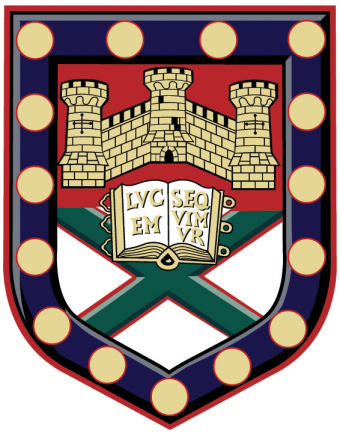
The evolution of resistance to multidrug antibiotic therapies.

Submitted by Mark Hewlett to the University of Exeter
as a thesis for the degree of Doctor of Philosophy in Biology
In September 2015

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree
by this or any other University.



Abstract

The purpose of this thesis is to explore the interaction between antibiotics at sub-lethal doses, and *E.coli*. Initially we focussed on pairwise antibiotic interaction, and the potential to exploit these interactions to minimise antibiotic resistance. In testing the hypothesis that antagonism will slow adaptation by reducing selection for resistance we determined that there are conditions in which this fails to be the case. We furthermore caution against treating drug interactions as anything other than a dynamic property of the bacteria-drug interaction, by showing that the relationship between two drugs may be both synergistic and antagonistic depending on a variety of factors. Whilst exploring the adaptive response to drug combinations we discovered a highly unusual effect of Doxycycline to act as a growth stimulant to *E.coli* AG100. Chapter 3 and 4 are devoted to determining the nature and mechanism of this stimulation, and analysing any potential genomic changes using whole genome re-sequencing. Having shown that dose response is not always a monotone function of increasing drug dose, in chapter 5 we also look at the dose response in a diffusive context, using a custom built imaging system to show the common non-monotonicity of disk diffusion type assays, that manifest themselves as bullseye patterns of growth. We use a mathematical model to explore the ecological and adaptive reasons for such patterns. Finally in chapter 6 we look at the coevolutionary history of phage and *E.coli* REL606 strains, by determining trade-offs caused by λ phage and the sole carbon source maltotriose both utilising the same porin (lamB) for cell entry.

Contents

List of Figures

List of Tables

Abbreviations

MIC: Minimum inhibitory concentration
IC_x: Inhibitory concentration (for x proportion of inhibition)
MPC: Mutant prevention concentration
MBC: Minimum bactericidal concentration
MSW: Mutant selection window
FICI: Fractional inhibition concentration indices
Dox: Doxycycline
Rif: Rifampicin
Ery: Erythromicin
Strep: Streptomycin
PK/PD: Pharmaco-kinetic/Pharmaco-dynamic
CLSI: Clinical laboratory standards institute
CFU: Colony forming units
OD: Optical density
α: Rate of adaptation
M9: M9 minimal media
LB: Lysogeny broth
SNP: Single nucleotide polymorphism
Indel: Insertion/deletion
Contig: Contiguous read

Acknowledgments

I would like to extend my sincere thanks to the many people people who have been of great assistance during my PhD.

- Firstly, to my supervisor, Rob Beardmore; not only for giving me the opportunity to work with him, but also for his guidance, support and advice over the last four years.
- To Ivana Gudelj, my academic mentor, for making lots of useful observations during lab meetings!
- To everyone in Lab 309 who shared their expertise, and were always willing to offer tips and improvements on experimental protocols; Ayari Fuentes-Hernandez, Francesca Fiegna, Jessica Plucain, Carlos Reding-Roman and Lisa Butt.
- To all those in the computational biology office who have patiently helped me to debug code, or analyse sequence data; Fabio Gori, James Harrison, Tom Laver, Christine Sambles and Rafael Peña-Miller
- To my family, and to all the good friends I have made whilst at Exeter, who never complain about standing me a pint!
- And finally to Charlotte, my wife who has been my constant support over the course of this degree, I could not have done it without you!

Chapter 1

Introduction

1.1 History of antibiotics

Antibiotics, in their first appearance in the early twentieth century, must have seemed to those original recipients, to be nearly magical in their futuristic ability to cure commonly fatal ailments with a simple injection. Whether Erlich's Salvarsan [90] in 1911; which can appropriately be described as the worlds first modern chemotherapeutic agent, to Prontosil, the first truly modern antibiotic in 1935, the health implications of these new drug classes was obvious and profound. In the decades following the release of benzylpenicillin in 1942, dozens of other antibiotics were being routinely used to treat disease and infection, a rapidity of drug discovery no doubt spurred on by the huge need arising from the injuries and spread of infectious disease due to the second world war. As more drug targets were identified, high throughput screening techniques quickly discovered large numbers of compounds that had bactericidal or bacteriostatic properties, yet with minimal side effects to humans. From 1942 -1952, six new classes of antibiotics were discovered and 12 new drugs were synthesised, leading US surgeons-general William H Steward to to famously state 'It is time to close the book on infectious diseases'. Although there is debate over the veracity of this quote [110], it is undeniable that the introduction of antibiotics marked the start of a medical revolution that might make such exaggerations understandable.

However, as was apparent even to the early pioneers of antibacterial drug development, the emergence of resistance was a foreseeable and inevitable consequence of exposure to these antibiotics. Sir Alexander Fleming was explicit in highlighting this risk when accepting his Nobel Prize (With Florey and Chain) in 1945, using the words 'If you use penicillin, use enough', in order to ensure that sufficient doses were used to remove all trace of an infection, leaving behind no resistant strains that may have benefitted from their exposure, to become more resistant.

It has been this principle that has guided pharmacological thought in the years since: although antibiotic resistance has been acknowledged as a problem, and more

recently has received a great deal of attention as a potential problem with world-altering consequences, the emphasis for scientific research has been in drug discovery (the so called 'arms race' between bacteria and medicine). The prospect of minimising the emergence of resistance has been predominately limited to using high doses or combinations of antibiotics in different classes to suppress resistant phenotypes. This method, however, is limited in scope due to the nature of the selective pressures, drug gradients within treatment, and competitive release. These themes, and the implications that they hold for antibiotic therapy design, will be discussed in more detail throughout this work.

The modern view of antibiotics is one of far less awe, most people in developed countries experience the contempt bred by familiarity, and have come to expect magic bullet treatments as par for the course. Therefore, when antibiotics fail, whether due to resistance, mis-diagnosis, or inappropriate treatment (for example, antibiotics being prescribed to treat viral infections, the public feels grieved at this failure of modern medicine. The concept of returning to a pre-antibiotic era, whereby surgical intervention would be almost certainly fatal, and diseases mostly eradicated in the west could once again kill thousands, is difficult to contemplate. Nonetheless, despite research priority being given to drug discovery, the pipeline for new classes of antibiotics has all but dried up, with truly novel antibiotics now a rare occurrence [67]. Furthermore, it is estimated by the World Health Organisation (WHO) that from it's clinical introduction, a new antibiotic has only two years of useful lifespan before resistance becomes globally prevalent. These findings are leading more and more people to question whether our guidelines for the usage of antibiotics is appropriate, and whether there are mechanisms by which selection pressures can be mediated, in order for resistance can be slowed, halted or even reversed.

1.2 Antibiotic testing and nomenclature

To ensure consistency we will now spend some time defining several core concepts that are used in drug interaction studies, and yet have some variability over the usage. For completeness, we now discuss some standard pharmacokinetic and pharmacodynamic (PK/PD) definitions and concepts. The Clinical Laboratory Standards Institute (CLSI) defines the Minimum Inhibitory Concentration (MIC) of an organism as the drug concentration in a 2-fold dilution plate at which no change in turbidity is seen by eye at 24 hours [25]. Although this definition has the convenience of being a cheap and low-tech measure of efficacy, it is prone to human error, and it ignores the potential for dosages with a less than total killing effect to act as reservoir for resistance mutations. This definition states the MIC to be used as meaning the lowest drug concentration to inhibit ***all*** growth, although it is also acknowledged that the use of this terminology is entirely dependent on the context

under which it was derived [84].

Other studies define the MIC as the lowest drug concentration that inhibits *any* growth (i.e, the smallest dose to have any observable effect [33] *et al.,*. Under this definition, the dose required for complete bacterial inhibition is called the mutant prevention concentration (MPC) and the doses falling between the MIC and MPC are referred to as the mutant selection window (MSW). Another concept that has been used previously is the minimum bactericidal concentration (MBC), or the concentration at which a 99.9% reduction (measure in CFUs) is seen [116]. This measure also suffers from inconsistent use, it having another definition as the minimum concentration required to prevent growth of subculture in drug-free media [60].

Assuming that the MIC is a fixed measure of the efficacy of a drug against each bacterial strain is erroneous. Environmental changes will effect the MIC, whether the concentration of available metabolites and extracellular resources, or temperature and pH. This is problematic for any comparison of MICs between laboratories and highlights the necessity of standardised protocols and language for referring to dose-dependent effects. Although the CLSI and EUCAST publish comprehensive protocols for the standardisation of MIC determination, and then define set breakpoints as meaning either a resistant or sensitive strain, it is made very clear by these organisations that these breakpoints are guidelines and can in no way be considered predictive or therapeutic doses. The broadness of the breakpoints even within the same species gives some evidence of the extent to which MIC can be affected by numerous factors.

According to the prevailing best practice, bacterial cells must spend as little time in the MSW as possible in order to minimise selection of resistant genotypes. This hypothesis has been challenged by several studies that have shown no correlation between time spent in the MSW and the emergence of resistance [4, 23, 22]. Additionally, as I shall discuss in more detail later, the idea that there is no bacterial growth simply because it is below the threshold detectable by the human eye (still the standard detection method for antibiotic resistance assays) may not be correct. Whether through the presence of persisters or slow growing phenotypes that remain at a population of $<10^5$, populations may recover, even from doses considerably higher than MIC.

With such confusion over the best use of terminology to accurately describe dose-dependent effects, it is perhaps even harder to quantify multi-drug effects in a consistent manner. Drug interactions are usually deemed to be synergistic, antagonistic, suppressive or indifferent (at a given time) using checkerboard assays [112] and one of the purposes of this document is to use bacterial growth kinetics to understand whether *in vitro* measures of drug interactions of antibiotic drugs are robust.

There are a number of physiological factors that can be measured to derive information about the growth and fitness of bacteria, including growth rate, final optical density and colony counting, to name a few common methods. It is important to realise that each of these measurements has benefits and drawbacks, for example, measuring growth rate is a standard measure of bacterial fitness but gives no information about the efficiency of the cell (yield) or the overall growth during a given time frame. This will be discussed in greater detail later, when measurement of *E.coli* grown in the presence of Dox shows that the growth rate is largely unchanged by increasing dose, rather it is the lag phase which is extended. For this reason we have chosen to use a combination of final optical density in order to assess yield, and area under the curve (AUC) as a proxy for growth rate, as it is a measure of all aspects of the growth kinetics, not just the exponential phase.

1.3 Structure and mechanism of antibiotics in this study

1.3.1 Doxycycline

Doxycycline is an antibiotic from the tetracycline family, and is an inhibitor of translation. Dox binds reversibly to the 16S ribosomal subunit, and competes with amino-acyl tRNA to bind to the ribosome. This competitive interaction means that Doxycycline has a bacteriostatic mode of action as it slows the production of proteins required for growth and division. However, it has been reported [31] that as little as a two-fold increase in Dox concentration above the MIC can cause the antibiotic to become lethal to cells, leading to a bactericidal mode of action. The mechanism behind this lethality is unclear, but it is thought to be due to rupturing of the cell membrane leading to lysis.

1.3.2 Rifampicin

Rifampicin is a semi-synthetic derivative of rifampycin, which was so named by the research team to discover it after the main character 'Riffifi' in a popular French book about a jewellery heist. Rifampicin is widely reported as being bactericidal, and works by inhibiting RNA polymerase, thereby reducing successful transcription. Although there exist other antibiotics that inhibit this process they do so by binding to the DNA template, and so are highly toxic to the host, i.e., humans. Rifampicin binds directly to the polymerase, and has been shown to be non-toxic at a four-fold higher concentrations than is required for complete inhibition in *E.coli*.

1.3.3 Erythromycin

Erythromycin is an antibiotic of the macrolide class, and similarly to Doxycycline, it inhibits translation to cause a bacteriostatic effect. Erythromycin binds to the 50S subunit of bacterial 70S ribosome, and inhibits, through a poorly understood mechanism, the translocation on aminoacyl-tRNA from the A binding pocket to the P site. This inhibited translocation means that the A site remains occupied, and so nascent polypepetide chains are terminated, creating short abortive peptides.

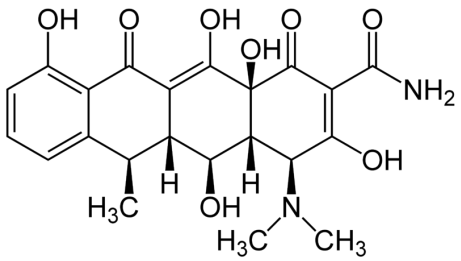


Fig. 1.1: The structure of Doxycycline. The usual form of commercial powdered Dox is as the hyclate salt.

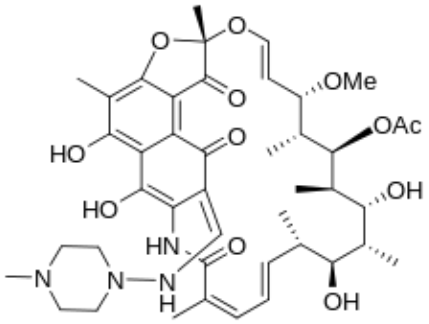


Fig. 1.2: The structure of Rifampicin.

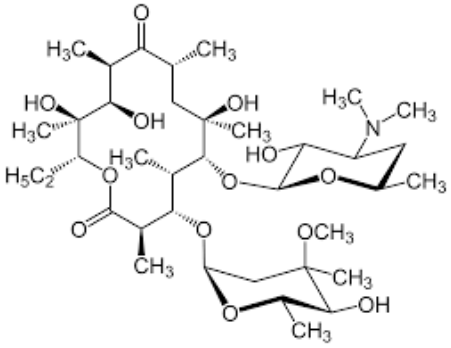


Fig. 1.3: The structure of Erythromycin.

1.4 Drug Combinations

To combat the evolution of antibiotic resistance, and to prolong the effective lifespan of these drugs, several treatment concepts have been developed to minimise the probability of such resistance evolving. One such concept is that of combinatorial drug treatments, whereby cocktails of two or more drugs are administered simultaneously in order to provide a more effective treatment regimen. Such combinations can be additive (in that the drugs act independently of each other, and do not interact), synergistic or antagonistic. However, it can be problematic defining the interactions of drugs, and the method used to determine this interaction will therefore be very important when basing a protocol on an assumed drug interaction.

The most common form of interaction test is the checkerboard or isobologram, whereby two-fold dilutions are made of each drug, one in the x-axis and one in the y-axis, usually in 96 or 384 well plates. When inoculated, the wells at which there is no visible bacterial growth following 24 hours incubation, are plotted, and the shape of the interaction will be used to estimate the interaction. A more quantitative assessment can be reached using fractional inhibition combination indices, (FICI measurements) from the same data.

Combinatorial treatments have come to be regarded as a vital tool in reducing the emergence of resistance, and in treating infections that are already characterised by antibiotic resistance, although there is increasing debate over whether combinatorial treatments reduce resistance by exploiting synergy, or by increasing the chance that an appropriate drug will be utilised against an unknown pathogen [113]. Under the assumption that synergy increases total efficacy compared to each individual drug given alone, the possibility arises that greater inhibition of bacteria can be achieved without increasing toxicity. Additionally, even if a bacterial population may have, or gain, resistance to one of the antibiotics, the second drug will be able to successfully clear the infection. In general it is assumed that drugs will be synergistically active if they are not competing for the same target, therefore an example of a typical synergistic combination would be penicillin (cell wall synthesis inhibition) and erythromycin (protein translation inhibition). The converse is also assumed, that is drugs must compete for the same target, their joint efficacy will be naturally reduced, and therefore the treatment is regarded as antagonistic [62]. Despite these guidelines, there are numerous reports of synergistic drugs utilising the same mechanism of action, and also of non-competitive drugs which show either additivity or antagonism, where one may expect synergy.

The idea that antagonistic drug combinations could be pressed into therapeutic service in order to combat the spread of antibiotic resistance is a fairly new concept [54], [119]. It has come about due to failure of synergistic treatments to limit the emergence of resistance as effectively as was hoped. It has been shown that in cases where a synergistic combination is deployed, the drug pair will exert a higher

selective pressure on the population, and thereby will lead more rapidly to a sweep of resistance than if the drugs were used in mono-therapy [13].

If antibiotic synergy can select for greater bacterial load, by providing a stronger selective pressure for resistance, can the opposite effect be exploited by using antagonism to select against resistance? It is this question that forms the genesis of this research, and in exploring the dynamic nature of drug interactions, and the adaptation of bacteria to sub-lethal concentrations of drug, we aim, initially, to test the hypothesis described by Hegreness that antagonism can minimise the rate of adaptation in bacteria [54]. This idea is highly clinically important, as new methods for controlling infection, whilst limiting resistance, are the apogee of infection control. However, there is the significant proviso of ensuring adequate patient outcome; a treatment that minimises the spread or emergence of resistance, but fails to clear an infection, is not a treatment that can be of use outside of the lab. Therefore, when looking at the properties of antibiotic combinations, both the effect on adaptation, and the overall inhibitory effect of the treatment must be considered.

1.5 Fitness gains from antibiotics

It is known that antibiotic molecules have different biological functions, dependent on dose. Given that most antibiotics in use today are either isolated, or derived from microbiological secretions, it is perhaps not surprising that these small molecules have roles beyond that of destroying competing strains of bacteria in the environment. Even taken logically, it should be evident that if this was the only purpose of antibacterial compounds, resistance should have been far more widespread, if not ubiquitous, immediately after clinical deployment. After all, many antibiotic compounds have been isolated from bacteria in their natural environments, and therefore presumably have ancient biological origins, and yet retained their effectiveness up until widespread usage as medicines became the norm. Therefore, do bacteria retain an advantage to antibiotic sensitivity, or are antibiotics not present in the environment in sufficient quantities to mediate resistance? If by gaining resistance to one or more antibiotics, a bacteria can protect itself from an external threat, then why is resistance not a universal phenotype? The standard explanation is that the emergence of resistance must carry a fitness cost, usefully expressed as a slowing of growth rate. Therefore, in an environment where the selective pressure of a drug is present, a resistant phenotype will have the survival advantage - better to grow slowly, than be killed by an antibiotic. When that selective pressure is removed however, that subpopulation will be outcompeted by their drug-sensitive fast-growing ancestors. Although this initially looks like good news for clinicians, who may be able to remove resistant populations altogether, simply by having a moratorium on specific drug uses, the reality is far more complex. Studies have shown that the rate

at which resistance is lost to a population is orders of magnitude slower than the speed at which it is gained [118]. Additionally, the acquisition of resistance to one antibiotic can promulgate further cross-resistance to other related compounds, or even to antibiotics from entirely different classes.

