Population genetics of rifampicin-resistant *Pseudomonas aeruginosa*

Danna R. Gifford

Department of Zoology

Somerville College
University of Oxford

Thesis submitted for the DPhil degree in Zoology

July 2014

Abstract

Population genetics of rifampicin-resistant *Pseudomonas aeruginosa*
Danna R. Gifford, Somerville College
Submitted for DPhil in Zoology
Trinity 2014

Antibiotic resistance is generally associated with a cost in terms of reduced competitive fitness in the absence of antibiotics. Despite this 'cost of resistance', the cessation of antibiotic treatment does not result in significant reductions in the prevalence of resistance. The maintenance of resistance, in spite of the costs, has been attributed to the rarity of reversion mutations, relative to compensatory mutations at other loci in the genome. However, the large size of bacteria populations, and the potential for migration, suggest that reversion mutations should occasionally be introduced to resistant populations. In this thesis, I show that additional mechanisms can prevent fixation of reversion mutations even if they do occur. Using an experimental evolution approach, with rifampicin resistance in *Pseudomonas aeruginosa* as a model system, I measured the costs of resistance in several environments and followed the adaptive dynamics of resistant populations where a sensitive lineage had invaded by migration. The results suggest that several additional mechanisms contribute to the maintenance of antibiotic resistance. Most rifampicin resistance mutations are not unconditionally costly in all environments, suggesting that migration between environments could maintain a resistant reservoir population. In environments where resistance is initially costly, the fixation of a revertant is not guaranteed, even if introduced through migration. Revertant fixation was impeded or prevented by clonal interference from adaptation in the resistant strain. Revertants that did successfully replace the resistant strain were forced to adapt to do so. Contrary to assumptions in the existing literature, fitness in the resistant strains was not recovered by general compensatory mutations, but instead by adaptive mutations specific to the environment. The data challenge several assumptions about the maintenance of antibiotic resistance: that resistance mutations are always costly, that the rarity of back mutations prevents the reversion of resistance, and that resistant strains recover fitness by compensatory mutations.

Acknowledgements

It is with deepest gratitude that I acknowledge the many people who made this thesis possible.

I would like to acknowledge the support and guidance of my supervisor, Dr Craig MacLean, who encouraged me to persevere with my project, maintaining a constant sense of optimism despite my occasionally pessimistic outlook. I benefited greatly from advice and comments given by my viva examiners, Professor Kevin Foster and Dr Nick Colegrave (University of Edinburgh). Early feedback on my project was given by my transfer committee, Professor Martin Maiden and Dr Andy Gardner (now at the University of St Andrews).

I would like to thank my colleagues: Mila Kojadinovic who taught me how to do DNA extractions, Macarena Toll-Riera for expertise in next-generation sequencing analysis, Tom Vogwill for an encyclopedic knowledge of rifampicin resistance evolution, Qi Qin for his work on transcription rate of *rpoB* mutants, and Karl Heilbron, Alvaro San Millan, Ros Whiteley, Victoria Furioso, and the theoretical/experimental evolution group at large (who are too many to name) for stimulating discussions. I would also like to thank Jonas Schluter and Sara Mitri for moral support during my 'writing-up' stage.

On a personal note, I would like to thank my friends from Somerville College, particularly Ruben, Carlo, Annie, and Hugh, for all the shenanigans over the last three years, and my partner James, for his patience and understanding during weekend experiments and 12-hour thesis-writing days. Finally, I would like to acknowledge the support of my parents: my Dad, for fostering in me a life-long interest in science; my Mom, for teaching me strength and integrity; and my step-parents, Marci and Cory, for their encouragement.

I also thank my sources of financial support for their generosity: the Natural Sciences and Engineering Research Council of Canada, the Clarendon Fund, Somerville College (Graduate Scholarship), and the Canadian Centennial Scholarship Fund.

Statement of contributions

I declare that this thesis is substantively my own work.
Contributions by others are listed below:

- A. H. Hall provided the data in Figure 3.1
- M. Kojadinovic assisted with DNA extractions (subsection 5.2)
- M. Toll-Riera designed the whole-genome analysis pipeline for sequence analysis (subsection 5.2)

1 | Introduction

The discovery of antibiotics has undoubtedly changed the course of human history, but the emergence of antibiotic resistance threatens to return humanity to the pre-antibiotic era. As the rate of resistance evolution eclipses the rate of novel drug discovery, we are rapidly running out of antibiotics that work (Overbye and Barrett 2005). It is therefore imperative that we understand the evolutionary forces driving the emergence and maintenance of resistance in order to prevent it from occurring. Using experimental evolution techniques, this thesis will address multiple facets of the problem of resistance to develop strategies for its prevention or reversal. The experiments will involve the bacterium *Pseudomonas aeruginosa* and the antibiotic rifampicin as a model system for antibiotic resistance.

1.1 Antibiotics and antibiotic resistance

Antibiotics are a group of compounds that have inhibitory or killing action against bacteria. The group encompasses a vast array of chemical structures with differing modes of action, and are generally assigned to a class based on their chemical structure (e.g. penicillins, aminoglycosides, glycopeptides, etc.). In general, antibiotics work by exploiting molecular pathways specific to bacteria, disrupting some aspect of cellular metabolism or preventing the synthesis of either DNA, RNA, proteins or cell walls.

Antibiotic resistance is the ability for bacteria to grow or persist in spite of the

presence of normally inhibitory or lethal concentrations of an antibiotic. As there are many different antibiotic modes of action, there are also many mechanisms of resistance (Dever and Dermody 1991). Resistance mechanisms generally fall into one of four groups: target alteration (e.g. mutation of binding site), metabolic pathway modification (e.g. skipping blocked steps), antibiotic degradation, or concentration reduction (e.g. efflux pumps or reduced permeability). Resistance mechanisms that degrade or alter antibiotic influx/efflux have often been co-opted from enzymes or other proteins with various functional roles (Wright 2007). Situational or behavioural mechanisms of resistance are also possible, observed in species that form biofilms or metabolically-inactive 'persister cells' (Singh *et al.* 2009; Høiby *et al.* 2010). Generally, resistance mechanisms are specific to each antibiotic; however, some provide 'cross-resistance' to multiple drugs. In particular, increased efflux pump expression, reduced cell permeability and biofilm production can confer resistance to multiple drugs.

The diversity of antibiotic modes of action, and the corresponding resistance mechanisms, have both arisen as the the product of millions years of microbial evolution (Wright 2007). Although antibiotics can kill bacteria, it is debated whether that is their original functional role (Yim *et al.* 2006; Davies and Davies 2010). The existence of resistance mechanisms, however, suggests they are at least sometimes used in cell-cell competition. Some resistance mechanisms have a long evolutionary history (Hall and Barlow 2004), while those that confer resistance to wholly-synthetic antibiotics have arisen more recently as a result of clinical use (Robicsek *et al.* 2006). The greatest diversity of antibiotic-producing and -resistant bacteria is in soil (D'Costa *et al.* 2006), with functional resistance elements being detected in 30,000 year old permafrost (D'Costa *et al.* 2011). Horizontal gene transfer has liberated resistance genes from the soil, and allowed them to move into pathogens (Courvalin 1994; Wright 2010).

Compared to the long evolutionary history of antibiotics, their exploitation by humanity is in its infancy, spanning just over 100 years. As with most other technological innovations, the development of antibiotics was partly mediated by humanity's proclivity for sex, war, and cheap food. Introduced in 1909, the first marketed compound with antibacterial activity was Salvarsan, one of the first effective treatments for syphilis (Aminov 2010). Demand for antibiotics during World War II led to the development of industrial production methods for penicillin (Neushul 1993). The increasing pressures on our food supply have initiated a shift in agricultural practices toward high-density farming, which prompted the use of antibiotics to prevent disease and stimulate growth (Misato *et al.* 1977). Since then, contemporary global annual consumption of antibiotics has reached as high as 10^8 kg (Wise 2002). As most antibiotics are naturally derived, use on this scale has uncovered the cache of resistance elements accumulated over the evolutionary history of bacteria. The current drug-discovery approach of altering known resistance compounds has not helped to slow the rate of resistance development, as pre-existing resistance scaffolds often require few modified to counteract these changes. However, even wholly-synthetic antibiotics are rendered useless by the evolution of novel resistance pathways, often by mechanisms that provide cross resistance to many different antibiotics simultaneously (Robicsek *et al.* 2006). Given the rapid generation time of bacteria, it is little wonder that resistance has developed to every natural and synthetic class of antibiotics within 10 years of their first clinical use (Walsh 2003).

1.2 Fitness costs of antibiotic resistance

Resistance to antibiotics was first observed shortly after their discovery, and the threat to treatment efficacy was immediately recognized (e.g. Demerec 1948).

However, initial fears were subdued by the observation that resistance imposes a cost, in terms of bacterial growth or competitive ability, when antibiotics are absent. Generally, resistance to antibiotics imposes large fitness costs in antibiotic-free environments (Sander *et al.* 2002). This is often because resistance involves changes to fundamental cellular processes or large-scale changes in cellular physiology (Dever and Dermody 1991). Costs of resistance have been demonstrated by numerous *in vitro* studies (reviewed in Andersson and Hughes 2010), but detecting resistance costs in natural bacterial isolates is more difficult, as the genetic basis for resistance may not be known. This is problematic because *in vitro* estimates of the cost may not correlate with actual costs in other environments (Andersson 2006). The cost of resistance lead to the assumption that resistant organisms would be replaced by sensitive ones if antibiotic treatment was stopped (Levin *et al.* 1997). This prediction was born out of population genetics models predicting the spread of beneficial mutations (such as reversion mutations in a resistant population) that have stood largely untested for nearly a century (Fisher 1922; Haldane 1927; Wright 1931). Although extensions and modifications of the original models take slightly different forms (reviewed in Patwa and Wahl 2008), at their core all models predict that the probability of fixation of reversion mutations restoring sensitivity should increase with the fitness cost of resistance. As fitness costs of resistance are usually large, and the reversion mutation is assumed to be the fittest possible mutant (if it was not, adaptation would already have selected for a fitter type), the prediction followed that the reversion mutation should always fix, should it occur. This prediction has largely proven to be false—once resistance arises, it is nearly impossible to eliminate. Large-scale antibiotic cessation trials have failed to detect any consistent reductions of resistance levels in streptococcal infections (Sepälä *et al.* 1997; Kataja *et al.* 1999; Arason *et al.* 2002; Bergman *et al.* 2004).

Similarly, there is a wealth of *in vitro* studies that have essentially failed to detect reversion of any significance (Andersson 2006). What little evidence that exists for reversion suggests that it should take a long time, on the order of years in real-life conditions (Andersson and Hughes 2010).

Why did the models fail to predict the maintenance of antibiotic resistance, despite its associated costs? The rarity of reversion mutations is usually invoked as the reason for the failure. The probability that a reversion mutation will occur is dictated by the per base pair mutation rate, which is on the order of 10^{-8} (Tago *et al.* 2005), assuming resistance arose from a single substitution, other mutations such as indels or duplications may be more difficult to reverse. In contrast, the number of mutations that can compensate for the fitness, without compromising resistance, is an order of magnitude higher (Poon *et al.* 2005). Theory and evidence suggest that the fitness benefit of compensatory mutations should increase with the magnitude of the fitness cost of resistance (Martin and Lenormand 2006; Barrick *et al.* 2010; Gifford *et al.* 2011), a feature not captured by probability of fixation models.

Although compensatory mutations have been extensively studied both clinically and *in vitro* (reviewed in Andersson and Hughes 2010), still little is known about many of their fundamental biological features. Many outstanding questions persist, particularly on the genetic basis for compensation. For example, are compensatory mutations specific to each resistance mutation within a gene, or are they general to the resistance gene? Do some resistance mutations have a greater potential for compensation? How does varying the availability of compensatory mutations affect the probability of reversion? Are compensatory mutations truly beneficial across all environments? The answers to these questions are potentially important for improving future treatment protocols, particularly as part of the process for novel drug discovery.

1.3 Experimental evolution approach

Previous models and experiments predicting the adaptive evolution of resistant populations treat the cost of resistance as a black-box, ignoring the specific mutations conferring resistance. Using rifampicin resistance in *Pseudomonas aeruginosa* as a model system, I will investigate whether this black-box approach to resistance is valid, and under what circumstances it might fail. To accomplish this, I will use an experimental evolution approach, using rifampicin resistance in the opportunistic pathogen *Pseudomonas aeruginosa* as a model system to study the cost and maintenance of resistance and the adaptive evolution of resistant strains. Experimental evolution involves propagating replicate populations across many generations in the face of a stressor to which the populations adapt (e.g. a new environment, a toxin, a predator or competitor, etc., Buckling *et al.* 2009). Rifampicin resistance in *P. aeruginosa* is a tractable model system for this project, due to the properties of each described below.

Rifampicin

Rifampicin is a semi-synthetic member of a class of antibiotics known as the rifamycins[1], which were first discovered in 1957 from soil bacteria collected in the Côte d'Azur. The natural rifamycins showed high activity against gram-positive bacteria, and some activity against gram-negative bacteria; however, they could not be given orally in their natural form, and so a number of semi-synthetic derivatives were produced (Maggi *et al.* 1966). The most effective of these derivatives was the *N*-amino-*N'*-methylpiperazine (AMP) derivative of rifamycin B, hence its name rif–amp–icin (Aronson 2004). Its primary clinical use

[1]The inspiration for the name 'rifamycin' was a French gangster film "*Du rififi chez les hommes*", a favourite of the discoverers (Aronson 2004).

has been against mycobacterium infections, such as tuberculosis (*Mycobacterium tuberculosis*) and leprosy (*Mycobacterium leprae* and *Mycobacterium lepromatosis*).

Rifampicin works by disrupting RNA transcription by binding within the DNA/RNA channel of RNA polymerase (RNAP) and preventing the elongation of RNA transcripts beyond 2-3 nt in length (Wehrli and Staehelin 1971; Campbell *et al.* 2001). RNAP itself is made up of multiple subunits that are strongly conserved across bacterial species: two of α (gene *rpoA*), and one each of β (*rpoB*), β' (*rpoC*), and ω (*rpoZ*). The high degree of RNAP sequence similarity across species means that rifampicin is broad-acting, and resistance mutations tend to be common across species (Campbell *et al.* 2001). Resistance to rifampicin most often occurs through changes to the β subunit through mutations in *rpoB*, which lower affinity for rifampicin (Wehrli 1983). 95% of resistance mutations in *M. tuberculosis* occur in the 'rifampicin resistance determining region' (RRDR), an 81 base-pair region spanning codons 507–533 of RpoB (Zaczek *et al.* 2009). Norcardia and related genera can also become resistant by chemically inactivating rifampicin, but this route to resistance is much less common (Tanaka *et al.* 1996; Baysarowich *et al.* 2008). Rifampicin resistance evolves rapidly under both laboratory and clinical conditions, as many small mutational changes (e.g. single base-pair substitutions and small insertions) in *rpoB* can confer resistance.

Rifampicin resistance is an ideal model system for studying the population genetics of antibiotic resistance. Owing to the highly conserved nature of the RNAP complex, results from one species are likely generalizable across others. This is advantageous, as experimental evolution with mycobacteria, which rifampicin is used to treat, would be impractical as their growth rate *in vitro* is too slow. Several *rpoB* mutations can give rise to resistance, with variable costs of resistance (Hall *et al.* 2010) that can be manipulated experimentally (Hall *et al.* 2011). An added advantage of rifampicin is that, unlike for many

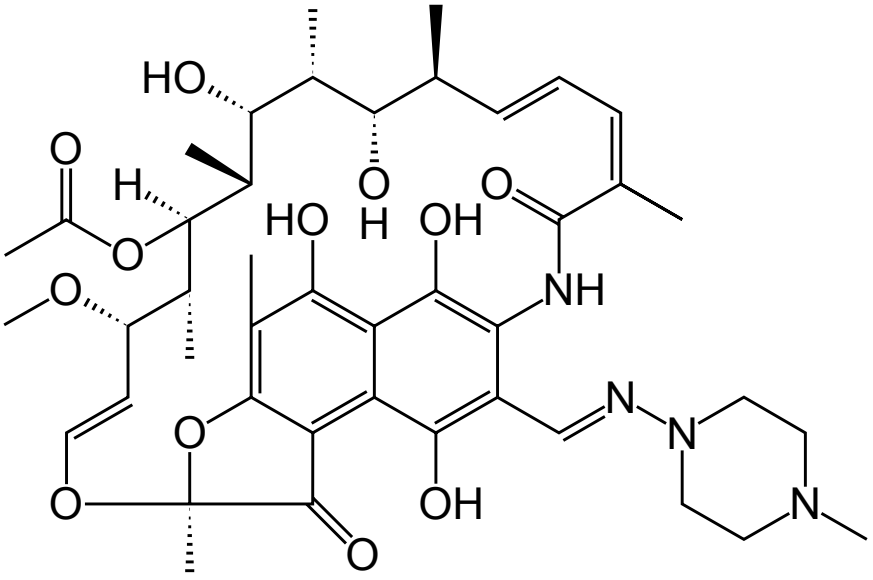

Figure 1.1: Structure of rifampicin.

antibiotic resistance genes, whole-genome sequencing of a number of clinical isolates has indicated probable locations for compensatory mutations (Brandis *et al.* 2012), which makes it easier to determine from sequence data whether a mutation is likely to be adaptive or compensatory.

Pseudomonas aeruginosa

P. aeruginosa is a bacterium commonly found in soil and water, but is also an opportunistic pathogen, responsible for many hospital acquired infections. *P. aeruginosa* infections are especially prevalent in patients with cystic fibrosis where it colonizes mucus retained in the lungs (Govan and Deretic 1996). *P. aeruginosa* is broadly drug resistant due to multi-drug efflux pump systems, particularly MexAB and MexXY, which become hyper-expressed via mutations when selected under antibiotic pressure (Poole 2001). *P. aeruginosa* also produces biofilms, which are inherently drug resistant due to their physical structure

(Drenkard and Ausubel 2002). These features make treating *P. aeruginosa* infections with conventional antibiotics difficult.

Despite being resistant to many antibiotics, *P. aeruginosa* infections are not presently treated with rifampicin, so studies with this bacterium permit the study of rifampicin resistance without interference from historical contingency. The mutational sites in *rpoB* that confer resistance in other species also confer resistance in *P. aeruginosa* (MacLean and Buckling 2009). Wild-type *P. aeruginosa* strains generally have a higher baseline minimal inhibitory concentration (MIC, 32 mg/l to 64 mg/l) than other species (e.g. *Staphylococcus aureus*: 0.002 mg/l, *Escherichia coli*: 8 mg/l, Williams and Piddock 1998), possibly due to efflux pumps or a permeability barrier (Yee *et al.* 1996).

1.4 Thesis outline

This thesis addresses several questions pertaining to population genetics and evolutionary ecology of rifampicin-resistant *P. aeruginosa* in the absence of rifampicin. The overarching theme of the thesis is that a black-box approach to the costs of resistance may be naïve, and that the genetic basis for resistance, the competitors a strain meets, and the environment they are both faced with, play an important role in shaping adaptive evolution of resistant populations. In chapter 2, I first describe the methods developed to conduct high-throughput fitness assays, which were required for all of the following experiments. Chapter 3 investigates the genotype-by-environment effects of rifampicin resistance mutations. Here, I discovered that two mutants with different genetic bases for resistance paid the same fitness cost. These mutants became the focus of chapters 4, 5, and 6. Because the two *rpoB* mutants paid the same cost of resistance, under black-box models, the two strains should have behaved the same. Chapter 4 determines the probability by which a rifampicin-sensitive strain is able to invade

populations of the two mutants. Chapter 5 investigates mechanisms behind a difference in their evolvability and explores the genetic basis for adaptation. Finally, chapter 6 looks how the difference in evolvability of a resistant resident population affects the fitness of sensitive invaders that successfully invade.

2 | Developing high-throughput fitness assays

Abstract

In experimental evolution, experimental results are generated in two phases: first, 'selection', where replicate populations adapt to novel conditions or stress, and second, 'fitness assay', where the realized adaptation of each population is measured. Presently, the logistics of the fitness assay are the major constraint on the number of replicate populations or treatments that can be studied in experimental evolution, which limits the scope of questions that can be answered with this approach. This methods chapter introduces two new protocols for 'high-throughput' fitness assays that were required for later chapters. One method uses a combination of fluorometry and luminometry (measuring fluorescence and light intensity) to measure the proportion of the population represented by a labelled strain versus total population size. Another method uses flow cytometry to measure the proportion of labelled and unlabelled strains. I discuss previous approaches to high-throughput measuring of fitness, and compare the components of these two methods with 'selective plating', the established protocol for measuring competitive fitness. Finally, I discuss the advantages and disadvantages of each method, and describe scenarios for when each assay would be most appropriate to use. Subsequent chapters in this thesis will reference to the protocols described here.

2.1 Introduction

Fitness is arguably the most fundamental concept of evolutionary biology, and of central importance in this thesis. Although stated in different ways, fitness is essentially defined as the success of an individual) in a population in contributing progeny to the following generation, relative to other individuals (or genotypes) in the same population (Orr 2009). Although difficult in natural populations, measuring fitness in bacterial microcosms is comparatively straight-forward. Two general approaches to measuring fitness have been commonly used: the first compares the maximum growth rate of pure cultures of different strains. Maximum growth rate is a measure of the 'adaptedness' of an individual to an environment and is a component of fitness (Bell 2008). However, although easy to measure, maximum growth rate is not always indicative of how strains will grow together, as other biological interactions (e.g. toxin production, social cheating, cross-feeding, etc.) may influence how two genotypes grow in the presence of each other. A better approach that better satisfies the definition of fitness is to grow genotypes together in the same microcosm and observe the change in their frequency over time. Traditionally, this has been accomplished by growing the population on selective agar media that permit growth of one type, but not the other. This method is labour- and resource-intensive to use for many genotypes simultaneously, which has limited the scope of classic experimental evolution experiments in bacteria to the order of tens of replicates, such as the 12 'long-term *E. coli* lines' (Lenski *et al.* 1991), which have now evolved for more than 50,000 generations (Wiser *et al.* 2013).

The experiments planned for my thesis required an order of magnitude more replicates than could be practically assayed using current methods, which required the development of a high-throughput assay for competitive fitness. In

this chapter, I first describe two classic methods of measuring fitness (relative maximum growth rate and selective plating), then develop two high-throughput methods: a microplate reader-based assay that uses fluorescence and luminescence as markers for frequency and total population density, and a flow cytometry-based method that uses fluorescence and cell counts for frequency and total population density. Using example data, I show that the three competition methods (selective plating, microplate reader and flow cytometer) all give comparable results, suggesting any method may be selected when appropriate. I then discuss the advantages of each method for experiments with different numbers of replicates.

2.2 Relative maximum growth rate

A simple approach to measuring the fitness of multiple mutant strains at once is to measure their maximum growth rate in pure culture, relative to the growth rate of a wild-type strain. Relative maximum growth rate (relative v_{max}) is estimated by measuring population density at several time points. This can be accomplished by plating on agar and counting colonies, or using a microplate reader and measuring the change in a chosen signal over time. Typically, optical density (a measure of culture turbidity) has been used to measure growth rates, but it suffers from a number of drawbacks: it cannot exclude dead cells, it is affected by cell size, and it has a narrow threshold at which values increase log-linearly over time.

Growth rate can be measured more sensitively by the detection of a fluorescence or luminescence genetic marker; the luminescence marker is particularly useful for measuring growth rates at low density, allowing the 'lag phase' of growth to be captured. Maximum growth rate is calculated either by fitting a log-linear regression to the data over a window of four to five time points, or

by fitting a growth model to the data, from which parameters such as lag time, growth rate and maximum density can be estimated. Optical density was used to measure growth rates of several carbon sources in chapter 3, as these strains did not bear a suitable marker for the other method. Chapter 6 uses the more precise luminescence-based approach, and also fits a growth model to determine growth rate.

2.3 Competitive fitness

Although measuring relative $v_{\max}$ allows many strains to be assayed simultaneously, it neglects interactions between different bacterial genotypes that may arise during evolution. Instead, it is better to measure fitness by competing strains together in a common environment. Direct competition against a common genotype is the best measure of fitness, as it directly relates to fitness defined as the 'relative contribution of progeny to the next generation' (Fox and Wolf 2006). However, measuring competitive fitness generally requires more effort than measuring growth rates. Competitive fitness is measured by growing a strain in the presence of a 'reference strain' and is calculated by measuring the change in proportion of the two strains from the initial culture (i) to the final culture (f). The competition assay involves (1) mixing initially dense cultures of the two strains together, (2) measuring the initial proportion of each type in the mixture, (3) diluting the mixture into fresh medium by a defined a factor D, (4) growing for a defined period, and (5) measuring the final proportion of each type. What differs between the following methods is how the initial and final proportions are measured.

Assaying by selective plating

To measure competitive fitness by selective plating, the two strains must be phenotypically distinguishable (e.g. through antibiotic resistance, auxotrophy, insertion of *lacZ* operon, etc.), which can be observed by plating on selective agar plates. Initial and final cultures are plated on solid selective media that distinguishes the two genotypes. For my assays, I used antibiotic resistance markers to either gentamicin, streptomycin or rifampicin, and agar plates containing these antibiotics as the selective medium. Competitive fitness is then calculated as the ratio of the number of doublings of each competitor during a fixed time period. For the fitness of 'strain A', relative to reference 'strain B', this calculation is (Bell 2008):

$$w_{AB} = \frac{\log_2 \left[\mathrm{A}_f / (D \times \mathrm{A}_i) \right]}{\log_2 \left[\mathrm{B}_f / (D \times \mathrm{B}_i) \right]}, \tag{2.1}$$

where the upper-case letters denote proportions of each type.

Assaying by microplate reader

Although sensitive, the selective plating method is costly in terms of time and materials required, and is generally not logistically suitable for simultaneous measurement of the fitness of many strains or in multiple environments. To overcome these limitations, I developed a novel competitive fitness assay using a microplate fluorometer/luminometer, a device that measures both fluorescence intensity (FI) and luminescence. The specific instrument used for my research was the filter-based fluorometer, the FLUOstar OPTIMA (BMG Labtech, Offenburg, Germany), although the method should apply equally to other capable devices.

