* polymyxin B-resistant mutants display enhanced exportal integrity. Journal of Bacteriology 196 (14), 2563–2577.

Poterra, C., 1999. Forging a link between biofilms and disease. Science 283, 1837–1839.

Poulsen, L. K., Lan, F., Kristensen, C. S., Hobolth, P., Molin, S., Krogfelt, K. A., 1994. Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. Infection and Immunity 62 (11), 5191–5194.

Qi, L., Li, H., Zhang, C., Liang, B., Li, J., Wang, L., Du, X., Liu, X., Qiu, S., Song, H., 2016. Relationship between antibiotic resistance, biofilm formation, and biofilm-specific resistance in *Acinetobacter baumannii*. Frontiers in Microbiology 7, 483.

Rainey, P. B., Buckling, A., Kassen, R., Travisano, M., 2000. The emergence and maintenance of diversity: Insights from experimental bacterial populations. Trends in Ecology and Evolution 15 (6), 243–247.

Rainey, P. B., Travisano, M., 1998. Adaptive radiation in a heterogeneous environment. Nature 394 (2), 69–72.

Rampersaud, R., Planet, P. J., Randis, T. M., Kulkarni, R., Aguilar, J. L., Lehrer, R. I., Ratner, A. J., 2011. Inerolysin, a cholesterol-dependent cytolysin produced by *Lactobacillus iners*. Journal of Bacteriology 193 (5), 1034–1041.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., Mcculle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., Forney, L. J., 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences 108, 4680–4687.

Reynolds, M. G., 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. Genetics 156 (4), 1471–1481.

Rodgers, K., Mcvey, M., 2016. Error-prone repair of DNA double-strand breaks. Journal of Cellular Physiology 231 (1), 15–24.

Roesch, L., Fulthorpe, R., Riva, A., Casella, G., Hadwin, A., Kent, A., Daroub, S., Camargo, F., Farmerie, W., Triplett, E., 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. The ISME Journal 1 (4), 283–290.

Rogers, J., Dowsett, A. B., Dennis, P. J., Lee, J. V., Keevil, C. W., 1994. Influence of temperature and plumbing material selection on biofilm formation and growth of *Legionella pneumophila* in a model potable water system containing complex microbial flora. Applied and Environmental Microbiology 60 (5), 1585–1592.

Römling, U., Balsalobre, C., 2012. Biofilm infections, their resilience to therapy and innovative treatment strategies. Journal of Internal Medicine 272, 541–561.

Rosche, W. A., Foster, P. L., 2000. Determining mutation rates in bacterial populations. Methods 20 (1), 4–17.

Roszak, D. B., Colwell, R. R., 1987. Survival strategies of bacteria in the natural environment. Microbiological Reviews 51 (3), 365–379.

Rousk, J., Bååth, E., 2011. Growth of saprotrophic fungi and bacteria in soil. FEMS Microbiology Ecology 78 (1), 17–30.

Rozen, D. E., De Visser, J. A. G., Gerrish, P. J., 2002. Fitness effects of fixed beneficial mutations in microbial populations. Current Biology 12 (12), 1040–1045.

Rozen, D. E., Habets, M. G. J. L., Handel, A., de Visser, J. A. G. M., 2008. Heterogeneous adaptive trajectories of small populations on complex fitness landscapes. PLoS ONE 3 (3), e1715.

Russell, R. R. B., 1972. Mapping of a D-cycloserine resistance locus in *Escherichia coli* K-12. Journal of Bacteriology 111 (2), 622–624.

Ryan, F. J., 1955. Spontaneous mutation in non-dividing bacteria. Genetics 40 (5), 726–738.

Rydberg, B., Lindahl, T., 1982. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-Lmethionine is a potentially mutagenic reaction. The EMBO Journal 1 (2), 211–216.

Ryder, V. J., Chopra, I., O'Neill, A. J., jan 2012. Increased mutability of staphylococci in biofilms as a consequence of oxidative stress. PloS one 7 (10), e47695.

Salzberg, S. L., Deicher, A. L., Kasif, S., White, O., 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Research 26 (2), 544–548.

Saumaa, S., Tover, A., Kasak, L., Kivisaar, M., 2002. Different spectra of stationary-phase mutations in early-arising versus late-arising mutants of *Pseudomonas putida*: Involvement of the DNA repair enzyme MutY and the stationary-phase sigma factor RpoS. Journal of Bacteriology 184 (24), 6957–6965.

Schaaper, R., Loeb, L., 1980. Depurination causes mutations in SOS-induced cells. Proceedings of the National Academy of Sciences 78 (3), 1773–1777.