1.6 Rate yield trade-offs and antibiotic resistance

Trade-offs have been widely theorised as being a primary driver of genetic diversity, allowing a population level flexibility to cope with changing environments, and conferring a more robust survival strategy, at the species level. Trade-offs have been described in many systems, including infectivity vs burst size in viruses, affinity vs transport rate in bacterial trans-membrane substrate transport, and growth vs maintenance for bacterial cells. The variables may differ between these examples, but the concept remains that same, that sacrificing fitness in one area will allow energy to be expended elsewhere, which may confer more advantages. It is important to note that the outcome of the trade-off, whether it is advantageous or not will be dependent on a range of external factors, and as the factors change, so too will the bacterial population adapt and evolve their trade-off responses.

It is also worth noting that although trade-off shapes have been used to predict increased diversity, *in vitro* data of such trade-offs is scarce, and attempts to measure trade-offs are complicated and often unclear. In this study we will use a very simple trade-off between maltotriose transport and lambda phage resistance in *E.coli* to determine whether a trade-off exists when these two species are co-evolved, and whether such a trade-off is a necessity in improving phage resistance, or whether bacteria can instead trade-up. This experimental setup is used because maltotriose can be supplied as the sole carbon source, which is taken into the cell via maltoporin, which is also the attachment site for phage lambda. This allows generation of a co-evolved library to be tested in varying resource concentrations and determine the relationship between resistance and cellular effeciency.

Additionally we will investigate the means by which external stressors, in this case the antibiotic Doxycycline, can mediate the trade-off between rate and yield in *E.coli*, and how the trade-off shape changes over time when in the presence of the drug. As Doxycycline is a bacteriostatic drug at medically relevant doses, it is important to understand how the antibiotic may alter the growth of microorganism, for example, if growth rate is slowed upon dosage, but over a longer timescale the balance of the trade-off is shifted in favour of an increased yield, then this may be a counter productive treatment strategy, only providing initial advantages, but in the longer term benefitting the bacterial population by increasing metabolic efficiency.

Table S1.1: *E.coli* strains used in this thesis

Strain	Adaptation	Source
AG100	Ancestral	S.B. Levy
$AG100_{1MIC}$	Dox MIC adapted	Evolved from AG100
$AG100_{2MIC}$	Dox 2xMIC adapted	Evolved from AG100
$AG100_{M9}$	Media adapted	Evolved from AG100
$AG100_{Str+}$	Streptomycin (1000xMIC) adapted	Evolved from AG100
MG1655	Ancestral	S.B. Levy
$MG1655_{1MIC}$	Dox MIC adapted	Evolved from MG1655
$MG1655_{2MIC}$	Dox 2xMIC adapted	Evolved from MG1655
$MG1655_{M9}$	Media adapted	Evolved from MG1655
REL606	*E.coli* B strain	J. Meyer
TB108	Ancestral	T. Bergmiller
eTB108	Ery MIC adapted	Evolved from TB108
BW25113	Ancestral	Keio collection
$BW25113_{\Delta manZ}$	Glucose transport knockout	Keio collection
$BW25113_{\Delta aceA}$	Isocitrate lysase knockout	Keio collection
$BW25113_{\Delta aceB}$	Malate synthase knockout	Keio collection
$BW25113_{\Delta icd}$	Citrate synthase knockout	Keio collection

1.7 Experimental Systems

The adaptation of bacterial cells to antibiotics will be studied, for the most part using a serial transfer protocol, with transfers occurring every 12 or 24 hours. Exceptions will be those experiments determining spatial response to antibiotics which will be carried out on M9 soft agar, and left for up to 72 hours for adaptation to occur. Liquid media based experiments will utilise 96 well plates in order to increase throughput and for ease of optical density measurements. The choice to use a minimal media is important, as in many cases we will be seeking to understand whether metabolism plays a part in antibiotic (or phage) resistance, and how trade-offs between various growth parameters can effect fitness. It was therefore necessary to be able to accurately control and quantify the amount of carbon resources initially provided to the growth environment. *E.coli* was used as the model organism in the following work, due to the ease and safety of use, the well understood genome and the availability of knockout libraries and tagged strains. The following strains will be used throughout this thesis.

1.8 Thesis Aims

This thesis aims to explore the stability of inhibitory effects of antibiotics, whether in combination or used singly, over clinically relevant timescales. The work shall be carried out using a variety of strains of *E.coli* derived from K12, chosen for it's short generation time, ease of molecular manipulation, and lack of safety concerns. The five chapters of this thesis will look at a number of physiological factors of bacterial

growth, and the potential effects of antibiotic treatment on them. Furthermore, it will assess how these factors are altered during the inevitable process of adaptation towards selective pressures. The first chapter will look at the stability of two-drug interactions, and at the robustness of claims of synergy and antagonism. This is an important area of research as antibiotic combinations and cocktails are increasingly being seen as a method for reducing the emergence and spread of multi-drug resistant isolates, however there is a dearth of research into the long term effects of combinatorial treatment on the selection for resistance.

In the second and third chapters we will be assessing the changes in inhibition that can occur when a simple dose-response assays is serially transferred and evolved for up to 10 days. Again, the relatively short evolutionary timescale of around one week is not an arbitrary length of time, rather it accurately reflects the average treatment length for infection with antibiotics. Although resistance measures are well standardised globally, there is a common shortfall in that such measurements are taken at only one time point, and therefore a promising looking treatment as assessed initially may turn out to be less efficacious over the course of the whole treatment window. We will show that in certain conditions, simple short term adaptation to a sub-MIC antibiotic challenge can produce bacteria that are fitter than their wild-type progenitors, and maintain a greater level of fitness even when the drug is removed. We will also demonstrate that specific drug sequences can cause this effect, and therefore there is the potential in real-world scenarios that poorly designed treatment regimens may be having the opposite of the desired effect. The third chapter will look at the genetic consequences of these effects, in an attempt to determine a mechanism by which antibiotic treatments can fail so dramatically.

In chapter four we will investigate the role that metabolic trade-off can play in influencing bacterial fitness. The model system that we shall use is based on a strain library of *E.coli* REL606 co-evolved with lambda phage. We will show that there exist multiple trade-offs within populations which can mediate many physiological factors, including phage resistance, lag phase length, growth rate and yield.

In the final chapter we will return to looking at the effects of antibiotics on bacteria, although in this case in a more ecological context. Experiments will be performed in a spatial structure and the diffusion of both antibiotic (from a point source) and glucose (from areas of low bacterial growth to high) will provide ecological niches that should favour bacteria with homogenous phenotypes. This work was brought about by the observation that when performing standard paper-disk diffusion assays, the resulting growth along the monotonically decreasing drug gradient was frequently non-monotonic. As such non-monotonicity has been observed in both spatially structured and homologous (shake-flask) experiments within the course of this thesis, these results would suggest that the assumption that increasing drug dosage will have a proportional decrease in bacterial viability is erroneous, and

further research is required to accurately determine the various roles of antibiotics at sub-MIC concentrations.

Chapter 2

Non-robust pharmacological measures: treatment with antagonistic combinations

2.1 Overview

In this chapter we test the robustness of drug interactions over a range of glucose concentrations, measurement times and antibiotic doses. Furthermore, we investigate the effect selection on the adaptation of bacteria in these combinatorial treatments.

1. We show that experimental protocol can determine the result of an interaction assay, simply by varying the time of measurement.
2. Model driven hypothesis suggests that drug interaction is not stable.
3. We further show that both the resource concentration and the drug concentration will alter the reports of synergy and antagonism.
4. We aim to test the hypothesis that antagonistic antibiotic combinations may slow the emergence of resistance.
5. Our results do not concur with the hypothesis that antagonism selects against resistance.
6. Dox, in monotherapy and in combination with Rif can have a stimulatory effect on *E.coli*.

2.2 Background

Antibiotic treatment, once considered a panacea for infectious disease, is increasingly failing. The prevalence of antibiotic resistance in diverse pathogenic species has been highlighted as a threat '*as severe as terrorism*' by the UK's Chief Medical Officer

in her most recent report [44]. The findings emphasise the risks posed to health worldwide and the hazards posed by a lack of new antibiotic development. However, simply investing in new drug design should not be seen as anything more than an ongoing and necessary measure in the treatment of microbial pathogens. Without a more complete understanding of the selective pressures responsible for the explosion in prevalence of drug-tolerant strains and species, any new antibiotic will surely lose its efficacy once it enters general medical usage.

Despite the fact that the existence of resistance strains became clear soon after antibiotic treatments became commonplace, the scientific and medical community was slow to recognise their importance [48]. Since that time, antibiotic resistance has become a widespread and pervasive problem with strains of almost all diseases losing some measure of susceptibility to a once efficacious treatment. This rise of resistance causes severe economic and health benefits due to longer hospital stays, higher mortality and the high cost of finding alternative treatments [100]. This economic cost is estimated to be between $5 and $24 billion per annum in the US alone [15].

However, the cost of resistance is measurable not only in human terms. Resistance to antibiotics generally confers a fitness cost on mutant bacteria, although this fitness cost may be small, and ameliorated with further mutational events [32]. This cost manifests itself as a slowed growth rate under non-drug conditions and it is hoped that we can exploit it to rationally design treatments that explicitly militate against resistance evolution [5] by 'reversing' resistance due to the negative selection pressure acting on these mutants in drug-free environments. It has been shown, however, that the time taken for a resistance mutation to disappear from a bacterial population is much longer than the time taken for such mutations to arise when treated with antibiotics [39]. Combined with the phenomenon of antibiotic persistence that is due to cells that do not have genetically-encoded resistance but which can survive by temporarily halting vegetative growth, the severe difficulties associated with controlling the spread of resistant strains of common infectious pathogens becomes apparent.

The evolutionary mechanisms of resistance include mutations of the bacterial chromosome and acquisition of genetic material through bacteriophages, plasmids and transposons [61]. The acquisition of resistance can be lateral (passed down generations) or horizontal (acquired from a different strain or species). The issue can be exacerbated by healthcare professionals as bacteria can carry multiple plasmids or be infected by phage harbouring multiple resistance-conferring transposons, thereby creating bacteria that are multi-drug resistant [69]. Even in the wild, bacteria harbour many defence systems that may offer protection against antibiotics, including numerous efflux pumps that can be up-regulated given the right environmental stimuli [77].

Since Erhlich first espoused his treatment philosophy in 1913, conventional wisdom in the treatment of infection or disease with antibiotics has been to 'hit early and hit hard'. This maxim was taken to extreme by the discoverer of the first commercial antibiotic (Prontosil), Gerhard Domagk, who treated his daughter's acute wound infection with a series of high, 10-gram doses, even before human trialling had been completed; luckily his daughter survived. Under-dosing is a more commonly-cited risk factor supporting the emergence of resistance [47, 37, 71]. Although Fleming's advice was based on the conceit of using high doses to entirely suppress mutant resistance populations, like other studies [76] it ignores the risks associated with providing such a strong selective pressure for resistance.

2.2.1 FIC measures

The standard protocol for establishing drugs interaction is using a checkerboard assay. The outcome of this has a visually-determined and, therefore, subjective component to it [112]. The FIC measurement has been reported as having poor agreement with other methods such as the E-test, with agreement of only 55% [45]. Synergy or antagonism is defined by this method as a Fractional Inhibitory Concentration (FIC) and the FIC index (FICI) is derived by calculating the proportional killing effect of a drug in combination when compared with its killing effect in monotherapy. Therefore:

$$FICI = \frac{C^a}{MIC^a} + \frac{C^b}{MIC^b} =: FIC^a + FIC^b. \tag{2.1}$$

Where C^a is the concentration of drug A required for a given level of inhibition when in combination, and MIC^a is the concentration required for the same inhibition in monotherapy. The FICI scores are normally defined as synergy, indifference or antagonism for $FICI < \frac{1}{2}$, $FICI \approx 1$ and $FICI > 4$, respectively [112], although these values have changed historically. There is uncertainty as to the value of results obtained using the arbitrary value of $FICI < 1/2$ defining synergy [81]; It has been argued that weakly-interacting drug pairs with a FICI score between 1/2 and 99/100 may still be considered significantly synergistic [52] and, given this degree of uncertainty, the test we adopt to test for synergy is described below.

The rationale behind multi-drug combination therapy is to simultaneously maximise drug efficacy and reduce resistance in bacterial populations. Another study suggests that not only are the pharmacodynamic properties of two single drugs insufficient to predict drug interactions, but drugs that have been empirically determined to suppress each others' mode of action can still offer protection against single drug resistance phenotypes [8]. A second study has suggested that the increased efficacy of combination treatments yields necessarily greater selection for resistance [54, 119]. As a corollary, the latter claims that antagonistic drug pairings select against re-

sistance. This is of little value if the reduced efficacy of the treatment means that patient care is compromised to an unacceptable level in order to reduce resistant evolution. And, as a result, we explicitly test the virtue of antagonism as a rationale for treatment design.

2.3 Does antagonism select against resistance?

The treatment of bacterial infections in the 1930-1940's with antibiotics created a revolution in medicine, with dramatically increased survival rates for a range of serious conditions [27]. Within just a few years, however, many strains of cultivable pathogens were showing signs of resistance to antibiotics *in vitro* [38]. The lack of new antibiotics in the development pipeline now [94] has dire implications for our future ability to combat bacterial infection. However, only seeking to address the scourge of resistance through the discovery of new drugs fails to address the underlying root causes of the emergence of resistance and how this can best be minimised with antibiotic therapy.

The use of combination therapies for the treatment of disease is not new. Used for over seventy years [97, 105, 26], drug cocktails are used to treat microbial infections, [109], viral infections [34] and cancers [96]. However, there are at least two rationales that support the use of combinations for treating disease. One is based on the observation that some drug pairings are more effective in combination than would be predicted from an equivalent dose of each monotherapy. This is termed synergy [26, 46, 24] and although used as the basis of increasing treatment efficacy, it is also believed to suppress antibiotic resistance adaptation [78].

As explained in [76], the basis of the theory behind this hypothesis is the following. Suppose p is the probability of a resistance allele arising in a pathogen population during treatment, if two drugs are used with independent targets, p^2 is then the probability of a resistance allele arising in a pathogen population during treatment in a combination therapy that uses two antibiotics, assuming that mutation is random across the genome. By the same argument, it would seem rational to use three drugs. However, a problem with this argument is the complete absence of any consideration of what selection for drug resistance may achieve. Indeed, it is axiomatic that greater selection for drug-resistance will be observed at higher dosages, the purpose of this work therefore is to begin to address question of how dosage mediates selection for resistance as part of a combination.

The second purpose of this research is to understand what other processes might lead to the loss of synergy during treatment with a combination of antibiotics. As a model we take two drugs used in the treatment of *Brucella melitensis*. Using simplified mathematical models of the modes of action of the drug pair *rifampicin*, itself used in the treatment of important human diseases like TB, and *doxycycline*, we

predict that antibiotic interactions, and therefore synergy, can depend on dosage and time. As we show, a combination that begins synergistic, may even be antagonistic later or *vice versa*.

Hegreness *et. al.,* suggest that antagonistic combinations of antibiotics may have unrecognised potential, due to the lower selection for resistance they postulate [54]. Imagine a bacteria under an antagonistic treatment; it would have a higher fitness (less inhibition) than the same bacteria treated with either of the drugs in mono-therapy. In such a case a mutation giving resistance to one of the antibiotics would reduce the bacterial fitness, by removing the antagonistic element of the treatment. Therefore there is a selective pressure to lose such mutations as fitness is negatively impacted. This is demonstrated in Figure 2.1, whereby the highest rate of adaptation is in the most synergistic region, with antagonistic displaying lower rates of adaptation than the single drug environments. This theoretical framework was tested in the paper mentioned, showing a slower adaptation when antagonistic combination were used. The figure below shows the results presented by Hegreness *et. al.,* [54], demonstrating the highest rate of adaptation in the most synergistic treatment region.

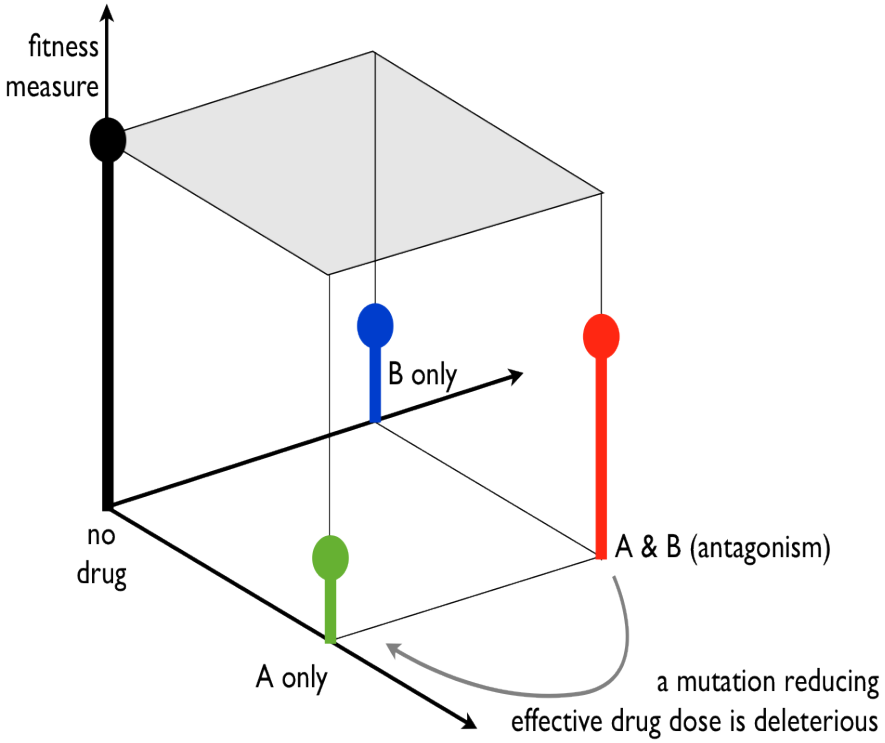


Fig. 2.1: Different drug pairs vary profoundly in their impact on the rate of adaptation. (A and B) For two pairs of drugs (A, synergistic ERY-DOX; B, antagonistic CIP-DOX), the initial level of inhibition is shown for a matrix of concentrations of the two drugs. The level of inhibition is defined as 1 - r/r0, where r is the growth rate of the population in the presence of antibiotics and r0 is the drug-free growth rate. The solid line corresponds to the line of 50% inhibition and is also shown on a linear scale in the insets. (C and D) The rate of adaptation for the drug combinations shown in A and B. The arrow in C points to a region of drug concentrations where the rate of adaptation for the synergistic ERY-DOX combination is accelerated relative to the single-drug treatments. This acceleration is surprising because the more expected outcome of combining drugs is for adaptation to slow down [54].