The proportion of one genotype is instead measured using a genetically-inserted fluorescence marker (in my experiments, either green or yellow fluorescent protein, GFP or YFP). The method is illustrated in Figure 2.1. A volume

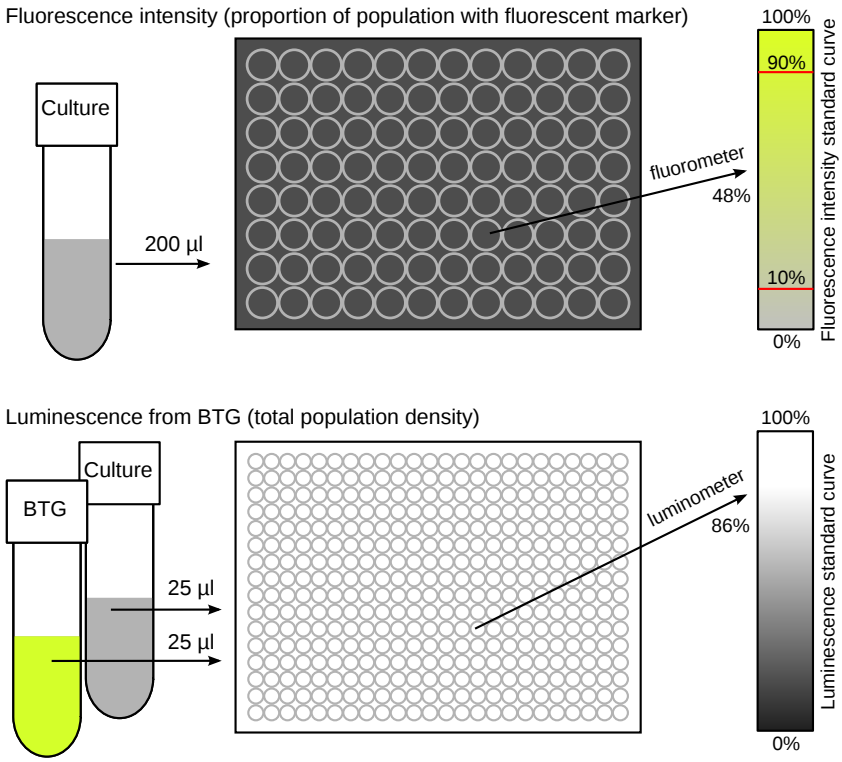

Figure 2.1: Illustration of microplate reader method for measuring competitive fitness, using a fluorescent genetic marker and BacTiter-Glo^{TM} (BTG) to measure labelled proportion and total population density.

of 200 µl of culture is added to the wells of a black opaque 96 well plate and placed in the microplate reader. The fluorescent proteins in the cells are excited by exposure to a given wavelength of light. The light emitted from the proteins is passed through a filter that permits only specified wavelengths of light, and the intensity of the light is then measured by the device. The excitation and emission wavelengths are dictated by the protein used (e.g. 510/390 nm for GFP and 540/500 nm for YFP). The FI produced by the competition cultures is then compared against a standard curve of known concentrations of the labelled strain, covering a range of 0% to 100%; however, the method is generally unreliable for proportions lower than 10% or greater than 90%.

Total population density is then measured either by OD (at 600 nm wave-

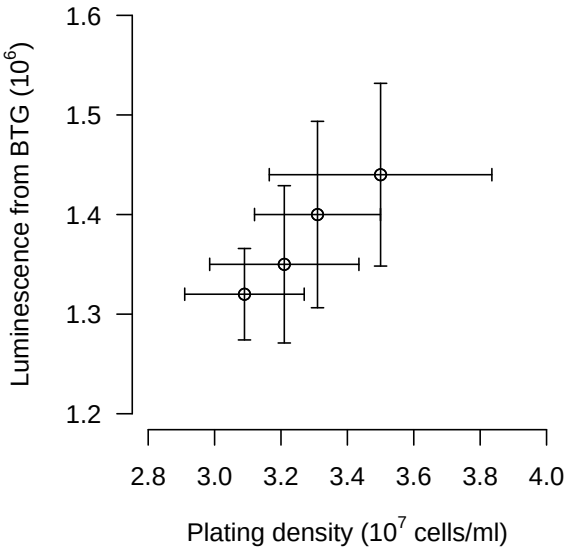

Figure 2.2: Luminescence from BacTitre-Glo™ (BTG), which provides a relative measure of total population ATP, accurately reflects population size measured by counting colonies plated on solid agar (circles are four different *P. aeruginosa* genotypes, means of $n = 5$ replicates ±SE on both axes).

length), or by the BacTiter-Glo™ (BTG) microbial cell viability assay kit (Promega, USA). For the BTG method, 25 µl of culture is treated with 25 µl of BTG in a white opaque 384 well plate, and incubated at 21–25 °C for five minutes. Treated cultures produce a luminescent signal proportional to the amount of ATP present, which is a proxy for population size. A standard curve of population density is also measured over the same 0% to 100% range as for FI. BTG is superior to OD, as it can measure low-density cultures (as low as 150 cells for *P. aeruginosa*, according to the BTG protocol), is specific to living cells and is unaffected by physiological changes in cell size.

BTG accurately reflects population size as measured by plating on solid agar (Figure 2.2, Spearman's rank correlation $\rho = 0.60$, $p < 0.01$). The fluorescence signal also accurately reflects either total population size in a pure culture

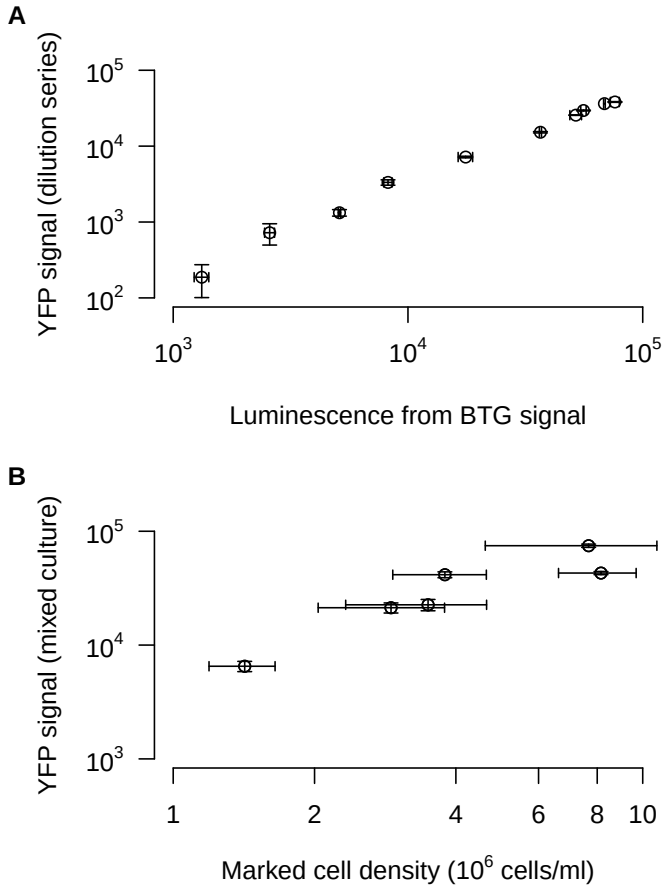

Figure 2.3: (A) In a dilution series, fluorescence signal (YFP) accurately reflects population size in a pure culture. (B) In mixed cultures, YFP accurately reflects the proportion of the population that is labelled with a marker, as measured by selective plating. All points means ± SE of $n = 3$ replicates.

(Figure 2.3A, Spearman's rank correlation $\rho = 0.99$, $p < 0.0001$) or proportion of marked cells as measured by plating on selective media (Figure 2.3B, Spearman's rank correlation $\rho = 0.77$, $p < 0.001$). However, the values and units of both FI and luminescence are arbitrary and depend on the microplate reader settings (e.g. read time per well, gain), so cultures must always be measured relative to a standard curve.

Having measured the proportion of cells with the marker, and the total

density, the proportion of the population initially made up by the unlabelled strain is then calculated as $\text{unlabelled}_i = \text{total}_i - \text{labelled}_i$, and similarly after the competition, $\text{unlabelled}_f = \text{total}_f - \text{labelled}_f$. Fitness of the unlabelled strain relative to the labelled strain is calculated as

$$w_{UL} = \frac{\log_2 \left[\text{unlabelled}_f / (D \times \text{unlabelled}_i)\right]}{\log_2 \left[\text{labelled}_f / (D \times \text{labelled}_i)\right]}, \tag{2.2}$$

where w is fitness, D is dilution factor, and i and f denote initial and final frequencies of each genotype.

As an entire 96-well plate can be read in approximately 3 minutes using the microplate reader, this method allows for assaying many strains or environments with high replication. However, the trade-off is lower precision of the measurements, particularly when there is a large difference in fitness between strains. This suggests that maximum growth rate is not an appropriate as a high-throughput fitness assay for this species.

Assaying by flow cytometry

A third method of measuring competitive fitness uses flow cytometry. The principal behind this assay is the same as for the microplate reader-based assay, also using a fluorescent marker to distinguish between cell phenotypes. However, rather than measure labelled proportion versus total population density, the unlabelled proportion is measured directly. To assay the population of cells, the flow cytometer passes a hydrodynamically-focused stream of liquid past a laser, illuminating cells as they pass the detector. The amount of light scattered in the forward (FSC) and side (SSC) direction is used to estimate the size of the cell. The FI emitted from the cell determines its phenotype, compared against pure culture populations containing only labelled or unlabelled cells.

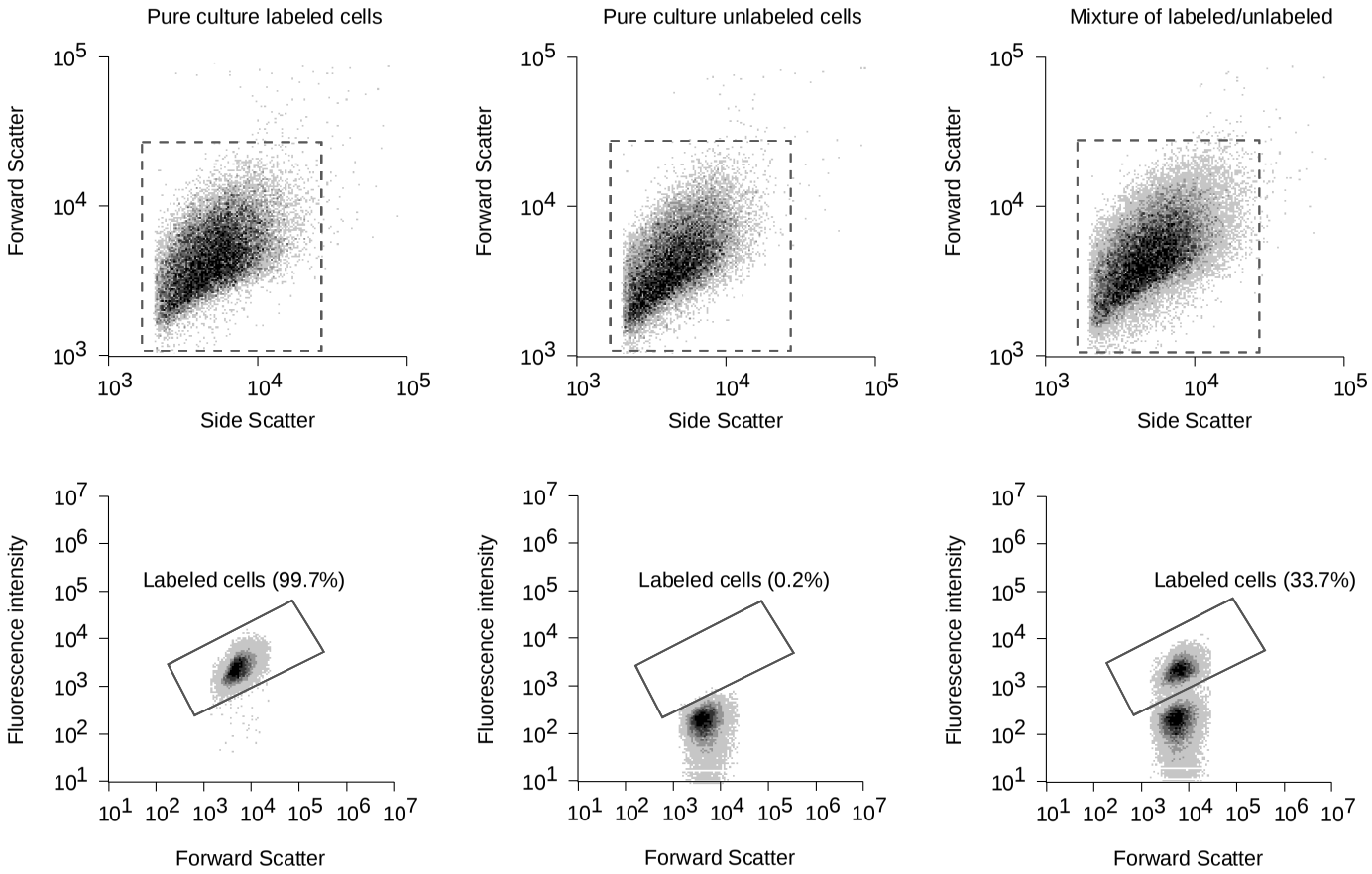

Figure 2.4: Example flow cytometry data used to measure competitive fitness. Each point represents a single event, triggered by the passage of a cell (or clump of cells) through the laser beam. A rectangular gate set in the Forward Scatter versus Side Scatter plots excludes cell clumps. Points in the Fluorescence intensity versus Forward Scatter plots include only single cells. Pure cultures of labelled and unlabelled cells were used to determine the percentage of fluorescently-labelled cells in the mixture.

Representative flow cytometry data are shown in Figure 2.4. For my assays, I used a 533±30 nm filter to distinguish either GFP or YFP from non-fluorescent cells (the emission spectra of GFP and YFP overlap too much to be discriminated from each other with this filter). A BD Accuri C6 flow cytometer (BD Biosciences, USA) was used, although comparable results should be attainable with other flow cytometry or fluorescence-activated cell sorting (FACS) equipment. To estimate the proportion of fluorescent cells in the sample, first a plot of FSC against SSC is made to eliminate occasions where multiple cells passed the detector simultaneously. A rectangular 'gate' is placed around the single-cell population. Second, a plot of FI against FSC is made for the single cells and a gate is drawn that includes only fluorescent cells in pure cultures (a paralellogram is used because size and FI are positively correlated). Finally, both gates are applied to populations with unknown mixtures of fluorescent and non-fluorescent cells. From these proportions, Equation 2.2 is used to calculate fitness.

Depending on cell density, 1–2 samples per minute can be processed, and from practice up to 400 samples may be measured in one session. Under good conditions (i.e. strong fluorescence production, cell density between 100 and 1500 cells/μl), little overlap is observed between the labelled and unlabelled populations. This results in greater accuracy over the microplate reader method, although fitness calculations remain inaccurate for small and large proportions. As an advantage, this method also obviates a standard curve, as individual cells, not populations, are assayed. Estimates of cell density by flow cytometry corresponds and BTG show good correspondence (Figure 2.5, Spearman's rank correlation $\rho = 0.78$, $p < 0.0001$).

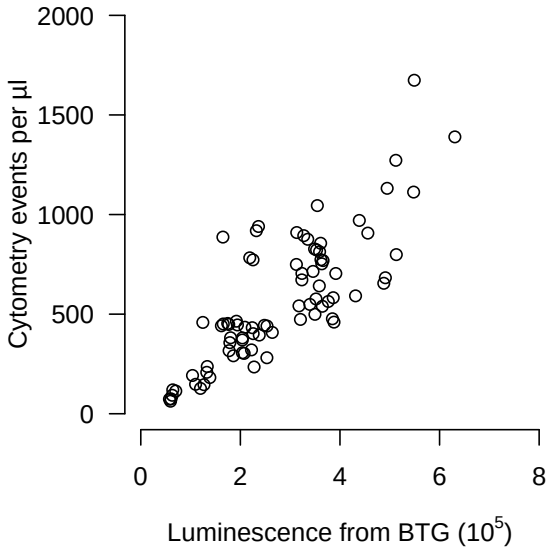

Figure 2.5: Population density measured by flow cytometry and BTG.

2.4 Cross-validation

Relative maximum growth rate versus competitive fitness

Relative maximum growth rate is often used as a proxy for fitness. Although, previous data has shown that relative $v_{\max}$ positiely correlates with competitive fitness (e.g. Lenski *et al.* 1998), Figure 2.6 shows this was not the case for a representative sample of 11 *P. aeruginosa* strains either grown in isolation or competed against a common ancestor in a complex medium [M9KB: 10.5 g/l M9 Broth (Biochemika), 10 g/l glycerol, 10g/l proteose peptone no. 3, 1 ml of 1M $MgSO_4$ after autoclaving]. The correlation coefficient between competitive fitness and relative $v_{\max}$ did not differ from zero (Spearman's rank correlation $\rho = -0.35$, $p = 0.29$).

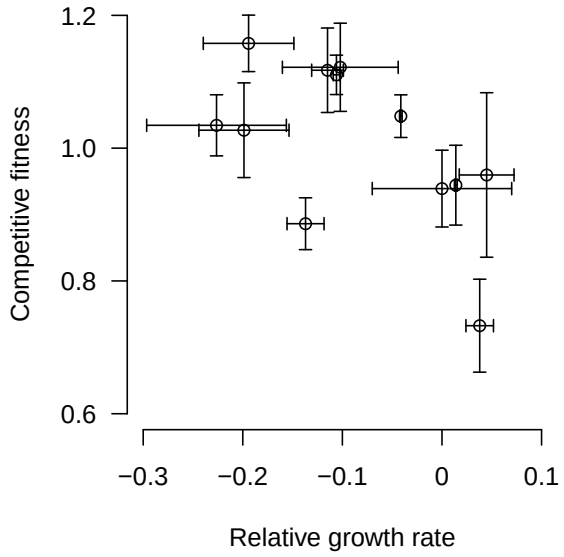

Figure 2.6: Relative growth rate versus competitive fitness for 11 *P. aeruginosa* strains grown in M9KB (means ± SE).

Competitive fitness methods

The three methods of measuring competitive fitness correspond well with one another. Figure 2.7 shows competitive fitness data for three *P. aeruginosa* strains (PA01, and rifampicin-resistant strains S536F and +P518, MacLean and Buckling 2009, see also chapter 3) against a reference strain, GFP-PA01 that produces green fluorescent protein and is gentamicin resistant (Choi and Schweizer 2006). Each of the fitness assays were conducted independently, showing the robustness of these methods.

2.5 Discussion

Here, I described and compared three methods for measuring the competitive fitness of bacterial genotypes. The two high throughput methods for measuring fitness demonstrated a marked increase in the number of genotypes that can be

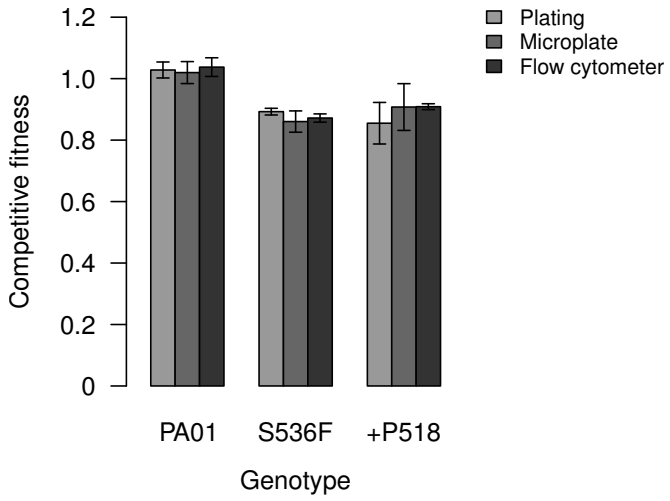

Figure 2.7: Comparison of competitive fitness measured by different methods. Competitions were against the GFP-PA01 reference strain on L-serine growth medium (see chapter 3, Table 3.2). Bars are means ± SE.

simultaneously assayed. The three methods provided comparable measures of the frequency of the marked genotype and total population size (Figures 2.2, 2.3 A and B, and 2.5). Further, fitness calculated from the components also showed good correspondence between the methods (Figure 2.7), suggesting each method can be used when appropriate. However, relative $v_{\max}$ did not predict competitive fitness for 11 *P. aeruginosa* strains in a complex medium (Figure 2.6). Relative $v_{\max}$ data should thus be corroborated with competitive fitness data.

There were advantages and disadvantages to each of the three methods of measuring competitive fitness. Selective plating can be used to measure the widest range of fitness values, as it is not constrained to the 10% detection limit of the two high throughput methods; if the percentage made up by one genotype goes below 10%, the number of cells plated on the selective media can be increased, which cannot be done for the two high throughput methods. However, this method constrains the size of experiments to the order of tens

of genotypes (assuming three replicates), which is not sufficient for large-scale experiments.

If the fitness assay can be designed to respect the 10% detection limit, either of the two high throughput methods can be used. The microplate reader method is appropriate for measuring the fitness of a fluorescently-labelled strain against an unlabelled strain. In principle, this method can be scaled to measure the fitness of hundreds of genotypes in one assay. However, as the proportion of unlabelled cells is calculated by subtracting the marked population density from total population density, this method becomes unreliable when the proportion of either strain is high as it can cause estimates of the unlabelled density to be outside the range of zero and one. As an improvement on the microplate reader method, the flow cytometry method eliminates the problem of estimating proportions outside the valid range by measuring the proportions of each strain directly. Due to the increased precision, the flow cytometry method supersedes the microplate reader method in the latter chapters of this thesis (the flow cytometry equipment was acquired after chapter 4 had been completed). However, time constrains make the reader method more practical for very large fitness assays, such as those involving more than about 400 measurements of fitness.

The high throughput fitness assays developed here were necessary for conducting competitive fitness assays on the scale required for my thesis, allowing the design of large experiments (and as an added benefit, also producing less waste than the selective plating method). Chapter 3 uses relative $v_{\max}$, selective plating and the microplate reader methods to measure the cost of rifampicin resistance in several different growth environments. Chapter 4 uses the microplate reader method to measure the fitness of several evolved rifampicin-resistant strains. Chapters 5 and 6 both use the flow cytometer method for measuring competitive fitness of evolved rifampicin-resistant and wild-type strains.

3 | Genotype-by-environment effects of rifampicin resistance

Abstract

Attempts to measure the costs of antibiotic resistance have usually only considered a single environment, but genotype-by-environment (G×E) interactions are known to affect fitness. In this chapter, I measured (1) the fitness cost of rifampicin resistance in several environments, and (2) the fitness benefit conferred by resistance at sub-inhibitory concentrations of rifampicin. I found that there are significant G×E interactions for fitness, with most *rpoB* mutants conferring an advantage in at least one environment, and no *rpoB* mutant being more or less costly than the others. Only a small concentration of rifampicin, roughly 1/64 the inhibitory concentration, was necessary to reduce wild-type growth rate by half, and 1/32 completely removed any cost associated with resistance. As bacteria migrate between environments, including those polluted with antibiotics, G×E effects on fitness may contribute to the maintenance of antibiotic resistance in reservoir populations of bacteria.

3.1 Introduction

Antibiotic resistance mutations generally impose a fitness cost in environments without antibiotics; for this reason, it was first thought that the cessation of

antibiotic use would eliminate resistance through natural selection (Levin *et al.* 1997; Lenski 1998). Resistance costs have usually been demonstrated *in vitro*, and occasionally *in vivo* (Andersson 2006). However, fitness costs have typically only been measured for one or two environments, which is problematic as genotype-by-environment interactions (G×E) are known to affect bacterial fitness (Remold and Lenski 2001). As resistant bacteria routinely experience different environments (Khachatourians 1998; Wright 2010), including those with residual concentrations of antibiotics that may select for resistance (Baquero *et al.* 2008; Kümmerer 2009a;b), a better understanding of G×E effects on fitness may help to explain the maintenance of antibiotic resistance.

In this chapter, I measured the fitness of several rifampicin-resistant strains of *Pseudomonas aeruginosa* on multiple carbon sources, and at sub-inhibitory rifampicin concentrations. Fitness of the different resistant mutants varied greatly by environment, with resistance even conferring a benefit in some environments. Further, I show that only residual amounts of rifampicin, several factors lower than MIC, are needed to eliminate the cost of rifampicin resistance for several resistance mutations. Together, these data suggest that environmental heterogeneity could, in theory, help to maintain antibiotic resistance in natural populations. Temporary passage through different environments, or an environment with a low antibiotic concentration, may help to maintain the prevalence of resistance mutations.

3.2 Genotype-by-environment costs of resistance

Rifampicin-resistant strain collection

The *P. aeruginosa* rifampicin-resistant strain collection used was developed as part of previous work (MacLean and Buckling 2009). Resistant strains were

isolated by plating spontaneous mutants from a large population of PA01 cells on solid M9KB agar plates [10.5 g/l M9 Broth (Biochemika), 10 g/l glycerol, 10 g/l proteose peptone no. 3, 1 ml of 1 M $MgSO_4$ after autoclaving] supplemented with 62.11 μg/ml rifampicin. The strains were then sequenced for mutations in *rpoB*. Table 3.1 lists the 12 strains used (PA01 wild-type and 11 rifampicin-resistant mutants, each bearing a single *rpoB* mutation). By convention, mutations are referred to by their resulting amino acid change in RpoB.

Table 3.1: Rifampicin-resistant mutants of *Pseudomonas aeruginosa* PA01 genetic background (all from MacLean and Buckling 2009).

Strain name	*rpoB* mutation	Residue change
PA01	none	none
Q152L	A455T	Glu→Leu
Q152R	A455G	Glu→Arg
S517L	C1550T	Ser→Leu
Q518L	A1553T	Glu→Leu
Q518R	A1553G	Glu→Arg
+P518	1552CGC1553[†]	Insertion of Pro
D521E	C1563G	Asp→Glu
D521V	A1562T	Asp→Val
H531R	A1592G	His→Arg
H531Y	C1591T	His→Tyr
S536F	C1607T	Ser→Phe

[†]Insertion at position 2 of codon 518; results in a CCG (Pro) at 518 and CAG (Glu) at 519.

Growth rates on different carbon sources[1]

Previous data (kindly provided by A. R. Hall) provided an estimate of relative fitness of the *rpoB* mutants, relative to their rifampicin-sensitive PA01 ancestor, on the substrates of a Biolog GN2 MicroPlate (Biolog, Inc., Hayward, CA, USA). Strains were revived from freezer stocks by growing for 24 h in M9KB at 37°C.

[1]As discussed in subsection 2.4, growth rate is not the best measure of fitness. However, I did not yet have access to fluorometry or flow cytometry equipment.