Schenk, M. F., Witte, S., Salverda, M. L. M., Koopmanschap, B., Krug, J., de Visser, J. A. G. M., 2015. Role of pleiotropy during adaptation of tem-1 β-lactamase to two novel antibiotics. Evolutionary Applications 8 (3), 248–260.

Schloss, P. D., Handelsman, J., 2006. Toward a census of bacteria in soil. PLoS Computational Biology 2 (7), e92.

Schoener, T. W., 1974. Resource partitioning in ecological communities. Science 185 (4145), 27–39.

Schreiber, S., Konradt, M., Groll, C., Scheid, P., Hanauer, G., Werling, H.-O., Josenhans, C., Suerbaum, S., 2004. The spatial orientation of *Helicobacter pylori* in the gastric mucus. Proceedings of the National Academy of Sciences 101 (14), 5024–5029.

Schumacher, M. a., Balani, P., Min, J., Chinnam, N. B., Hansen, S., Vulić, M., Lewis, K., Brennan, R. G., 2015. HipBA-promoter structures reveal the basis of heritable multidrug tolerance. Nature 524 (7563), 59–64.

Selander, R. K., Leven, B. R., 1980. Genetic diversity and structure in *Escherichia coli* populations. Science 210 (4469), 545–547.

Seoane, J., Yankelevich, T., Dechesne, A., Merkey, B., Sternberg, C., Smets, B. F., 2011. An individual-based approach to explain plasmid invasion in bacterial populations. FEMS Microbiology Ecology 75 (1), 17–27.

Shakil, S., Khan, R., Zarrilli, R., Khan, A. U., 2008. Aminoglycosides versus bacteria–a description of the action, resistance mechanism, and nosocomial battleground. Journal of Biomedical Science 15 (1), 5–14.

Shapiro, B. J., Polz, M. F., 2014. Ordering microbial diversity into ecologically and genetically cohesive units. Trends in Microbiology 22 (5), 235–247.

Simões, M., Pereira, M. O., Vieira, M. J., 2005. Effect of mechanical stress on biofilms challenged by different chemicals. Water Research 39 (20), 5142–5152.

Slatkin, M., 1985. Gene flow in natural populations. Annual Review of Ecology and Systematics 16, 393–430.

Smith, J. M., Smith, N. H., O'Rourke, M., Spratt, B. G., 1993. How clonal are bacteria? Proceedings of the National Academy of Sciences 90 (10), 4384–4388.

Smith, S. A., Dunn, C. W., 2008. Phyutility: A phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24 (5), 715–716.

Sniegowski, P. D., Gerrish, P. J., 2010. Beneficial mutations and the dynamics of adaptation in asexual populations. Proceedings of the Royal Society B-Biological Sciences 365, 1255–1263.

Snyder, L., 1995. Phage-exclusion enzymes: A bonanza of biochemical and cell biology reagents? Molecular Microbiology 15 (3), 415–420.

Sommer, L. M., Alanin, M. C., Marvig, R. L., Nielsen, K. G., Høiby, N., von Buchwald, C., Molin, S., Johansen, H. K., 2016. Bacterial evolution in PCD and CF patients follows the same mutational steps. Scientific Reports 6 (28732).

Sonti, R. V., Roth, J. R., 1989. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. Genetics 123 (1), 19–28.

Spear, G. T., French, A. L., Gilbert, D., Zariffard, M. R., Mirmonsef, P., Sullivan, T. H., Spear, W. W., Landay, A., Micci, S., Lee, B. H., Hamaker, B. R., 2014. Human alpha-amylase present in lower-genital-tract mucosal fluid processes glycogen to support vaginal colonization by *Lactobacillus*. Journal of Infectious Diseases 210 (7), 1019–1028.

Stalder, T., Top, E. M., 2016. Plasmid transfer in biofilms: a perspective on limitations and opportunities. npj Biofilms and Microbiomes 2, 16022.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Steenackers, H. P., Parijs, I., Foster, K. R., Vanderleyden, J., 2016. Experimental evolution in biofilm populations. FEMS Microbiology Reviews 40 (3), 373–397.

Sternberg, C., Christensen, B. B., Johansen, T., Nielsen, A. T., Andersen, J. B., Givskov, M., Molin, S., 1999. Distribution of bacterial growth activity in flow-chamber biofilms. Applied and Environmental Microbiology 65 (9), 4108–4117.

Stewart, P. S., 1996. Theoretical aspects of antibiotic diffusion into microbial biofilms. Antimicrobial Agents and Chemotherapy 40 (11), 2517–2522.