This is a compelling picture, which would indicate that antagonism could be used to minimise the emergence of resistance. There remains the practical matter of ensuring that the antagonistic treatment is still sufficiently potent to act as an effective treatment; there being little point in minimising resistance if it comes at the cost of treatment failure. Equally, this fitness landscape is making assumptions that mutations providing resistance are the predominant mechanism for resistance, whereas other mechanisms, such as up-regulation of efflux, could provide protection against both drugs. Likewise, this hypothesis takes no account of cross-resistance, which would undermine the argument for using antagonism to select against resistance. It may be the case, therefore, that other drug combinations with less robust antagonistic profiles may not reliably slow adaptation.

2.4 Antagonism and synergy are not well defined

The utility of antibiotic for treatment can only be guaranteed if it is robust or stable through time. Indeed, it has been noted that synergy can be lost as the antibiotic molecules decay due to either host metabolism or instability of the molecules themselves, resulting in an 'inversion' of selection for resistance [92] because drugs can decay into metabolites that promote pathogen growth. Before being able to test the hypothesis that antagonistic treatments may slow the rate of adaptation to antibiotics, it is important that the interaction between the two agents is robust and stable, otherwise a bias is introduced. We have demonstrated through a simple mathematical model that synergy can invert to antagonism, and thereby determined to test the interaction of Rif-Dox in a variety of conditions to determine the robustness of the interaction.

2.4.1 Testing the robustness of antibiotic synergy and antagonism

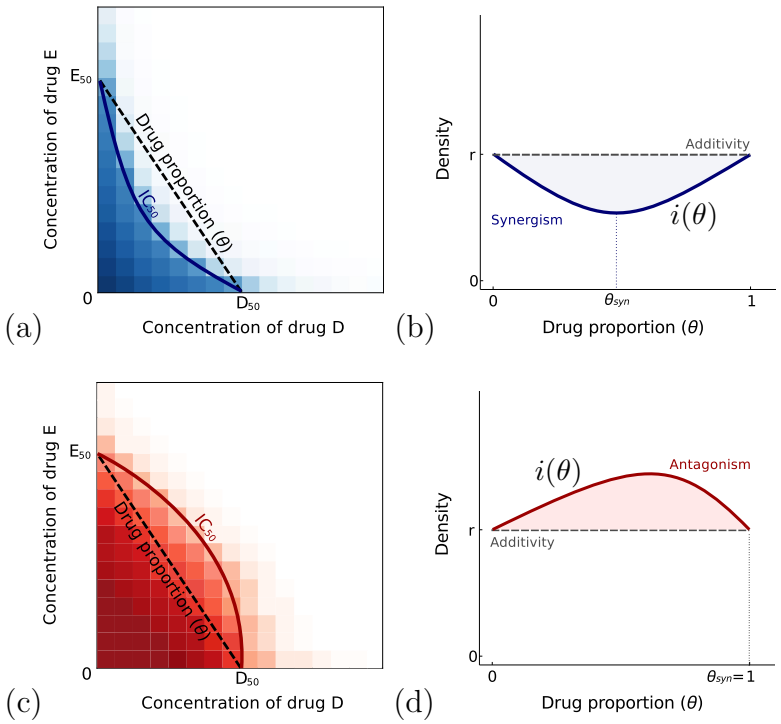


Fig. 2.2: The drug interaction profile is related to a 'checkerboard' diagrams shown in (a) and (c). The concentration of both drugs is given on the x and y axes, bacterial growth inhibition (or population density or some other fitness measure) is then plotted on the z axis. The contour of all concentrations that reduce this measure by half is an *isobole* here denoted IC_{50} and figures (a) and (c) show two checkerboard plots viewed from above. Basal concentrations of both drugs that achieve the same inhibitory effect in this illustration are D_{50} and E_{50}, θ then parameterises the equidosage line between these two values. The fitness measure evaluated along this line is shown in (b) and (d) and we define the degree of interaction based on this curve, this is $i(\theta)$. We say the interaction is *synergistic* when the drug proportion that minimises $i(\theta)$ satisfies $0 < \theta < 1$ as in (b), we denote the resulting value by θ_{syn}. In (d) the drugs are said to be *antagonistic* as $i(\theta)$ is maximised by a drug combination.

Establishing synergy, both *in vivo* and *in vitro* can be difficult. For example, different clinical strains of MRSA [12] show characteristics associated with both synergy and antagonism when treated with vancomycin plus rifampin depending on the times they were measured, whether 24h or 48h. However, there has been no explicit recommendation not to treat MRSA with this combination, just that caution should be exercised when deploying them in the clinic [104]. Furthermore, even for the most common interaction measure, the FICI (see page 28 for mathematical definition), there is a lack of consistency in applying breakpoints for defining an interaction as synergistic, antagonistic or indifferent. Authors have previously used widely varied

interpretations of FICI scores, from $< 1 = synergy, > 1 = antagonism$ [45] to the more common definition of $< 0.5 = synergy, 0.5{-}4 = additivity, > 4 = antagonism$ [81]. This situation has led to the possibility of drug pairs being simultaneously classified as, for example, synergistic and additive, depending on what cut-off is used. A further problem with FICI scores as a means of quantifying interaction is that the design of the checkerboard assay can affect the reported interaction. This was observed as long ago as 1987, where the assignment of interaction was different, depending on whether a checkerboard method was used with 2-fold dilutions or customised dilutions within a smaller range of concentration [55].

To begin with, suppose that $B(t; A, B)$ denotes the density of a bacterial population at time t, measured in h, when grown in an environment containing two antibiotics at initial concentration A and B in $\mu g/ml$. These two values will be called 'basal concentrations' in the remainder. To ensure a fair comparison is made of the efficacy of different drug combinations, it is necessary that each basal drug concentration, A and B, is normalised to achieve equal inhibitory effect over some fixed period of time, T say. We therefore assume that each monotherapy achieves a factor-r reduction in bacterial density relative to a drug-free environment by time T. This condition can be written

$$B(T; A, 0) = B(T; 0, B) = rB(T; 0, 0).$$

This definition is arbitrary because T could represent any time during the entire course of treatment, nevertheless it will be important to have one time over which the two drugs can be compared in a fair manner. In practise, we take T to be either $12h$ or $24h$. B will increase with increasing T, up to the point at which resources become exhausted and the cells enter stationary phase, at which point the increase in B will stop.

With A and B as basal drug concentrations, we now introduce a parameter that allows us to combine both drugs whilst maintaining a constant effective dosage. This parameter, denoted throughout by θ, takes values between zero and one inclusive and it can be used to represent any drug combination along the so-called *equidosage line* (see figure 2.2). For each value of θ, the actual concentration of drug deployed to the environment of the bacteria is then $\theta \cdot A\, \mu g/ml$ of one drug and $(1-\theta) \cdot B\, \mu g/ml$ of the other.

The degree of interaction between drugs is usually defined in terms of the deviation from a neutral interaction, derived either using Bliss independence or Loewe additivity (see Definition 2 and 3, page 34-35) (for an extensive discussion of this topic see [49]). Here we will use the latter and define synergism based on the so-called *drug interaction profile*

$$i(\theta) = B(T; \theta A, (1-\theta)B),$$

defined for $0 \leq \theta \leq 1$.

A measure of the convexity and concavity of $i(\theta, T)$ obtained by fitting a quadratic,

$$q(\theta) = \alpha\theta^2 + \beta\theta + \gamma,$$

can be used to assess the nature of the drug interaction. Significant *positivity* (obtained using a t-test) of α indicates *synergy*, negativity indicates antagonism. If the density data is significantly nonlinear as a function of θ, meaning $\alpha \neq 0$, the fitted quadratic can be used to robustly estimate the drug proportion that maximises and minimises bacterial density at each time. This proportion is given by one of $\theta = 0, 1$ or $-\beta/(2\alpha)$ depending on which value is the lowest of $q(0), q(1)$ or $q(-\beta/(2\alpha))$. Provided $-\beta/(2\alpha)$ lies between 0 and 1, an approximate upper bound on the confidence interval for this optimal value can be found from a t-test that returns confidence intervals for α, β and γ. Throughout we will refer to the test described in this paragraph as the 'α-test' and it is implemented using the regression facilities in the Statistics Toolbox of Matlab.

Definition 1 (IC_x)**.** *The inhibitory concentration-x (IC_x) of a bacteriostatic antibiotic used to treat a bacterial population is the minimal concentration required (stated in $\mu g/ml$) to reduce the population density at 24h to x% of the growth observed when no antibiotic is administered. (In practise the IC_x concentration will be calculated as the mean concentration required to inhibit x% of normal drug free growth at a given time across several experimental replicates.)*

Definition 2 (the Loewe drug interaction profile: $i(\theta)$; synergy, antagonism and additivity)**.** *A drug interaction is said to be* synergistic *if, for all θ between zero and one exclusive, the effect of the drugs combined is greater than the sum of effects produced by each drug separately:*

$$i(\theta) < \theta \cdot i(1) + (1 - \theta) \cdot i(0). \tag{2.2}$$

It follows by construction that $i(1) = i(0) = rB(T; 0, 0)$ and so (2.2) asks that treatment efficacy is greatest when drugs are combined. Note that inequality (2.2) is necessarily satisfied if $i(\theta)$ is a concave function of θ. The drug interaction is said to be antagonistic *if the reverse inequality applies in (2.2); this will apply if $i(\theta)$ is a convex function of θ. Under these assumptions the drug interaction is said to be* additive *if $i(\theta)$ is a constant and therefore independent of θ. Examples of convex and concave $i(\theta)$ can be seen in (2.2).*

Bacterial inhibition due to drugs is measured here over a time interval of total length T hours. On occasion we will therefore introduce time, T, explicitly into the interaction profile and write the latter as a time-dependent interaction $i(\theta, T)$.

Definition 3 (Bliss independence)**.** *Bliss independence assumes that drugs in combination act independently of one another, therefore two drugs both inhibiting 50% of growth, would in combination inhibit growth by 75% (*$1 - 05 * 05 = 0.75$*). Positive or negative deviations from this independence would presume synergy or antagonism between the drugs.*

The issue of the robustness of measurements of antibiotic interaction are important when the results of such tests may influence clinical regimens, or dictate best practice. Interaction must be considered in light of environmental context, such as carbon resource level or drug concentration, as well as over a period of time. An interaction that is not consistent within these parameters may still be useful for combating infection, or slowing resistance, but the range of interactions must be understood if the drug pair is to be used effectively.

2.4.2 Mathematical models of synergy loss

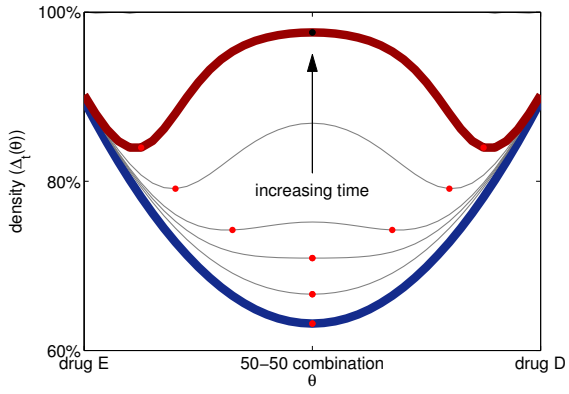


Fig. 2.3: Solving equation (2.3) and plotting population density against drug proportion shows that a short-term synergistic combination (blue) can appear to be antagonistic later because of a change in population structure due to selection for resistance (red). Note $\Delta_t(\theta) := S(t) + R(t)$ is shown on the y-axis.

2.4.3 Synergy loss from population structure and selection for drug-resistance

It is known that changes in population structure can lead to the loss of synergy in drug interaction profiles because of selection for drug-resistant alleles. To see this, consider the following simple example of the phenomenon, by examining a theoretical model taken from [95]. The following is the logistic growth equation modified to include antibiotics:

$$\frac{d}{dt}S = S(1-(S+R)) - (k(\theta)+\mu)S, \tag{2.3a}$$

$$\frac{d}{dt}R = R(1-(S+R)) + \mu S, \tag{2.3b}$$

where $0 < S(0) \ll 1$ and $R(0) = 0$, we therefore begin with susceptible cells but no resistant ones. S represents drug sensitive cells and R is resistant cells. μ denotes the change from S to R representing adaptation to the antibiotic. This model therefore allows for growth and adaptation to the antibioitc. If μ is set to 0, then the smile-frown inversion does not occur and synergy is maintained. Here $k(\theta) = 1 + \theta(1-\theta)$ represents the antibiotic synergy because it is a quadratic function minimised when $\theta = 1/2$. figure 2.3 shows the population densities that result from this model plotted as a function of θ for increasing values of time t.

For short times (2.3a-b) exhibits synergy because growth is suppressed most by the combination where $\theta = 1/2$, so the plot of $\Delta_t(\theta)$ has the convex, U-shaped 'smile' shown in blue in Figure 2.3(b). At later times, but only provided $\mu > 0$, the shape of the profile changes and now density is *greatest* for the 50-50 combination

and lowest for the 'monotherapies', where $\theta = 0$ and $\theta = 1$. So the plot of $\Delta_t(\theta)$ now exhibits a near-concave, W-shaped 'frown' consistent with antagonism having its maximal value at $\theta = 1/2$, as shown in red in Figure 2.3(b). The passage from synergy to antagonism is called the *smile-frown transition* [95].

2.4.4 Antagonism and synergy at different dosages in a pathway

Metabolic adaptation has been proposed as one reason why synergy and antagonism might be found within a population of human cells [101, 103]. However, we propose that the mode of action of rifampicin and doxycycline can lead to the reporting of both synergy and antagonism as the interaction of this drug pair, simply by changing the concentration of drugs used. Rif and Dox were chosen for this study due poor agreement in the literature as to the nature of the interaction between the two. Furthermore, this combination is an important clinical combination as it is the first line treatment for *Brucellosis mellentisis*.

$$X_{in} \rightarrow X \xrightarrow[\downarrow P']{\overset{A}{\downarrow} E_1} P \xrightarrow[\downarrow Y']{\overset{B}{\downarrow} E_2} Y$$

Fig. 2.4: Rifampicin is represented by the inhibitor 'A' and doxycycline is represented by the inhibitor 'B' where transcription and translation are viewed as enzymatic processes acting in serial. E_1 and E_2 represent RNA polymerase and the ribosome, therefore they will form complexes with the input material (wither DNA or RNA, represented by X and P)

A conceptual model of the action of rifampicin and doxycycline is depicted in figure 2.4. The former is a transcriptional inhibitor and the latter is a translational inhibitor and, at least at their coarsest scale, we can choose to view these as inhibitors acting on two processes in serial, as represented in the latter figure. Here a substrate X_{in} is convected physically (through the periplasm, for example) into a region (the cytoplasm, for example) where, now denoted X, it can be processed by two pathways represented *en bloc* by two single enzymes, E_1 and E_2, representing, in this model, enzyme complexes RNA-polymerase and the ribosome, which will both form intermediate complexes with their respective inputs (i.e DNA or RNA). The former enzyme processes X to create a product P (RNA) and a by-product, P' that is processed no further by the model. However, P is then processed by E_2 to form a further product Y (protein) and a by-product Y' (for example, abortive transcripts). The rate of production per unit time of the full product Y is the rate of flux, J, through this pathway and this is inhibited by drugs at concentrations A

and B that bind competitively to E_1 and E_2, forming complexes that cannot process any of the substrates. Thus, J is a function of (A, B) so we write $J(A, B)$ to make this explicit.

To see how $J(A, B)$ has the potential to express different interactions at different dosages, consider the following kinetic model whereby all elements of the model are assumed to satisfy the law of mass action.

$$\begin{aligned}
\frac{d}{dt}X &= -k_1 X \cdot E_1 + k_{-1} \cdot (XE_1) + X_{in}, && \text{(2.4a)}\\
\frac{d}{dt}E_1 &= -k_1 X \cdot E_1 + k_{-1}(XE_1) + k_2(XE_1) - k_3 A \cdot E_1 + k_{-3}(AE_1), && \text{(2.4b)}\\
\frac{d}{dt}(XE_1) &= k_1 X \cdot E_1 - k_{-3}(XE_1) - k_2(XE_1), && \text{(2.4c)}\\
\frac{d}{dt}A &= -k_3 A \cdot E_1 + k_{-3}(AE_1), && \text{(2.4d)}\\
\frac{d}{dt}(AE_1) &= k_3 A \cdot E_1 - k_{-3}(AE_1), && \text{(2.4e)}\\
\frac{d}{dt}P &= k_2(XE_1) - k_4 P \cdot E_2 + k_{-4}(PE_2), && \text{(2.4f)}\\
\frac{d}{dt}E_2 &= -k_4 P \cdot E_2 + k_{-4}(PE_2) - k_6 B \cdot E_2 + k_{-6}(BE_2) + k_5(PE_2), && \text{(2.4g)}\\
\frac{d}{dt}(PE_2) &= k_4 P \cdot E_2 - k_5(PE_2) - k_{-4}(PE_2), && \text{(2.4h)}\\
\frac{d}{dt}B &= -k_6 B \cdot E_2 + k_{-6}(BE_2), && \text{(2.4i)}\\
\frac{d}{dt}(BE_2) &= k_6 B \cdot E_2 - k_{-6}(BE_2), && \text{(2.4j)}\\
\frac{d}{dt}Y &= k_5(PE_2), && \text{(2.4k)}
\end{aligned}$$

where a bracketed expression, as in (YZ), denotes a complex between molecules Y and Z. k_1 and k_{-1} are the association and dissociation rates between X and E_1. k_2 is the non-reversible rate for the production of P from the the complex E_1X. k_3 and k_{-3} are the rates for the association and dissociation of the antibiotic A and E_1. Similarly k_4 and k_{-4} are rates for the formation (and reverse reaction) of E_2P, k_5 is the rate for production of Y. Finally k_6 and k_{-6} are the rates for the formation (and reverse) of the antibiotic B complex with E_2.

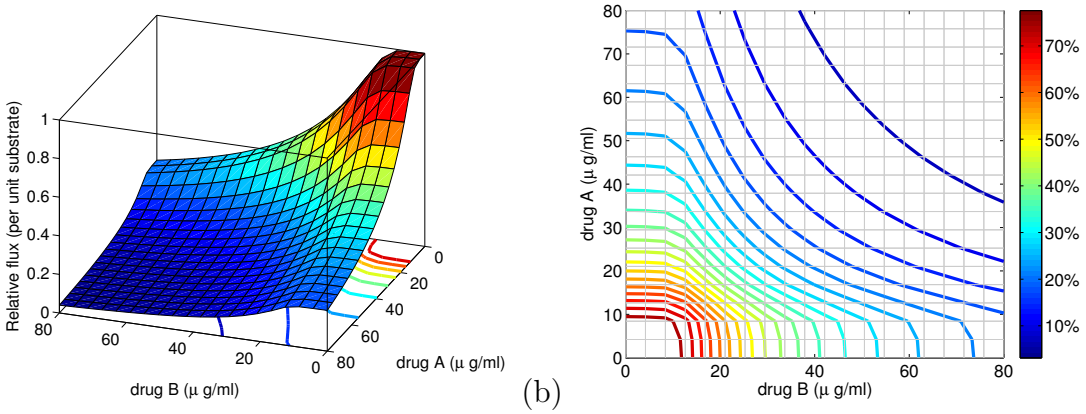


Fig. 2.5: (a) This plot of $J(A, B)/J(0, 0)$ as a surface is a checkerboard whose contours (isoboles) are shown in (b). The geometry of these contours is consistent with antagonism at low dosages but synergy at high dosages.