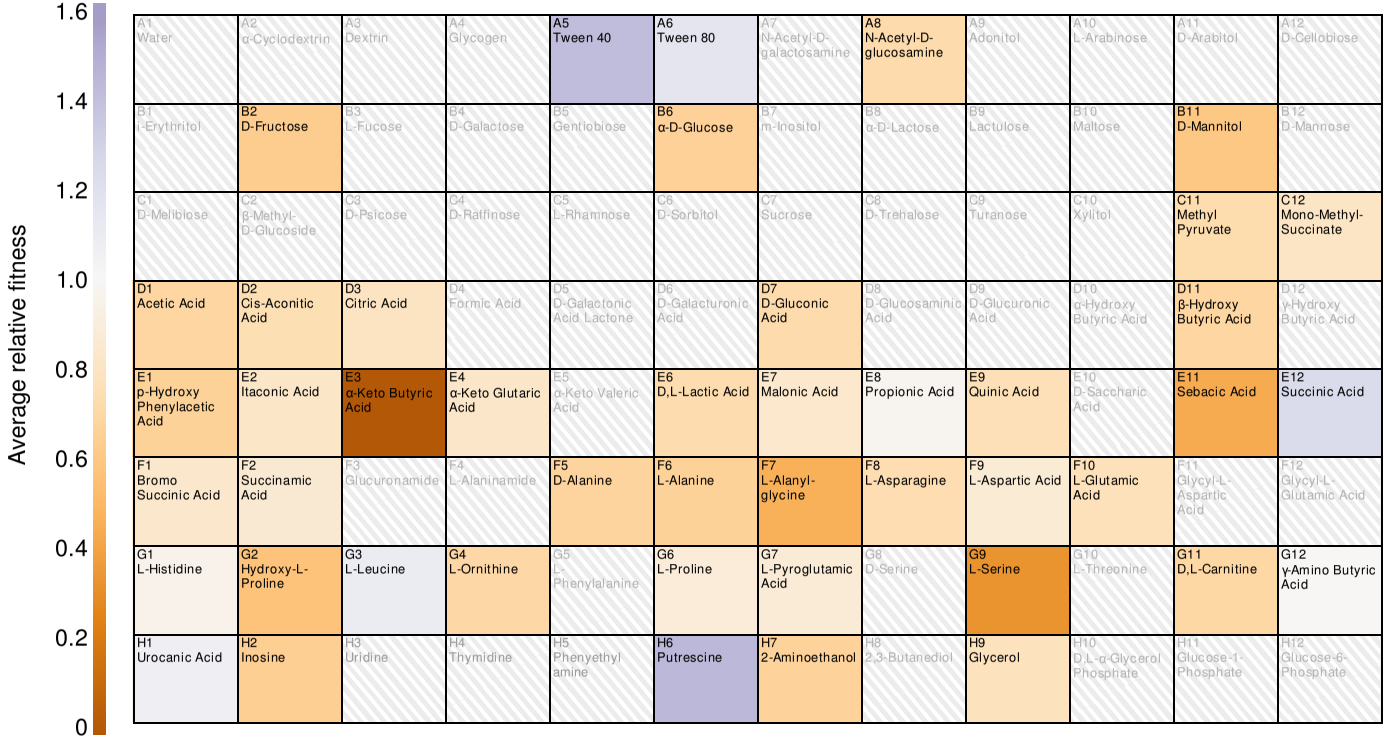

Figure 3.1: Average relative fitness, measured by optical density (OD), of rifampicin-resistant mutants on Biolog substrates ($n = 10$, mutant +P518 was not included). Environments where resistance was on average costly are shown in shades of orange, whereas environments where resistance was on average beneficial are shown in shades of purple. A grey hatch indicates fitness was not measured on that substrate due to absence of detectable growth of the PA01 wild-type (OD < 0.15). Raw data kindly provided by A.R. Hall.

The overnight cultures were resuspended in M9 broth (10.5 g/l M9 salts) and starved for a minimum of 2 hours. 150 µl of M9 broth was added to each well of the Biolog GN2 microplate to reconstitute the carbon sources. The strains were diluted 1/100 into the microplate from the resuspended culture. Fitness was estimated as optical density (OD), relative to the PA01 ancestor, after 24 h at 37 °C. The mutants were able to grow on 44 out of 95 substrates, with variable average fitness (Figure 3.1).

From the set of 44 viable substrates, I chose five substrates (α-D-glucose, glycerol, sebacic acid, L-leucine and L-serine) for further study. These substrates were chosen because the fitness costs of rifampicin resistance were variable when measured on the GN2 Biolog plate. A complex medium (M9KB) was also chosen for comparison. Concentrations of the carbon substrates were based on the GN2 MicroPlate (Table 3.2). The growth media consisted of these carbon substrates in M9 minimal medium (10.5 g/l M9 Broth and 1 ml of 1M $MgSO_4$ after autoclaving). Each strain was inoculated from freezer stocks into 1 ml of M9KB and allowed to grow for 24 hours at 37°C. Overnight cultures were then centrifuged at 13,000 rpm for 2 minutes, and the supernatant was removed. Pelleted cells were resuspended in M9 minimal medium, and then diluted 1/100 into the culture medium containing the carbon substrates. Cultures were grown in 96 well plates at 37°C, with the position of the *rpoB* genotype and carbon source randomized to minimize edge effects. Three replicates of each genotype by environment pair were measured over three blocks. To measure growth rate, OD was read once per hour for 24 hours. Maximum growth rate was measured as the maximum slope resulting from a log-linear regression over a window of five data points. Relative maximum growth rate, v_{rel}, was calculated from the maximum growth rate of the mutant and the PA01 wild-type, $v_{\text{rel}} = v_{\text{mutant}}/v_{\text{PA01}} - 1$.

Individual *rpoB* mutations caused different effects in different environments

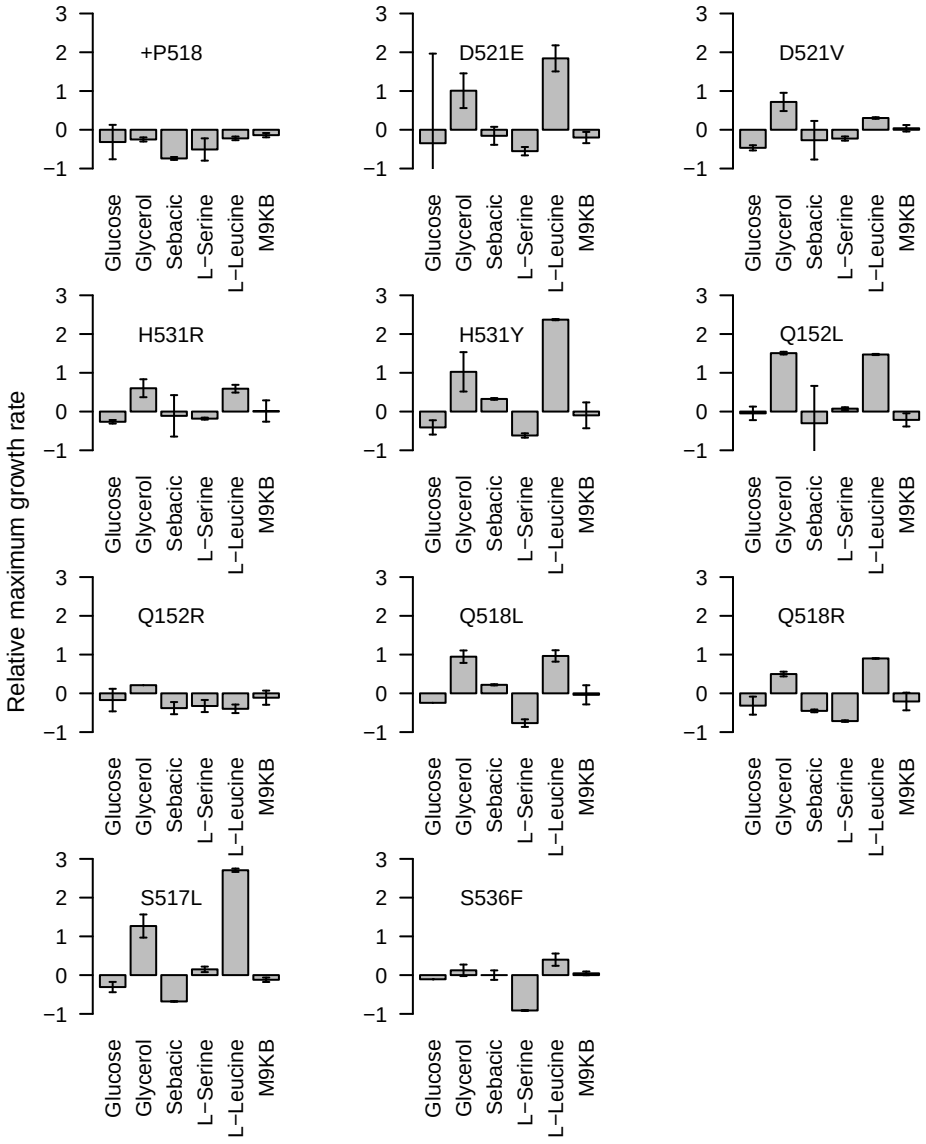

Figure 3.2: Maximum growth rate of *rpoB* mutants relative to PA01, measured on several carbon substrates (means ± SE of $n = 3$ replicates).

Table 3.2: Concentrations of carbon sources in which the effects of *rpoB* mutations on growth rates were measured. Data are the proprietary property of Biolog, Inc. (kindly provided by B.R. Bochner, personal communication).

Carbon source	Concentration (g/l)
Glycerol	2.50
L-Leucine	3.75
α-D-Glucose	12.50
Sebacic acid	2.50
L-Serine	3.75

(Figure 3.2) and no mutation had significantly higher fitness rank across all environments than any other (Friedman rank-sum test: $\chi^2_{11} = 15.6$, $p = 0.16$). There was a significant genotype-by-environment interaction (Type III ANOVA: genotype $F_{11,55} = 1.87$, $p = 0.06$, environment $F_{5,55} = 15.2$, $p < 0.0001$, G×E interaction $F_{55,144} = 1.87$, $p = 0.002$). +P518 was the only mutation that was unconditionally deleterious across all environments. Relative maximum growth rate was only weakly correlated between most environments (excepting L-leucine and glycerol). Of the five substrates tested, the *rpoB* mutants tended to have higher growth rates than PA01 in L-leucine and glycerol, and lower growth rates in the others (Figure 3.3). The lowest average growth rate was observed in L-serine.

Effect of resistance on competitive fitness in L-serine

As the resistant mutants had the lowest average growth rate in L-serine, I elected to measure competitive fitness in this environment. Five rifampicin-resistant strains (+P518, Q518L, D521E, H531Y and S536F) were competed against a common competitor, the PA01 ancestral strain bearing two genetic markers: gentamicin resistance and the production of yellow fluorescent protein (YFP-PA01, Choi and Schweizer 2006). Strains were first inoculated from freezer

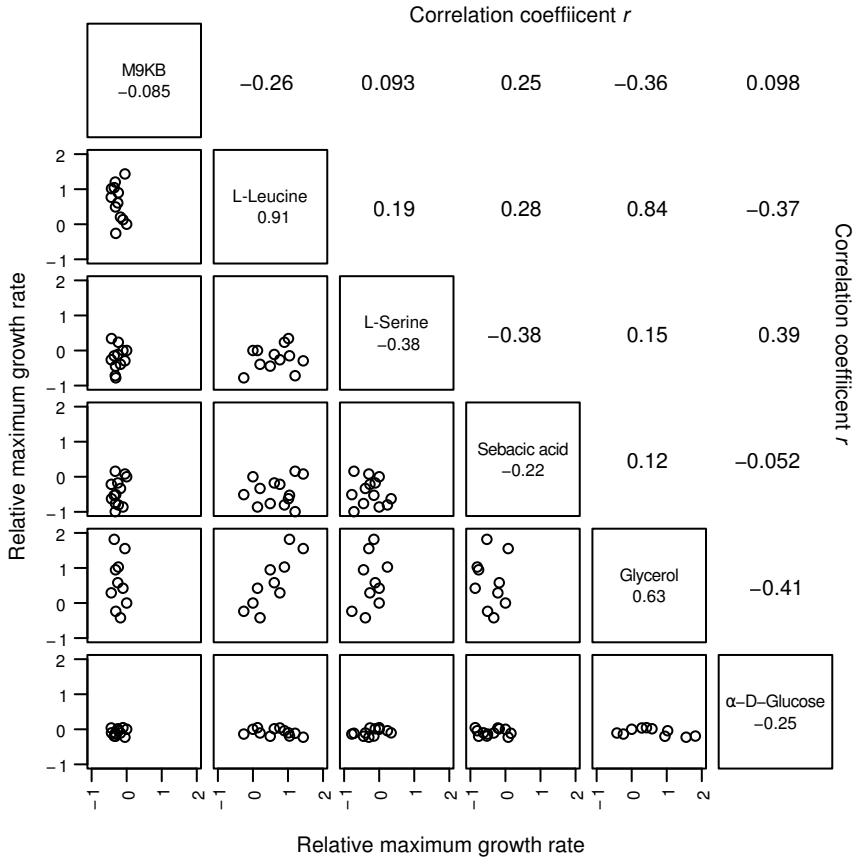

Figure 3.3: Correlation of relative maximum growth rates across environments (data from Figure 3.2). Mean relative maximum growth rate for each environment across all genotypes is given on the diagonal ($n = 3$ replicates).

stocks into 1 ml of M9KB and allowed to grow for 24 hours at 37°C. Overnight cultures were then centrifuged at 13,000 rpm for 2 minutes, and the supernatant was removed. Pelleted cells were resuspended in M9 serine, and then mixed in a 1:1 volumetric ratio with the YFP-PA01 competitor strain. The mixtures were then diluted 1/100 into 200 µl M9 serine (M9 salts with 3.75 g/l L-serine) in a 96 well plate. Five replicate competitions were measured per *rpoB* mutant. Competitive fitness was measured by the selective plating method, using antibiotic resistance markers to distinguish the genotypes (see chapter 2). The initial mixtures were then plated on selective M9KB agar, containing either 64 mg/l of rifampicin or

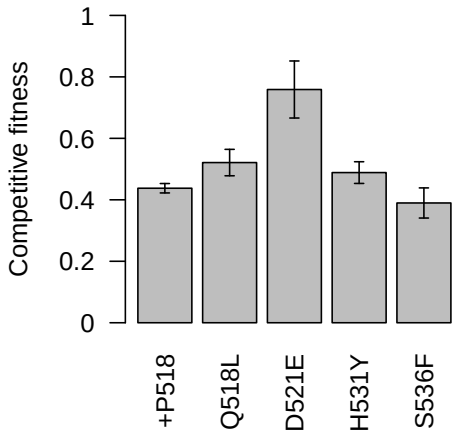

Figure 3.4: Competitive fitness of *rpoB* mutants in L-serine medium (means of $n = 5$ replicates ± SE).

40 mg/l gentamicin. The mixed cultures were grown for 24 h, and the final cultures were again plated on the same selective media.

Rifampicin resistance was associated with a large reduction in competitive fitness when grown on L-serine (Figure 3.4). Of the five mutants tested, +P518 and S536F had the largest fitness cost when grown in L-serine. This pair of mutants had no significant difference in fitness cost ($w_{\mathrm{S536F}} = 0.401 \pm 0.0269$ SE, $w_{\mathrm{+P518}} = 0.387 \pm 0.0640$ SE, Welch's t-test: $t_{6.87} = 0.20$, $p = 0.42$).

3.3 Cost of resistance eliminated by sub-MIC rifampicin

Although resistance mutations often impose a fitness cost in antibiotic-free environments, the presence of a low concentration of the antibiotic, even several times lower than the MIC, can make resistance beneficial. To determine the effect of sub-MIC rifampicin concentrations, I first measured the growth rate of the PA01 ancestral strain at several sub-MIC concentrations in isolation. PA01 was cultured overnight from freezer stocks in 1 ml of M9KB, and then diluted into M9KB with varying rifampicin concentrations (0, 0.5, 1, 2, 4, 8, 16, and 32

mg/l). PA01 was grown at these rifampicin concentrations overnight at 37 °C in a 96 well plate, and OD was measured once per hour. From a sliding window of five points, maximum growth rate was calculated (Figure 3.5A). Although the MIC of rifampicin for this strain is 64 mg/l (MacLean and Buckling 2009), a rifampicin concentration of 2 mg/l produced a statistically significant reduction in growth rate. To determine the benefit of resistance at these sub-MIC levels, I competed the YFP-PA01 strain against the 11 *rpoB* mutants at four concentrations of rifampicin (0, 2, 4, and 8 mg/l, Figure 3.5B). Strains were revived from freezer stocks as above, and mixed at a 1:1 ratio before inoculation into the rifampicin-containing media. The competitions were grown for 24 h at 37 °C. Competitive fitness was measured by growing the sensitive and resistant strains together in the presence of rifampicin. Initial and final proportions were measured using the microplate reader method described in chapter 2, using fluorescence intensity to measure the proportion of the population made up by the YFP-PA01 strain.

3.4 Discussion

Antibiotic resistance has long been thought to impose a fitness cost on the lineages that bear them. Here, I have shown that the cost of resistance for *rpoB* mutants is a product of the interaction between genotype and growth environment. Further, only small concentrations of an antibiotic are needed to select for resistance in the presence of the sensitive genotype. Together, this suggests that temporary passage through different environments, or an environment with a low antibiotic concentration, may help to maintain the prevalence of resistance mutations. Theoretical work has suggested that genotype-by-environment interactions can maintain genetic variation in natural populations (Gillespie and Turelli 1989). Although natural environments are thought to play a role in maintaining antibi-

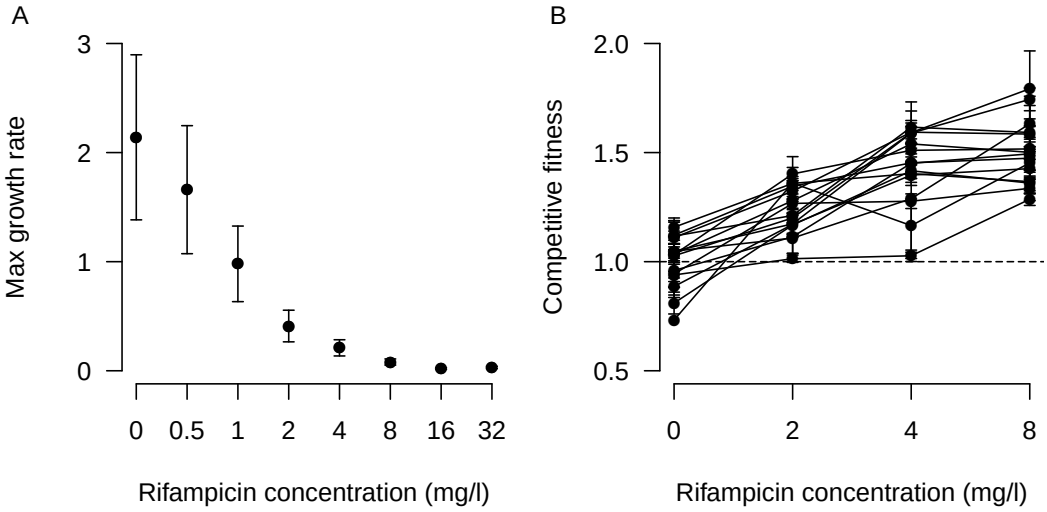

Figure 3.5: (A) Growth rate of PA01 at increasing rifampicin concentrations (means of $n = 6$ replicates $\pm$ SE). (B) Competitive fitness of the 11 *rpoB* mutants at increasing rifampicin concentrations (means of $n = 10$ replicates, $\pm$ SE).

otic resistance (Martinez 2009), to date, little is known about how resistance costs vary by environment (but see Remold and Lenski 2001; Hall 2013).

The results demonstrated that resistance need not always impose a fitness cost: all mutants had a growth rate greater than or equal to than the PA01 strain in at least one environment, except +P518, which was unconditionally detrimental. Aside from +P518, no one *rpoB* mutant tended to be better or worse than the others across environments. This suggests that the multiple environments that strains commonly experience may influence the demographics of resistance mutations. Consistent with this finding, previous studies that found rifampicin resistance emerging as a side-effect of another selective pressure. Rodríguez-Verdugo *et al.* (2013) showed that certain rifampicin resistance mutations confer a benefit at under high temperature and low glucose conditions. Other mutations may also confer an advantage in nutrient-poor conditions, particularly in the

absence of amino acids. Studies in *E. coli* have demonstrated that some *rpoB* mutations cause bacteria to mimic the stringent response, which in nutrient-poor medium down-regulates genes involved in rapid growth and an up-regulates biosynthesis genes (such as for amino acids, Zhou and Jin 1998). Instead, these few *rpoB* mutants behave stringently even in nutrient-rich conditions. Of the 11 mutants assayed in this chapter, S536F (S531F in *E. coli*) behaves stringently, which may explain why its growth rate was not significantly reduced in most environments tested (Figure 3.2).

Are extensive genotype-by-environment interactions expected for resistance to antibiotics other than rifampicin? The extent of the G×E interaction observed for *rpoB* mutants could be due to the mechanism of rifampicin resistance. As mentioned in the Introduction, RpoB is a component of RNA polymerase, which transcribes DNA into mRNA. Mutations in *rpoB* are known to reduce the rate of mRNA transcription, but also to alter the global profile of gene transcription (Qi Qin, unpublished data). However, a significant genotype-by-environment interaction was observed for several chloramphenicol, kanamycin and tetracycline resistant *E. coli* mutants across four environments (Remold and Lenski 2001), suggesting that G×E interactions on resistance costs could be general.

Recognizing the unique opportunity presented by having two resistant mutants with equally large costs but different bases of resistance (S536F and +P518), I chose to focus on these mutants in L-serine for the remainder of my thesis. The large fitness cost imposed by these mutations when grown on L-serine presents a convenient model system for testing predictions regarding the maintenance of antibiotic resistance. Given the large fitness cost of being resistant, a 'revertant' strain restoring the wild-type sensitive phenotype should rapidly fix if it should appear in populations of either of these mutants. As both of these mutants

have approximately the same cost of resistance, revertants should fix with equal probability in both backgrounds. This prediction is tested in chapter 4. S536F and +P518 are also used to investigate the genetic basis of evolvability on the novel L-serine environment in chapter 5. Finally, the effect that these strains have on the adaptation of an invading antibiotic-sensitive strain is investigated in chapter 6.

4 | Evolutionary reversals of rifampicin resistance†

† Published as: Gifford, D. R. and MacLean, R. C. (2013), Evolutionary reversals of antibiotic resistance in experimental populations of *Pseudomonas aeruginosa*. Evolution, 67: 2973–2981. doi: 10.1111/evo.12158

Abstract

Antibiotic resistance mutations are accompanied by a fitness cost, and two mechanisms allow bacteria to adapt to this cost once antibiotic use is halted. First, it is possible for resistance to revert; second, it is possible for bacteria to adapt to the cost of resistance by compensatory mutations. Unfortunately, reversion to antibiotic sensitivity is rare, but the underlying factors that prevent reversion remain obscure. Here, we directly study the evolutionary dynamics of reversion by experimentally mimicking reversion mutations–sensitives–in populations of rifampicin-resistant *P. aeruginosa*. We show that, in our populations, most sensitives are lost due to genetic drift when they are rare. However, clonal interference from lineages carrying compensatory mutations causes a dramatic increase in the time to fixation of sensitives that escape genetic drift, and mutations surpassing the sensitives' fitness are capable of driving transiently common sensitive lineages to extinction. Crucially, we show that the constraints on reversion arising from clonal interference are determined by the potential for compensatory adaptation of the resistant population. Although the cost

of resistance provides the incentive for reversion, our study demonstrates that both the cost of resistance and the intrinsic evolvability of resistant populations interact to determine the rate and likelihood of reversion.

4.1 Introduction

Antibiotic resistance mutations are usually accompanied by a fitness cost, expressed in terms of reduced competitive ability, transmission rate and virulence relative to sensitive strains (Andersson and Levin 1999; Andersson 2006). There are two possible evolutionary responses to this cost of resistance when populations of resistant bacteria encounter an antibiotic-free environment, as occurs during transmission between hosts or when antibiotic use is halted. First, bacterial populations can eliminate the cost of resistance by reverting to antibiotic sensitivity. Second, bacterial populations can adapt to the cost of resistance by fixing second site compensatory mutations that fully or partially recover the cost of resistance without compromising antibiotic tolerance (Levin *et al.* 2000; Reynolds 2000; Nagaev *et al.* 2001; Arason *et al.* 2002; Maisnier-Patin *et al.* 2002; Comas *et al.* 2012). Unfortunately, compensation occurs much more commonly than reversion, both *in vivo* and *in vitro* (Andersson 2003; Andersson and Hughes 2010; 2011), but the underlying population genetic reasons for this are not completely understood.

Three key factors likely constrain reversion. First, reversion mutations are rare relative to compensatory mutations; it has been estimated that there are an average of 11-12 compensatory mutations per resistance mutation (Poon and Chao 2005) although some may have as many as 35 or more (Björkman *et al.* 2000). The abundance of compensatory mutations is likely to impose an especially important constraint on reversion because compensatory mutations show epistatic interactions with resistance mutations, such that reversion muta-

tions incur a fitness cost once bacterial populations have acquired compensatory mutations. Second, genetic drift will lead to the stochastic extinction of reversion mutations when they are at a low frequency. Finally, competition between reversion mutations and compensatory mutations (i.e. clonal interference) may constrain reversion via two mechanisms. Compensatory mutations that do not fully recover the cost of resistance are expected to delay, but not prevent, the fixation of reversion mutations (De Visser and Rozen 2006). Mutations that recover fitness to higher than wild-type levels, on the other hand, are expected to prevent the fixation of reversion mutations (Gerrish and Lenski 1998). Here, we define mutations in the resistant background that surpass the current fitness of the sensitive type as 'super-compensatory' mutations[1] (note that this does not preclude adaptation in the sensitive type, only that the rate of such adaptation is lower than the rate of compensatory adaptation). Although it is clear that the relative rarity of reversion mutations helps contribute to the maintenance of antibiotic resistance, the impact of drift and clonal interference on reversion has not been investigated.

The goal of this manuscript is to investigate how drift and clonal interference affect the evolutionary dynamics of reversion using an experimental model system based on the evolution of rifampicin resistance in *Pseudomonas aeruginosa*. Rifampicin is a potent broad-spectrum antibiotic that binds to a highly conserved domain on the β-subunit of RNA polymerase (*rpoB*), preventing the elongation of RNA transcripts. Resistance to rifampicin in many bacteria, including *P. aeruginosa*, results from mutations in the rifampicin binding domain that compromise the ability of rifampicin to bind to its target (Wehrli 1983). Rifampicin resistance mutations in *P. aeruginosa* tend to impose a fitness cost

[1]Later, I learned that these are adaptive mutations, but at the time this manuscript was published, the sequencing data was unavailable.

in the absence of antibiotics, most likely resulting from compromised RNA polymerase function (Hall *et al.* 2011). Previous work has shown that adaptation to the cost of rifampicin resistance in *P. aeruginosa* usually occurs by compensatory mutations, often at second sites in *rpoB*, but complete reversion to antibiotic sensitivity is also possible (Hall *et al.* 2010).