Stewart, P. S., 1998. A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. Biotechnology and Bioengineering 59 (3), 261–272.

Stewart, P. S., Franklin, M. J., 2008. Physiological heterogeneity in biofilms. Nature Reviews Microbiology 6 (3), 199–210.

Stewart, P. S., Zhang, T., Xu, R., Pitts, B., Walters, M. C., Roe, F., Kikhney, J., Moter, A., 2016. Reaction-diffusion theory explains hypoxia and heterogeneous growth within microbial biofilms associated with chronic infections. npj Biofilms and Microbiomes 2, 16012.

Stoodley, P., Sauer, K., Davies, D. G., Costerton, J. W., jan 2002. Biofilms as complex differentiated communities. Annual Review of Microbiology 56, 187–209.

Sun, S., Selmer, M., Andersson, D. I., 2014. Resistance to β-lactam antibiotics conferred by point mutations in penicillin-binding proteins PBP3, PBP4 and PBP6 in *Salmonella enterica*. PLoS ONE 9 (5), e97202.

Sutherland, I. W., 2001. The biofilm matrix - an immobilized but dynamic microbial environment. Trends in Microbiology 9 (5), 222–227.

Suzuki, T., Ohsumi, S., Makino, K., 1994. Mechanistic studies on depurination and apurinic site chain breakage in oligodeoxyribonucleotides. Nucleic Acids Research 22 (23), 4997–5003.

Tamrakar, R., Yamada, T., Furuta, I., Cho, K., Morikawa, M., Yamada, H., Sakuragi, N., Minakami, H., 2007. Association between *Lactobacillus* species and bacterial vaginosis-related bacteria, and bacterial vaginosis scores in pregnant Japanese women. BMC Infectious Diseases 7 (1), 128.

Tavare, S., 1986. Some probabilities and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17, 57–86.

Tett, A., Pasolli, E., Farina, S., Truong, D. T., Asnicar, F., Zolfo, M., Beghini, F., Armanini, F., Jousson, O., De Sanctis, V., Bertorelli, R., Girolomoni, G., Cristofolini, M., Segata, N., 2017. Unexplored diversity and strain-level structure of the skin microbiome associated with psoriasis. npj Biofilms and Microbiomes 3 (1), 14.

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., DeBoy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., Madupu, R., Brinkac, L. M., Dodson, R. J., Rosovitz, M. J., Sullivan, S. A., Daugherty, S. C., Haft, D. H., Selengut, J., Gwinn, M. L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K. J. B., Smith, S., Utterback, T. R., White, O., Rubens, C. E., Grandi, G., Madoff, L. C., Kasper, D. L., Telford, J. L., Wessels, M. R., Rappuoli, R., Fraser, C. M., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". Proceedings of the National Academy of Sciences 102 (39), 13950–13955.

Thomas, C. M., Nielsen, K. M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nature Reviews Microbiology 3 (1740-1526 (Print)), 711–721.

Thomas, J. A., Welch, J. J., Lanfear, R., Bromham, L., 2010. A generation time effect on the rate of molecular evolution in invertebrates. Molecular Biology and Evolution 27 (5), 1173–1180.

Thompson, J. R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D. E., Benoit, J., Sarma-Rupavtarm, R., Distel, D. L., Polz, M. F., 2005. Population genotypic diversity within a natural coastal bacterioplankton population. Science 307, 1311–237.

Tillman, D., 1994. Competition and biodiversity in spatially structured habitats. Ecology 75 (1), 2–16.

Toh, H., Weiss, B. L., Perkin, S. A. H., Yamashita, A., Oshima, K., Hattori, M., Aksoy, S., 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Research 16 (2), 149–156.

Tolker-Nielsen, T., Molin, S., 2000. Spatial organization of microbial biofilm communities. Microbial Ecology 40 (2), 75–84.

Torres-Cruz, J., van der Woude, M. W., 2003. Slipped-strand mispairing can function as a phase variation mechanism in *Escherichia coli*. Journal of Bacteriology 185 (23), 6990–6994.

Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R., Cooper, V. S., 2013. Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. Proceedings of the National Academy of Sciences 110 (3), E250–E259.

Truglio, J. J., Croteau, D. L., van Houten, B., Kisker, C., 2006. Prokaryotic nucleotide excision repair: The UvrABC system. Chemical Reviews 106, 233–252.

Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C., Segata, N., 2017. Microbial strain-level population structure & genetic diversity from metagenomes. Genome Research 27 (4), 626–638.