Rather than solve equation (2.4) analytically in steady-state, a single numerical solution suffices to illustrate our point, as shown in figure 2.5. The latter illustrates that the nonlinear kinetic model (2.4) can predict simultaneously the existence of an antagonism at low dosages while also displaying synergy at higher dosages. It is often assumed that drugs that both competitively inhibit the same target will have an antagonistic interaction, whereas separate sequential targets will produce synergy. However, at low dose an explanation for antagonism as predicted by the model is that as flux is still very high, adding small amount of drug B to drug A (or vice versa) has no noticealbe effect on growth - for this reason the lowest dose curves are almost right angles, giving an antagonistic interaction. As the dose increases, the central section of each curve becomes more linear (additive).

2.4.5 Empirical evidence that the rif-dox interaction is different at different doses

When Rif-Dox combination was tested (using 5 treatments; both mono-therapies, as well as 25:75, 50:50 and 75:25 combinations) at the basal concentration achieving a 50% inhibition, the resulting interaction is not stable over the course of a 24 hour growth period. In brief, strains of *E.coli* AG100 were grown in M9 minimal media supplemented with glucose at 0.2% and casamino acids (0.1%), as well as a single or combinatorial drug treatment. Full methods are available in Appendix 1.1 (pg 172-173). As Figure 2.6 shows, what starts as an antagonistic combination inverts to synergy within 18 hours. For low sub-MIC concentrations tested this result was repeated at IC_{60}. However, when concentration was increased to IC_{70} and IC_{80}, the inversion of antagonism becomes less clear, with the interaction ending in additivity.

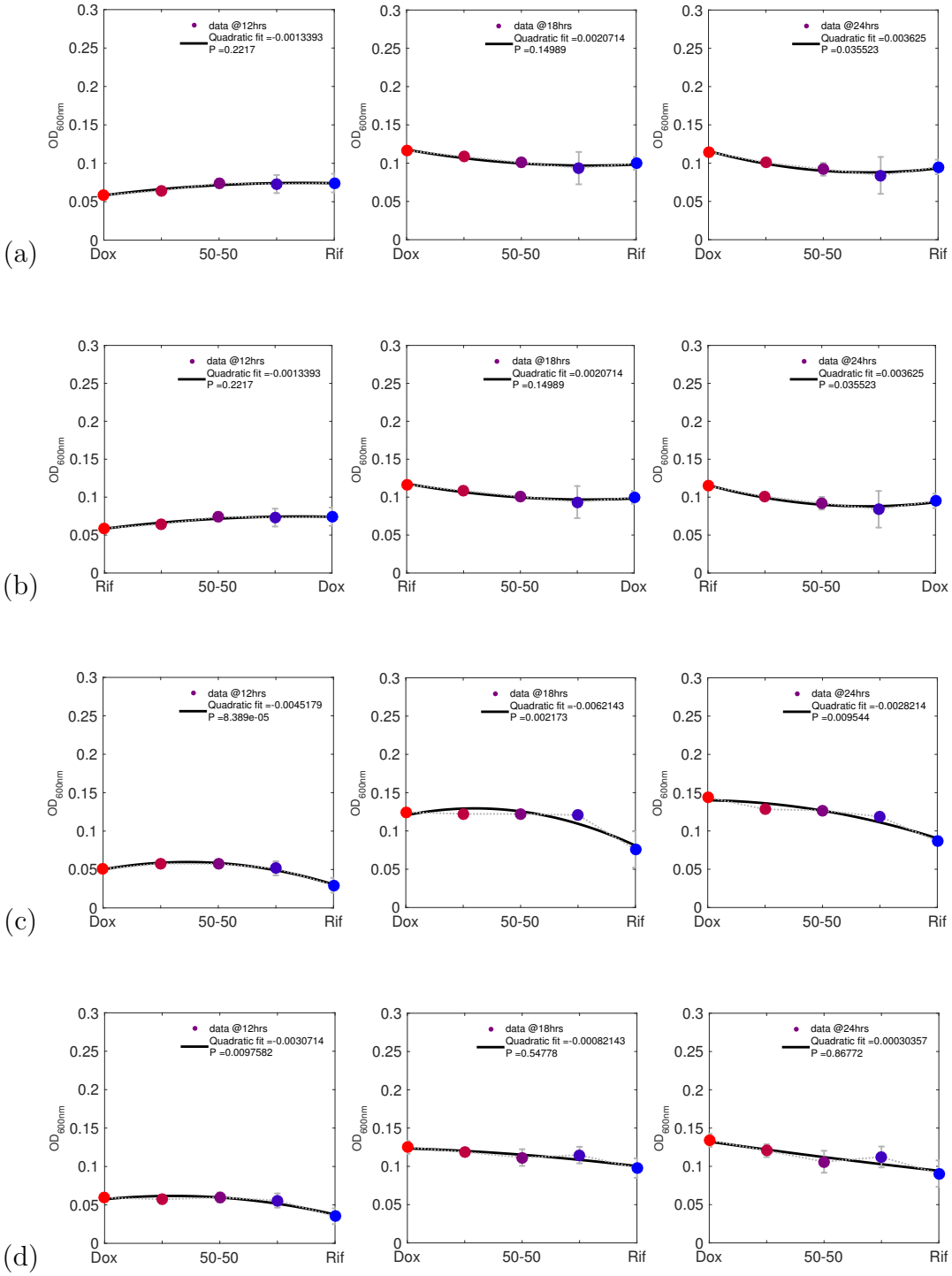


Fig. 2.6: (a) Drug interaction between Rif and Dox at varying basal mono-therapy concentration at IC_{50}. (b) at IC_{60} (c) at IC_{70} (d) at IC_{80}. 0.2% glucose minimal media was used in each case. Data point colours describe proportion of Dox - Rif (Blue - Red).

2.4.6 Empirical evidence that the Rif-Dox interaction is different at different glucose concentrations

The available resources for bacterial growth also play a role in the interaction of the drugs with the bacterial populations. Realistically, within nature the concentration of resource is certain to vary over time, as a result of diffusion of nutrients and

variable intake of energy (for example, resource concentration available to bacteria in the human gut will peak after mealtimes, and fall to a minimum during the night). In order to assess whether the non-robustness of drug interaction occurs at different resource concentrations, the test was repeated with higher glucose concentrations (Appendix 1.1, pg. 173). Having seen an inversion of synergy at low dose, in a 0.2% glucose minimal media, we tested the same drug combination and organism, using a range of glucose concentrations from 0.4% to 1%. Figure 2.7. High glucose environments seem to stabilise the initial antagonism, as can be seen in the 1% glucose background. However, in other glucose concentrations, there is no inversion to synergy. It is important to point out at this point the limitations of the method for determining interaction based on a quadratic fit, namely that unless the drug pair is accurately calibrated, the quadratic fit may return a statistically significant fit indicating synergy, and yet not actually be showing an increased efficacy compared to the mono-therapies. This is seen in Figure 2.7(b); in the right most plot, the p-value of the quadratic fit indicates significant synergy, but because the *E.coli* has adapted more rapidly to the Dox, every combinatorial treatment is actually less inhibitory than the Rif mono-therapy. For a drug to be truly synergistic, the combination must be more inhibitory than both the mono-therapies.

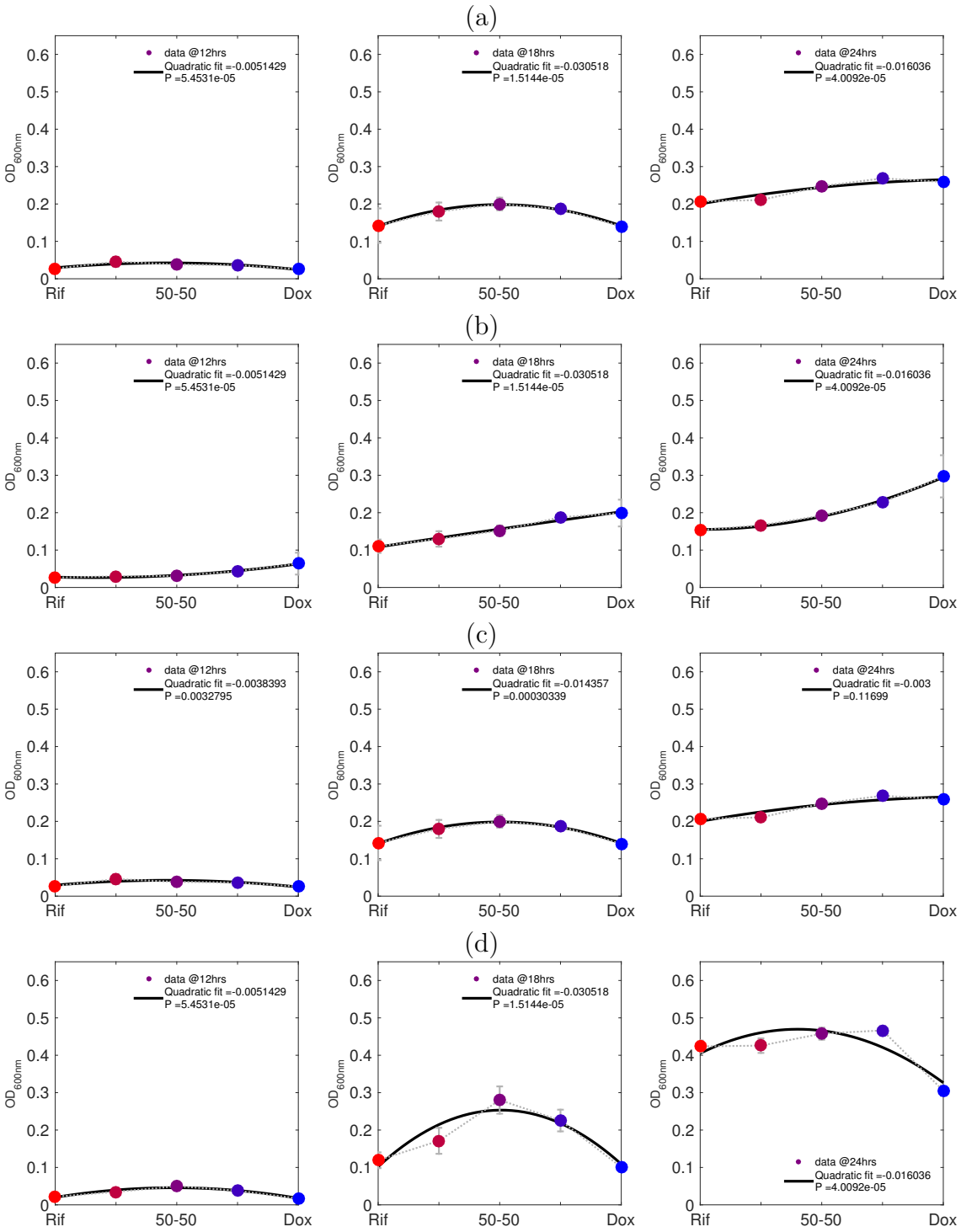


Fig. 2.7: (a) Drug interaction between Rif and Dox at varying concentrations of glucose in minimal media at 0.4% glucose. (b) at 0.6% glucose. (c) at 0.8% glucose. (d) at 1% glucose. Basal IC_{50} mono-therapy dose was used in each case.

The changing nature of the interaction between rifampicin and doxycycline is clear in Figure 2.6 and Figure 2.7: when the basal concentration of the drugs or the available glucose concentration is changed, the interaction can be altered and be reported as either synergy or antagonism. At low doses, the drug interaction can be said to invert from antagonism to synergy and this can occur within a 24h period. This pattern, however, is not seen at either the higher dosages or higher glucose

concentrations that we tested. The same results are observed in a secondary pair of predicted antagonistic drugs, Ery-Rif, although at a higher carbon concentration (1% glucose), with the interaction being antagonistic initially (with a significant p-value according to our quadratic fit test) and over 24 hours inverting to synergy (again with a statistically significant negative quadratic fit).

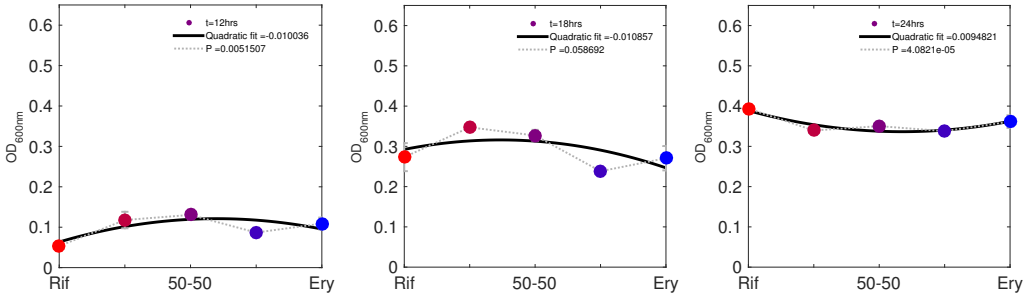


Fig. 2.8: Drug interaction between Rif and Ery, using basal drug concentrations of IC50.

As these results demonstrate, the interaction between these two antibiotics is not stable within a 24 hour growth period, inverting from antagonism to synergy. If this interaction is serially transferred for one week, replenished with fresh media every day, it will be possible to determine if the drug interaction is dynamic, or whether the interaction stabilises. Using just a small section of the drug pair interaction profile, we evolved bacteria in combination treatments using small proportions Rifampicin with larger proportions of Doxycycline, as shown (Figure 2.9).Due to the short timescales involved in these experiments it is unlikely that selection is responsible for the non-robustness of these measures of interaction. It is possible that the dynamics of growth determine this feature, for example a trade-off between rate and carrying capacity would favour high growth rates at short timescales (i.e antagonism), and yet at a later time point would favour the slower growing monotherapies, allowing them to develop higher final density (e.g synergy).

2.5 Antagonism increases adaptation in Rif-Dox

Even at sub-inhibitory concentrations, (IC_{80}) Rifampicin mono-therapy was effective in removing all measurable trace of the bacteria by optical density within 2 days, this is most likely due to the dilution effect of the transfer of around 1% by volume. Populations treated with monotherapy doxycycline however, recovered very quickly to the growth kinetics of the drug free control. Therefore, the intuitive expectation would be that if Rifampicin (a very effective drug) is added in small quantities to a less effective drug (Doxycycline) the overall efficacy of that treatment should be improved. However, knowing that Rif-Dox is an antagonistic combination, we were able to test whether antagonism is capable of suppressing the rate of adaptation with respect to that of Doxycycline. Because Rifampicin is effective at clearing bacteria,

a subsection of the antagonistic curve will be evolved using small quantities of Rif as an adjuvant. Proportions of Dox:Rif of of $99:1, 95:5$ and $90:10$ will be used (Figure 2.9) to ensure that the quantity of Rif does not eliminate all the bacteria. The experiment was continued for 8 days, using a 24 hour serial transfer protocol, details available in Appendix 1.1 (pg. 176).

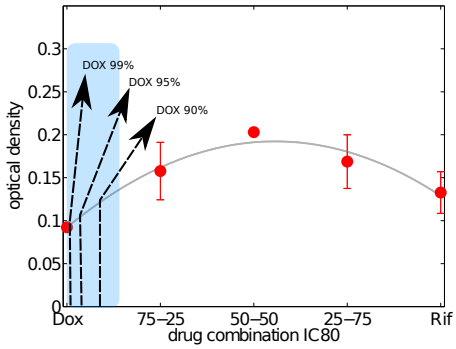


Fig. 2.9: Example data showing an antagonistic interaction. The shaded area represents the proportions of Rif/Dox to be used for the adaptive serial transfer experiment. This region of the curve is to be used because higher proportions of Rif lead to bacterial clearing within 2 days, due to the dilution effect of the serial transfers.

2.5.1 Initial antagonism leads to overgrowth

The following data (Figures 2.10 and 2.11) quantify the stimulatory effect of both the Dox mono-therapy and the Rif-Dox combination. Figure 2.10 shows the yield of the combination treatments, normalised to the yield of the Dox mono-therapy. It shows that whilst all Rif-adjuvants treatments fail to inhibit growth, only the combination containing the largest proportion of Rif (10%) has a stimulatory effect that is significantly greater than the Dox mono-therapy (right-tailed students *t-test*, $p = 0.02$).

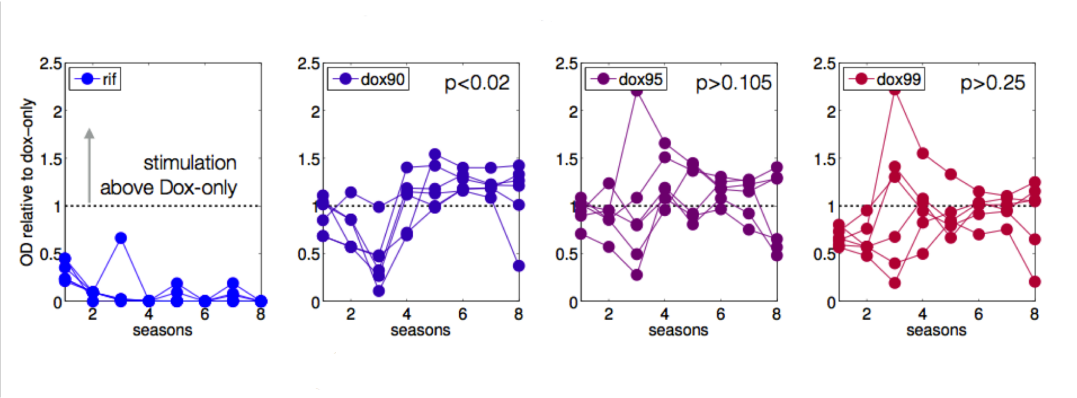


Fig. 2.10: Density of bacteria over 8 seasons (1 season = 24 hours) normalised to the average growth in Dox monotherapy. The combination with the highest proportion of Rif (10%) is the ony condition that shows stimulation above the growth seen in Dox mono-therapy ($p = 0.02$). The left hand panel shows the Rif only therapy, and it can be seen that in most replicates there is total bacterial clearance within 3 seasons. Data colours show the proportion of Rif/Dox from blue to red.

Figure 2.11 shows the results from the same experiment but normalised to the growth of the drug free control. As it clearly shows, all treatments used, with the exception of Rif mono-therapy, have clear and significant stimulatory effect on the final cell density of *E.coli*, within six seasons (in this experiment one season is 12 hours) of adaptation ($p0.001$ in each case).