The conventional method that has been used to study reversion and compensation is to challenge populations of antibiotic resistant bacteria with adapting to the cost of resistance *in vitro* and *in vivo* (reviewed in Andersson and Hughes 2010). Following some period of selection, the fitness and resistance of evolved populations are then determined relative to suitable control populations to test for both compensatory adaptation and reversion. Although this approach has consistently shown that compensation tends to occur more commonly than reversion, it does not provide any insight into whether reversion is rare simply because of a paucity of reversion mutations or because of a low probability of fixation of reversion mutations due to drift and clonal interference.

To study reversion, we used an experimental model system in which populations of two rifampicin-resistant strains of *P. aeruginosa* were seeded with individual cells of a genetically-marked rifampicin-sensitive strain, effectively mimicking a spontaneous reversion mutation. We then followed the evolutionary dynamics of sensitive and resistant lineages over approximately 300 generations in a highly replicated selection experiment. This experimental approach has two important benefits. First, because we seeded populations with a sensitive strain, we could be confident that any lack of reversion was not simply attributable to a low rate of supply of reversion mutations. Second, because the sensitive that we introduced was genetically marked, we were able to accurately follow the dynamics of sensitive lineages, even when sensitives were extremely rare (frequency of less than 1 in 2×10^5). Using this approach, we showed that

despite the large cost of resistance, most sensitives were lost due to genetic drift when they were initially rare. The main effect of clonal interference was to dramatically increase the time to fixation of reversion mutations, although in some cases reversion mutations that escaped genetic drift were ultimately eliminated by competing super-compensatory mutations. Crucially, we found that the ultimate probability of fixation and time to fixation, following drift and clonal interference, were determined by the intrinsic evolvability of the resistant population. As the resistant mutants we studied had virtually indistinguishable fitness costs with respect to their common sensitive ancestor, this suggests that the ability to compensate for resistance, and not the cost associated with resistance, may be more important in determining the probability of reversion. Intriguingly, we found evidence that the nature of evolvability differed in these two strains: compensatory mutations were seemingly more common in one resistant lineage while super-compensatory mutations were more common in the other.

4.2 Materials and Methods

Strains

We used two rifampicin-resistant strains of *P. aeruginosa*, which are both mutants from the same PA01 clone. Resistant mutants were isolated on solid M9KB agar plates [10.5 g/l M9 Broth (Biochemika), 10 g/l glycerol, 10 g/l proteose peptone no. 3, 1 ml of 1 M MgSO4 after autoclaving] supplemented with 62.11 μg/ml rifampicin (previously described in MacLean and Buckling 2009). Sequencing determined that both strains had mutations in the *rpoB* gene, one a single base pair substitution (C1607T, corresponding to amino acid change S536F) and the other a three base pair insertion (1552CGC1553, a proline insertion between 517 and 518, +P518). For convenience, we will refer to these mutants by

their amino acid changes. We used rifampicin-sensitive PA01 strains marked with either green fluorescent protein and gentamicin resistance (GFP-PA01), or yellow fluorescent protein and streptomycin resistance (YFP-PA01) as the sensitive competitor strains (Choi and Schweizer 2006). GFP-PA01 and the two rifampicin-resistant strains were used to determine the cost of resistance and in the selection experiments, while YFP-PA01 was used as a reference strain to measure competitive fitness (see below).

Plate reader-based competitive fitness assays

To enable us to conduct competitive fitness assays with many strains, we developed a plate reader-based fitness assay using a fluorescent marker to quantify the proportion of the population made up by a marked competitor, versus total population density. Competitive fitness of a non-fluorescently-labelled "strain of interest" was measured against YFP-PA01, standardized by the competitive fitness of wild-type PA01. Fluorescence intensity (FI) was measured using a combined fluorometer/luminometer (FLUOstar OPTIMA, BMG Labtech, Offenburg, Germany), with filters optimised for detecting the peak of the YFP emission spectrum (excitation $= 540 \pm 5$ nm, emission $= 500 \pm 5$ nm). YFP was chosen over GFP as the fluorescent marker for these competition assays to reduce detection of endogenous fluorescent proteins produced by *P. aeruginosa* (such as pyoverdin), which have emission spectra that overlap with GFP (Cox and Adams 1985; Tsien 1998). Total population density of each culture was measured using BacTiter-Glo™ (Promega, USA), which measures total population ATP using a bio-luminescent enzyme. We mixed 25 µl of culture with 25 µl of prepared assay reagent and measured total luminescence intensity (LI) using the same fluorometer/luminometer. FI and LI were measured relative to a standard curve (i.e. a dilution series of a pure culture of the YFP-PA01 strain), to give

standardized total population density (T) and standardized labelled density (L). Proportion of the population made up by unlabelled strains U was then calculated as $U = T - L$. Fitness of the unlabelled strain relative to the labelled strain was then calculated by comparing the initial (i) and final (f) proportions of the unlabelled and labelled strains, or $w_{UL} = \log_2(U_f/U_i)/\log_2(L_f/L_i)$. Preliminary work showed that competitive fitness measured using this method correlates strongly (with slope indistinguishable from one) with fitness measured from colony counts on selective media (slope $= 1.01 \pm 0.04$, $F_{1,8} = 619$, $R^2 = 0.98$, $p < 0.0001$).

Cost of resistance

To measure the cost of resistance, we directly competed our resistant mutants against rifampicin-sensitive GFP-PA01 in minimal culture medium supplemented with serine as the carbon and nitrogen source [M9-serine: 10.5 g/l M9 Broth 1 ml of 1 M $MgSO_4$, and 3.75 g/l (35.7 mM) L-serine]. Pre-competition cultures were prepared by inoculating single colonies of the resistant and sensitive strains into M9KB broth and incubating these cultures overnight at 37 °C with constant shaking at 200 rpm (overnight density 109 cells/ml). Pre-competition cultures of the resistant strains and the sensitive strain were then mixed at a 10 to 1 volume to volume ratio and diluted by 10^{-4} into M9-serine to set up five replicate competition cultures that were incubated overnight at 37 °C. We determined the initial and final titre of strains by plating samples of cultures before and after competition on M9KB agar supplemented with 64 mg/L rifampicin (to select for the resistant strain) and 40 mg/L gentamicin (to select for the resistant strain). We computed the relative fitness of the resistant strain ($w_{\text{resistant}}$), relative to the sensitive strain, as the log-ratio of the number of doublings. The cost

of resistance is defined as the difference between the sensitive strain's fitness (standardized at 1) and the resistant strain's fitness.

Introduction of sensitive strain

To study the dynamics of reversion, we inoculated a genetically-marked sensitive strain (GFP-PA01) into populations of S536F and +P518. To initiate the experiment, we inoculated 384 populations of each genotype (in eight 96-well microtitre plates) of 5.25×10^4 cells of S536F and +P518 into 200 µl M9-serine. Each rifampicin-resistant population received an average of 0.37 cells of the rifampicin-sensitive GFP-PA01, which was achieved by serial diluting an overnight culture of GFP-PA01 grown in liquid M9KB. Preliminary experiments demonstrated that the number of sensitive cells per population followed a Poisson distribution, therefore 118.75±9.05 s.d. populations were expected to receive at least one sensitive cell. [The number of populations follows a binomial distribution $B \sim (p = 1 - e^{-0.37}, n = 384)$, see section 4.6 for full calculations]. A statistical determination of the number of populations 'seeded' with the sensitive strain was necessary as there was no method at our disposal sensitive enough to detect the small number of sensitive cells in the presence of the rifampicin-resistant cells. By the Poisson distribution, the number of sensitive cells seeded into each population was expected to be either one ($n = 98$), two ($n = 18$), or three ($n = 3$)–no populations were expected to receive four or more sensitive cells. This represents an average initial sensitive frequency of 2.2×10^5, if weighted by the distribution of sensitives. We estimated that, by chance alone, there could be a difference of approximately 10.22 successful sensitive inoculations between the two rifampicin-resistant mutants (see section 4.6), meaning that any difference in outcomes detected between S536F and +P518 must differ significantly from this amount. The 265 populations that did not

receive a sensitive were selected alongside those that did to determine the rate and extent of compensation possible in these resistant genetic backgrounds.

Selection experiment

All 384 populations were then transferred over 300 generations in a 24-hour serial transfer protocol. From each population, 1 µl was transferred by pin replicator to 200 µl of fresh medium (dilution factor $D = 0.005$). Throughout the selection experiment, final population size was on the order of 10^7 after 24 hours. A subset of populations (96 for each resistant genotype) was stored at −80 °C.

Approximately every six to seven generations, we used simple phenotypic assays to test for the fixation of the sensitive (minimum detectable frequency of 1 in 2×10^5 or 0.0005%) or the resistant strain (minimum detectable frequency of 1 in 4×10^3 or 0.025%), assuming that the loss of one strain was equivalent to the fixation of the other. Although the frequencies at which we define fixation are arbitrary, they were chosen to be stringent so that our assays represent true fixations rather than simply failing to detect either strain present in the population. Presence of the GFP-PA01 rifampicin-sensitive strain was assayed by pin replicating 1 µl of the undiluted overnight culture on a M9KB agar plate with 40 mg/l gentamicin. Presence of the rifampicin-resistant strains was assayed by pin replicating 1 µl of the 1/200 dilution of the overnight culture on a M9KB agar plate with 64 mg/l rifampicin (Figure 4.1). Differences in the proportion of sensitives present in both rifampicin-resistant genetic backgrounds were assessed using a two-sample test for equality of proportions (without Yates' correction), function `prop.test()` in R version 2.14.1 (R Development Core Team 2011).

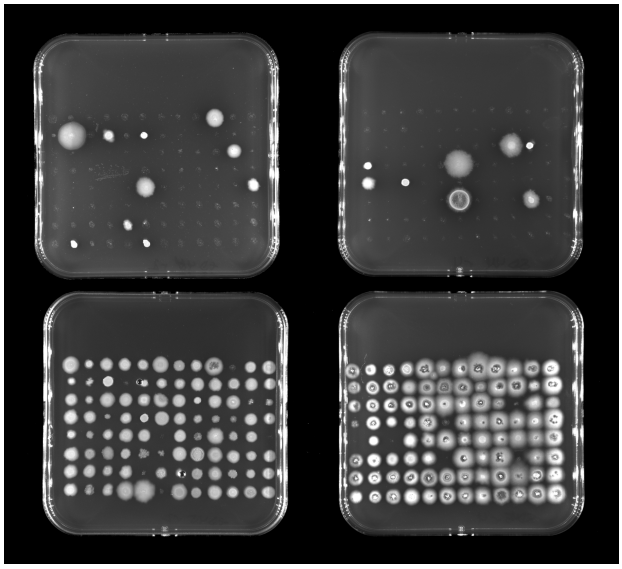

Figure 4.1: Phenotypic assay of fixation. Upper plates contain 40 mg/L gentamicin and are selective for GFP-PA01. Lower plates contain 64 mg/L rifampicin and are selective for both rifampicin-resistant mutants (left: +P518, right: S536F). Fixation of the sensitive invader is scored if growth occurs on the upper plate, but not on the corresponding lower plate.

Frequency of sensitive strain

In a subset of five populations of each resistant genotype, we tracked the frequency of the sensitive while populations were polymorphic for resistance (generations 0–150). We measured the frequency of the sensitive strain and the total population size using the same method as for the plate reader-based competition assays, except we measured FI in two replicates and LI in four replicates.

Compensatory ability

To measure the ability of S536F and +P518 to acquire compensatory mutations, we measured the competitive fitness (against YFP-PA01) of a subset of populations that, by chance, either received no sensitive cells or where the sensitive never

reached detectable frequency. Although some of this subset may have been seeded with the sensitive strain, the fact that it never reaches detectable frequency makes it unlikely to have had any effect on the adaptation of the resistant strain, and hence we treat these populations as though they were pure cultures of the resistant strain.

4.3 Results

Fitness of fluorescently-labelled PA01 reference strains

Using a selective plating competition assay, the two fluorescently-labelled strains did not differ significantly in fitness from wild-type PA01 (PA01 fitness versus GFP: $w_{\text{PA01}} = 0.975 \pm 0.036$ SE, $t_4 = -0.7$, $p > 0.2$; PA01 vs YFP: $w_{\text{PA01}} = 0.987 \pm 0.032$ SE, $t_4 = -0.4$, $p > 0.3$). We found no significant difference in competitive fitness when the two fluorescently-labelled strains were directly competed(GFP versus YFP: $w = 0.979 \pm 0.069$ SE, t-test against mean of 1: $t_4 = -0.3$, $p > 0.35$). Preliminary work has shown that fitness measurements of different strains against GFP and YFP are well-correlated (slope $= 0.98 \pm 0.03$ SE, $F_{1,8} = 850$, $R^2 = 0.98$, $p < 0.0001$), hence GFP and YFP can be used interchangeably for measuring competitive fitness.

Cost of rifampicin resistance

To test for a fitness cost associated with resistance, we directly competed rifampicin-sensitive GFP-PA01 against two rifampicin-resistant mutants (S536F and +P518) in M9-serine culture medium. We found that in this environment, these resistance mutations were associated with large and indistinguishable fitness costs when compared to their sensitive PA01 ancestor ($w_{\text{S536F}} = 0.401 \pm 0.0269$ SE, $w_{\text{+P518}} = 0.387 \pm 0.0640$ SE, Welch's t-test: $t_{6.87} = 0.20$, $p = 0.42$).

Probability of escaping genetic drift

Population genetic theory predicts that even highly beneficial mutations are likely to be lost from populations due to genetic drift when they are initially rare (reviewed in Patwa and Wahl 2008). In agreement with theory, we found that most sensitives fail to reach a detectable frequency: we identified sensitives above our detection threshold of 1 in 4×10^3 in 38/118 populations of S536F (proportion $= 0.32 \pm 0.043$ SE) and 54/118 populations of +P518 populations (proportion $= 0.46 \pm 0.046$ SE). We found that the number of populations containing sensitives rapidly decreased, as we would expect if rare sensitive populations were lost due to stochastic extinction (Figure 4.2). The proportions of detectable sensitives in each rifampicin-resistant background differed significantly ($\chi^2_{12} = 4.56, p = 0.033$); however, the observed difference ($d = 54 - 38 = 16$) did not differ significantly from the expected difference in the number of wells receiving at least one sensitive ($d = 10.22, p = 0.211$, see section 4.6), hence we cannot conclude that there was a difference in the probability of escaping drift.

Is this high rate of loss of sensitive lineages due to drift surprising? Haldane (1927) famously predicted that the likelihood that a weakly beneficial mutation survives stochastic extinction when it is rare is equal to $2s$, for mutations with small s. In the context of fixing sensitive mutants in resistant populations, s is the fitness advantage of being sensitive, which is inversely proportional to the cost of resistance: $s = 1/w_{\text{resistant}} - 1$ (because $w_{\text{sensitive}} = 1$). For our resistant strains, s cannot be considered small and so the revertant lineages should have been expected to survive; accounting for measurement error in the cost of resistance, s for the sensitive strain is 1.505 ± 0.168 SE in the presence of S536F and 1.66 ± 0.439 SE in the presence of +P518. However, more recent results suggest that drift also depends on either final population size and bottleneck

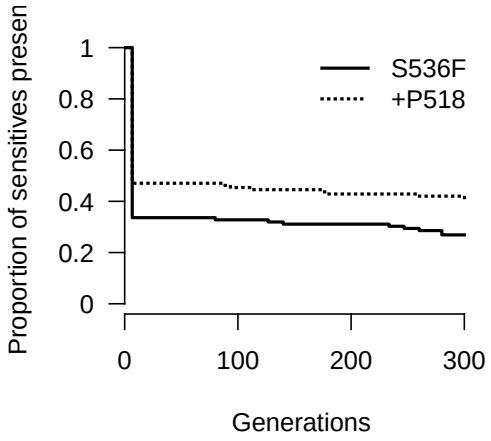

Figure 4.2: Survival of reversion mutations. This figure shows the proportion of populations ($n = 118$) with detectable sensitives (sensitive frequency > 0.0005%) in selection lines founded by the S536F (solid line) and +P518 strains (dotted line).

size (Levin *et al.* 2000; Wahl *et al.* 2002). Levin *et al.* 2000 showed an increase in the probability of fixation with an increase in population size and bottleneck size. Wahl *et al.* 2002 extends Haldane's model assuming reproduction by binary fission and recurrent populations bottlenecks and estimates that the fixation probability of a beneficial mutation is maximized with a dilution factor of $D \approx 0.135$, more than an order of magnitude larger than our $D = 0.005$, which may have contributed to the loss of sensitives in our populations. Under this model, we can generate predictions for the probability of escaping drift that range from $0.38 - 0.47$ for S536F and from $0.34 - 0.59$ for +P518. These predicted ranges overlap with the 95% confidence intervals of our observed probabilities of escaping drift: from $0.24 - 0.41$ for S536F and from $0.37 - 0.55$ for +P518. We therefore cannot reject the predictions from this model for either genotype.

Time to fixation

In the absence of clonal interference, beneficial mutations that escape genetic drift are expected to fix rapidly as a result of natural selection. For example, the passage time between when a back mutation to sensitivity becomes detectable (i.e. frequency > 0.0005%) and when it becomes fixed (i.e. frequency > 99.975%) under the conditions of our experiment is expected to be 8-15 generations: $t = \ln(F_t/F_0)/s$, where $F_t = f_t/(1-f_t)$ and $F_0 = f_0/(1-f_0)$ for arbitrary initial and final frequencies f_0 and f_t (Bell 2008). In contrast to this prediction, we found that sensitives that became fixed had passage times in excess of 100 generations, representing a 10-fold increase in passage time relative to our theoretical expectation (Figure 4.3). Intriguingly, we found that the time to fixation differed significantly between resistance mutations (Welch's t-test: $t_{59.8} = -4.32$, $p < 0.0001$), with a mean time to fixation of sensitives in 110.8 ± 8.4 SE generations for S536F and 157.8 ± 6.9 SE generations for +P518.

One potential explanation for the large discrepancy between expected and observed time to fixation is that clonal interference slows the fixation of the sensitive strain by effectively reducing the selection coefficient associated with sensitivity. To test this possibility, we followed the dynamics of reversion mutations in five populations of each resistant genotype where sensitives eventually became fixed and calculated the selection coefficient associated with the reversion mutations. Sensitives initially increased in frequency from 0.002% to about 50% within approximately 30 generations, followed by a slow approach to a frequency of 100% over the next 120 generations (Figure 4.4 A-B). This represents a decline in the selection coefficient associated with reversion mutations of $s = 1.5$ at the outset of the experiment to approximately $s = 0.01$ during the following 250 generations (Figure 4.5).

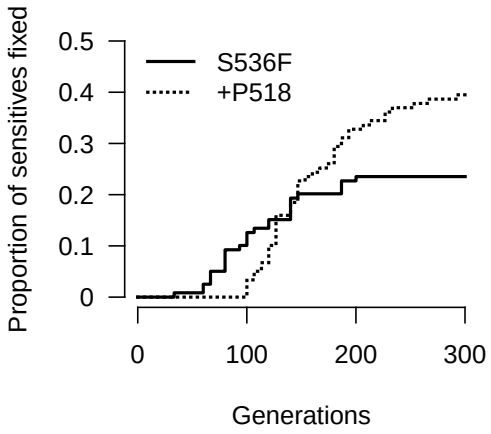

Figure 4.3: Fixation of sensitive lineages. This figure shows the proportion of populations ($n = 118$) where the sensitive had replaced the resistant strain (resistant frequency $< 99.975\%$) in populations founded by S536F (solid line) and +P518 (dotted line).

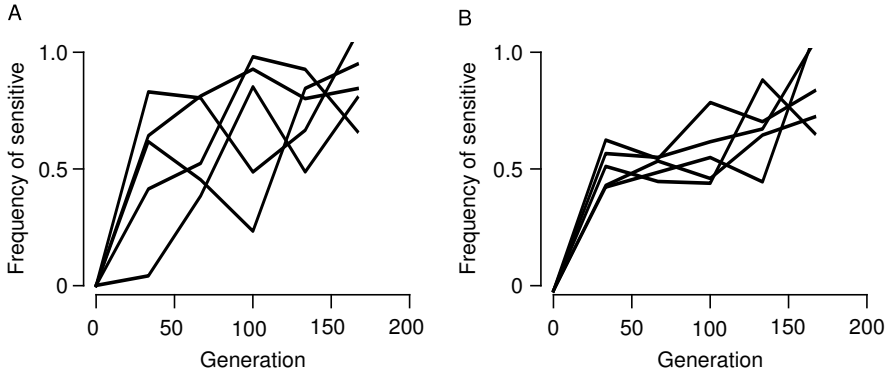

Figure 4.4: Dynamics of reversion. This figure shows the frequency of a random subset of sensitives that were eventually fixed by selection ($n = 5$) in populations of the S536F (A) and +P518 (B) resistant mutants. Frequency of sensitives through time was marked by a rapid increase to approximately 0.5 and then a gradual increase to 1, showing more variability in S536F populations (A) than +P518 populations (B).

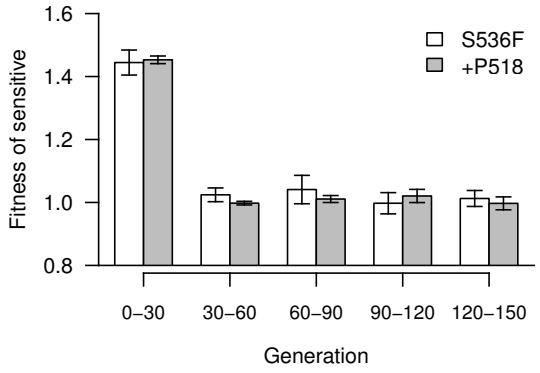

Figure 4.5: Change in advantage of sensitivity through time, calculated from the frequency of the sensitive lineage, relative relative to the resistant lineage (see Figure 4.4, mean ± SE; $n = 5$) for each time frame in populations founded by S536F (white bars) and +P518 (grey bars).

Likelihood of fixation

Surviving genetic drift does not guarantee that a beneficial mutation will be fixed by natural selection in an asexual population because clonal interference allows strongly beneficial mutations to drive weakly beneficial mutations extinct. In our experiment, we found that a minority of sensitives that escaped genetic drift at the outset of the experiment were ultimately driven extinct, suggesting that the occurrence of super-compensatory mutations increasing fitness beyond the fitness of the current sensitive type prevented the fixation of sensitives (Figure 4.2). The likelihood of extinction due to clonal interference was twice as high for sensitives in S536F as +P518, with 13/38=0.34 losses of sensitives that escaped drift in S536F populations and 9/54=0.17 losses of sensitives in +P518 populations, with marginal statistical significance (two-sample test for equality of proportions with continuity correction: $\chi^2_{12} = 3.77$, $p = 0.052$). However, we note the low power associated with test due to the small number of sensitives observed

(41-55%); to achieve power of 80%, we would have needed to track 96 sensitives in each resistant background.

This intriguing difference suggests that the two resistant mutants differ in their ability to generate the super-compensatory mutations that are required to drive established sensitive lineages extinct. To test this hypothesis, we measured the fitness of a random sample of 42 out of the 265 evolved populations of each resistant genotype in which the sensitives never achieved a detectable frequency (i.e. $> 0.0005\%$). Consistent with our earlier results (i.e. Figure 4.2), we found that selection for compensation was very efficient, as evolved populations fully recovered the large costs associated with resistance. However, the gains in fitness were asymmetric as predicted, with S536F compensating significantly more than +P518, with fitnesses of 1.19 ± 0.037 SE and 1.02 ± 0.059 SE, respectively (Figure 4.6, Welch's t-test: $t_{48.5} = -2.35$, $p = 0.023$).

Spontaneous reversion of resistance

We assayed for spontaneous reversion events by tracking the loss of the resistant type in the absence of the introduced sensitive. We did not observe any loss of resistance due to spontaneous reversion mutations in the 265 populations of resistant bacteria that we tracked for 300 generations. However, this does not preclude the possible occurrence of true spontaneous reversions that did not reach detectable frequency.

4.4 Discussion

In this study, we experimentally mimicked a mutation reversing antibiotic resistance in populations carrying costly resistance mutations and then tracked the evolutionary dynamics of the sensitive and resistant lineages to determine how drift and compensatory adaptation influence the likelihood of fixation of a

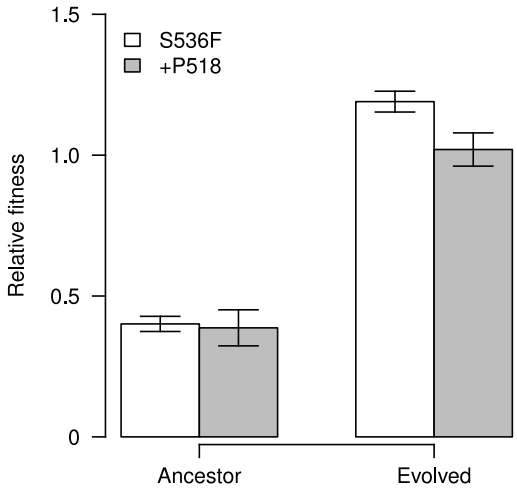

Figure 4.6: Evolvability of resistant mutants. This figure shows the relative fitness of the ancestral resistant mutants used to initiate the selection experiment (mean ± SE, $n = 5$) and the average fitness of 42 evolved populations derived from each resistant mutant. Note that for this assay we only sampled evolved populations in which sensitive lineages were never detected.

reversion mutation. We deliberately designed two features of this experiment to make it a 'best-case scenario' for the fixation of reversion mutations. First, we made the cost of resistance very large, which should maximize the likelihood of reversion by minimizing the likelihood that reversion mutations are lost due to genetic drift when they are initially rare. Second, we mimicked the appearance of reversion mutations before resistant populations had the opportunity to compensate for the cost of resistance, which should minimize the constraints on reversion imposed by clonal interference between reversion and compensatory mutations.

The majority of sensitives that we introduced were rapidly lost due to genetic drift, supporting a model in which the likelihood of a beneficial mutation (in

this case, a reversion mutation) escaping drift is determined by the selection coefficient associated with the beneficial mutation (Wahl *et al.* 2002). Most sensitives that escaped stochastic loss at the start of the experiment were eventually fixed by natural selection, but both the time to fixation and the probability of fixation differed substantially between populations of different resistant mutants. The extent to which the resistant strains compensated for the cost of resistance played a key role in determining the dynamics of the sensitive fixation. This implies that the propensity to ameliorate the cost associated with resistance may be just as important as the magnitude of the cost itself in determining the dynamics of reversion.