Tsukayama, D. T., Loon, H. J. V., Cartwright, C., Chmielewski, B., Fluit, A. C., Werken, C. V. D., Verhoef, J., 2004. The evolution of *Pseudomonas aeruginosa* during antibiotic rotation in a medical intensive care unit: The RADAR-trial. International Journal of Antimicrobial Agents 24 (4), 339–345.

Tyerman, J. G., Ponciano, J. M., Joyce, P., Forney, L. J., Harmon, L. J., 2013. The evolution of antibiotic susceptibility and resistance during the formation of *Escherichia coli* biofilms in the absence of antibiotics. BMC Evolutionary Biology 13 (1), 22.

Vásquez, A., Jakobsson, T., Ahrné, S., Forsum, U., Molin, G., 2002. Vaginal *Lactobacillus* flora of healthy Swedish women. Journal of Clinical Microbiology 40 (8), 2746–2749.

Verstraelen, H., Swidsinski, A., 2013. The biofilm in bacterial vaginosis: Implications for epidemiology, diagnosis and treatment. Current Opinion in Infectious Diseases 26 (1), 86–89.

von Schillde, M. A., Hörmannsperger, G., Weiher, M., Alpert, C. A., Hahne, H., Bäuerl, C., Van Huynegem, K., Steidler, L., Hrncir, T., Pérez-Martínez, G., Kuster, B., Haller, D., 2012. Lactocepin secreted by *Lactobacillus* exerts anti-inflammatory effects by selectively degrading proinflammatory chemokines. Cell Host and Microbe 11 (4), 387–396.

Wakeley, J., 1998. Segregating sites in Wright's island model. Theoretical population biology 53 (2), 166–174.

Wang, H., Ding, S., Wang, G., Xu, X., Zhou, G., 2013. In situ characterization and analysis of *Salmonella* biofilm formation under meat processing environments using a combined microscopic and spectroscopic approach. International Journal of Food Microbiology 167 (3), 293–302.

Wang, Z.-L., Zhang, D.-Y., Wang, G., 2005. Does spatial structure facilitate coexistence of identical competitors? Ecological Modelling 181 (1), 17–23.

Webster, N. S., Negri, A. P., Base, S., 2006. Site-specific variation in antarctic marine biofilms established on artificial surfaces. Environmental Microbiology 8, 1177–1190.

Weinreich, D. M., Watson, R. A., Chao, L., 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59 (6), 1165–1174.

Weller, C., Wu, M., 2015. A generation-time effect on the rate of molecular evolution in bacteria. Evolution 69 (3), 643–652.

Wentland, E. J., Stewart, P. S., Huang, C. T., McFeters, G. A., 1996. Spatial variations in growth rate within *Klebsiella pneumoniae* colonies and biofilm. Biotechnology progress 12 (3), 316–21.

Wertz, J., Isaacs-Cosgrove, N., Holzman, C., Marsh, T. L., 2008. Temporal shifts in microbial communities in nonpregnant African-American women with and without bacterial vaginosis. Interdisciplinary Perspectives on Infectious Diseases 2008, 181253.

Whitchurch, C. B., Tolker-nielsen, T., Ragas, P. C., John, S., Whitchurch, C. B., Tolker-nielsen, T., Ragas, P. C., Mattickl, J. S., 2002. Extracellular DNA required for bacterial biofilm formation. Science 295 (5559), 59–60.

Whitlock, M. C., 2003. Fixation probability and time in subdivided populations. Genetics 164 (2), 767–779.

Whitman, W. B., Coleman, D. C., Wiebe, W. J., Jun, N., Whitman, W. B., Colemant, D. C., William, J., 1998. Prokaryotes : The unseen majority. Proceedings of the National Academy of Sciences 95 (12), 6578–6583.

Wijker, C. A., Wientjes, N. M., Lafleur, M. V. M., 1998. Mutation spectrum in the *lacI* gene, induced by gamma radiation in aqueous solution under oxic conditions. Mutation Research 403, 137–147.

Wilder, C. N., Allada, G., Schuster, M., 2009. Instantaneous within-patient diversity of *Pseudomonas aeruginosa* quorum-sensing populations from cystic fibrosis lung infections. Infection and Immunity 77 (12).

Williams, I., Venables, W. A., Lloyd, D., Paul, F., Critchley, I., 1997. The effects of adherence to silicone surfaces on antibiotic susceptibility in *Staphylococcus aureus*. Microbiology 143 (7), 2407–2413.

Wira, C. R., Fahey, J. V., Sentman, C. L., Pioli, P. A., Shen, L., 2005. Innate and adaptive immunity in female genital tract: Cellular responses and interactions. Immunological Reviews 206, 306–335.