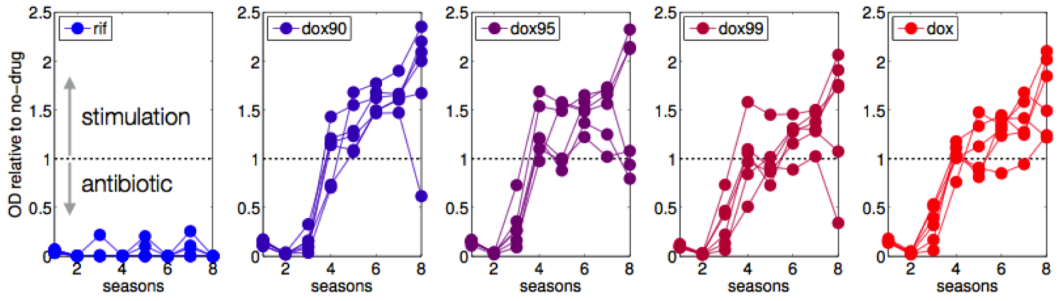


Fig. 2.11: Growth of bacteria in combination treatments, normalised to drug free growth. I can be seen that Rifampicin clears all bacteria in most replicates, whereas all combinations that include Dox have a stimulatory effect compared to drug free growth in most replicates. As in (Figure 2.10) the colours denote the proportion of Rif:Dox from blue to red.

.

2.5.2 Rate of adaptation increases with antagonism

The hypothesis described by Hegreness *et. al.,* is that due to the deleterious effect on fitness of a mutation protecting the cell form one of a pair of combination antibiotics, antagonism will select against resistance. The use of such a treatment is problematic, as the primary purpose of antibiotic treatment must be to treat infection, with prevention of resistance being an important secondary concern. Therefore, an antagonistic treatment, even if completely successful at preventing the emergence of resistance, is useless as a therapy if it fails to reduce bacterial load. Our results however, have shown that the measurement of interaction between drugs will effect the result, i.e., measuring the interaction at 12 or 24 hours can produce opposite interactions. Furthermore, we have shown that in one specific pair, a clinically used combination that is reported as antagonistic in our model organism, adaptation will lead to the most antagonistic combinations having the greatest yield of bacteria within on week. Using the method for quantifying the rate of adaptation from the paper cited above, we can calculate the rate at which the bacteria is adapting (in terms of time taken to reach half the maximum change in growth rate). Results show that treatments with Rif-adjuvant are adapting fastest, in general. Drug free environments and Rif-monotherapy show no adaptation at all, either because of very limited selection pressures (drug free) or because there is no detectable replication in the Rifampicin treatments, so adaptation is zero of necessity. There is a weak correlation between increasing antagonism and increasingly adaptation, leading to a complete failure of the antagonistic treatment, to the point that the drug combination becomes stimulatory, rather than inhibitory (Figure 2.12).

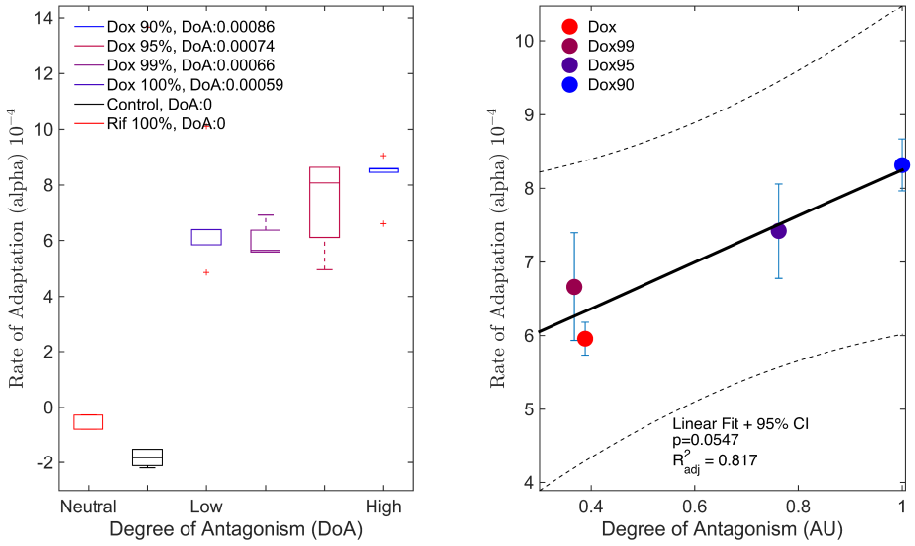


Fig. 2.12: As proportion of doxycycline in IC80 treatment increases, antagonism decreases (pairwise right-tailed t-test between Dox mono-therapy and Dox 90% $p = 0.03$). The left hand panel shows the degree of antagonism, calculated by the difference in density between every condition at season 8, and the mean density of Dox and Rif mono-therapies at the end of season 1. The right hand panel shows the same data for only the Dox containing conditions, with degree of antagonism normalised to the most antagonistic treatment (Dox 90%). Rate of adaptation is calculated using the formula explained in Appendix 1.1 (pg. 178-179).

2.6 Summary

We have used mathematical models and one-day antibiotic interaction assays to establish that the interaction of transcriptional and translational inhibitors can differ between dosages when used to treat *E.coli*. However, the observation of changes in interaction within one day is contingent not only on the presence of antibiotic efflux genes, (such as the multi-drug efflux pump *acrAB*), but whether changes in antibiotic interaction are observed in one strain can depend on the environmental state. By this we mean that antibiotic interactions can be manipulated by changing the concentration of a limiting carbon source.

2.6.1 Effective treatment design using rifampicin and doxycycline

It has been predicted that antagonism selects less for drug-resistant bacterial strains than do synergistic ones [54]. Their idea states that point mutations that yield increases to one of the drugs used as part of the combination treatment have greater fitness-increasing effects in synergistic backgrounds than in antagonistic ones. However, the authors of the aforementioned study have not addressed whether antibiotic antagonism can be used to define an effective anti-bacterial treatment.

In order to begin to remedy this omission, with the caveat in mind that antagonism may not be stable throughout the treatment, we sought to test this with the

following protocol: we implemented a serial transfer protocol to extend a 12h treatment of AG100 to 96h, with two transfers per day (12h per transfer), four different basal dosages and fresh medium used at the start of each transfer.

From this data, we formulate the following hypothesis: Doxycycline monotherapy exhibits greater selection for resistance than Rifampicin, but when small quantities (up to 10%) of Rif are added as an adjuvant, the combination produces the greatest number of bacterial cells over the 96h period. The only treatments that produce no significant bacterial growth throughout the treatment is the rifampicin monotherapy. This leads to total bacterial clearance, despite being a sub-MIC initial dose. Despite the observed antagonism at season 1 of these two drugs, and moreover in disagreement with the theory of [54], these combination therapies have produced more bacteria across all replicates than either of the two mono therapies, exhibiting a greater level of adaptation. The nature of the adaptation that occurs over timescales of 96 hour will be explored in Chapter 4, albeit for a slightly different experiment.

Chapter 3

Stimulatory effect of sub-MIC concentration of Doxycycline

3.1 Overview

Following on from results obtained in Chapter 2 (Figure 2.11), we see that Dox can have a stimulatory effect on the growth of *E.coli* at doses that are initially completely inhibitory. We go on to assess the mechanisms and causes of this stimulation. The main findings of this chapter are summarised below:

1. The overgrowth observed is strain specific to AG100 (a K12-derived strain from S.B Levy).
2. There is no apparent cost of resistance to the growth rate or cell density of the adapted AG100 populations.
3. Although MG1655 adapts to higher concentrations of Dox, it does not overgrow in the same manner as AG100.
4. A novel trade off between lag and yield is observed in the Dox stimulated populations.
5. Sequential treatment with Strep followed by Dox can confer the overgrowing phenotype.

3.2 Introduction

It is readily acknowledged that antibiotic resistance is a major global problem, with an extremely limited novel drug development pipeline meaning that there have been dire warnings made about the future of infectious disease control. The last new antibiotic was discovered in 1997, and not introduced clinically until 2012, however resistance had already been reported by 2006 [63]. This gives an indication as to the

scope of the problem, with resistance emerging so rapidly pharmaceutical companies are unwilling to invest large sums in discovering drugs that either lose efficacy very quickly, or are withheld by doctors to minimise resistance. It is therefore imperative to understand the factors that lead to the emergence of resistance, and any possible strategies that can be used to prevent it's spread. In the previous chapter we saw that drug interactions (often used with the express purpose of reducing resistance) can be unstable based on environmental conditions and timescale. We now suggest that there is a similar paucity of research into the effects of sub-MIC concentrations as a selective pressure, and that such an understanding will be critical in devising strategies for the minimisation of antibiotic resistance. This is due to the fact that agricultural and medical uses, as well the presence of naturally occurring antibiotics in the environment, concentration gradients will ensure that bacteria experience sub-lethal doses of antibiotics. This alongside the potential for antibiotics to change bacterial phenotype, for example through acting as signalling molecule, as well as it's genotype through selective pressure, increases the speed at which resistance emerges and spreads [50].

The dose-response assay is one of the most basic tests that can be performed on cultivable bacteria to determine the concentration on a given antibiotic (or antiseptic, toxin, etc.,) required to inhibit the growth of that bacterial population. There are numerous protocols that all follow the same principles; that an increasing concentrations of drug are inoculated with bacteria and grown under specific conditions. The lowest concentration tested which results in no visible turbidity (usually assessed at 24 hours of growth, although this varies depending on the species begin tested) is regarded as the Minimum Inhibitory Concentration, or MIC.

Despite the fact that this test is ubiquitously used to determine effective dosages, as well as to assess antibiotic resistance, (when compared to internationally collated samples) [42] it provides no information about the robustness of the treatment to adaptation. It likewise ignores the potential for the selection of resistant mutants at sub-MIC concentrations. Figure 3.2 is a schematic representation that demonstrates that even at sub-lethal antibiotic doses there is still a selective pressure being exerted upon a resistant subpopulation, at any point where the inhibition of the antibiotic reduces the wild-type fitness to a level below that of the mutant [6]. Whilst it may seem like an unimportant factor in drug resistance, for example during disease treatment whereby a dose considerably in excess of the MIC is likely to be prescribed, it is worth considering in more detail. Firstly, however high the dose given is, temporal and spatial variations will mean that the drug is not distributed evenly throughout all compartments of the body, or even cellular compartments. This means that antibiotic gradients will exist, and so at some point the range of drug at any given location will fall into this sub-MIC selective window, Furthermore,

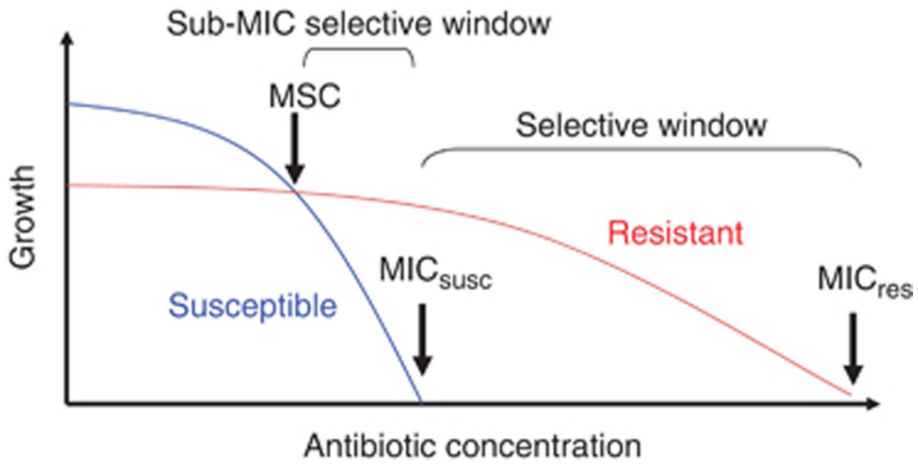


Fig. 3.1: In a region whereby a resistant mutant has a higher fitness than the wild-type strain, a selective pressure will act and increase the frequency of the mutant, even at drug concentrations that are not sufficient to eliminate the susceptible strain. This sub-MIC selective window is often overlooked as a source of resistance. Figure reproduced from FEMS microbiology reviews [6]

the degradation of the antibiotic, whether by active processes intracellularly, or by thermodynamic properties, will again lead to drug gradients, which must include these concentrations. Lastly, although it is an understandable impulse to associate the cost of antibiotic resistance with human health, there are numerous other reservoirs of resistance which may play equally important roles in the emergence of drug-resistant strains, including agricultural use, and environmental contamination (often as a result of excretion of antibiotics into water courses, either from human or animal recipients of antibiotics).

It is with these factors in mind that we decided on the importance of determining the impact of evolution on the shape of a dose-response profile. This information is strangely lacking from the literature, and may therefore be an overlooked clinical factor in preventing the emergence of resistant phenotypes.

3.3 Stimulatory effects of Doxycycline

3.3.1 Evolving an exponential phase dose-response

In order to establish the effect of treatment duration on MIC, a dose response for *E.coli* AG100 in Doxycycline was set up as described in Appendix 2.1 (Materials and Methods), and was propagated for eight 12-hour seasons, through serial transfers into fresh media. Each transfer represents a bottleneck allowing only 1% of the population to survive, and therefore selective pressure should ensure the most resistant cells will come to dominate the population, leading to an increase in MIC.

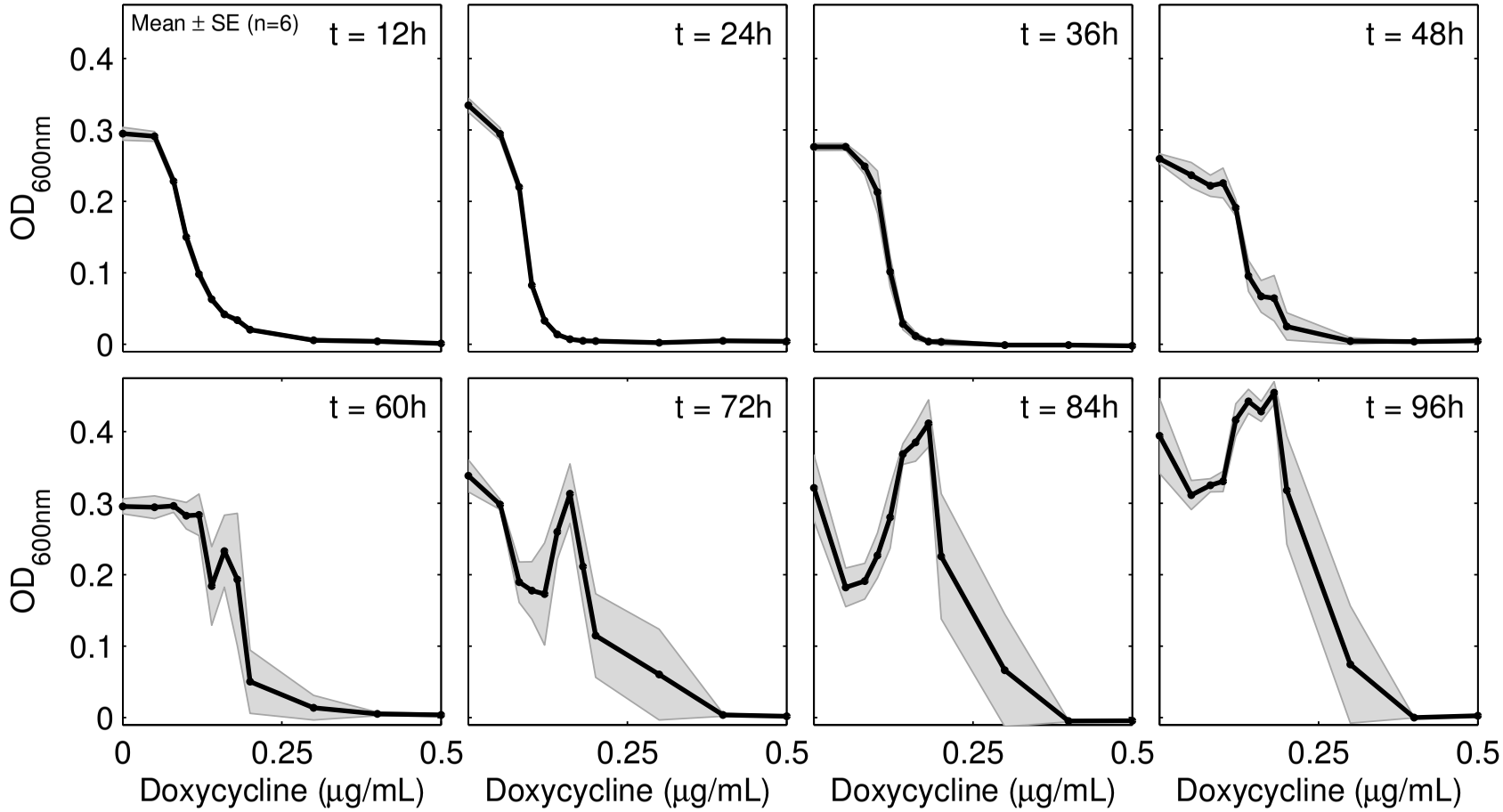


Fig. 3.2: Evolution of an AG100 dose response curve to Doxycycline, in exponential phase, shown by growth curves in Appendix 2.2 (pg. 201). Season one shows a standard monotonic decrease in absorbance as the drug concentration increases. After adaptation, however, there is a drug concentration that strongly mediates high growth, to a level exceeding the drug free control.

As Figure 3.2 shows, the dose-response follows a conventional sigmoidal shape initially, with an MIC of $\tilde{0}.22\mu$g/ml. However, after three seasons the response starts to lose in monotonic form , and by the end of the experiment there is formed a 'hotspot' range of concentrations at which the antibiotic is having a stimulatory effect on the bacterial population size. Despite having observed such stimulatory behaviour previously with Dox and Rif-Dox mixed treatment,this is an unexpected result in the context of commonly held assumptions about the monotonicity of dose response; it is strongly in disagreement of the principle of proportionality between dose and inhibition. In addition, such a result can not be simply explained as the result of the selective pressure at the inhibitory (but still viable) concentrations of Dox causing the emergence of resistance. In a resistance population it is possible that in the absence of a fitness cost, the bacteria could attain the same cell density as the drug-free wild type, but for the density to exceed that of the control indicates that adaptation to Doxycycline has increased the cellular efficiency, and stimulated growth (as measured by final OD).

In order to quantitatively assess the non-monotonicity of growth over the course of the experiment, a Hill function was fit to the data, and the data was subtracted from the subsequent fit estimate. Therefore where the data does not fit to the model, the result will be non-zero. Any concentrations which lead to an overgrowth will have a positive value; Figure 3.3 demonstrates that the non-monotonicity of the dose-response curve develops over the course of the experiment between the concentrations 0.1μg/ml and 0.25μg/ml of Dox. The raw data and Hill-fit that this test was derived from can be found in Appendix 2.2 (Figure B.1), with the methodology described in Appendix 2.1, pg. 180.