Maintenance of resistance

When is resistance likely to revert? Given that the cost of resistance provides the incentive for reversion, intuition suggests that the cost of resistance should play an important role in determining the likelihood of reversion (Andersson and Hughes 2010). Integrating the results of our experiment into the broader literature on compensatory adaptation suggests that, paradoxically, the cost of resistance may play only a minor role in the probability of reversion. Our results agree with the earlier work on streptomycin resistance by Levin *et al.* (2000), which demonstrated that introduced sensitives were more likely to be lost stochastically if introduced into compensated instead of uncompensated resistant populations. Levin *et al.* (2000) demonstrated that compensated resistant strains could prevent the fixation of an introduced sensitive strain with higher competitive fitness, putatively by increasing the stochastic effect of random sampling. Contrary to our findings, they found that introduced sensitives rapidly rose to high frequency in uncompensated resistant populations without hindrance from novel compensatory mutations–this is likely because the sensitive

strain was introduced at roughly four times the density of our experiments, making the changes in frequency more deterministic. One of the most striking results of our study is the finding that sensitives that escape drift always took a long time to achieve fixation, with an average fixation time about ten times longer than the predictions of a simple model (Bell 2008) that ignores compensatory mutations (Figure 4.3). The cause of this discrepancy is clear: sensitives fixed very slowly because compensatory adaptation resulted in a rapid decline in the selection coefficient associated with the reversion mutation from 0.5 to 0.01 (Figure 4.4), implying that both the +P518 and S536F mutants are able to easily acquire compensatory mutations with approximately equivalent benefits to reversion mutations. Is this surprising? The model of compensation that has emerged from mechanistic studies of compensatory adaptation (Kirschner and Gerhart 1998; Maisnier-Patin *et al.* 2002) is that compensatory mutations increase fitness by fully or partially recovering the biophysical defects associated with deleterious mutations, including resistance mutations. This biophysical model of compensation predicts that the benefits associated with compensatory mutations will likely covary with those of reversion mutations. The implication of this is that the likelihood of reversion could be largely independent of the cost of resistance; hopefully future studies will test this hypothesis directly.

Compensatory adaptation and evolvability

Given that the rate and likelihood of reversion depends critically on the ability of resistant mutants to evolve compensatory adaptations, it is important to understand the factors that determine the evolvability of resistant mutants. It is often argued that evolvability is determined by broad features of protein and genome organization, such as robustness, redundancy and modularity (Wagner 2008; Tokuriki and Tawfik 2009; Woods *et al.* 2011). Our results clearly show

that subtle sequence variation can also have a profound impact on evolvability, at least in the short term (see also MacLean *et al.* 2010). Specifically, our results imply that +P518 has access to more compensatory mutations than S536F, while S536F has access to better compensatory mutations than +P518[2]. A meta-analysis of compensatory adaptation shows that there are an average of approximately 11-12 compensatory mutations per resistance mutation (Poon and Chao 2005), possibly because compensatory mutations tend to be found in close proximity to deleterious mutations in the three dimensional structure of proteins (Kirschner and Gerhart 1998). Given that the opportunity for compensation is clearly limited, it is not surprising that resistance mutations differ in their evolvability. In the future, we plan to use a combination of detailed fitness measurements combined with whole-genome sequencing to understand the causes of this variation in evolvability.

4.5 Summary

This study challenges the long-held notion that it is the rarity of reversion mutations alone that impedes the reversion of antibiotic resistance and instead emphasizes the importance of the intrinsic evolvability of resistant populations for determining the dynamics of reversion. Future work needs to focus on understanding how the cost of resistance affects the likelihood of reversion and how the molecular mechanisms of resistance determine the ability to compensate for the cost of resistance. In a more applied context, we suggest that resistance management strategies should take information on both the short term costs and long-term evolvability associated with resistance mechanisms into account (MacLean 2010).

[2]Chapter 5 will later show that this hypothesis was incorrect.

4.6 Appendix to Chapter 4[3]

Expectation of number of wells receiving revertants

Here, we are interested in determining the expected absolute difference in the number of 'successful' wells in the two rifampicin-resistant genetic background, that is, those that received at least one revertant cell at the outset of the selection experiment. We need to know this expected difference in order to know if any observed differences in the number of revertants escaping genetic drift is due a biological feature or simple random chance.

To introduce revertants into the rifampicin-resistant background, we diluted an overnight culture of revertants (i.e. high bacterial density) to a density of $0.0185/\mu$l and inoculated 20 μl of the diluted culture to 384 wells containing resistant bacteria; this resulted in an average of 0.37 cells/well, which we determined empirically. The number of cells in each 20 μl inoculum thus follows a Poisson distribution with mean $\lambda = 0.37$ cells. The percentage of populations receiving one or more cells is given by the complement of the probability mass function when the number of cells introduced, k, is zero; this represents the area of the right-tail of the Poisson distribution:

$$\begin{aligned} 1 - \Pr(X = k) &= 1 - \lambda^k e^{-\lambda}/k! \\ 1 - \Pr(X = 0) &= 1 - 0.37^0 e^{-0.37}/0! \\ &\approx 0.309 \end{aligned}$$

On average, this should result in $384 \times 0.309 \approx 118$ wells receiving one or more cells. However, there is going to be variation in the number of wells receiving one or more cells. Let X_i, $i = 1, 2, \ldots, 384$, be the number of revertant cells

[3]Published as online supplemental text to the published version of this chapter.

received by well i. To find the number of wells receiving one or more revertant cells ($X_i \leq 1$), let Z_i represent the failure or success of receiving at least one revertant cell, $i = 1, 2, \ldots, 384$, such that

$$Z_i = \begin{cases} 0 & \text{for} \quad X_i = 0 \\ 1 & \text{for} \quad X_i \geq 1 \end{cases}$$

Each Z_i follows a Bernoulli distribution with probability of success $p = 1 - e^{-0.37}$. As we inoculated the dilution containing revertants into 384 wells in total for each rifampicin-resistant genotype, we let Q_i represent the number of successful introductions of revertants in each resistant background ($j = 1, 2$, for S536F and +P518). Because $Q_i = \sum_i Z_i$ the distribution of Q_j is Binomial($n = 384, p = 1 - e^{-0.37}$). The distributions of Q_j are identical because the inoculations were performed from the same culture dilution for both rifampicin-resistant strains, and independent because the inoculations were done sequentially.

Now we can define the expected difference of interest as the expected absolute difference between any two draws from Q_1 and Q_2, or in other words $E[\Sigma|Q_1 - Q_2|]$. We took two approaches to estimating this difference, which gave the same final result. In the first approach, we simulated 10^7 random draws from Binomial($n = 384, p = 1 - e^{-0.37}$) and found the mean of the absolute difference using the following code in R 2.14.1 (R Development Core Team 2011),

```
q1=rbinom(size=384, n=1e7,prob=1-exp(-0.37))
q2=rbinom(size=384, n=1e7,prob=1-exp(-0.37))
diff=q1-q2
mean(abs(diff))
var(abs(diff))
```

Using this approach, we found point estimates for the mean $\hat{\mu} = [|Q_1 - Q_2|] \approx 10.22$ and variance $\hat{s}^2 = [|Q_1 - Q_2|] \approx 59.76$.

For the second approach, we used the normal approximation for the binomial distribution to approximate the distribution of $Q_1 - Q_2$ as a normal distribution with mean μ and variance σ^2. The mean of this distribution is simply the difference of the means of Q_1 and Q_2, hence $\mu = 0$. The variance of $Q_1 - Q_2$ is given by $\text{Var}[Q_1 - Q_2] = \text{Var}[Q_1] + \text{Var}[Q_2] - \text{Cov}[Q_1, Q_2]$. Given that Q_1 and Q_2 are identical and independent random variables, $\text{Cov}[Q_1, Q_2] = 0$ and $\text{Var}[Q_1] = \text{Var}[Q_2]$, hence $\sigma^2 = 2np(1-p) = 2 \times 384(1 - e^{-0.37})(e^{-0.37})$. The absolute value $|Q_1 - Q_2|$ is therefore approximated by a half-normal distribution (a special case of the folded-normal distribution with $\mu = 0$), whose mean and variance are defined relative to its corresponding normal distribution (Leone *et al.* 1961):

$$E[|Q_1 - Q_2|] = \sigma\sqrt{2/\pi}$$
$$\text{Var}[|Q_1 - Q_2|] = \sigma^2(1 - 2/\pi)$$

Given μ and σ^2 from the normal distribution approximation, $E[|Q_1 - Q_2|] \approx 10.22$ and $\text{Var}[|Q_1 - Q_2|] \approx 59.62$. The observed difference between the number of revertants that escaped drift was 16. This value lies well within the above half-normal distribution ($p = 0.211$), and suggests that our observed difference in the number of revertants lost to drift is not significantly different from the chance difference in the number of wells that received at least one revertant at the outset of the experiment.

5 | Epistasis constrains evolvability of two *rpoB* mutants on L-serine

Abstract

Evolvability is a measure of the ability to produce adaptive genetic variation, upon which natural selection can act. Theoretical models of adaptive evolution suggest that genotypes with the same initial fitness and mutation rate should have the same evolvability; for antibiotic resistant bacteria, this is measured as the fitness cost imposed by resistance relative to the sensitive wild-type strain. In contradiction of these models, here I show that two *P. aeruginosa rpoB* mutants (S536F and +P518) with the same costs of resistance have different evolvability when adapting to an environment containing L-serine as the sole carbon source. The data suggest that the higher evolvability of S536F was due to a combination of increased mutation rate and magnitude epistasis between beneficial mutations and the *rpoB* locus. Fitness gains were not compensatory for the *rpoB* defect, but specific to L-serine. The greatest gains in fitness were associated with mutations in a single regulatory gene, PA2449, with a previously demonstrated role in L-glycine and L-serine metabolism.

5.1 Introduction

What constrains the ability for an organism to adapt to a new environment? Minimalist models often applied to compensatory evolution of antibiotic-resistant genotypes assume that the size of beneficial mutations available is proportional to the fitness deficit of the resistance mutation (Martin and Lenormand 2006). Previous results have shown that evolvability increases with the size of the fitness deficit caused by *rpoB* mutations in *E. coli* (Barrick *et al.* 2010). Chapter 4 produced a contradictory result: two *rpoB* mutants, S536F and +P518, with essentially equal costs of resistance had dramatically different ability to adapt to L-serine as a sole carbon source (Figure 4.6). Building on this result, in this chapter I investigated the genetic basis for adaptation to M9-serine to determine if the difference in evolvability was due to fixing more mutations, different mutations, or the same mutations different effects on fitness. A difference in evolvability between two bacterial genotypes can be caused by differences in mutation rate, by epistasis between the ancestral genotype and new mutations that arise, or by an interaction between the two (Colegrave and Buckling 2005; Colegrave and Collins 2008). In asexual organisms, increases in mutation rate increase the rate of adaptation, up until a point where multiple mutations simultaneously present in different backgrounds interfere with each other (De Visser *et al.* 1999). If mutation rates differ, strong population bottlenecks should have a greater inhibitory effect on the adaptation of the genotype with lower mutation rate. Unlike differences in mutation rate, epistasis alters the topology of the fitness landscape on which the genotypes are adapting (Colegrave and Collins 2008). How epistasis affects evolvability depends which form of epistasis dominates, namely sign epistasis or magnitude epistasis. Sign epistasis occurs when a mutation confers a benefit in one background, but a cost in another, and affects

the ruggedness of the landscape, leading to the fixation of a different suite of mutations in each background. Magnitude epistasis, on the other hand, reduces the magnitude of the fitness benefit, and hence the height of adaptive peaks on the fitness landscape. Common mutations should fix in both backgrounds (although if epistasis causes selection for these mutations to weaken due to the smaller effect, other mutations may also fix in the background with the lower average effect).

For *rpoB* mutants, there are two potential routes for improving fitness in M9-serine: compensate for the *rpoB* defect or improve the ability to consume L-serine. Previous studies have found compensatory mutations in *rpoA* and *rpoC*, and that some *rpoB* mutations can also compensate for each other (Reynolds 2000; Hall *et al.* 2011; Comas *et al.* 2012). The difference in evolvability could arise because of sign or magnitude epistasis between compensatory mutations and the *rpoB* genotype, if some *rpoB* defects are harder to correct. However, there is also likely to be ample potential for adaptation to L-serine metabolism, as it is not a preferred carbon source for *P. aeruginosa* (Boulette *et al.* 2009). How the *rpoB* genotype could affect the fitness benefit adaptive mutations is unclear, as the genetic basis for adaptation to L-serine is not yet known.

To investigate these mechanisms, I selected for single beneficial mutations in both *rpoB* backgrounds by tracking the frequency of a fluorescent marker in a serial transfer regime, where the selective sweep of a beneficial mutation was assumed to be linked to fixation or loss of the marker. For lineages where a beneficial mutation appeared to have fixed, I performed whole-genome sequencing and competitive fitness assays to determine which mutations contributed to increased fitness. To test for the effects of mutational supply, I used two bottleneck sizes during selection, and used a fluctuation assay to measure relative mutation rates of the two *rpoB* ancestors. Finally, I measured the competitive

fitness of the mutations across several environments to determine whether they were general compensatory mutations or adaptive to M9-serine. Among the many genes mutated, I found two genes in common in both *rpoB* backgrounds. Mutations appeared to be specific to M9-serine metabolism, rather than generally compensatory. S536F tended to fix more mutations in a gene with a known function in L-serine metabolism, although a bias inadvertently introduced into the sequencing makes this conclusion tentative until further work is done.

5.2 Methods

Growth medium and strains

The selection experiment was performed in minimal culture medium supplemented with L-serine as the carbon source [M9-serine: 10.5 g/l M9 Broth 1 ml of 1 M $MgSO_4$, and 3.75 g/l (35.7 mM) L-serine]. We used two rifampicin-resistant strains of *P. aeruginosa*; in RpoB, these correspond to amino-acid changes S536F and +P518. The two rifampicin-resistant mutants were tagged with a YFP genetic marker using the mini-Tn7 insertion protocol (with delivery vector pUC18T-mini-Tn7T-Gm-eyfp, accession no. DQ493879, Choi and Schweizer 2006). This marker was initially associated with a slight fitness cost that was only significant in +P518 (0.054 ± 0.025, $t_7 = -3.18$, $p = 0.016$) but not S536F (0.010 ± 0.004, $t_7 = -2.13$, $p = 0.12$).

Long-term fitness trajectories

In chapter 4, S536F and +P518 evolved over the course of 300 generations in M9-serine. I measured the fitness of these populations after approximately 67, 134, 201 and 268 generations (corresponding to transfers 10–40), using the flow

cytometry method (subsection 2.3). Evolved lineages were competed against YFP-PA01 at a 20% to 80% initial ratio.

Selection for single beneficial mutation

A short-term selection experiment in M9-serine was performed to select for individual beneficial mutations. Populations were initiated as a 1:1 mixture by volume of YFP-labelled and unlabelled *rpoB* mutant cells. The presence of the marker allows for the detection of selective sweeps; if the marker fixes in the population, this suggests that the marker has 'hitch-hiked' along with a beneficial mutation. Similarly, if the marker is lost, it suggests that a beneficial mutation has occurred in the unmarked lineage. Thirty populations were transferred to 200 µl of fresh M9-serine every 24 h for 14 transfers in total. To modulate mutational supply, I used two dilution factors $1/D$: low ($D = 200$) and high ($D = 20,000$). These dilution factors require a minimum number of daily generations of $\log_2(200) \approx 7.64$ or $\log_2(20,000) \approx 14.28$ generations/day to prevent population extinction due to repeated dilution, hence the $D = 20,000$ populations experienced nearly twice as many generations of evolution. A third dilution factor ($D = 2 \times 10^6$, requiring 20.93 generations/day) was abandoned after several populations went extinct. The frequency of marked and unmarked cells were assayed on every second day by flow cytometry using a 10 µl sample, corresponding to roughly $10^3 - 10^4$ cells. I used 90% and 10% as the threshold for fixation or loss of the marker. I froze a 100 µl population aliquot every second day at −80 °C.

DNA extraction and quality assessment

I extracted genomic DNA (gDNA) of isolates from populations that had fixed for one of the two strains (i.e. marked or unmarked with YFP). Based on the

marker trajectory data, lineages from this selection experiment were chosen for sequencing. A single clone was isolated from stocks frozen on the day that it passed the threshold, to reduce the chance of obtaining clones with multiple mutations. DNA extractions were performed with the Promega Wizard Genomic DNA Purification Kit (Promega, UK). First, single colonies were isolated from LB agar plates and grown overnight in liquid LB. 1 ml of dense culture was centrifuged at maximum speed (13,000 rpm) for 2 min on a microcentrifuge. The supernatant was removed and the pelleted cells mixed with 600 µl nuclei lysis solution. Cells were then incubated for 5 min at 80°C and then cooled to room temperature. RNA in the cell lysate was degraded by adding 3 µl of RNase solution, mixing by inverting and incubation for 15-60 min at 37°C. Proteins were precipitated by adding 200 µl protein precipitate solution to the cell lysate and vortexing at maximum speed for a minimum of 30 s, followed by incubation on ice for 10 min and centrifuging at maximum speed for 10 min. 400 µl of supernatant containing DNA was recovered by pipetting and mixed with 600 µl isopropanol, followed by centrifuging at maximum speed to pellet the gDNA. The isopropanol was removed by pouring and replaced with 600 µl 70% ethanol. The gDNA pellet was washed by gentle pipetting. The gDNA was again pelleted by centrifuging at maximum speed for 2 min, after which ethanol was removed by pipetting. gDNA pellets were dried for 1 h and then rehydrated overnight in 100 µl to 300 µl of elution buffer (EB), depending on pellet size.

The extracted gDNA was assessed for quality. gDNA degradation was assessed by migrating approximately 75 ng on a 0.7% agarose gel with SYBR Safe (Invitrogen) for 30 min at 100 V. Bands were checked for smearing, which would indicate degradation. The presence of salt and protein in the samples was detected using a NanoDrop 2000 small-volume spectrophotometer (Thermo Fisher Scientific). Absorbance at 230 nm, 260 nm, and 280 nm was measured,

and ratios are used to calculate acceptable levels of proteins (260 nm/280 nm) and salts (260 nm/230 nm). Guideline acceptable ratio for proteins ranges from 1.8-2.0, and for salts ranges from 2.0-2.2. However, these guidelines assume an average of the ratios for pure nucleotide solutions, and GC-rich *P. aeruginosa* gDNA is likely to have lower ratios due to the lower relative abundance of adenine, which has the highest absorbance at 260 nm. Therefore, I accepted protein ratios as low as 1.6 and salt ratios as low as 1.7.

gDNA concentration was determined using the QuantiFluor dsDNA System (Promega, UK), which binds fluorescent proteins to DNA to determine concentration relative to a standard curve of known Lambda phage DNA concentration. Stock solutions of the QuantiFluor dye were prepared according to the protocol instructions. gDNA is first diluted to 1:100 and 1:1000 in TE buffer, then 100 µl is mixed with then 100 µl QuantiFluor dye in a black 96 well microtitre plate. After 5 min of incubation, fluorescence intensity is read at 490 nm/510 nm emission/excitation. After determining gDNA concentration, the samples were diluted to approximately 50 ng/µl in 100 µl in EB, according to the specifications of the sequencing centre. The concentration was re-assessed using the QuantiFluor protocol. gDNA was stored at −20°C until delivery to the sequencing centre, and was delivered on dry ice to prevent degradation in transit.

Sequencing and analysis

Whole genome sequencing was performed by the Wellcome Trust Centre for Human Genetics (Oxford, UK) using the Illumina hiseq 2000 platform with 100 bp paired-end reads. Initial filtering of the reads was done using NGS QC Toolkit (Patel and Jain 2012). 5' or 3' ends were trimmed if the Phred quality score was less than 20. Reads were discarded if they were < 50 bp after trimming, if $> 2\%$ of bases were ambiguous, or if more than 20%

of bases had a Phred score < 20. The filtered reads were mapped to the *P. aeruginosa* PA01 reference genome (NC_002516.2) using BWA. I processed the mapped reads to increase the quality of the variant calling: 1) reads with multiple best hits were discarded; 2) duplicated reads were discarded using `MarkDuplicates` from the Picard package (http://picard.sourceforge.net); 3) reads around indels were locally realigned using `RealignerTargetCreator` and `IndelRealigner` from the GATK package to correct for misalignment; and 4) mate pairs were sorted using `FixMateInformation` in the Picard package. Variant calling was performed using two tools: GATK's `Unified Genotyper` (DePristo *et al.* 2011) and Samtools's `mpileup` (Li *et al.* 2009). VCFtools (`vcf-annotate`, Danecek *et al.* 2011) and GATK toolkit (`VariantFiltration`, DePristo *et al.* 2011) were used to filter the raw variants for strand bias, end distance bias, base quality bias, SNPs around gaps, low coverage and erroneously high coverage. High quality variants not filtered were annotated using SnpEff (Cingolani *et al.* 2012).

Structural variants were detected using three approaches. First, `BreakDancer` (Chen *et al.* 2009) was used to predict deletions, insertions, inversions, and translocations using deviations in the separation or orientation of mapped read pairs. `Pindel` (Ye *et al.* 2009) was used to infer deletions, short insertions, long insertions, inversions, tandem duplications, and breakpoints using a split-read approach (the output of `BreakDancer` was also fed to Pindel to improve its output). Finally, `Control-FREEC` (Boeva *et al.* 2011) was used to detect copy number variants (CNVs). `Control-FREEC` finds CNVs using depth-of-coverage (normalized by GC-content). Regions of low mappability were excluded by supplying `Control-FREEC` with mappability tracks generated by `gem-mappability` (GEM library, Marco-Sola *et al.* 2012).

Competitive fitness assays

To determine the fitness benefit of the sequenced mutations, competitive fitness of the sequenced lineages was measured in M9-serine, using the flow cytometry method (see chapter 2, subsection 2.3). I mixed marked and unmarked cells at an initial frequency of 20% evolved to 80% *rpoB* ancestor (hence a neutral mutation would have fitness of 1), and quantified fitness according to the change in frequency across a 24-hour growth period. Evolved populations were revived from freezer stocks by pin replicating into fresh medium, growing for 24 h, and again transferring into fresh medium for another 24 h, to ensure that strains could acclimate to the growth medium.

To measure mutational pleiotropy, competitive fitness was also measured in M9 with three additional carbon sources: M9-glycine (4.01 g/l L-glycine, to match the carbon molarity of M9-serine), M9-glucose (12.5 g/l α-D-glucose) and M9KB (10 g/l glycerol, 10 g/l proteose peptone no. 3). These data were used to assess whether mutations were adaptive (advantageous in M9-serine only) or compensatory (advantageous in all environments).

Mutation rates

Mutation rates were measured using a Luria-Delbrück fluctuation test. Thirty independent cultures of each genotype were grown in 200 µl of LB broth for 24 h at 37°C. To estimate total population size (N_t), I sampled 10 µl from each of the 30 populations. The dense cultures were mixed, then diluted to 10^{-5}, 10^{-6} and 10^{-7} in ten independent dilution series, and three replicates of 20 µl of each dilution were plated on solid LB agar. Colonies were counted after 12 h growth at 37 °C. To estimate the number of mutations, I plated 170 µl of dense cultures on to selective plates: 30 plates of LB agar with 256 mg/L of streptomycin or 240 mg/L gentamicin, each concentration representing 8 ×

MIC for these strains. Resistant colony counts were determined after 36 h of growth for streptomycin. After 36 h and 60 h of growth, no gentamicin-resistant colonies were observed. To estimate mutation rate for streptomycin resistance, I used the program Fluctuation AnaLysis CalculatOR (FALCOR, Hall *et al.* 2009) implementation of the Ma-Sandri-Sarkar Maximum Likelihood Estimator method (MSS-MLE, Sarkar *et al.* 1992), which is regarded as the best method for fluctuation test analysis (Foster 2006). The estimated mutation rates were corrected by a factor $(z-1)/(z\ln z)$ where $z = 170/200$, due to plating only 170 µl of 200 µl from each well (Foster 2006). To estimate the mutation rate for gentamicin resistance, the p_0 method was used, which uses the proportion of cultures without mutants (p_0) to estimate mutation rate (Rosche and Foster 2000). An 95% upper confidence limit for the proportion of cultures without a mutant was calculated using a binomial test, with $n = 30$ trials corresponding to the number of cultures tested. Mutation rate is then calculated as $(-\ln p_0)/N_t$.

5.3 Results

Long-term fitness trajectories

Competitive fitness measured at several time points from the long-term lineages from chapter 4 showed that fitness reached a plateau after approximately 150 generations (Figure 5.1).

Marker trajectory-predicted selective sweeps

The population marker trajectories are shown in Figure 5.2. Based on the fixation or loss of the marker, logistic regression suggested a significant interaction effect on the number of sweeps (Figure 5.3, $p < 0.001$) between *rpoB* genotype and dilution factor D, with most occurring in $D = 200$ for S536F and in $D = 20,000$

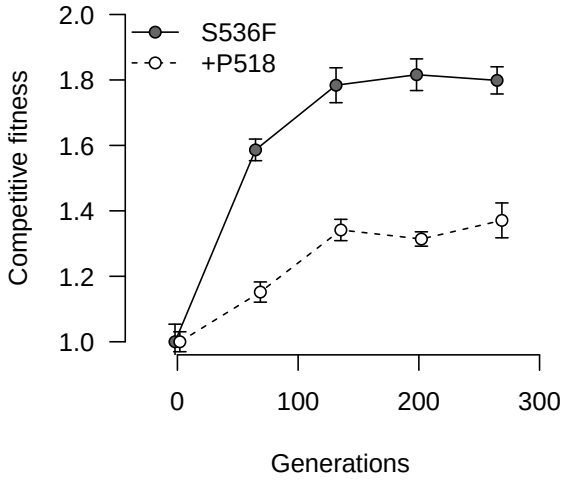

Figure 5.1: Fitness of long-term lineages (from chapter 4) relative to their *rpoB* ancestor.

for +P518. However, the sequencing data (see following section) showed that some of the lineages where the YFP marker was lost were not associated with any mutations.