Wira, C. R., Fahey, J. V., White, H. D., Yeaman, G. R., Givan, A. L., Howell, A. L., 2002. The mucosal immune system in the human female reproductive tract: Influence of stage of the menstrual cycle and menopause on mucosal immunity in the uterus. In: Glass, S., Aplin, J., Guidice, L., Tabibzadeh, S. (Eds.), The Endometrium. Francis & Taylor, New York.

Witkin, S. S., Mendes-soares, H., Linhares, I. M., Jayaram, A., Ledger, W. J., Forney, J., 2013. Influence of vaginal bacteria and D - and L -lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer : Implications for protection against upper genital tract infections. mBio 4 (4), e00460–13.

Wolcott, R. D., Rhoads, D. D., Bennett, M. E., Wolcott, B. M., Gogokhia, L., Costerton, J. W., Dowd, S. E., 2010. Chronic wounds and the medical biofilm paradigm. Journal of Wound Care 19 (2), 45–53.

Wolf, D. P., Sokoloski, J. E., Litt, M., 1980. Composition and function of human cervical mucus. Biochimica et Biophysica Acta 630 (4), 545–558.

Wolfson, J. S., Hooper, D. C., 1989. Fluoroquinolone antimicrobial agents. Clinical Microbiology Reviews 2 (4), 378–424.

Wolin, E. A., Wolin, M. J., Wolfe, R. S., 1963. Formation of methane by bacterial extracts. The Journal of Biological Chemistry 238 (8), 2882–2886.

Wolrath, H., Forsum, U., Larsson, P. G., Borén, H., 2001. Analysis of bacterial vaginosis-related amines in vaginal fluid by gas chromatography and mass spectrometry. Journal of Clinical Microbiology 39 (11), 4026–4031.

Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceeding of the 6th International Congress of Genetics 1, 356–366.

Wright, S., 1941. On the probability of fixation of repciprocal translocations. American Naturalist 75 (761), 513–522.

Xu, K. D., McFeters, G. A., Stewart, P. S., 2000. Biofilm resistance to antimicrobial agents. Microbiology 146 (3), 547–9.

Xu, K. D., Stewart, P. S., Xia, F., Mcfeters, G. A., Huang, C.-t., 1998. Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. Applied and Environmental Microbiology 64 (10), 4035–4039.

Yachi, S., Loreau, M., 1999. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. Proceedings of the National Academy of Sciences 96 (4), 1463–1468.

Ziv, N., Brandt, N. J., Gresham, D., 2013. The use of chemostats in microbial systems biology. Journal of Visualized Experiments 80, e50168.

Zozaya-hinchliffe, M., Lillis, R., Martin, D. H., Ferris, M. J., 2010. Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. Journal of Clinical Microbiology 48 (5), 1812–1819.

Zuckerkandl, E., Pauling, L. B., 1962. Molecular disease, evolution and genic heterogeneity. In: Kasha, M., Pullman, B. (Eds.), Horizons in Biochemistry. Academic Press, New York.

Appendix A: Supplemental methods for Chapter 3

A.1 Experimental evolution

In this study we evolved *Escherichia coli* K12 MG1655 metapopulations that experienced varying degrees of spatial structure. Each metapopulation spanned twelve wells in a 96 well plate, with each well containing one of twelve subpopulations. To manipulate the degree of spatial structure experienced by the evolving metapopulations, we varied the rate of migration between these twelve subpopulations. We tested four degrees of spatial structure: complete structure, high structure, low structure, and no structure. In the following document we provide more detail into the procedure used to generate and maintain these metapopulations during the 150 generations of evolutions. The entire procedure, starting with the inoculation of the ancestors from the freezer, was repeated three times.

A.1.1 Preparing the inocula

The metapopulations were inoculated with equal proportions of two genotypes (the cycloserine sensitive ancestor and a spontaneous cycloserine resistant mutant) that we could differentiate by plating on selective and non-selective media. Both genotypes were inoculated from a freezer stock into 5mL of the glucose limited media. These cultures were incubated overnight at 37°C with shaking at 185 rpm. The cultures were then diluted by a factor of 10^{-6} in 0.85% NaCl and then 100μL of the diluted was plated on non-selective (for the cycloserine sensitive genotype) and selective (for the resistant genotype) agar media. These two plates were incubated overnight at 37°C. Following this incubation, 12 colonies were randomly selected from each of the two plates and resuspended in fresh media (Figure A.1A) The resulting colony suspensions were then standardized by OD at 600nM. Pairs of resistant and sensitive genotypes were combined together creating twelve colony suspension mixtures, each of which contained roughly equal proportions of the sensitive and one resistant genotypes (Figure A.1A). The twelve colony suspension mixtures where then diluted 1:10 in fresh media and then 15μL was inoculated into 185μL of fresh media in four wells of a 96 well plate, one each for the four metapopulation types (complete structure, high structure, low structure, and no structure). The 96 well plate containing the four metapopulation was then incubated for 24 hours at 37°C with shaking at 185 rpm. Metapopulations were then serially transferred for 150 generations.