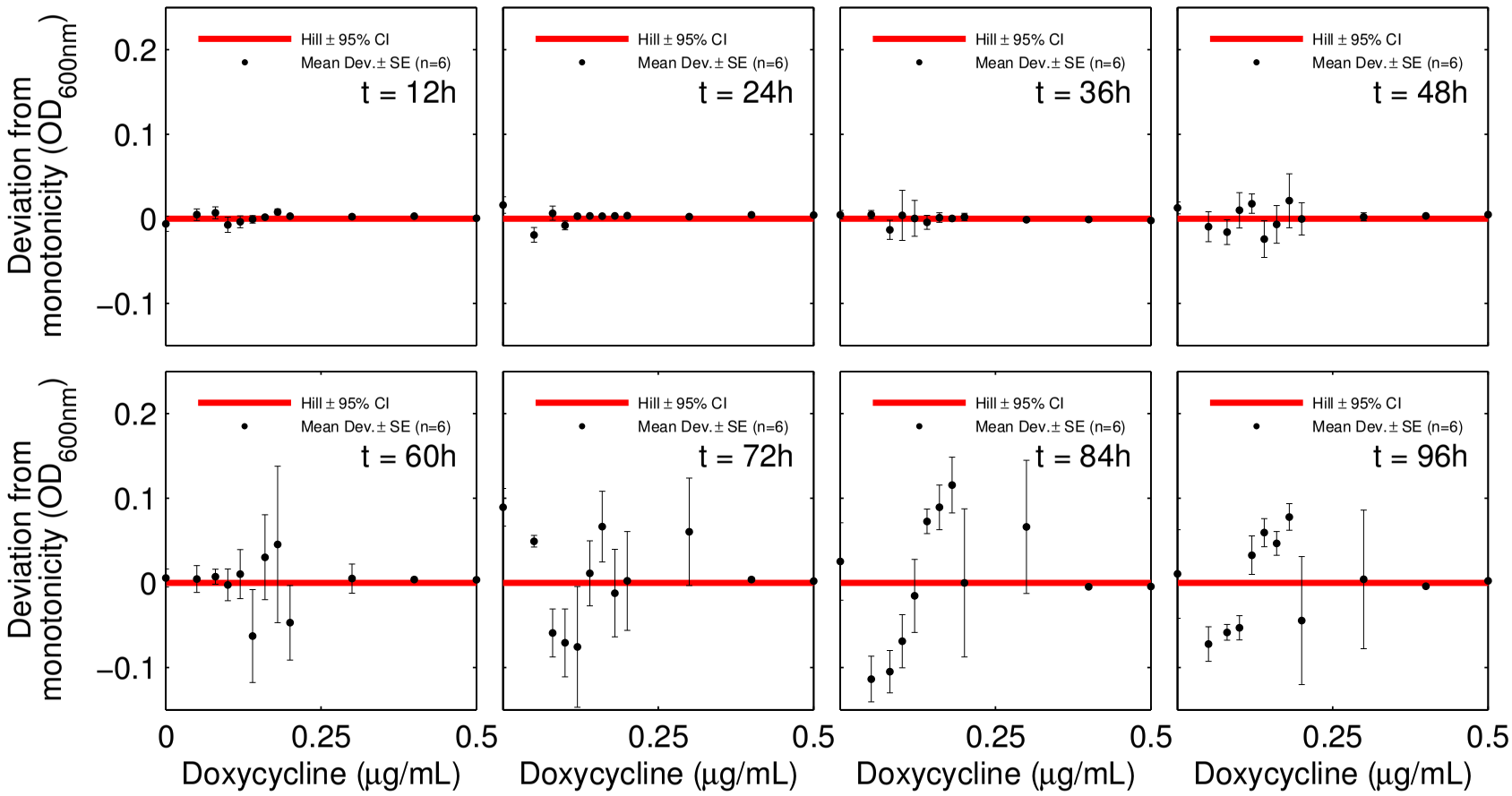


Fig. 3.3: The difference between the estimated Hill-fit and the data gives a measure of non-monotonicity that increases throughout the experiment

Initially the dose-response was measured at 12 hour seasons, rather than the more traditional 24 hour dose response. The reason for this is that as *E.coli* has such a short generation time, in order to select for growth rate, the transfers were made during exponential phase, to provide an evolutionary bottleneck which would favour rapidly growing individuals. To determine whether this was artificially creating this striking overgrowth phenomena, the same dose response adaptive experiment was performed at growth-arrest phase at 24 hours. It can be seen (Appendix 2.2; Figure B.4 and Figure B.5) that the overgrowing phenotype is still clearly evident when yield is selected for, rather than rate. As will be discussed in detail later in this chapter, we hypothesis that as Dox concentration increases, rather than growth rate decreasing, the lag phase is increasing, but growth rate is unchanged (Figure 3.20 and Figure 3.19). In order to ensure that future tests were consistent and to simplify the experimental logistics, we decided to use 24 hour seasons for the remainder of this study, as selection for either exponential or stationary phase populations gave rise to the same phenotype.

To confirm that this unexpected result was the result of a stimulation of growth by Dox, and not a change in cell size or conformation that may lead to an over-reported OD value, the experiment was repeated with a larger number of replicates at just three concentrations (0μg/ml, 0.25μg/ml, 0.5μg/ml) and recorded the optical density for 10 seasons Figure 3.4. The effect was consistent over the course of the ten transfers, and colony counting (see Appendix 2.1; materials and methods) was performed on sample replicates from the final season (*n=3*). CFU counts confirmed a higher viable cell number in the overgrowing MIC environment, compared to the media-only control. In AG100, no growth was seen at the 2xMIC condition. The same larger number of replicates was also grown at 1 and 2xMIC (0.3/ml and 0.6/ml respectively) to determine whether this effect is replicable in a closely related *E.coli* strain. Unexpectedly, the growth dynamics observed were very different, with MG1655 adhering much more closely to the expected dose response. We hypothesise, on the basis of genetic differences between these strains that the reduction in sigma factor produced by AG100 is responsible for this difference in phenotype, however this will be investigated in greater detail on page 80. Inhibition was monotone with increasing drug, with no significant overgrowth at any concentration (Appendix 2.2; Figure B.6); in addition the lag phase length was changed by a lesser degree. MG1655 did adapt more successfully to higher antibiotic concentrations however, with growth seen at 2xMIC in 6 replicates (*n=16*), compared to just 1 replicate in AG100. Other *E.coli* strains were also tested under the same experimental setup, and did not display the overgrowth phenotype of AG100 (Appendix 2.2, pg.195-197).

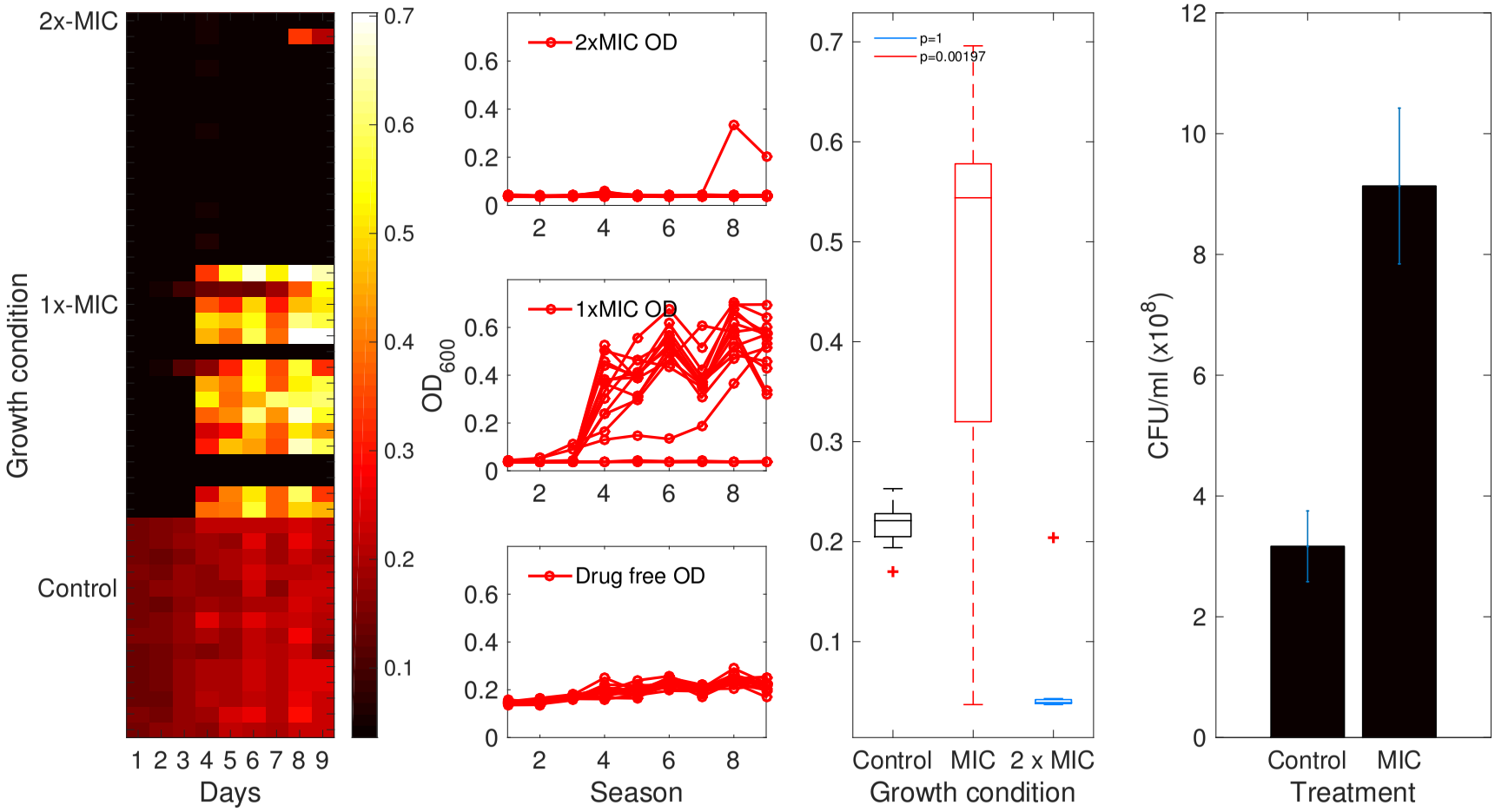


Fig. 3.4: AG100 is shown to overgrow significantly compared to control, by optical density readings, and CFU counts

3.3.2 Effect of selection pressure on adaptation rate

In order to better understand the properties of this overgrowth we applied a measure of the rate of adaptation to assess whether an increasing selective pressure was correlated with an increase in growth rate, and therefore a more rapid adaptation. Because Dox is not bactericidal drug, and has no direct effect on DNA damage, the mutational rate of the cell should remain unchanged despite the drug-challenge. Numerous other processes can affect phenotype, however, including differential gene regulation, selection and increased translocational events [65], and by whatever mechanism, such a striking 'overgrowth' phenotype suggests considerable adaptation. The rate of adaptation (α) is defined as the time taken for the half maximal change in growth rate over the course of the experiment. The calculation on this measurement is described in Figure 3.5. This method of assessing adaptation rate was devised by Hegreness *et., al* in [54] and is effectively a second derivative of the growth rate. Although initial observations show a greater change in lag phase than growth rate as the concentration of Dox increases, the growth rate is still somewhat effected over the course of the experiment, with slight increase seen in the overgrowing samples.

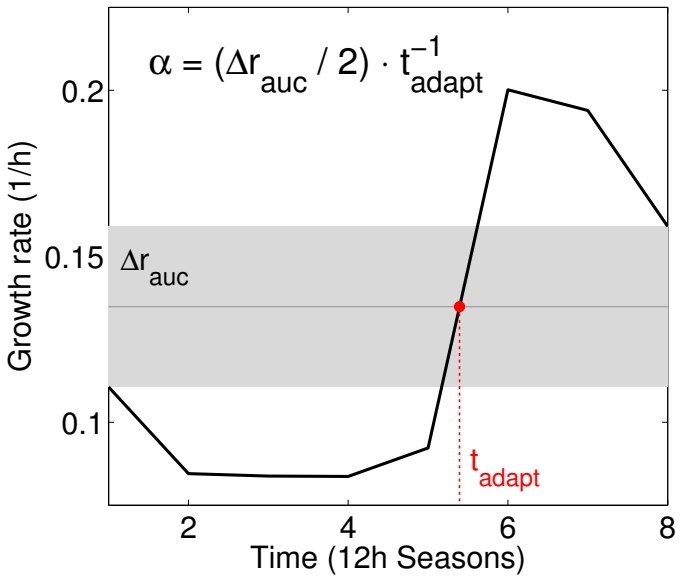


Fig. 3.5: Representative data plot showing the cumulative change in growth rate over the 8 seasons of the experiment. The rate of adaptation (α) is calculated using the formula in the plot. The grey shaded area shows the change in growth rate (rate as measured by area under the curve (AUC).

Several methods were tried to extract accurate growth rate information from the growth curves of the evolved dose-responses shown previously. As described in the material and methods section, linear, exponential and logistic models were fit to the growth curves, and the best fit (as measured by adjusted-R^2) was used to estimate the maximum slope of the curve. However in this method, poor data fits caused by inhibited populations not reaching stationary phase, gave wide variations in rate estimates. To more accurately estimate a rate parameter we instead used the area under the curve method, where $a^n < a^{n+1} < a^{36}$ are OD readings separated by 20 minutes:

$$\sum_{n=1:36} (a^{n+1} - a^n)\left(\frac{f(a^{n+1}) + f(a^n)}{2}\right) \tag{3.1}$$

This AUC is then divided by the final density, and used as a proxy for growth rate, and unlike other methods such as exponential slope measurement, it incorporates all aspects of the growth kinetics, such as the lag phase and stationary phase. This method gave considerably more feasible rate estimates, and is the method to be used in the following calculations of the adaptation rate (α). When this adaptation rate calculation was performed on the data from exponential phase transfer, the rate of adaptation appeared greatest in a hotspot region overlapping highly with the highest OD concentrations (Figure 3.7). It is a further demonstration of the necessity of understanding the changes to dose-response over time, as the sub-MIC selective window is selecting most strongly for rapid adaptation to the antibiotic.

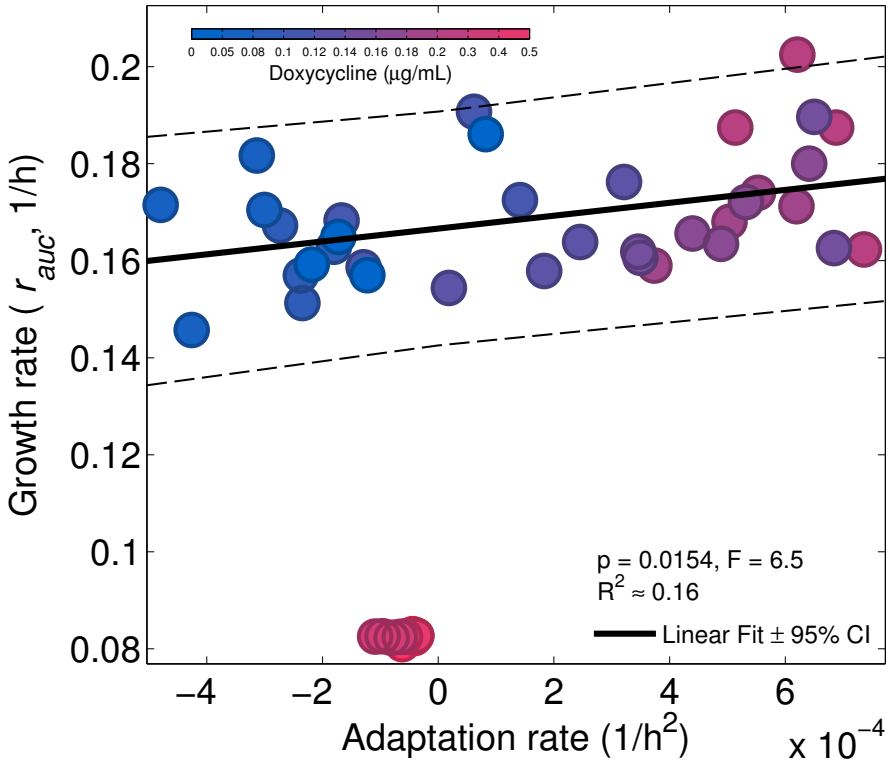


Fig. 3.6: Growth rate shows a slight increase with increasing Doxycycline, but the larger change is in adaptation rate, which increases with [Dox], until the point at which the concentration of doxycycline is again inhibitory.

There is a positive correlation between rate and α, which is unsurprising, as the α figure is the differential of the change in growth rate (Figure 3.6). What is more revealing is, as Figure 3.6 shows, the growth rate increase is only marginal (shown as a function of increasing Dox in Appendix 2.2; Figure B.3, and yet the increase in both cell density and adaptation rate is significantly increased. It is suggestive of another factor being involved in the adaptation to Dox, that is not related to the growth rate, but instead is effecting a different physiological variable such as lag phase. The change in rate over the course of the experiment (used to calculate the rate of adaptation) is shown in Appendix 2.2; Figure B.2 for all replicates and conditions. These results are consistent with the adaptation rate data for 24 hour transfers, although with the slight caveat that in a longer growth window, the growth rate does increase more noticeably with increasing Dox, as measured on season 4. This is likely due to populations reaching stationary phase in 24 hours, giving higher AUC values. This is a potential limitation of using AUC as a method for measuring growth rate, but the underlying pattern of higher adaptation at the Dox 'overgrowth' hotspot remains true (Appendix 2.2; Figures B.7, B.2, B.9, B.10).

Previous studies have shown that the acrAB/tolC efflux system is a vital part of *E.coli* defence against Dox, with the operon containing these efflux pump genes undergoing a duplication during adaptation to sub-lethal doses [13]. When a dose

response was adapted, as above, using the acrAB deleted AG100A strain, there is no overgrowth, and no non-monotonicity of dose response observed over four transfers (Appendix 2.2; Figure B.11). This shows that although cells lacking the efflux genes can still increase their resistance (the MIC moves to the right), efflux is necessary for the overgrowth phenotype to develop.