Sequenced mutations

We sequenced isolates from populations that fixed for one of the two strains (i.e. marked or unmarked with YFP): from the $D = 200$ treatment, 14 from S536F and six from +P518; from the $D = 20,000$ treatment, four from S536F and 14 from +P518. The mutations found are listed in Table 5.1. Most lineages fixed only one mutation, with three S536F lineages fixing two (these were excluded from analyses of competitive fitness). A single gene, PA2449, had the highest frequency of mutations. The sequencing revealed that in 10 +P518 populations where the marker was lost (two from $D = 200$ and eight from $D = 20,000$), there was no associated beneficial mutation (dashed red lines, Figure 5.2). These lineages were excluded from the analysis.

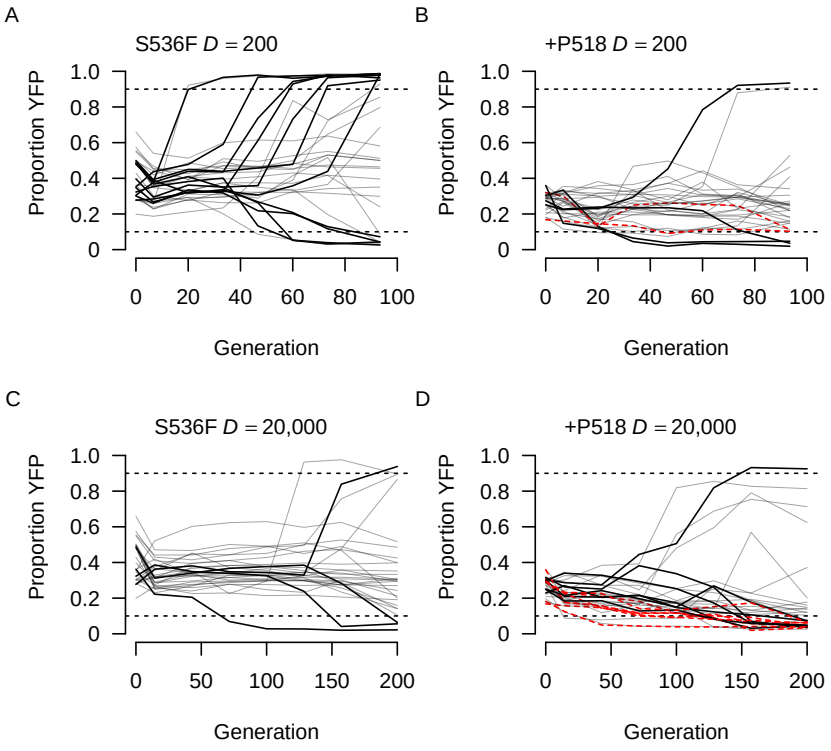

Figure 5.2: Evolutionary trajectories of populations starting from a mixture of YFP-labelled and unlabelled cells with dilution factor $D = 200$ for (A) S536F and (B) +P518 and $D = 20,000$ for (C) S536F and (D) +P518. Solid black lines: one or more mutations found. Dashed red lines: sequenced, but no mutations found. Grey lines: not sequenced.

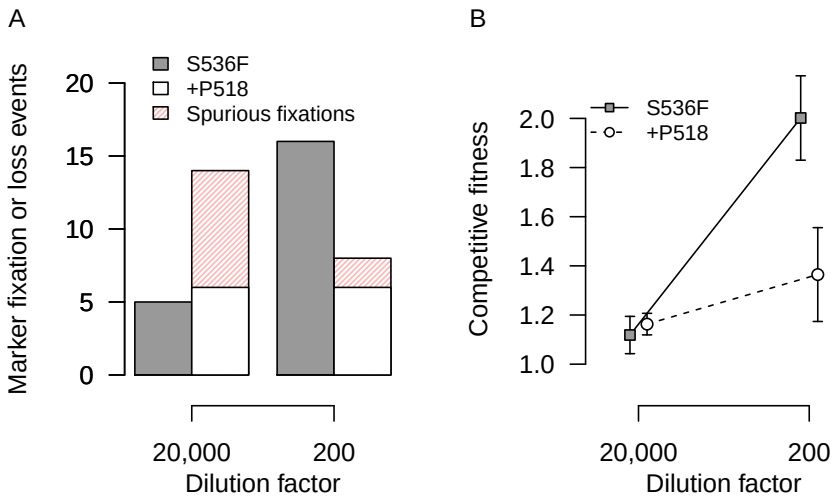

Figure 5.3: (A) Number of times the fluorescent marker was fixed or lost (i.e. surpassed 90% or 10% of the population). Spurious fixations are events where the marker fixed, but sequencing revealed no mutation was present. (B) Competitive fitness of the winning lineage, where the marker was either fixed or lost from the population (spurious fixations removed).

Table 5.1: Mutations found in rifampicin-resistant lineages adapted to M9-serine, relative to ancestral S536F and +P518.

Gene	Mutation[a]	Protein function[b]	*rpoB*	Treatment
Single mutants				
PA1975	ΔALGA157	contains putative amino acid sensory domain	+P518	$D = 20,000$
PA1975	V265G		+P518	$D = 20,000$
PA1976 (*ercS'*)	F323SfsX22[c]	sensor kinase	+P518	$D = 20,000$
PA2449	R77C	transcriptional regulator (TyrR- or EBP-like)	S536F	$D = 200$
PA2449	R77C		S536F	$D = 200$
PA2449	L143V		S536F	$D = 200$
PA2449	P181L		S536F	$D = 200$
PA2449	P181L		S536F	$D = 200$
PA2449	P181S		S536F	$D = 200$
PA2449	R183C		S536F	$D = 200$
PA2449	A189V		S536F	$D = 200$
PA2449	H193D		S536F	$D = 200$
PA2449	P263L		S536F	$D = 200$
PA2449	S308P		S536F	$D = 200$
PA2449	Q397L		S536F	$D = 200$
PA2449	G22D		+P518	$D = 200$
PA2449	G22D		+P518	$D = 200$
PA2449	I184V		+P518	$D = 200$
PA2449	R77C		S536F	$D = 20,000$
PA2449	P263L		S536F	$D = 20,000$
PA2459	L22CfsX11[d]	hypothetical protein	+P518	$D = 20,000$
PA2620 (*clpA*)	I362TSRV	ATP-binding protease component	+P518	$D = 200$
PA3521	S146L	putative outer membrane efflux protein (RND family)	+P518	$D = 20,000$
PA4314 (*purU1*)	V14G	formyltetrahydrofolate deformylase	S536F	$D = 200$
PA4696	G449S (*ilvI*)	acetolactate synthase large subunit	+P518	$D = 20,000$
PA5181	P338L	predicted oxidoreductase	+P518	$D = 20,000$
intergenic	a3106864g		S536F	$D = 200$
Double mutants				
PA0278	M45T	hypothetical protein (predicted permease)	S536F	$D = 20,000$
PA2449	Q397L	transcriptional regulator (TyrR- or EBP-like)		
PA1975	V228A	contains putative amino acid sensory domain	S536F	$D = 20,000$
intergenic	c942581t			
PA2449	Q397L	transcriptional regulator (TyrR- or EBP-like)	S536F	$D = 200$
intergenic	a223677g			

D–treatment dilution factor; [a]Lower case characters are nucleotide substitutions, upper case are amino acid substitutions, both relative to PA01 reference genome (NC_002516.2); [b]From Pseudomonas Genome Database (Winsor *et al.* 2011); [c]Single nucleotide deletion causing F323S and premature stop codon at position 345; [d]Single nucleotide insertion causing L22C and premature stop codon at position 33.

Competitive fitness of single mutants

The competitive fitness of all single mutants was measured in M9-serine, M9-glycine, M9-glucose and M9KB (Figure 5.4). In M9-serine, average fitness of S536F lineages carrying PA2449 mutations was 1.92 ± 0.17 SE; only one lineage received a single mutation in another gene (PA1975), which had a fitness of 1.04 ± 0.02 SE. In +P518, lineages with PA2449 mutations had higher average fitness than those with mutations in other genes, but without statistical significance due to small sample size (1.43 ± 0.31 SE vs. 1.17 ± 0.04 SE, $t_{1.03} = -0.84$, $p = 0.55$). Mean fitness of PA2449 mutants in both backgrounds roughly corresponded to the fitness plateau in Figure 5.1.

In M9-glycine, evolved S536F lineages carrying PA2449 mutations had higher average fitness than the ancestor (1.12 ± 0.05 SE, $t_{10} = 2.3$, $p = 0.046$), but were significantly deleterious in M9-glucose (0.75 ± 0.04 SE, $t_{10} = -5.55$, $p < 0.001$) and M9KB (0.94 ± 0.03 SE, $t_{10} = -2.82$, $p = 0.02$). Among PA2449 mutations in S536F, fitness in M9-serine and M9-glycine had a significantly negative rank-correlation (Spearman's $\rho = -0.67$, $p = 0.04$), suggesting a trade-off in the ability to use either carbon source. In +P518 lineages, while carrying a mutation in a gene other than PA2449 conferred a benefit in M9-serine ($t_8 = 3.27$, $p = 0.01$), it was deleterious in M9-glycine (0.86 ± 0.03 SE, $t_8 = -5.3$, $p = 0.001$), but effectively neutral in M9-glucose (0.99 ± 0.05 SE, $t_8 = -0.56$, $p = 0.59$) and M9KB (0.99 ± 0.03 SE, $t_8 = -0.27$, $p = 0.8$). Mutational pleiotropy, measured by the variance in fitness across environments, was greatest for PA2449 mutations ($s^2 = 0.255$ vs 0.020, F-ratio test: $F_{31,47} = 0.0791$, $p < 0.0001$).

Mutation rates

Total population size sampled for the fluctuation assay was $14.1 \pm 1.19 \times 10^9$ SE for S536F and $13.7 \pm 1.07 \times 10^9$ SE for +P518 (two-sample t-test: $t_{17.5} = 0.26$,

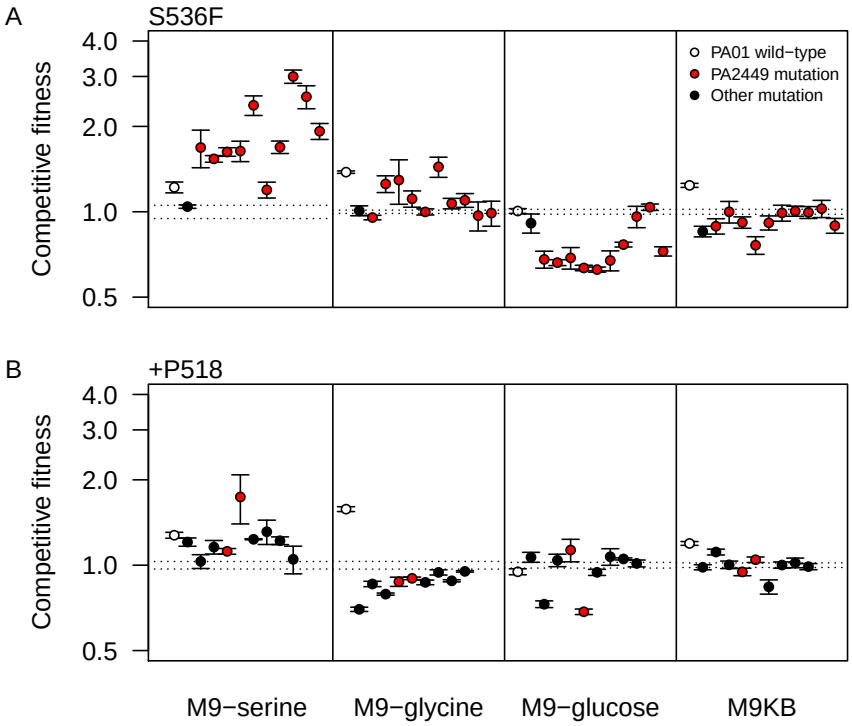

Figure 5.4: Pleiotropy was measured as the variance in competitive fitness across four environments, of the single beneficial mutations fixed in the S536F (A) and +P518 (B) backgrounds. (Points: means ± SE, $n = 4$ replicates; horizontal bars: SE of ancestral fitness, $n = 10$ replicates).

$p = 0.6$). For the LB with streptomycin plates, the mutation rate per 10^9 cells for S536F was 0.64 (95% CI: 0.50-0.79) and for +P518 was 0.43 (95% CI: 0.33-0.55). The difference in mutation rates is statistically significant (two-sample t-test: $t_{58} = 5.28$, $p < 0.0001$. This corresponds to an approximately 50% greater mutation rate to streptomycin resistance for S536F over +P518. No resistant colonies formed on the LB with gentamicin plates, placing an 95% upper confidence limit on the proportion of colonies without a mutant at $p_0 = 0.116$ (obtained from drawing zero mutations out of 30 trials from a binomial distribution). This corresponds to an upper limit on the mutation rate

to gentamicin resistance for S536F and +P518 at 0.157 and 0.150 per 10^9 cells, respectively.

5.4 Discussion

In this chapter, I investigated the difference in evolvability of two *rpoB P. aeruginosa* strains, S536F and +P518, by selecting for and sequencing single beneficial mutations. Evolvability to an environment containing L-serine as the sole carbon source seems to be dictated by magnitude epistasis between the *rpoB* locus and the mutations responsible for adaptation, and to some extent, by the supply of beneficial mutations. However, a bias in the data collection complicates the comparison of the whole genome sequence data between lineages evolving from the two *rpoB* genotypes. Here I will discuss the interpretations that can be made from this data, and the further steps I am taking to rectify the problems with the analysis.

Basis for evolvability

The recovery of fitness in these strains was made through adaptation, not compensation. Mutations in a single regulatory gene, PA2449, were associated with the greatest gains in fitness in both *rpoB* backgrounds, but the average beneficial effect was greater in S536F. This would suggest that magnitude epistasis constrains evolvability in this system; however, the number of lineages sequenced in each dilution treatment was heavily skewed toward S536F lineages in $D = 200$ and +P518 lineages in $D = 20,000$. It is unclear, therefore, whether the increased frequency and higher average fitness benefit of PA2449 mutations in S536F is due to genetic background or dilution factor treatment. The sampling bias was introduced by choosing lineages for sequencing based on the YFP marker exceeding the threshold for fixation or loss, which was an attempt to avoid

sequencing lineages with either no or more than one mutations. However, several sequenced +P518 lineages had no mutations anyway; sequencing an equal number of lineages from each treatment would have been a better approach.

Despite the sampling bias, the available evidence still provides some insight into evolvability in M9-serine. By manipulating dilution factor (D) and measuring the mutation rate to streptomycin resistance, I found that mutational supply may also contribute to the difference in evolvability. Strain S536F had a roughly 50% greater mutation rate than +P518, suggesting with an increased opportunity for S536F to sample large-effect beneficial mutations. The effects of higher mutation rate are evident from the sequencing data, where S536F fixed PA2449 mutations even in the face of a strong bottleneck ($D = 20,000$), but +P518 did not. Why S536F has a higher mutation rate is unclear, but it may be because some RpoB mutations, including S536F, cause RNA polymerase to behave stringently, even in the absence of environmental cues for stress conditions (Zhou and Jin 1998). This response shifts gene expression toward amino acid synthesis and away from rapid growth, and has also been associated with elevated mutation rates (Wright 1996).

Although the difference in mutation rate suggests +P518 would have a longer waiting time for beneficial mutations, mutational supply alone cannot account for the difference in evolability previously observed in the long-term experiment (Figure 5.1). Typical for adaptive evolution experiments (Elena and Lenski 2003), the fitness of the adapting lineages from both *rpoB* mutants appeared to reach a plateau after roughly 150 generations, rather than continuing to increase for the duration. If mutational supply had been the only limiting factor, +P518 lineages should have had ample time to catch up, as the additional generations of selection should have negated the higher mutation rate of S536F. Instead, the average effect of arising beneficial mutations in each *rpoB* background must

be inherently different, either because the set of beneficial mutations differs or because the same set of beneficial mutations has a different magnitude of effect on fitness (i.e. magnitude epistasis). Although a greater variety of genes was hit with mutations in the +P518 background, this result is again confounded by biased sampling from the $D = 20,000$ treatment, hence I cannot conclude that the set of beneficial mutations is different. Some support is evident for magnitude epistasis: of the two unique PA2449 mutations that fixed in +P518, both had lower fitness than the average of PA2449 mutations in S536F (although, clonal interference in S536F lineages could also enable it to fix the 'best' PA2449 mutations, due to higher mutational supply).

Characteristics of sequenced mutations

Contrary to expectation, no mutations fixed in the genes known to compensate for rifampicin resistance (i.e. *rpoABC*). Instead, the sequencing revealed mutations in two genes in both *rpoB* backgrounds (PA2449 and PA1975), and in a smattering of other genes in +P518. The two common genes are believed to be involved in the uptake or metabolism of amino acids.

Mutations in gene PA2449 occurred with the greatest frequency in both genetic backgrounds, and were associated with the largest gains in fitness. Sequence similarity suggests that PA2449 is an RpoN/σ^{54} enhancer binding protein (EBP) involved in the regulation of expression of several pathways, including phenazine biosynthesis, quorum sensing, and the enzymes for L-serine and L-glycine metabolism (Figure 5.5, Lundgren *et al.* 2013). PA2449 has high protein sequence similarity (41%) to another *P. aeruginosa* EBP, PhhR, involved in L-tyrosine and L-phenylalanine metabolism (Palmer *et al.* 2010), including the highly conserved central RpoN interaction domain. Bacterial EBPs are required for the activation of RpoN-dependent transcription: RpoN directs RNAP to bind

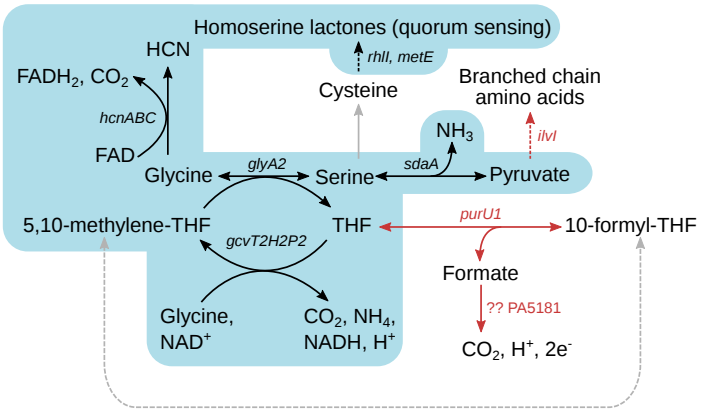

Figure 5.5: Simplified L-serine metabolic pathway. Genes under suggested PA2449 control are enclosed in blue. Other genes with sequenced mutations shown in pink. Dotted arrow indicates multiple reaction steps. Pathway data from KEGG (Kanehisa and Goto 2000; Kanehisa *et al.* 2014) and Lundgren *et al.* (2013).

to the promoter sequence, forming an energetically-stable closed complex (RNAP and double-stranded DNA); interactions with the EBP alter the conformation of the RpoN/RNAP holoenzyme into an open complex (RNAP and antisense strand of DNA) to permit transcription (Bush and Dixon 2012).

The function of PA2449 as an EBP, and the sites of the mutations that fixed, suggests sign or magnitude epistasis with *rpoB* is a real possibility. Of the 13 PA2449 mutations fixed, two were in the RpoN interaction domain and six were within a 24-residue region upstream (Figure 5.6). To speculate on the mechanism of interaction, the insertion of a bulky proline residue at 518 could disrupt the interaction between PA2449 and RpoB and thus reduce the beneficial effect of the mutations. However, no structure for PA2449 is available, so this speculation cannot yet be confirmed. Another possibility is that the beneficial effect of PA2449 mutations depends on transcription rate. S536F is known to have a similar transcription rate to the PA01 wild-type, which is unusual for

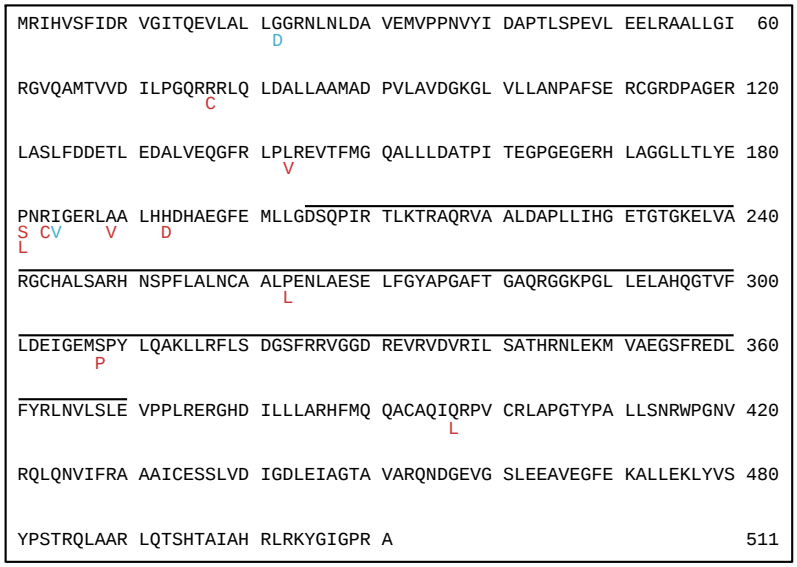

Figure 5.6: PA2449 protein, showing amino acid substitutions (S536F background in red, +P518 in blue). Solid bar shows RpoN interaction domain.

rpoB mutants (Qi Qin, unpublished data). As PA2449 knockouts are incapable of growth on L-glycine or L-serine (Lundgren *et al.* 2013), it is unlikely that the mutations found in this study are loss of function mutations. The trade-off between fitness in L-serine and L-glycine suggests that the mutations alter expression of genes in the pathway differently, rather than increasing expression across the board. To clarify the role of PA2449 mutations, I will be analyzing 60 whole genome sequences from the long-term selection experiment (which was also conducted at $D = 200$) in chapter 4. With the additional sequencing data, I plan to investigate the transcription of genes for the enzymes in the L-serine metabolism pathway using qPCR and RNA Seq, to better understand the function of these mutations. I will then make allelic replacements of PA2449 mutations into each *rpoB* background, as a more rigorous test of epistasis.

The second gene mutated in both *rpoB* backgrounds was PA1975. PA1975 is predicted to contain a FIST-N sensory domain, which may bind amino acid ligands for transport into the cell (Borziak and Zhulin 2007). The three mutations

found in this gene fall within this domain. Genes containing FIST-N domains are often found adjacent to amino acid metabolism and transport genes; PA1974 is a predicted porin, and could also be related to the uptake of L-serine. These genes should be considered candidates for further study into amino acid uptake and metabolism. PA1975 mutations were not associated with the same large fitness gains as PA2449, however, and too few mutations were found to test for epistasis.

Although none of the other mutated genes occurred in both genetic backgrounds, possible roles in L-serine metabolism can be speculated for three other genes that were mutated only in +P518, as they link into the PA2449-regulated pathway (Figure 5.5). PA4696 (*ilvI*, acetolacate synthase large subunit) catalyses the first step in the synthesis of branched-chain amino acids from pyruvate. Elevated expression of PA4696 has previously been found in a cystic fibrosis lung-adapted *P. aeruginosa* (Hoboth *et al.* 2009). However, the PA4696 mutant had a modest (5%) fitness benefit in M9-serine and M9-glucose, but was deleterious in M9-glycine, where the ability to synthesize amino acids would also seem to be advantageous. Mutations in genes PA4314 (*purU1*) and PA5181 may help to reduce concentrations of the by-products of L-serine catabolism. *purU1* converts 10-formyl-THF into THF and formate (the anion of formic acid). PA5181 is an oxidoreductase, and based on homology is likely a formate dehydrogenase, breaking formate into CO_2, a proton and two electrons. No other connections to L-serine metabolism could be speculated for the other genes.

Mutational pleiotropy

By the genes hit with mutations, adaptation, rather than compensation, was the route to higher fitness followed by the *rpoB* mutants. Supporting this, I found little evidence of increased fitness across multiple environments, suggesting that

the mutations did not directly compensate for the faulty *rpoB*. Perhaps due to its role as a regulatory gene, mutations in PA2449 exhibited a greater degree of pleiotropy than those in other genes, conferring a weak benefit on the related carbon source M9-glycine, but being strongly deleterious in M9-glucose. A strong negative rank-correlation suggests that there may be a trade-off between improved ability to consume L-serine and L-glycine, potentially by altering the expression of the individual enzymes making up the metabolic pathway differently. The absence of a benefit on M9-glucose and M9KB suggests that these mutations do not compensate for the *rpoB* defect. In contrast, the mutations in other genes tended to have no appreciable benefit in M9-glycine, but were on average less costly in the other environments. None of these mutations was unconditionally beneficial in M9-glucose or M9KB, suggesting they were also adaptive, not compensatory (although for most, no role in L-serine metabolism could be identified). A similar result was previously observed for *E. coli rpoB* mutants adapted to a complex undefined medium, where fitness increased in the selective environment but not across other environments (Hall 2013).

Conclusions and future directions

The results demonstrate that *rpoB* mutants can adapt to environments in different ways. As amino acids are abundant in cystic fibrosis sputum (Palmer *et al.* 2007), and most *P. aeruginosa* infections are environmentally-acquired (Burns *et al.* 2001; Speert *et al.* 2002), evolvability to amino acids may play a critical role in the ability to establish a chronic infection. This has been previously identified in chronic *P. aeruginosa* infections, which showed an increase in transcription of genes for amino acid and fatty acid metabolism (Hoboth *et al.* 2009). Whether the ability to catabolize L-serine matters for *P. aeruginosa* pathogenicity has yet to be determined. However, in *Campylobacter jejuni*, mutants defective in

L-serine metabolism were incapable of establishing infection in an animal model (Velayudhan *et al.* 2004), suggesting a reliance on L-serine is not out of the question. L-serine represents about 10% of the available amino acid pool in CF sputum (Palmer *et al.* 2007), and mutants incapable of consuming their preferred amino acid L-alanine increase their consumption of L-serine (Boulette *et al.* 2009), thus L-serine metabolism could be potentially relevant to *P. aeruginosa* infections.

The results of this chapter inspire two further research questions. The first is on the nature of genes mutated during adaptive evolution. Are the greatest gains in fitness made through structural mutations to metabolic enzymes, or through regulatory genes, like PA2449? Although it is generally easier to increase the expression of an enzyme, rather than improve its function, does this tactic lead to the greatest gains in fitness? The pleiotropy of PA2449 mutations demonstrated that altering regulatory genes can have deleterious knock-on effects in other environments, suggesting that adaptation to a relatively static environment might involve regulatory genes and adaptation to a fluctuating environment might involve structural mutations to metabolic enzymes. This could be tested by adapting populations to multiple fluctuating environments and using whole-genome sequencing to look for changes. The second question concerns whether a mutation ever be truly compensatory, in that it raises fitness unconditionally in all antibiotic-free environments. Compensation has been assumed to be mode of adaptation for antibiotic resistant strains (e.g. Lenski *et al.* 1998; Reynolds 2000; Barrick *et al.* 2010), has only been conclusively demonstrated in a handful of studies (Björkman *et al.* 2000; Poon and Chao 2005). Antibiotic resistance mutations, such as those for rifampicin resistance, can change bacterial metabolism and gene expression in dramatic and unpredictable ways. Although the classic rifampicin resistance-compensating mutations are thought to be found

in *rpoA*, *rpoB*, and *rpoC*, it remains to be shown whether these mutations increase fitness in many environments, or only the ones they were selected in.