A.1.2 Manipulating the degree of spatial structure

The degree of spatial structure experience by the evolving metapopulations was manipulated by varying the migration rates between the 12 subpopulations in each metapopulation. All metapopulations were subcultured into fresh media daily (1:128 dilution, providing 7 generations per day), however the protocol differed depending on the degree of spatial structure in question. For the completely structured metapopulations no migration occurred between the subpopulations. Instead, each subpopulation

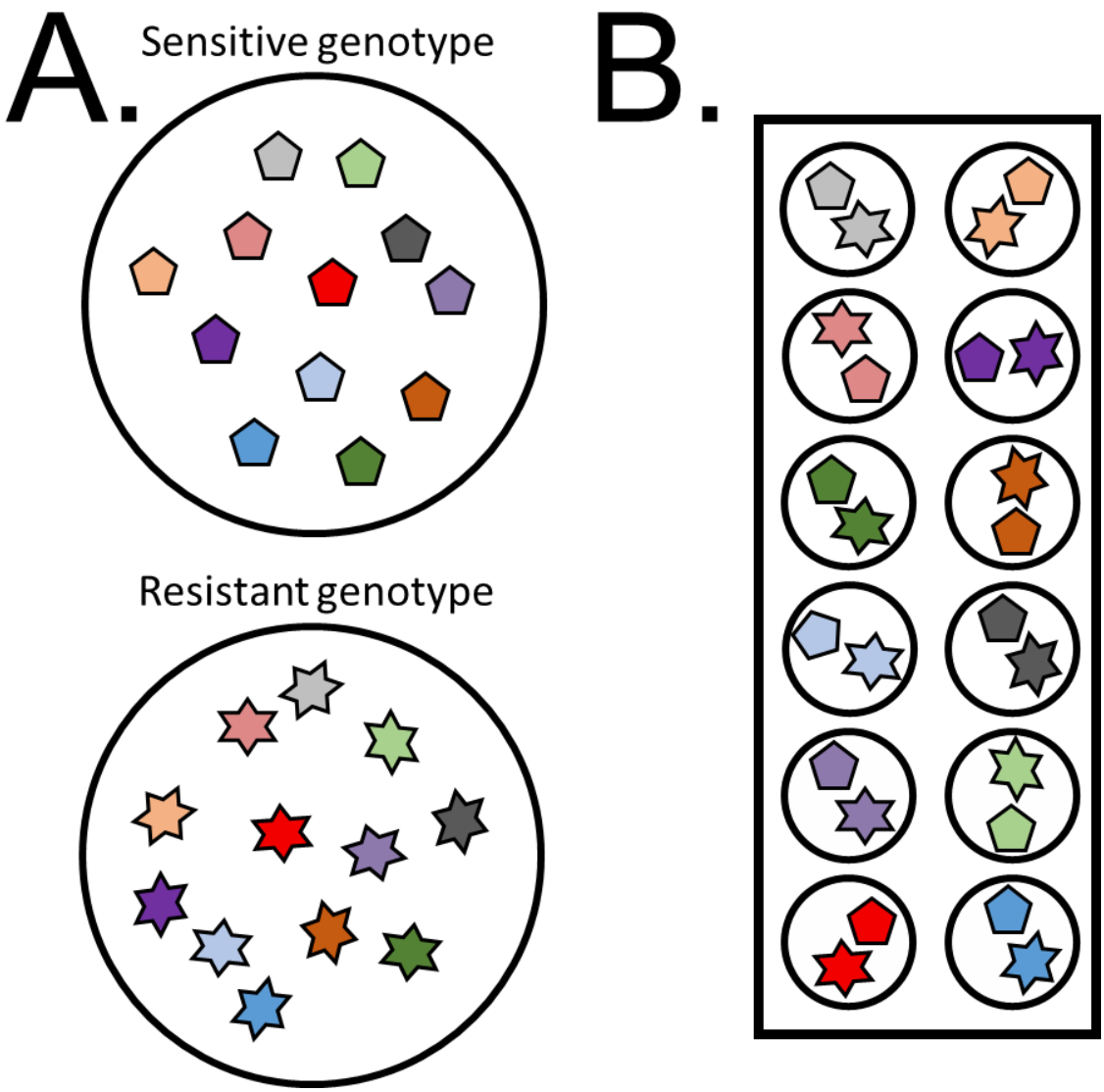

Figure A.1: Twelve sensitive and twelve resistant colonies were randomly selected (A), standardized by OD at 600nM, and then combined in equal volumes. These twelve colony mixtures were used to inoculate four sets of 12 subpopulations (B)