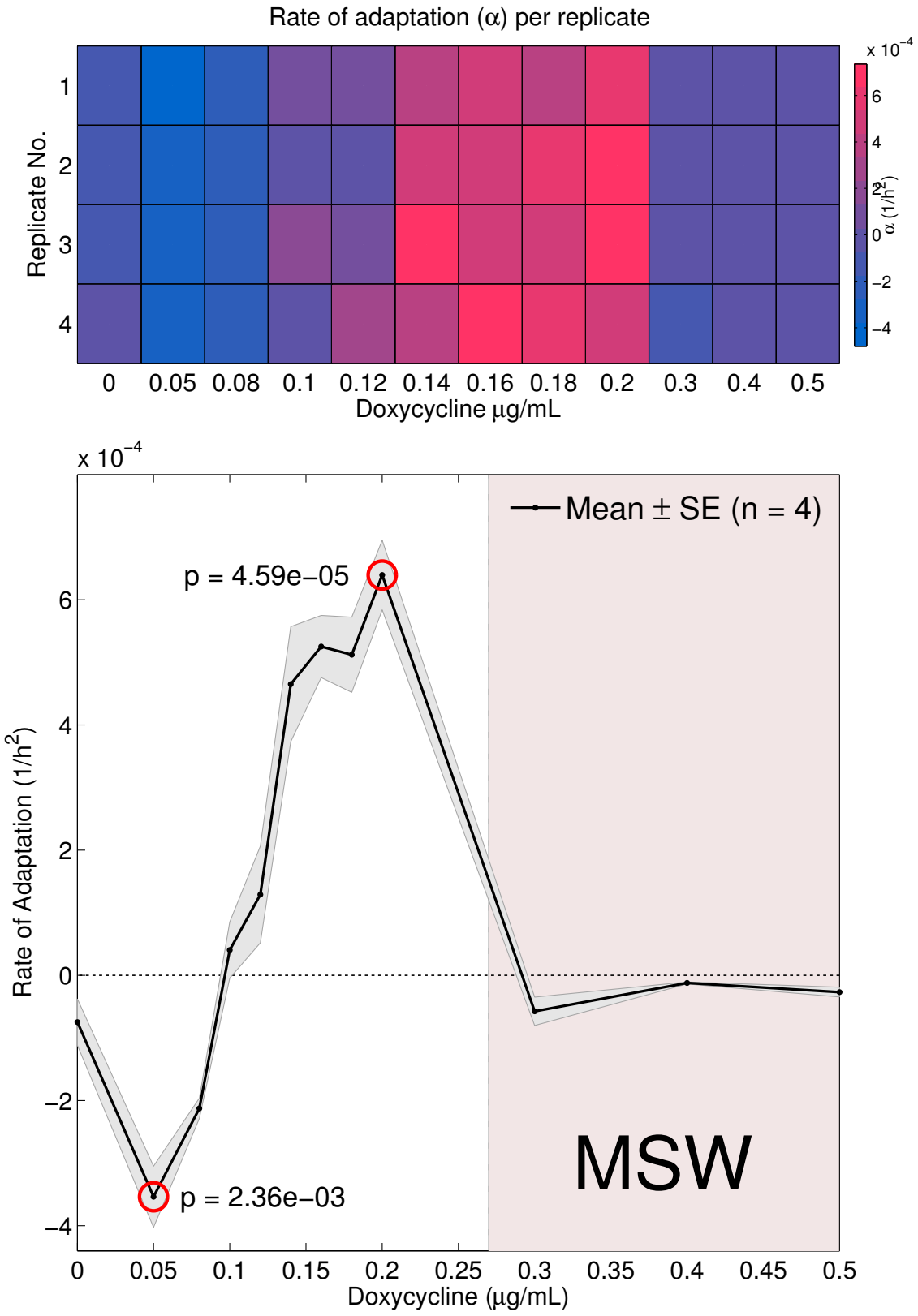


Fig. 3.7: The rate of adaptation does not increase monotonically with increasing selective pressure ([Dox]) but instead shows a 'hotspot' concentration of Doxycycline at around 1x MIC, at which the rate of adaptation is highest. The shaded area shows the mutant selection window (MSW) whereby the only surviving cells will be antibiotic resistance.

3.4 Cost of resistance in overgrowth Dox treatments

3.4.1 No cost of resistance in AG100

The question of the longevity of resistance once the selective pressure has been removed has long been asked, as logic would dictate that such adaptations that are advantageous in one context must be costly in a different environmental condition, otherwise that adaptation would be permanently acquired. Mutational events that can confer resistance can be chromosomal, for example modification of the target molecule, or can involved accessory elements, such as enzymes to inactivate the antibiotics [72]. However the resistance arises, there is compelling logic behind the assumption of a fitness cost, whether due to the increased energy expenditure of producing and maintaining accessory elements, or because a modified target organism, for example increased hydrophobicity in a binding pocket, may negatively effect the primary function of the enzyme, as well as inhibiting drug-binding [7]. As there are still high frequencies of bacteria which remain antibiotic sensitive, despite millions of years of natural microbial exposure, as well as significant levels of antibiotic exposure from human sources over the last 70 years, it would imply that most resistance mechanisms are costly to the bacteria in the absence of antibiotic. Meta-analysis of fitness costs (usually measured by growth rate compared to the wild-type) has shown that in general this is true, with most drug classes reducing growth rate regardless of species tested [82] [107] [16]. Despite a large number of studies aimed at discovering the costs of resistance, it should be cautioned that the absence of an observable cost in laboratory settings does not guarantee that there are no real world conditions in which a significant cost may manifest itself [7], with the inverse situation also being true.

The situation is further confounded by alternative explanations for the maintenance of resistance, of which there are three potential mechanisms. In the simplest case, the resistance may be genuinely cost free, alternatively secondary mutations may be able to compensate for the fitness costs of the resistance mechanism [17]. The final explanation is that a costly genetic change may be maintained in the genome by genetic linkage with a different genetic marker under an alternative selection pressure (this could be a second antibiotic with a separate mechanism of action), meaning that the costly resistance could effectively piggyback on the selection of the secondary marker. These compensatory mutations are difficult to discover, as the complexity of gene regulation and expression is not fully understood even in the most well characterised bacteria, and so it is remains challenging to differentiate between a cost-free resistance, and a cost-compensated resistance phenotype [82].

We were interested to see if AG100 would display a significant cost of resistance when the environment does not necessitate such protection against antibiotics. As

we have already established that the presence of Dox in the environment had a minimal effect on growth rate (Figure 3.6) as measured by AUC, the cost of resistance would be measured by cell number, in the same manner that the original overgrowth was detected and measured. AG100 adapted to MIC Dox for 10 days was transferred to drug-free media, and propagated for a further 10 days, with 1% serial transfers. Dox-adapted AG100 was also propagated in Dox containing media at the MIC, and wild-type cells were grown in both drug free media and Dox media as both positive and negative controls. The full methodology can be found in Appendix 2.1, pg. 182. As Figure 3.8 shows, after 10 days in drug free media, the growth of the Dox-adapted population is considerably higher than both the positive control, and the drug-challenged resistant population. This would imply that not only is there is no cost to resistance, the adaptation to Dox has conferred a fitness benefit to the strain which is maintained, and in fact increases, in the absence of selection.

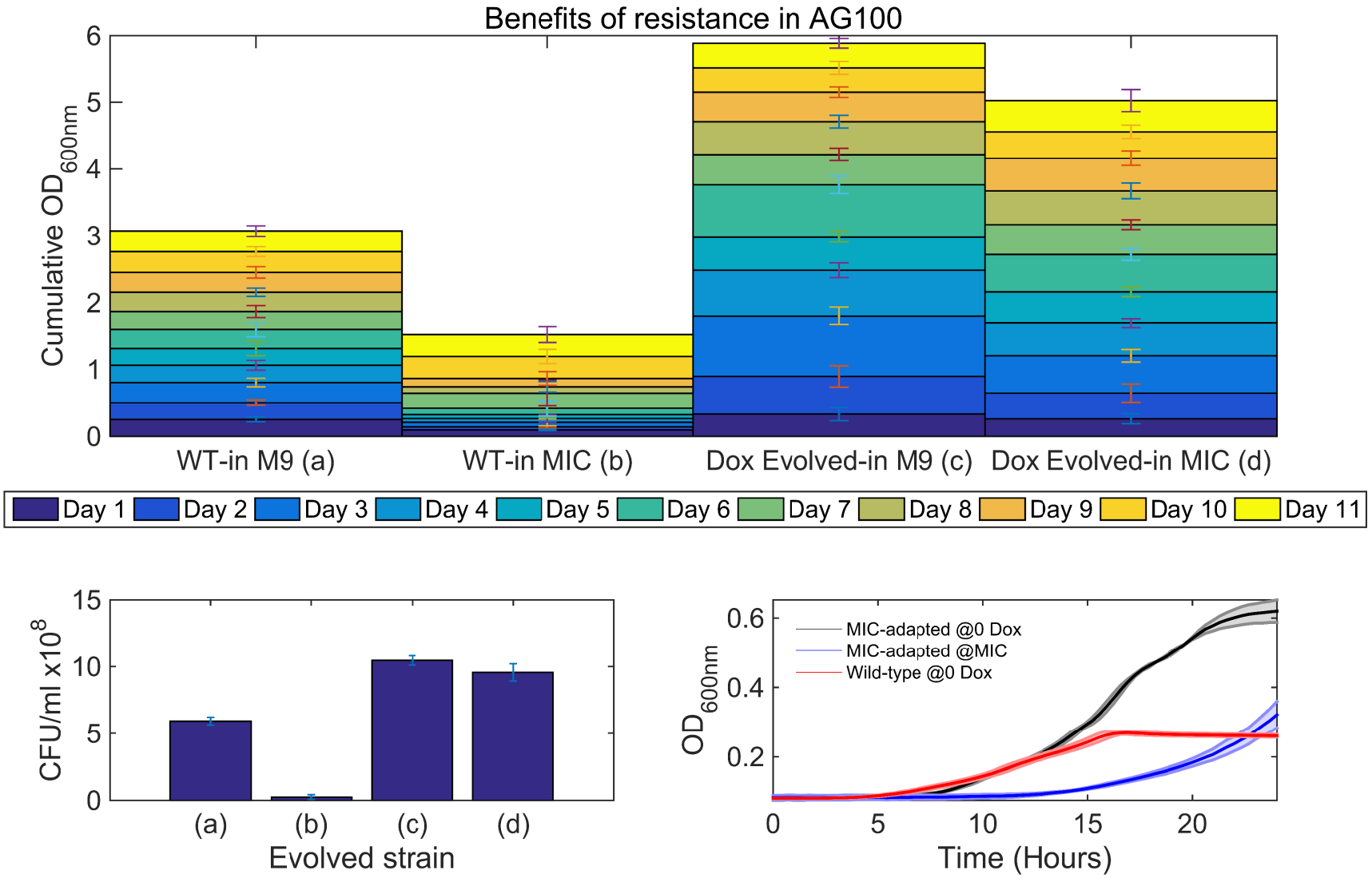


Fig. 3.8: Cumulative OD (top plot) shows that even when the selective pressure is removed, the biomass growth of AG100 exceeds that of both the environment that it is adapted to (Dox) and that of the drug free control. Even by season 11 whereby the dox-adapted growth (in M9) is at it's lowest level, it is still higher than the drug free control (p=*0.00075*)The bottom left plot confirms this overgrowth with CFU count data, taken from 3 replicates of Day 8 growth. Example growth curves are shown in the bottom right.

3.4.2 Cost of resistance present in MG1655

MG1655, despite being very closely related to AG100 and sharing over 99% genetic homology, does not show the same overgrowth when adapted to Dox. However, MG1655 adapts more rapidly to higher concentrations, showing growth at both 1xMIC and 2xMIC within a week of adaptation. Although this adaptation does not confer a boost to cell density, it is interesting that such similar strains can show markedly different responses to drug-challenge. It was therefore of further interest to see if such high-dose resistance would also come with no cost, as in AG100. The same experiment was performed, and as can be seen in Figure 3.9, there is no similarity in response whatsoever. Although the growth of the Dox-adapted population is slightly higher when grown in drug free media, this can not be described as a lack of cost, due to the fact that the cell density reached is still considerably lower than that of the wild-type control. From previous studies we already know that MG1655 adaptation to Dox involves a large genomic duplication of the *mar* operon, responsible for the acrAB/tolC efflux system [13]. Such a large duplication (approximately 15% of the genome) may be the cause of this cost of resistance, as the energetic cost of maintaining a larger genome would mean less cellular resources are dedicated to cell division. The genetic differences between strains will be explored in greater detail shortly, in order to determine whether a basis for the difference in response to Dox can be elucidated. This experiment was carried out exactly as previously described (Appendix 2.1, pg.180).

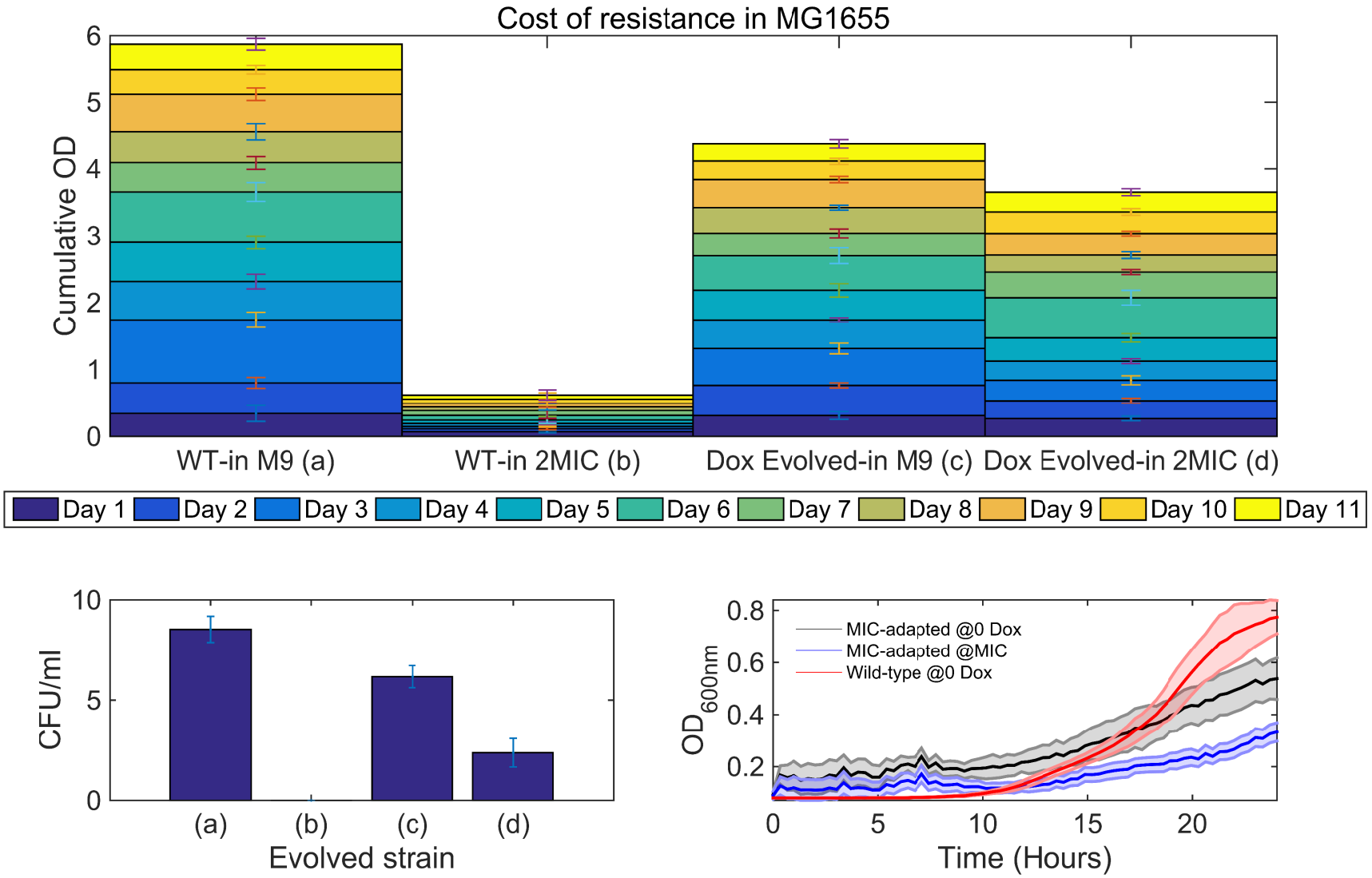


Fig. 3.9: Cumulative OD (top plot) shows that for MG1655 when a selective pressure is imposed, the biomass growth decreases compared the drug free control. Additionally, when the selective pressure is removed the growth increases, suggesting a cost of resistance for this strain. The bottom left plot confirms the OD data with CFU counts, taken from 3 replicates of Day 8 growth. Example growth curves are shown in the bottom right.

3.5 New trade-off present between lag and yield

3.5.1 Doxycycline mediated Lag-Yield trade off (LYTO)

Doxycycline has been previously described as both bacteriostatic and bactericidal in it's mode of action [43] depending on dose. Surprisingly that study found that that as little as a two fold difference in dose could mean the difference between a growth-arrest and a killing effect of the antibiotic. Despite this, Doxycycline is predominantly described as bacteriostatic, as are the other tetracycline class antibiotics. The logic behind this assignation is that because the mode of action is reliant on inhibiting translation, a bacteria that is suffering from this drug stress will still be able to grow (in terms of producing proteins required for replication) but will do so at a slower rate, due to expending energy on the production of abortive transcription products [43].

Bacteria can protect themselves from the inhibitory effects of antibiotics, particularly bacteriostatic drugs, by assuming a drug-tolerant state in which growth and replication is dramatically reduced, or stopped. These non-replicating bacterial cells can then tolerate very high doses of drug, as the level of translation required to maintain in this state is very low, depriving the antibiotic of it's target. The downside of this strategy is that tolerant cells can have a lower drug resistance as there is far smaller chance of acquiring beneficial mutations, either *de novo* or through horizontal gene transfer. However, these mechanisms are merely possible hypotheses to explain the change in growth kinetics observed, a more detailed analysis of the growth kinetics in this experiment will be required.

I had observed during performing dose-response assays using Doxycycline that the growth rate remained largely unchanged regardless of the sub-MIC concentration of drug used (Fig 3.6). The reduction in bacterial load at the end of a given time frame was instead due to a lengthening of lag, proportional to the amount of drug used. Given that in the previous chapter it has been shown that antibiotic treatments, (Doxycycline as well as Doxycycline/Rifampicin mixtures) can lead to stimulation of growth in certain conditions, we were interested to see if there was a trade off between the increase in lag-phase, and the stimulation of growth. Bacteria have been shown to substantially up-regulate genes involved in ribosomal assembly and translation [31], it is therefore a possible mechanism of resistance, that a long lag phase during which ribosomal production is increased, gives protection against Dox by reducing the competitive inhibition for the amino-acyl tRNA binding domain.