6 | Here's to the losers: Evolvable residents force invaders to adapt

Abstract

Classical population genetics models suggest that high fitness genotypes should dominate the evolutionary dynamics of asexual populations. However, contemporary research is demonstrating that the long-term evolvability of genotypes also plays an important role in determining the outcome of adaptation, because genotypes with low fitness but high evolvability can replace contending lineages with high fitness and low evolvability. Here, I investigated how the evolvability of an initially abundant and low fitness resident influences the evolution of an invader with higher fitness. To accomplish this, I experimentally invaded *P. aeruginosa* PA01 into rifampicin-resistant populations bearing high fitness costs, and measured fitness at time of fixation. Invader fitness was greater when invading into resident populations with high evolvability than those with low evolvability. I investigate two mechanisms for the increased adaptation. First, clonal interference from large-effect beneficial mutations in the highly-evolvable resident could out-compete small-effect beneficial mutations arising in the invader. Second, the rapid adaptation of the highly-evolvable resident could lead to environmental degradation, which in turn could accelerate the spread of beneficial mutations in the invader. The results suggest that invasion

models that treat the resident population as an unchanging parameter are naïeve. Beyond ecology, this also has implications for predicting evolutionary outcomes in mixed microbial communities, which is particularly relevant to chronic clinical infections.

6.1 Introduction

The classic model of evolution in asexual populations is periodic selection, characterized by infrequent sweeps of beneficial mutations in otherwise clonal populations. However, recent work has cast doubt on this paradigm. Multiple beneficial mutations are known to occur in different lineages and must compete for fixation, the process known as clonal interference (De Visser and Rozen 2006), which allows only the fittest mutations to fix. Next-generation sequencing of entire adapting populations has also demonstrated extensive hitchhiking of mutations fixing together in the same genetic background (Lang *et al.* 2013). Under these conditions, a key determinant of the evolutionary fate of competing lineages is their 'evolvability', defined as the extent to which they can adapt to a new environment (Díaz Arenas and Cooper 2013). Recent results have shown that a strain with lower initial fitness but greater evolvability can become the eventual 'winner' of an evolutionary contest between competing strains (Woods *et al.* 2011), drawing into question the validity of established population genetics models that suggest the fittest strain should always win (Patwa and Wahl 2008).

Although evolvability is known to be important for the lineage that ultimately wins, little attention has been paid to the 'losers' of adaptation. Invasive species are a classic eco-evolutionary example involving losers, where an invader that enters a resident population at low frequency may drive the resident extinct. Previous work has shown that invader and resident may interact to drive each other's evolution, and that the extent to which the resident can adapt determines

its invasibility (Sakai *et al.* 2001). In chapter 4, I invaded a sensitive lineage into populations containing one of two rifampicin-resistant *rpoB* mutants. Due to the high cost of rifampicin resistance, the invader ought to have rapidly swept to fixation, but this did not occur (Figure 4.3). As the *rpoB* mutants could adapt to equal or surpass the fitness of the invader (Figure 4.6), and the invader was more successful in low-evolvability populations (Figure 4.3), it is likely that the invaders also needed to adapt in order to fix, despite starting at higher fitness.

Here, I ask whether resident evolvability influenced the fitness of the successful invaders, and propose two mechanisms through which this could occur. The first mechanism is via clonal interference[1] from the adapting resident. A resident that produces offspring that surpass invader fitness will prevent wimpy invader offspring from fixing. I evaluated this mechanism by measuring the fitness of invaders that fixed in both *rpoB* backgrounds and comparing the distributions of their competitive fitness. The second mechanism is via environmental degradation, where the adapting resident depletes resources more rapidly, increasing selection for mutations that improve the trait. As chapter 5 demonstrated that adaptation was most likely through increased L-serine consumption, resource depletion is the likely mechanism of environmental degradation. For this mechanism to work, several factors have to be determined: (1) does environmental quality affect growth rate of the ancestor, (2) does the presence of an adapted resident have the same effect, and (3) does environmental degradation increase the strength of selection on beneficial mutations (Agrawal and Whitlock 2010).

[1] 'Competitor interference' would be more apt, as the invader and residents are not necessarily clones, but I am recycling the term for convenience.

6.2 Materials and methods

Strains and growth conditions

The strains used were the same as for chapter 4 (rifampicin-sensitive PA01 and GFP-PA01, and rifapmicin-resistant S536F and +P518), and an additional rifampicin-sensitive luciferase-tagged PA01 strain (lux-PA01, with insertion of *luxCDABE* operon and gentamicin resistance using the mini-Tn7 insertion protocol Choi and Schweizer 2006). Using the terminology developed in the introduction, GFP-PA01 is the 'invader' and the two *rpoB* mutants are the 'residents' (RS: S536F and RP: +P518). Lux-PA01 is used to test the effects of different competitors on growth rate because the lux marker allows more sensitive determination of growth rate in mixed populations, but it is otherwise equivalent to GFP-PA01.

Selection experiment

The selection phase of the experiment was described in subsection 4.2. To recapitulate, individual cells ($n < 4$) of the invader were introduced into resident populations ($N_0 = 5.25 \times 10^4$). Presence of the invader was tracked until it fixed or was lost from the population (frequency $\geq 99.975\%$ or $< 0.0005\%$). Invaders were stored at -80 °C in glycerol stocks on the day they were recorded as fixed.

Competitive fitness assays

From glycerol stocks (20 each from invaders fixing in RS and RP), I first isolated an individual colony on agar medium. Colonies were inoculated into 200 µl M9-serine and were grown for 24 hours at 37°C to enable the cells to physiologically adapt. I then diluted the culture 1 µl into 200 µl M9-serine and grew for a further 24 hours at 37 °C. Following the second growth period, I mixed the

evolved strains with the non-adapted genotype of the lineage they co-adapted with, at a 20% to 80% volume ratio. The mixtures were diluted 1 µl into 200 µl of M9-serine and grown for 24 hours at 37 °C. Competitive fitness was measured using the flow cytometry method (subsection 2.3).

Clonal interference

To test for increased clonal interference, I compared the distributions of fitnesses of invader lineages that had fixed in the presence of the residents. Assuming a gamma distribution (as predicted by Rozen *et al.* 2002), I used the invader fitness data to fit rate and shape parameters. Three characteristics of the fitness distribution indicate greater clonal interference: an increase in mean fitness, a right-shifted lower bound, and a decrease in distribution skew (Gerrish and Lenski 1998; Rozen *et al.* 2002). The shift in the lower bound and the change in shape arise because the fitness distributions of adapting invader and resident do not completely overlap (conceptually depicted in Figure 6.1). A large degree of overlap prevents low-fitness invader offspring from fixing, and as a result the invaders that fix are disproportionately-sampled from the upper portion of the contending distribution, decreasing skew. In contrast, a small degree of overlap between the distributions allows more sampling from the left-hand side of the contending distribution, resulting in a relatively unchanged fixed distribution. This contradicts the findings of Rozen *et al.* (2002), which predicted little change in shape with increased clonal interference, but that work assumed that the contending distribution was the same for all individuals (i.e. perfect overlap), which is true for clonal populations but not true for a higher-fitness invader entering a population at low initial frequency.

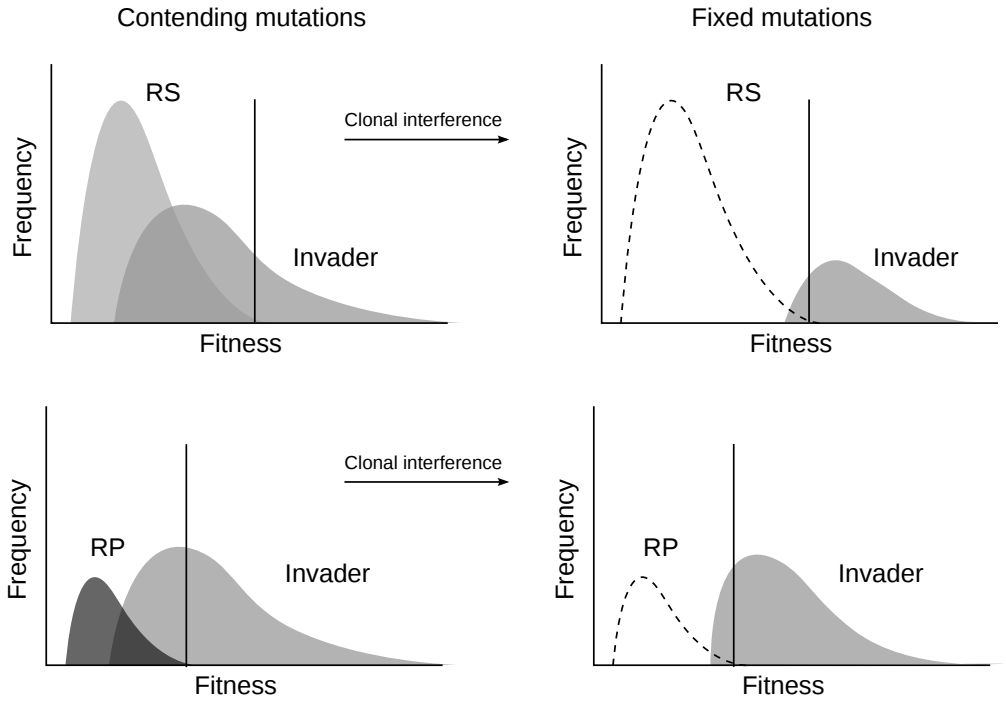

Figure 6.1: Conceptual explanation effect of clonal interference on the fitness distribution of fixed invaders. The RS distribution overlaps the invader distribution to a greater extent than the RP, causing a higher probability of loss of small-effect mutations in the invader background.

Environmental degradation

To determine whether adapting residents reduced environmental quality, I first determined what effect a degraded environment had on the lux-PA01 tester strain. I experimentally degraded the environment by reducing the concentration of L-serine and measured the maximum growth rate of the lux-PA01 tester strain ($v_{\max}$) in the presence of the ancestral resident (20% tester to 80% resident). I chose luminescence as a marker for measuring the growth rate of a single strain in a mixed population because there is little signal interference from the resident. I manipulated environmental degradation by reducing the L-serine concentration to 75%, 50%, 25%, and 10% of the selection environment in

200 µl medium. To enhance signal detection, I used a white opaque 96-well microtitre plate. Total luminescence (captured in 3 second intervals for each well) was measured every 5 minutes for the first 8 h, then every 30 minutes for a further 16 h using the FLUOstar OPTIMA (BMG Labtech, Offenburg, Germany). Maximum growth rate was estimated by fitting Gompertz growth curves to log-transformed luminescence using the 'grofit' package (Kahm *et al.* 2010) in R 3.0.1 (R Core Team 2013). Gompertz growth curves provided a better fit than logistic curves, but fitting either returned similar parameter estimates, as previously demonstrated by Shi and Xia (2003). I transformed L-serine concentration as the measure of environmental degradation, $E = [1/(\%\text{ L-serine}) - 1]/10$ to linearize the relationship with growth rate (the constant divisor is to put E on a convenient scale). As a consequence, degradation ranges from 0 (a pristine environment) to infinity (completely devoid of resources). Using this transformation, I calculated linear regression coefficients from the model

$$v_{\max} = a_E + b_E I_R + c_E E, \tag{6.1}$$

where E is environmental degradation, a_E is the global intercept, b_E is the resident effect on the intercept, c_E is the global slope of E on $v_{\max}$, and I_R is an indicator variable for resident (1 for RS, 0 for RP). Following these results, growth rate of the tester strain in the presence of the adapted resident was then used to calculate the degree of environmental degradation. The tester was mixed together with evolved residents (20% tester to 80% resident) in the 100% concentrated selection environment. Growth rate was again measured using the above method. Effective environmental degradation was calculated using the regression coefficients calculated above. To determine the effect of environmental degradation on fitness, I measured the competitive fitness of the successful invaders in a pristine ($E = 0$) and degraded ($E = 0.9$) environment, with the same method as in subsection 6.2.

6.3 Results

Resident evolvability increases adaptation of invader

To determine if highly evolvable residents can accelerate the evolution of a invader strain, I measured the fitness of invaders that fixed in the presence of either resident. The average time to fixation indicates that invader fixation was slowed by clonal interference from adapting resident offspring: the ancestral invader should have fixed within 15 generations, but the average time to fixation was 110.8 ± 8.4 SE generations in RS populations and 157.8 ± 6.9 SE generations for RP (see chapter 4). The fitness of invaders that fixed was significantly increased by the presence of RS over RP (1.71 ± 0.064 SE versus 1.41 ± 0.126 SE: two-sample t-test $t_{26.9} = -2.1$, $p = 0.048$). Although average time to fixation differed in the two residents, the number of generations until the fixation of the invader was only a marginally-significant co-variate for invader fitness (Figure 6.2A; ANCOVA, resident genotype: $F_{1,36} = 8.41$, $p = 0.006$, generations until fixation: $F_{1,36} = 3.67$, $p = 0.06$).

Putative mechanism behind accelerated evolution

To determine why resident evolvability increased adaptation in the invader, I considered two mechanisms: clonal interference from superior beneficial mutations in the RS resident, and environmental degradation caused by adaptation to the environment. To test for increased clonal interference, I compared the distributions of fitnesses of invader lineages that had fixed in the presence of the residents. The fitness distributions of the invader lineages in the presence of the two residents were significantly different (Kolmogorov-Smirnov test

[2]This is not a distribution of fitness effects of single mutations, as fixed invaders may possess more than one beneficial mutation.

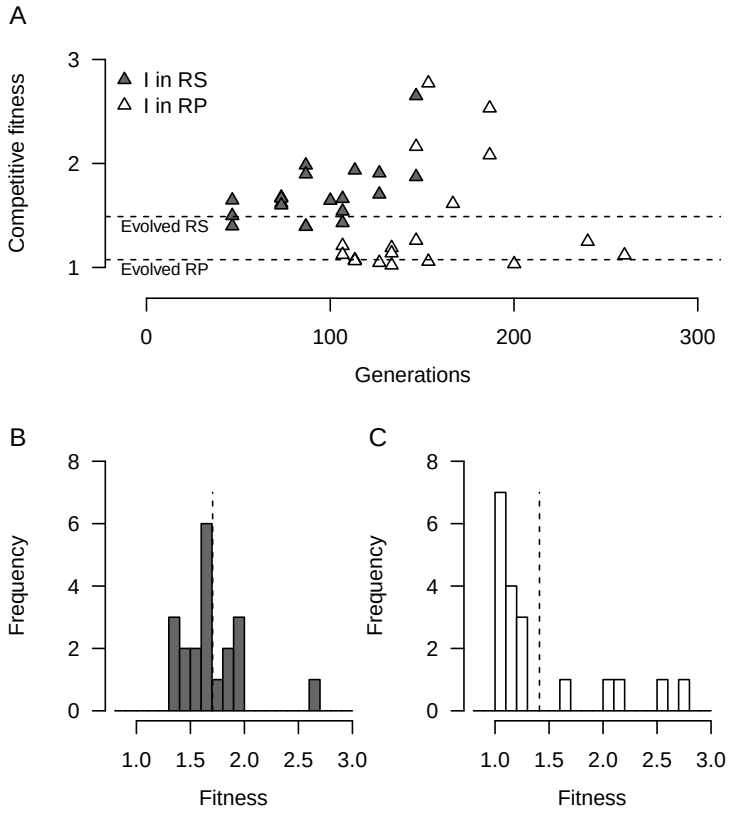

Figure 6.2: Evolvability of competitors forces increased adaptation. A. Fitness of successful invaders (I) was increased by resident evolvability (RS: resident S536F, RP: resident +P518). The skew of the fitness distribution[2] differed between the two residents; B. RS and C. RP. Vertical dashed bar is mean.

$D = 0.74$, $p < 0.0001$; Figure 6.2B and C). I estimated shape and rate parameters for the two distributions by fitting them to a gamma distribution (invader in RS: shape $= 41.80 \pm 13.2$ SE, rate $= 24.51 \pm 7.76$ SE; invader in RP: shape $= 8.92 \pm 2.84$ SE, rate $= 6.32 \pm 2.07$ SE). The shape parameter of the invader in RS distribution was significantly larger, indicating greater symmetry (Welch's t-test: $t_{20.3} = 2.44$, $p = 0.027$). The minimum, or left-hand bound, of each distribution also corresponded to the maximum fitness gain realized by the resident adapted in pure culture. To test for the effect of environmental degradation on growth rate, I performed a control experiment observing the

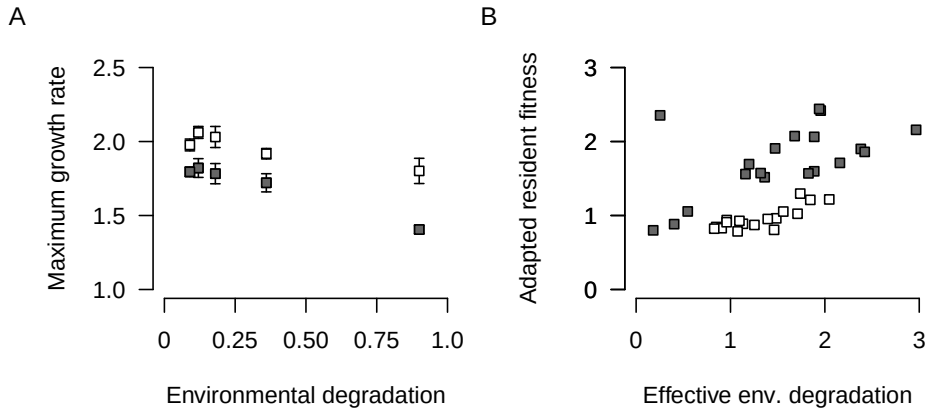

Figure 6.3: A. Growth rate of the tester strain correlates with manipulated environmental quality, when grown in the presence of the ancestral resident. B. Effective environmental degradation calculated from regression coefficients (see text). Grey: tester in RS, white: tester in RP.

effect of manipulated environmental degradation on growth rate of the lux-tagged tester strain in the presence of the ancestral residents. Degradation was statistically associated with decreased maximum growth rate and resident genotype (Figure 6.3A, ANCOVA, degradation: $F_{1,7} = 7.6$, $p = 0.03$, resident: $F_{1,7} = 14.5$, $p = 0.007$). Presence of an adapted resident had a similar effect as environmental degradation on the tester strain's growth rate (linear regression: intercept $= 1.90$, slope $= -0.43$, RS resident $= 0.01$, $F_{2,40} = 24.3$,$p < 0.0001$, Adjusted $R^2 = 0.53$). With the estimated regression coefficients, I calculated the effective environmental degradation caused by the adapted resident using Equation 6.1. Although average effective environmental degradation was higher in the presence of evolved RS residents (RS: 1.52 ± 0.12 vs. RP: 1.31 ± 0.06), the effect was not significant (Wilcoxon rank-sum: $W = 118$, $p = 0.18$). The effective degradation caused by adapted RS had higher variance ($F_{16,18} = 0.24$, $p = 0.006$).

For environmental degradation to lead to increased invader adaptation, it

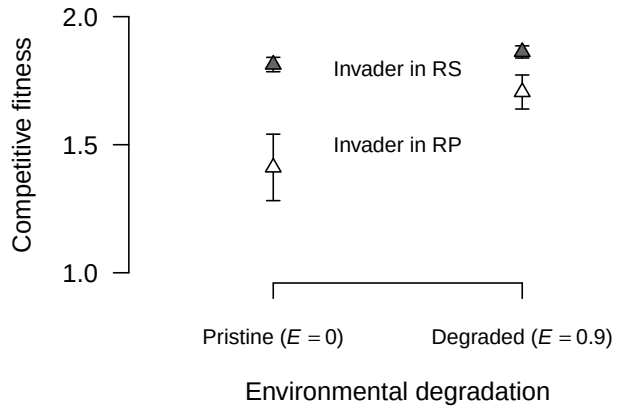

Figure 6.4: Fitness of fixed invaders in pristine (no degradation) and degraded environments.

should increase the strength of selection (Agrawal and Whitlock 2010). To test this, I measured the competitive fitness of the evolved invader lineages against the ancestral residents in an artificially degraded environment. I found that fitness of the evolved invader strains was significantly greater in the degraded environment than the selection environment (Figure 6.4, Type II ANOVA, competitor genotype: $F_{1,76} = 4.32$, $p = 0.041$, environment: $F_{1,76} = 43.39$, $p < 0.0001$).

6.4 Discussion

Invasive species models have investigated effect of life history traits on invasion success (e.g. propagule number, invader fitness (Sakai *et al.* 2001; Hayes and Barry 2008), yet few, if any, consider the role of resident evolvability in invasion dynamics. Although chapter 4 demonstrated that resident evolvability helps to resist invasion, this chapter suggests that the invaders that successfully invade a highly-evolvable resident are ultimately better adapted to their environment. The data suggest that clonal interference is primarily responsible for the increase in invader adaptation. Although previous results have demonstrated the importance

of evolvability in experimental evolution (Woods *et al.* 2011), our results are the first, to our knowledge, to demonstrate that the evolvability of the lineage lost to selection can influence the outcome of adaptation.

Evidence for clonal interference comes from the fitness distribution of the fixed invader strains. As shown in Figure 6.1, differences in the amount of overlap of the mutant fitness distributions between the invader and resident leads to increased sampling from the right-hand side of the distribution in invaders invading RS. Theory (e.g. Rozen *et al.* 2002; Park and Krug 2007) and data (e.g. Miralles *et al.* 1999; De Visser and Rozen 2006) have previously demonstrated that clonal interference can lead to increased adaptation, but this is the first evidence that clonal interference from competitors that are initially lower in fitness also matters. The data from Figure 6.2A was the most compelling evidence of clonal interference: fixed invaders needed to surpass the adapting residents' fitness, which lead to more adaptation in invaders invading RS populations. At this stage, it is not possible to say whether the difference in invader adaptation was due to multiple beneficial mutations, or single mutations with larger effect. Whole-genome sequence data from fixed invaders would be needed to answer this question; an analysis of 30 fixed invaders (15 fixed in the presence of each resident) is currently under way to address this. It is also important to note that successful invaders are unlikely to have all reached their maximum potential fitness in L-serine. Although clonal interference from the RS resident was stronger than from the RP resident, it is likely weaker than the clonal interference the invader would experience if it had adapted to the environment on its own. In predicting how resident evolvability would affect the shape of the fitness distribution of fixed invaders, I have neglected any clonal interference effect from the presence of multiple adapted invader lineages. I believe this is reasonable because the invader was introduced at low frequency

(probabilistically less than four cells, subsection 4.2) and at any rate, the effect should be independent of resident genotype.

Environmental degradation is thought to be an important driver of adaptive evolution. The idea that adapting competitors change each other's environment is not new: Maynard Smith (1976) proposed the idea in a definition of the Red Queen effect for co-evolution: 'each evolutionary advance made by one species in an ecosystem is experienced as a degradation of the environment by other species, and consequently if a species is to survive it must evolve continuously and rapidly'. Environmental stress has previously been shown to be important for invasion success in natural populations (Gilchrist and Lee 2007). There was some evidence that adapting residents caused environmental degradation, but the amount was not statistically different between RS and RP (Figure 6.3). However, environmental degradation caused by the adapted RS resident was higher, and had significantly higher variance, so environmental degradation cannot be completely ruled out as a mechanism. Further, environmental degradation was associated with a higher strength of selection (Figure 6.4), suggesting that environmental degradation could lead to a faster spread of beneficial mutations. Curiously, this would suggest that fitness in a pristine environment could be maximized more rapidly by adapting to a degraded version of the environment; this was also found for a structured-population evolutionary rescue model (Uecker *et al.* 2014). However, degraded environments could also select for other mutations, so mean fitness across multiple lineages may not increase. This effect is likely to occur for fairly simple environments (e.g. for increasing salt concentration, Bell and Gonzalez 2011) but further experimental tests would be needed to determine its generality. A caveat, however, is that these results are merely correlational; a proper test of the role of environmental degradation would be to manipulate this variable directly,

and allow invaders to adapt. Here, I assumed that nutrient depletion was the greatest source of degradation. This assumption is based on the sequencing data from chapter 5, which showed that adaptation was generally due to a mutation in a regulatory gene, PA2449, involved in L-serine metabolism. Although generally the accumulation of waste products and active toxification are important determinants of environmental quality, they were unlikely to be an important factor in this experiment. Although L-serine catabolism produces NH_3 as a waste product (Hofmeister and Buckler 1993), M9 media are phosphate-buffered. Even if all L-serine was consumed, most NH_3 would be buffered to the less-toxic ammonium ion (NH_4+), with only 3% (or 1×10^3 mol/l) remaining as NH_3 (given the pH and temperature conditions, Emerson *et al.* 1975). Active toxification is a possibility, as *P. aeruginosa* is notorious for producing a wide variety of anti-competitor compounds, including bacteriocins, and phenazines, which are regulated by PA2449. However, the killing efficacy of these compounds declines with strain relatedness (Inglis *et al.* 2011; Bucci *et al.* 2011; Schoustra *et al.* 2012), making active toxification less likely to play a role, as the invader and residents share a common genetic background.

These results argue against classic adaptive dynamics models that only consider the interactions between mutants that are currently present in a population–the interaction of mutants with potential future competitors must also be considered. Likewise, predictions based on a fixed environment are unlikely to be broadly useful, as adaptation can lead to a degradation of environmental quality that can drive increases in selection pressure. Ultimately, the interactions between communities and their dynamic environment are what determine evolutionary outcomes. These findings are especially relevant to predicting the evolutionary course of chronic bacterial infections, such as in the lungs of individuals with cystic fibrosis, where mixed communities are likely and

mutators with potentially increased evolvability are commonly found (Harrison 2007; Oliver *et al.* 2000).