was simply subcultured into fresh media every day (diluted 1:10 in fresh media and then 15μL was inoculated into 185μL of fresh media). While in the no structure metapopulations, the 12 subpopulations were combined together in equal volumes. This migrant mixture was then used as the inoculum to subculture the 12 subpopulations (diluted 1:10 in fresh media and then 15μL was inoculated into 185μL of fresh media. In the low and high structure metapopulations, the subpopulations were first subcultured into fresh media (diluted 1:10 in fresh media and then 15μL was inoculated into 175μL of fresh media). Additionally, the 12 subpopulations were combined together in equal volumes to create a migrant mixture. This migrant mixture was diluted in fresh media by a factor of 10^{-3} (low structure) or 10^{-5} (high structure) and then μL of the diluted migrant mixture was added to each subpopulation (bringing the total volume to 200μL in each well). This provides a migration rate of 2,000 individuals per day between any two subpopulations in the low structure and 20 individuals per day in the high structure.

Appendix B: Supporting information for Chapter 5

In Chapter 5 we showed that biofilm populations accumulated kanamycin and rifampicin resistant variants prior to antibiotic treatment. One possible explanation for this increase is that these resistant mutants were selected for in the biofilm despite the absence of the antibiotic. However, we argue that it is more likely that the increase was driven by a neutral process of spontaneous mutation. Here, we present simple calculations using our own empirical estimates of the mutation rates towards kanamycin and rifampicin resistance to demonstrate that it is plausible that spontaneous mutations drove the observed accumulation of resistant cells.

Two different mutation rates toward resistance to each antibiotic were considered, one for exponentially growing cells and one for stationary phase cells. It is likely that members of a biofilm population experience some combination of the two rates depending on their location in a biofilm. Cells near the top have first access to nutrients in the bulk medium and are therefore likely to resemble those in exponentially growing liquid cultures. Whereas cells near the bottom experience limited growth due to a lack of nutrients and are likely to more closely resemble cells in stationary phase liquid cultures. For this reason, we have calculated both the number of generations of exponential growth required and the number of hours in stationary phase that would be required to explain the observed accumulation of resistant clones. It is likely that some combination of these two estimates best describes the experimental data.

B.1 Kanamycin

We observed that the frequency of kanamycin resistant (Kan^r) mutants increased from 2.31 x 10^{-8} Kan^r mutants per CFU at the time of inoculation, to 1.07 x 10^{-6} Kan^r mutants per CFU after 15 days of antibiotic-free cultivation.

If we assume that the increase from day 0 to day 15 was linear, then Kan^r mutants increased in frequency at a rate of $6.98*10^{-8}$ new Kanr mutants per CFU per day. In order for mutation alone to explain this increase, the mutation rate would need to exceed this value.

B.1.1 Analysis using the mutation rate during exponential growth

Rate of increase observed in biofilms: 6.98 x 10^{-8} new Kan^r mutants/CFU/day

Rate of mutation to kanamycin resistance: 9.48 x 10^{-8} new Kan^r mutants/CFU/gen

$$\frac{6.98 * 10^{-8}\, new\, Kan^r\, mutants\, per\, CFU\, per\, day}{9.48 * 10^{-8}\, new\, Kan^r\, per\, CFU\, per\, gen} = 0.74\, gen\, per\, day$$

B.1.2 Analysis using mutation rate during stationary phase

Rate of increase observed in biofilms: 6.98 x 10^{-8} new Kan_r mutants/CFU/day

Rate of mutation to kanamycin resistance: 6.38 x 10^{-7} new Kan_r mutants/CFU/day

$$\frac{6.98 * 10^{-8}\, new\, Kan^r\, mutants\, per\, CFU\, per\, day}{6.38 * 10^{-7}\, new\, Kan^r\, per\, CFU\, per\, gen} = 0.11$$

Kan^r mutants increased in the biofilms at approximately 1/10 the rate of mutation during stationary phase.

B.2 Rifampicin

We observed that the frequency of rifampicin resistant (Rif^r) mutants increased from 5.81 x 10^{-8} Rif^r mutants per CFU at the time of inoculation, to 2.46 x 10^{-6} Rif^r mutants per CFU after 15 days of antibiotic-free cultivation.