Conclusions

A theme that has emerged organically from this thesis is that theoretical models of adaptive evolution in their current form are too naïve to be broadly useful for predicting the adaptive dynamics of resistant populations. In chapter 3, I demonstrated that migration between different environments can negate the costs of resistance, and may be a contributing factor in the maintenance of antibiotic resistant reservoir populations. In chapter 4, I demonstrated that the fixation of a reversion mutation is not guaranteed, even if it does occur, due to the effects of genetic drift and clonal interference from adaptation in the resistant strain. This challenges the notion that the rarity of reversions is alone responsible for the maintenance of antibiotic resistance. Although not initially a focus of the chapter, I also found that resistant genotypes with the same initial cost of resistance could differ in their ability to adapt to their environment. Following up on this result, in chapter 5, I investigated the genetic basis for adaptation to an environment where L-serine was the only carbon source, and found that magnitude epistasis caused by different *rpoB* mutations was the most likely constraint on evolvability. This challenges recent theoretical models (Martin and Lenormand 2006) and experiments (Barrick *et al.* 2010; Gifford *et al.* 2011) of adaptive evolution predicting that initial fitness determines the magnitude of the effect of new beneficial mutations. The gains in fitness were made not through general compensatory mutations, but to a regulatory gene specific to the environment; this was surprising, given the initially large

cost of resistance, compensatory mutations were expected to provide a large benefit. Previous experimental evolution studies of the recovery of fitness of resistant strains have assumed that resistant strains compensate for the costs of resistance, rather than adapt, and a common criticism of many such papers is that compensatory adaptation is fundamentally different from 'real adaptation'. My results suggest that adaptation should not be discounted for resistant strains, even if the costs of resistance are significant. Finally, in chapter 6, I demonstrated that the evolvability of a resident, not just its initial fitness, can influence the fitness of invaders that successfully invade a population. Despite an initial fitness advantage on the part of the invader, clonal interference from the resistant population was sufficient to force invaders adapt to the environment, rather than relying on their initial advantage. Although recent studies have demonstrated how evolvability can determine the 'winners' of natural selection, these results are the first to demosntrate that the evovlability of evolutionary 'losers' matters as well.

Together, these results suggest that a 'black-box' approach to predicting adaptive evolution is unlikely to capture all the important details. A better understanding of how genotypes, environments, and communities interact is needed to build truly predictive models of adaptive evolution. Fortunately, advances in high-throughput genomics and fitness assays are beginning to allow manipulation of all of these variables in an experimental context. The results of chapter 5 in particular demonstrate where the empirical data is lacking. Although many experiments have claimed evidence of compensation, we actually have little understanding of the genetic basis for the recovery of fitness in resistant populations. Are there compensatory mutations that can compensate in all environments? Further work is also needed to determine whether compensation is more likely than adaptation, and indeed whether the two processes should

be treated differently at all. Finally, high-throughput genomics will allow us to determine whether any general trends exist for adaptation, specifically, whether adaptation proceeds through changes to regulatory genes or structural changes to metabolic enzymes. The results of chapter chapter 5 suggest that changes to regulatory genes can have large effects on fitness, but may also cause greater pleiotropy.

An alternate theme for this thesis is 'many secondary mechanisms help to maintain the prevalence of antibiotic resistance'. It has clearly emerged that the reversion of resistance is even less likely than originally predicted from theoretical models. We cannot necessarily rely on either reversion mutations or the migration of sensitive strains from source populations to supplant resistant populations once they have become established (chapter 4). Migration between different environments may reduce or eliminate the costs of resistance (chapter 3). Even when resistance is initially costly, adaptation can allow resistant strains to overcome the costs (chapter 5), which can force sensitive strains to wait for beneficial mutations in order to successfully invade (chapter 6). This increases the likelihood that the evolution of resistance will eventually render all current antibiotics ineffective.

References

Agrawal, A. F., and M. C. Whitlock, 2010 Environmental duress and epistasis: how does stress affect the strength of selection on new mutations? Trends in ecology & evolution **25**: 450–458.

Aminov, R., 2010 A brief history of the antibiotic era: lessons learned and challenges for the future. Frontiers in Microbiology **1**: doi: 10.3389/fmicb.2010.00134.

Andersson, D. I., 2003 Persistence of antibiotic resistant bacteria. Current Opinion in Microbiology **6**: 452–456.

Andersson, D. I., 2006 The biological cost of mutational antibiotic resistance: any practical conclusions? Current Opinion in Microbiology **9**: 461–465.

Andersson, D. I., and D. Hughes, 2010 Antibiotic resistance and its cost: is it possible to reverse resistance? Nature Reviews Microbiology **8**: 260–271.

Andersson, D. I., and D. Hughes, 2011 Persistence of antibiotic resistance in bacterial populations. FEMS Microbiology Reviews **35**: 901–911.

Andersson, D. I., and B. R. Levin, 1999 The biological cost of antibiotic resistance. Current Opinion in Microbiology **2**: 489–493.

Arason, V. A., A. Gunnlaugsson, J. A. Sigurdsson, H. Erlendsdottir, S. Gudmundsson, *et al.*, 2002 Clonal spread of resistant pneumococci despite diminished antimicrobial use. Microbial Drug Resistance **8**: 187–192.

Aronson, J., 2004 When I use a word.... The British Journal of General Practice **54**: 559.

Baquero, F., J.-L. Martínez, and R. Cantón, 2008 Antibiotics and antibiotic resistance in water environments. Current opinion in biotechnology **19**: 260–265.

Barrick, J., M. Kauth, C. Strelioff, and R. Lenski, 2010 *Escherichia coli rpoB* mutants have increased evolvability in proportion to their fitness defects. Molecular biology and evolution **27**: 1338–1347.

Baysarowich, J., K. Koteva, D. W. Hughes, L. Ejim, E. Griffiths, *et al.*, 2008 Rifamycin antibiotic resistance by ADP-ribosylation: structure and diversity of Arr. Proceedings of the National Academy of Sciences **105**: 4886–4891.

Bell, G., 2008 *Selection: the mechanism of evolution*. Oxford University Press, Oxford, UK, 2nd edition.

Bell, G., and A. Gonzalez, 2011 Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. Science **332**: 1327–1330.

Bergman, M., S. Huikko, M. Pihlajamäki, P. Laippala, E. Palva, *et al.*, 2004 Effect of macrolide consumption on erythromycin resistance in *Streptococcus pyogenes* in Finland in 1997–2001. Clinical infectious diseases **38**: 1251–1256.

Björkman, J., I. Nagaev, O. Berg, D. Hughes, and D. I. Andersson, 2000 Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. Science **287**: 1479–1482.

Boeva, V., T. Popova, K. Bleakley, P. Chiche, J. Cappo, *et al.*, 2011 Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics **28**: 423–425.

Borziak, K., and I. B. Zhulin, 2007 FIST: a sensory domain for diverse signal transduction pathways in prokaryotes and ubiquitin signaling in eukaryotes. Bioinformatics **23**: 2518–2521.

Boulette, M. L., P. J. Baynham, P. A. Jorth, I. Kukavica-Ibrulj, A. Longoria, *et al.*, 2009 Characterization of alanine catabolism in Pseudomonas aeruginosa and its importance for proliferation in vivo. Journal of bacteriology **191**: 6329–6334.

Brandis, G., M. Wrande, L. Liljas, and D. Hughes, 2012 Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. Molecular Microbiology **85**: 142–151.

Bucci, V., C. D. Nadell, and J. B. Xavier, 2011 The evolution of bacteriocin production in bacterial biofilms. The American Naturalist **178**: E162–E173.

Buckling, A., R. C. Maclean, M. A. Brockhurst, and N. Colegrave, 2009 The *Beagle* in a bottle. Nature **457**: 824–9.

Burns, J. L., R. L. Gibson, S. McNamara, D. Yim, J. Emerson, *et al.*, 2001 Longitudinal assessment of Pseudomonas aeruginosa in young children with cystic fibrosis. Journal of Infectious Diseases **183**: 444–452.

Bush, M., and R. Dixon, 2012 The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. Microbiology and molecular biology reviews **76**: 497–529.

Campbell, E. A., N. Korzheva, A. Mustaev, K. Murakami, S. Nair, *et al.*, 2001 Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell **104**: 901–912.

Chen, K., J. W. Wallis, M. D. McLellan, D. E. Larson, J. M. Kalicki, *et al.*, 2009 BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nature methods **6**: 677–681.

Choi, K.-H., and H. P. Schweizer, 2006 mini-Tn7 insertion in bacteria with single *att*Tn7 sites: example *Pseudomonas aeruginosa*. Nature Protocols **1**: 153–161.

Cingolani, P., A. Platts, M. Coon, T. Nguyen, L. Wang, *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly **6**: 80–92.

Colegrave, N., and A. Buckling, 2005 Microbial experiments on adaptive landscapes. BioEssays **27**: 1167–1173.

Colegrave, N., and S. Collins, 2008 Experimental evolution: experimental evolution and evolvability. Heredity **100**: 464–470.

Comas, I., S. Borrell, A. Roetzer, G. Rose, B. Malla, *et al.*, 2012 Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nature Genetics **44**: 106–110.

Courvalin, P., 1994 Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. Antimicrobial agents and chemotherapy **38**: 1447.

Cox, C. D., and P. Adams, 1985 Siderophore activity of pyoverdin for *Pseudomonas aeruginosa*. Infection and Immunity **48**: 130–138.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, *et al.*, 2011 The variant call format and VCFtools. Bioinformatics **27**: 2156–2158.

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

D'Costa, V., K. McGrann, D. Hughes, and G. Wright, 2006 Sampling the antibiotic resistome. Science **311**: 374–377.

Demerec, M., 1948 Origin of bacterial resistance to antibiotics. Journal of Bacteriology **56**: 63–74.

DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, *et al.*, 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics **43**: 491–498.

Dever, L., and T. Dermody, 1991 Mechanisms of bacterial resistance to antibiotics. Archives of internal medicine **151**: 886–895.

Drenkard, E., and F. Ausubel, 2002 Pseudomonas biofilm formation and antibiotic resistance are linked to phenotypic variation. Nature **416**: 740–743.

Díaz Arenas, C., and T. Cooper, 2013 Mechanisms and selection of evolvability: experimental evidence. FEMS Microbiology Reviews **37**: 572–582.

D'Costa, V. M., C. E. King, L. Kalan, M. Morar, W. W. L. Sung, *et al.*, 2011 Antibiotic resistance is ancient. Nature **477**: 457–461.

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

Emerson, K., R. C. Russo, R. E. Lund, and R. V. Thurston, 1975 Aqueous ammonia equilibrium calculations: effect of pH and temperature. Journal of the Fisheries Research Board of Canada **32**: 2379–2383.

Fisher, R., 1922 On the dominance ratio. Proceedings of the Royal Society of Edinburgh **42**: 321–341.

De Visser, J. A. G. M., and D. E. Rozen, 2006 Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. Genetics **172**: 2093–2100.

De Visser, J. A. G. M., C. W. Zeyl, P. J. Gerrish, J. L. Blanchard, and R. E. Lenski, 1999 Diminishing returns from mutation supply rate in asexual populations. Science **283**: 404–406.

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

Fox, C. W., and J. B. Wolf, editors, 2006 *Evolutionary genetics: Concepts and case studies*. Oxford University Press.

Gerrish, P. J., and R. E. Lenski, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102/103**: 127–144.

Gifford, D. R., S. E. Schoustra, and R. Kassen, 2011 The length of adaptive walks is insensitive to starting fitness in *Aspergillus nidulans*. Evolution **65**: 3070–3078.

Gilchrist, G. W., and C. E. Lee, 2007 All stressed out and nowhere to go: does evolvability limit adaptation in invasive species? Genetica **129**: 127–132.

Gillespie, J. H., and M. Turelli, 1989 Genotype-environment interactions and the maintenance of polygenic variation. Genetics **121**: 129–138.

Govan, J., and V. Deretic, 1996 Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. Microbiological Reviews **60**: 539–574.

Haldane, J. B. S., 1927 A mathematical theory of natural and artificial selection, part V: selection and mutation. Proceedings of the Cambridge Philosophical Society **23**: 838–844.

Hall, A., 2013 Genotype-by-environment interactions due to antibiotic resistance and adaptation in Escherichia coli. Journal of evolutionary biology **26**: 1655–1664.

Hall, A. R., V. F. Griffiths, R. C. MacLean, and N. Colegrave, 2010 Mutational neighbourhood and mutation supply rate constrain adaptation in *Pseudomonas aeruginosa*. Proceedings of the Royal Society B **277**: 643–650.

Hall, A. R., J. C. Iles, and R. C. MacLean, 2011 The fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase. Genetics **187**: 817–822.

Hall, B., and M. Barlow, 2004 Evolution of the serine β-lactamases: past, present and future. Drug resistance updates **7**: 111–123.

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

Harrison, F., 2007 Microbial ecology of the cystic fibrosis lung. Microbiology **153**: 917–923.

Hayes, K. R., and S. C. Barry, 2008 Are there any consistent predictors of invasion success? Biological Invasions **10**: 483–506.

Hoboth, C., R. Hoffmann, A. Eichner, C. Henke, S. Schmoldt, *et al.*, 2009 Dynamics of adaptive microevolution of hypermutable Pseudomonas aeruginosa during chronic pulmonary infection in patients with cystic fibrosis. Journal of Infectious Diseases **200**: 118–130.

Hofmeister, A. E. M., and R. G. D. L. W. Buckler, 1993 L-Serine and L-threonine dehydratase from *Clostridium propionicum* Two enzymes with different prosthetic groups. European Journal of Biochemistry **215**: 341–349.

Høiby, N., T. Bjarnsholt, M. Givskov, S. Molin, and O. Ciofu, 2010 Antibiotic resistance of bacterial biofilms. International Journal of Antimicrobial Agents **35**: 322–332.

Inglis, R. F., P. G. Roberts, A. Gardner, and A. Buckling, 2011 Spite and the scale of competition in *Pseudomonas aeruginosa*. The American Naturalist **178**: 276–285.

Kahm, M., G. Hasenbrink, H. Lichtenberg-Fraté, J. Ludwig, and M. Kschischo, 2010 grofit: fitting biological growth curves with R. Journal of Statistical Software **33**: 1–21.

Kanehisa, M., and S. Goto, 2000 KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research **28**: 27–30.

Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi, *et al.*, 2014 Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic acids research **42**: D199–D205.

Kataja, J., P. Huovinen, M. Skurnik, H. Seppälä, *et al.*, 1999 Erythromycin resistance genes in group A streptococci in Finland. Antimicrobial agents and chemotherapy **43**: 48–52.

Khachatourians, G. G., 1998 Agricultural use of antibiotics and the evolution and transfer of antibiotic-resistant bacteria. Canadian Medical Association Journal **159**: 1129–1136.

Kirschner, M., and J. Gerhart, 1998 Evolvability. Proceedings of the National Academy of Sciences **95**: 8420–8427.

Kümmerer, K., 2009a Antibiotics in the aquatic environment–a review–part I. Chemosphere **75**: 417–434.

Kümmerer, K., 2009b Antibiotics in the aquatic environment–a review–part II. Chemosphere **75**: 435–441.

Lang, G. I., D. P. Rice, M. J. Hickman, E. Sodergren, G. M. Weinstock, *et al.*, 2013 Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature .

Lenski, R., 1998 Bacterial evolution and the cost of antibiotic resistance. International microbiology: the official journal of the Spanish Society for Microbiology **1**: 265–270.

Lenski, R., J. Mongold, P. Sniegowski, M. Travisano, F. Vasi, *et al.*, 1998 Evolution of competitive fitness in experimental populations of E. coli: What makes one genotype a better competitor than another? Antonie Van Leeuwenhoek **73**: 35–47.

Lenski, R. E., M. R. Rose, S. C. Simpson, and S. C. Tadler, 1991 Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. The American Naturalist **138**: 1315–1341.

Leone, F., L. Nelson, and R. Nottingham, 1961 The folded normal distribution. Technometrics **3**: 543–550.

Levin, B. R., M. Lipsitch, V. Perrot, S. Schrag, R. Antia, *et al.*, 1997 The population genetics of antibiotic resistance. Clinical Infectious Diseases **24**: S9–S16.

Levin, B. R., V. Perrot, and N. Walker, 2000 Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics **154**: 985–997.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, *et al.*, 2009 The sequence alignment/map format and SAMtools. Bioinformatics **25**: 2078–2079.

Lundgren, B., W. Thornton, M. Dornan, L. Villegas-Peñaranda, C. Boddy, *et al.*, 2013 The Gene PA2449 is Essential for Glycine Metabolism and Pyocyanin Biosynthesis in *Pseudomonas aeruginosa* PAO1. Journal of bacteriology .

MacLean, R., 2010 Predicting epistasis: an experimental test of metabolic control theory with bacterial transcription and translation. Journal of Evolutionary Biology **23**: 488–493.

MacLean, R., A. Hall, G. Perron, and A. Buckling, 2010 The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. Nature Reviews Genetics **11**: 405–414.

MacLean, R. C., and A. Buckling, 2009 The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. PLoS Genetics **5**: e1000406 doi:10.1371/journal.pgen.1000406.

Maggi, N., C. Pasqualucci, R. Ballotta, and P. Sensi, 1966 Rifampicin: a new orally active rifamycin. Chemotherapy **11**: 285–292.

Maisnier-Patin, S., O. G. Berg, L. Liljas, and D. I. Andersson, 2002 Compensatory adaptation to the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. Molecular Microbiology **46**: 355–366.

Marco-Sola, S., M. Sammeth, R. Guigó, and P. Ribeca, 2012 The GEM mapper: fast, accurate and versatile alignment by filtration. Nature methods **9**: 1185–1188.

Martin, G., and T. Lenormand, 2006 The fitness effect of mutations across environments: A survey in light of landscape fitness models. Evolution **60**: 2413–2427.

Martinez, J. L., 2009 The role of natural environments in the evolution of resistance traits in pathogenic bacteria. Proceedings of the Royal Society B: Biological Sciences **276**: 2521–2530.

Maynard Smith, J., 1976 What determines the rate of evolution? The American Naturalist **110**: 331–338.

Miralles, R., P. J. Gerrish, A. Moya, and S. F. Elena, 1999 Clonal interference and the evolution of RNA viruses. Science **285**: 1745–1747.

Misato, T., K. Ko, and I. Yamaguchi, 1977 Use of antibiotics in agriculture. Advances in Applied Microbiology **21**: 53–88.

Nagaev, I., J. Björkman, D. I. Andersson, and D. Hughes, 2001 Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. Molecular microbiology **40**: 433–439.

Neushul, P., 1993 Science, government and the mass production of penicillin. Journal of the history of medicine and allied sciences **48**: 371–395.

Oliver, A., R. Cantón, P. Campo, F. Baquero, and J. Blázquez, 2000 High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science **288**: 1251–1253.

Orr, H. A., 2009 Fitness and its role in evolutionary genetics. Nature Reviews Genetics **10**: 531–539.

Overbye, K., and J. Barrett, 2005 Antibiotics: Where did we go wrong? Drug Discovery Today **10**: 45–52.

Palmer, G. C., K. L. Palmer, P. A. Jorth, and M. Whiteley, 2010 Characterization of the Pseudomonas aeruginosa transcriptional response to phenylalanine and tyrosine. Journal of bacteriology **192**: 2722–2728.

Palmer, K. L., L. M. Aye, and M. Whiteley, 2007 Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. Journal of bacteriology **189**: 8079–8087.

Park, S.-C., and J. Krug, 2007 Clonal interference in large populations. Proceedings of the National Academy of Sciences **104**: 18135–18140.

Patel, R. K., and M. Jain, 2012 NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One **7**: e30619.

Patwa, Z., and L. M. Wahl, 2008 The fixation probability of beneficial mutations. Journal of the Royal Society Interface **5**: 1279–1289.

Poole, K., 2001 Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. Journal of Molecular Microbiology and Biotechnology **3**: 255–264.

Poon, A., and L. Chao, 2005 The rate of compensatory mutation in the DNA bacteriophage ΦX174. Genetics **170**: 989–999.

Poon, A., B. Davis, and L. Chao, 2005 The Coupon Collector and the Suppressor Mutation. Genetics **170**: 1323–1332.

R Core Team, 2013 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Development Core Team, 2011 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Remold, S. K., and R. E. Lenski, 2001 Contribution of individual random mutations to genotype-by-environment interactions in Escherichia coli. Proceedings of the National Academy of Sciences **98**: 11388–11393.

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

Robicsek, A., G. Jacoby, and D. Hooper, 2006 The worldwide emergence of plasmid-mediated quinolone resistance. The Lancet Infectious Diseases **6**: 629–640.

Rodríguez-Verdugo, A., B. S. Gaut, and O. Tenaillon, 2013 Evolution of Escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. BMC evolutionary biology **13**: 50.

Rosche, W. A., and P. L. Foster, 2000 Determining mutation rates in bacterial populations. Methods **20**: 4–17.

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

Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, *et al.*, 2001 The Population Biology of Invasive Species. Annual Review of Ecology and Systematics **32**: 305–332.

Sander, P., B. Springer, T. Prammananan, A. Sturmfels, M. Kappler, *et al.*, 2002 Fitness cost of chromosomal drug resistance-conferring mutations. Antimicrobial Agents and Chemotherapy **46**: 1204–1211.

Sarkar, S., W. Ma, and G. H. v Sandri, 1992 On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. Genetica **85**: 173–179.

Schoustra, S. E., J. Dench, R. Dali, S. D. Aaron, and R. Kassen, 2012 Antagonistic interactions peak at intermediate genetic distance in clinical and laboratory strains of *Pseudomonas aeruginosa*. BMC microbiology **12**: 40.

Sepälä, H., T. Klauskka, J. Vuopio-Vaskila, *et al.*, 1997 The effect of changes in the consumption of macrolide antibiotics on erythromycin resistance in group A streptococci in Finland. New England Journal of Medicine **337**: 441–6.

Shi, B., and X. Xia, 2003 Changes in growth parameters of Pseudomonas pseudoalcaligenes after ten months culturing at increasing temperature. FEMS microbiology ecology **45**: 127–134.

Singh, R., P. Ray, A. Das, and M. Sharma, 2009 Role of persisters and small-colony variants in antibiotic resistance of planktonic and biofilm-associated Staphylococcus aureus: an in vitro study. Journal of Medical Microbiology **58**: 1067–1073.

Speert, D. P., M. E. Campbell, D. A. Henry, R. Milner, F. Taha, *et al.*, 2002 Epidemiology of Pseudomonas aeruginosa in cystic fibrosis in British Columbia, Canada. American journal of respiratory and critical care medicine **166**: 988–993.

Tago, Y., M. Imai, M. Ihara, H. Atofuji, Y. Nagata, *et al.*, 2005 *Escherichia coli* mutator *δpolA* is defective in base mismatch correction: the nature of *in vivo* DNA replication errors. Journal of Molecular Biology **351**: 299–308.

Tanaka, Y., K. Yazawa, E. R. Dabbs, K. Nishikawa, H. Komaki, *et al.*, 1996 Different rifampicin inactivation mechanisms in Nocardia and related taxa. Microbiology and immunology **40**: 1–4.

Tokuriki, N., and D. S. Tawfik, 2009 Protein dynamism and evolvability. Science **324**: 203–207.

Tsien, R. Y., 1998 The green fluorescent protein. Annual Review of Biochemistry **67**: 509–544.

Uecker, H., S. P. Otto, and J. Hermisson, 2014 Evolutionary Rescue in Structured Populations. The American Naturalist **183**: E17–E35.

Velayudhan, J., M. A. Jones, P. A. Barrow, and D. J. Kelly, 2004 L-serine catabolism via an oxygen-labile L-serine dehydratase is essential for colonization of the avian gut by Campylobacter jejuni. Infection and immunity **72**: 260–268.

Wagner, A., 2008 Robustness and evolvability: a paradox resolved. Proceedings of the Royal Society B: Biological Sciences **275**: 91–100.

Wahl, L. M., P. J. Gerrish, and I. Saika-Voivod, 2002 Evaluating the Impact of Population Bottlenecks in Experimental Evolution. Genetics **162**: 961–971.

Walsh, C., 2003 Where will new antibiotics come from? Nature Reviews Microbiology **1**: 65–70.

Wehrli, W., 1983 Rifampin: mechanisms of action and resistance. Review of Infectious Diseases **5**: S407–S411.

Wehrli, W., and M. Staehelin, 1971 Actions of the rifamycins. Microbiology and Molecular Biology Reviews **35**: 290–309.

Williams, K. J., and L. Piddock, 1998 Accumulation of rifampicin by Escherichia coli and Staphylococcus aureus. Journal of Antimicrobial Chemotherapy **42**: 597–603.

Winsor, G. L., D. K. Lam, L. Fleming, R. Lo, M. D. Whiteside, *et al.*, 2011 Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. Nucleic acids research **39**: D596–D600.

Wise, R., 2002 Antimicrobial resistance: priorities for action. Journal of Antimicrobial Chemotherapy **49**: 585–586.

Wiser, M. J., N. Ribeck, and R. E. Lenski, 2013 Long-Term Dynamics of Adaptation in Asexual Populations. Science **342**: 1364–1367.

Woods, R. J., J. E. Barrick, T. F. Cooper, U. Shrestha, M. R. Kauth, *et al.*, 2011 Second-order selection for evolvability in a large *Escherichia coli* population. Science **331**: 1433–1436.

Wright, B. E., 1996 The effect of the stringent response on mutation rates in Escherichia coli K-12. Molecular microbiology **19**: 213–219.

Wright, G., 2007 The antibiotic resistome: the nexus of chemical and genetic diversity. Nature Reviews Microbiology **5**: 175–186.

Wright, G., 2010 Antibiotic resistance in the environment: a link to the clinic? Current opinion in microbiology **13**: 589–594.

Wright, S., 1931 Evolution in Mendelian populations. Genetics **16**: 97–159.

Ye, K., M. H. Schulz, Q. Long, R. Apweiler, and Z. Ning, 2009 Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics **25**: 2865–2871.

Yee, Y. C., B. Kisslinger, L. Y. Victor, and D. J. Jin, 1996 A mechanism of rifamycin inhibition and resistance in *Pseudomonas aeruginosa*. Journal of Antimicrobial Chemotherapy **38**: 133–137.

Yim, G., H. Huimi Wang, and J. Davies, 2006 The truth about antibiotics. International journal of medical microbiology **296**: 163–170.

Zaczek, A., A. Brzostek, E. Augustynowicz-Kopec, Z. Zwolska, and J. Dziadek, 2009 Genetic evaluation of relationship between mutations in *rpoB* and resistance of *Mycobacterium tuberculosis* to rifampin. BMC Microbiology **9**: 10 doi:10.1186/1471–2180–9–10.

Zhou, Y. N., and D. J. Jin, 1998 The *rpoB* mutants destabilizing initiation complexes at stringently controlled promoters behave like “stringent” RNA polymerases in *Escherichia coli*. Proceedings of the National Academy of Science of the United Statesof America **95**: 2908–2913.