If we assume that the increase from day 0 to day 15 was linear, then Rif^r mutants increased in frequency at a rate of 1.60*10^{-7} new Rif^r mutants per CFU per day. In order for mutation alone to explain this increase, the mutation rate would need to exceed this value.

B.2.1 Analysis using the mutation rate during exponential growth

Rate of increase observed in biofilms: 1.60 x 10^{-7} new Rif^r mutants/CFU/day

Rate of mutation to rifampicin resistance: 3.54 x 10^{-8} new Rif^r mutants/CFU/gen

$$\frac{1.60 * 10^{-7}\, new\, Rif^r\, mutants\, per\, CFU\, per\, day}{3.54 * 10^{-8}\, new\, Rif^r\, per\, CFU\, per\, gen} = 4.52\, gen\, per\, day$$

B.2.2 Analysis using mutation rate during stationary phase

Rate of increase observed in biofilms: 1.60 x 10^{-7} new Rif^r mutants/CFU/day

Rate of mutation to rifampicin resistance: 6.58 x 10^{-8} new Rif^r mutants/CFU/day

$$\frac{1.60 * 10^{-7}\, new\, Rif^r\, mutants\, per\, CFU\, per\, day}{6.58 * 10^{-8}\, new\, Rif^r\, per\, CFU\, per\, gen} = 2.43$$

Rif^r mutants increased in the biofilms at approximately 2.5 times the rate of mutation during stationary phase.

Appendix C: Copyright Permissions

Included below are the appropriate copyright permissions for the previously published works that appear in Chapters 2 and 6 of this dissertation.

SPRINGER NATURE

Dr. Michael France
University of Idaho
Institute for Bioinformatics and Evolution Studies

April 26, 2018

Dear Dr. France,

Re: Your permission request dated: April 11, 2018,
your reference: Dr. Michael France ("Licensee"),
our reference: Springer Customer Service Center GmbH ("SCSC")

Request for permission to use: your chapter "Effects of spatial structure and reduced growth rates on evolution in bacterial populations" from Rampelotto, Pabulo H (Ed.): **Molecular Mechanisms of Microbial Evolution (in prep. – scheduled for publication September 10, 2018)**
ISBN 978-3-319-69077-3
("Licensed Material") for your forthcoming thesis to be published by the University of Idaho.

This permission covers:

- non-exclusive print and electronic rights only
- the English language
- distribution in the following territory: Worldwide
- copyright fee: gratis
- for the avoidance of doubt, derivative rights are excluded from this licence

This permission is subject to the following conditions:

1. An acknowledgement is made as follows: 'AUTHOR, TITLE, published [YEAR] [publisher - as it appears on our copyright page] reproduced with permission of SCSC'
2. This permission does not apply to quotations, figures or tables from other sources which may be part of the material to be used. If the book acknowledgements or credit line on any part of the material you have requested indicates that it was reprinted or adapted by SCSC with permission from another source, then you should also seek permission from that source to reuse the material.
3. This permission is only valid if no personal rights, trademarks, or competitive products are infringed;
4. This permission is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published (with ISBN) , permission needs to be reobtained.
5. This permission includes use in an electronic form after an embargo of 6 month, provided it is an author-created version of the thesis on his/her own website and his/her university's repository, including UMI (according to the definition on the Sherpa website: http://www.sherpa.ac.uk/romeo/);
6. The Licensee may not make the Licensed Material available in a manner intended to allow or invite a third party to download, extract, redistribute, republish or access the Licensed Material as a standalone file.
7. This permission is personal to you and may not be sublicensed, assigned, or transferred by you without SCSC's written permission, unless specified and agreed with SCSC.

8. Delivery and use of the material must be in a format that retains the integrity of the text. Figures, illustrations, and tables may be altered minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of SCSC.
9. We do not routinely require a complimentary copy of your product on publication; however we reserve the right to receive a free copy on request at a later stage.
10. We may terminate this licence with immediate effect on written notice to you for any reason. Upon termination of this licence for any reason, all rights in this licence shall immediately revert to SCSC.
11. This licence shall be governed by, and shall be construed in accordance with, the laws of the Federal Republic of Germany.

Yours sincerely

Springer Nature Permissions Team

On behalf of SCSC

Alex Collery: Apr 5, 2018 4:01:16 PM

Dear Mr. Michael France,

ASM authorizes an advanced degree candidate to republish the requested material in his/her doctoral thesis or dissertation free of charge, provided that proper credit is given to the original ASM publication. If your thesis, or dissertation, is to be published commercially, then you must reapply for permission. Please contact us if you have any questions.

Thank you.
ASM Journals
journals@asmusa.org