As can be seen in Figure 3.10, when *E.coli* AG100 is grown in range of Doxycycline concentrations, as per a standard dose-response assay (see Appendix 2.1; methods and materials), growth is a monotonically decreasing function of drug concentration at 24 hours. However, if the same assay is measured at a later time point (in this instance at 40 hours), the growth no longer decreases monotonically

with antibiotic, instead there is a distinct region in which cell number (as measured by OD_{600nm} and CFU counting) increases as Dox increases. The right-hand plots give a stark example of how in cases where the bacteria suffers an initial inhibition due to the drug (with a significantly longer lag-phase), this can improve the efficiency of the cell, and allow the population to reach an optical density some 50% higher than the stress-free cells. Such a trade off between lag-phase length and cell yield (forthwith referred to as a LYTO) has not been previously reported and is interesting and potentially clinically important, as cells that adopt a drug tolerant strategy will eventually find themselves in a region in which the concentration of the drug has fallen to stimulatory levels. This is inevitable due to the fact that as the drug is metabolised or excreted, or as it diffuses through multiple cellular compartments, there will exist a gradient of concentration over time. Potentially then, application of an antibiotic could lead to the stimulation of subpopulation of bacteria that already exhibit drug-tolerant like behaviours. It is also important to note that although the 0.4μg/ml Dox concentration growth curves at 24 and 48 hours are shown as representative of the trade-off, there are a number of concentrations that show the same behaviour, as can be seen in Figure 3.11.

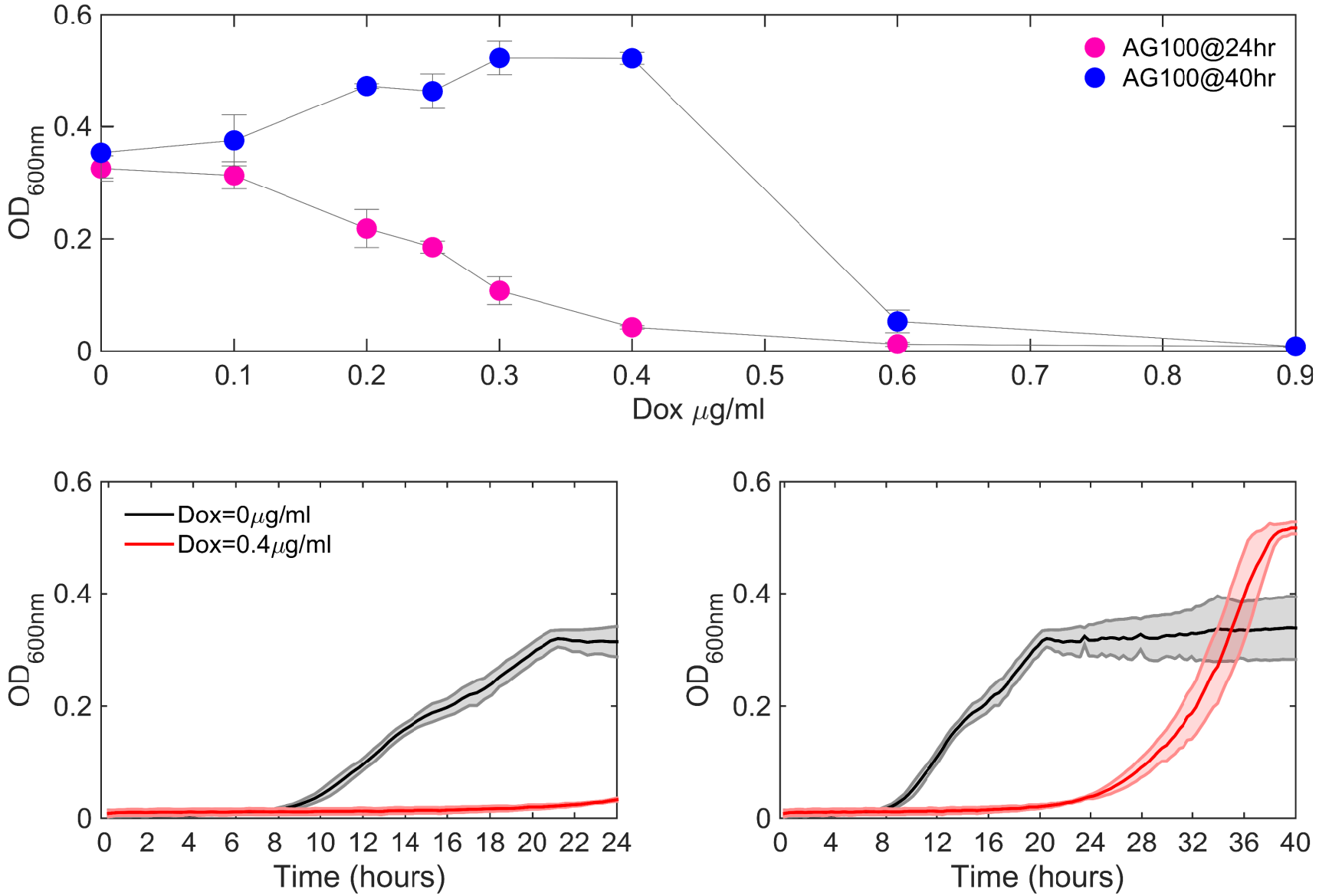


Fig. 3.10: The top plot shows the same dose response curve at T=24hr (pink) and T=40hrs (blue). The bottom plots make it clear that Doxycycline is increasing the lag phase of the bacteria, but that there is a significant trade off between lag and yield, as density at 40 hours is significantly higher than drug free control. Example growth curves at 24hr (left) and 40hr (right) are shown on the right. Potential mechanisms for the increase in yield will be discussed in more detail later in this chapter.

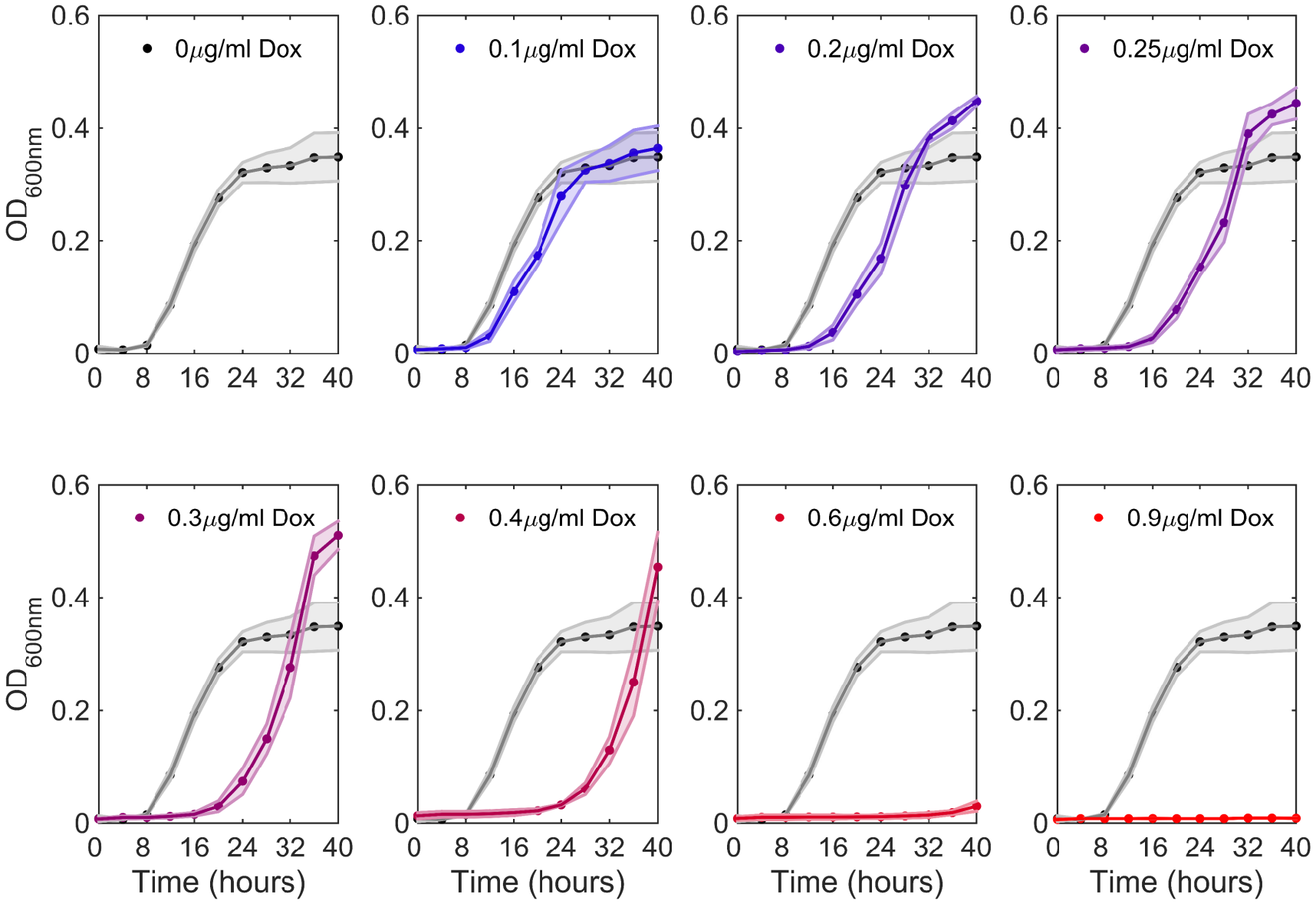


Fig. 3.11: All doses plotted against control, a LYTO is clear in the doses corresponding to the hotspot observed in the adaptive study

3.5.2 LYTO is strain specific

In contrast to these initial results when the same experiment was carried out using *E.coli* MG1655, very little evidence of trade-off was seen between the lag-phase length and the cell yield (Figure 3.13); although a longer exposure time of 40 hours did slightly increase the MIC, there was only weak evidence of stimulatory behaviour at the lowest Dox concentration (Figure 3.14). Furthermore, the increase in lag phase length is not as pronounced as in AG100, and does not seem to confer the same protection to the cells. As Figure 3.14 shows, the rate of growth decreases with increasing Dox concentration; this is as would be expected as Dox is competing for ribosomal binding, and inhibiting protein production. The experiment was repeated for a further five strains (BW25113, ΔaceA, ΔaceB, Δicd and ΔmanZ), again with no observation of such a trade-off (Appendix 2.2; Figure B.13) The rationale behind choosing these five strains to test will be explained in more detail later.

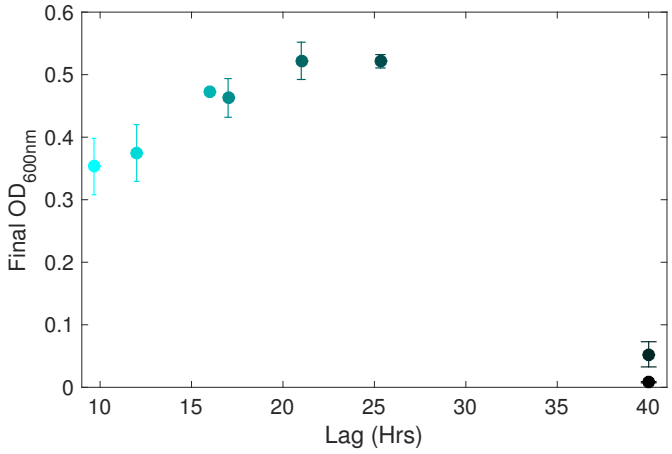


Fig. 3.12: Plot of Lag against Yield for growth curves plotted in Fig 3.11 showing that as lag increases with increasing drug (darker blue dots) the yield also increases.

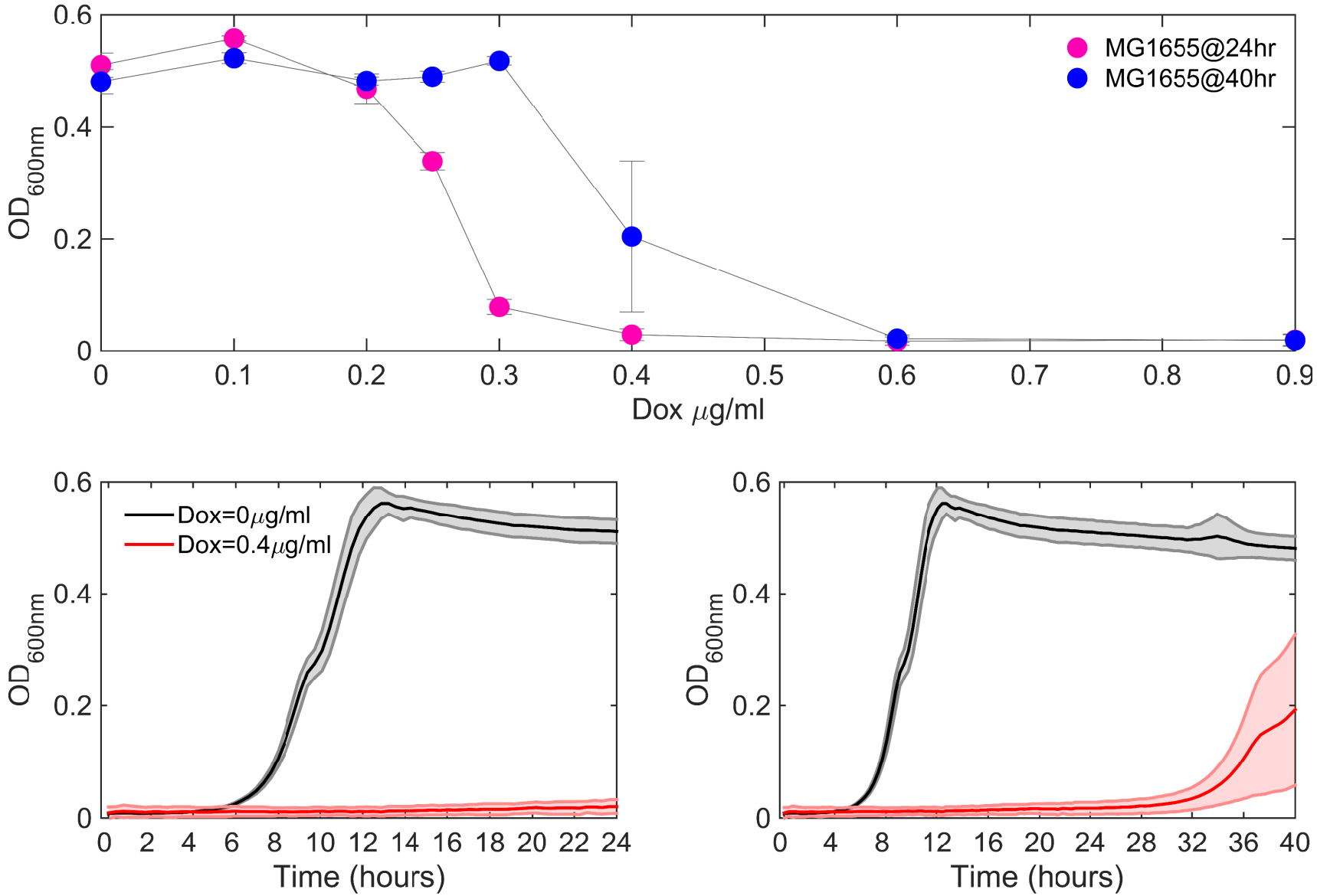


Fig. 3.13: There is no evidence of a lag-yield trade off in MG1655, with the longer lag phase giving rise to the same density as the drug free control.

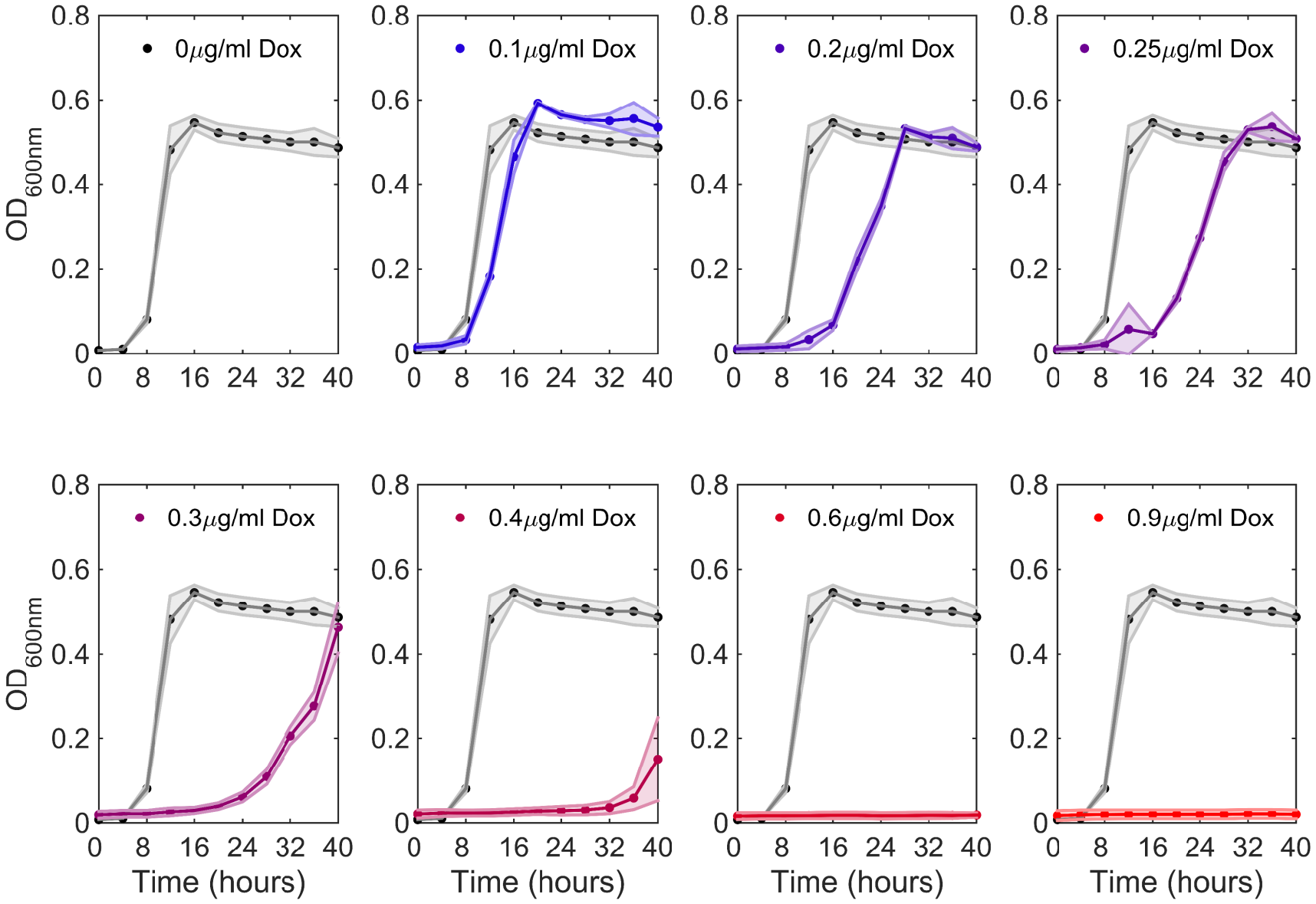


Fig. 3.14: There is no evidence of a lag-yield trade off in MG1655, with the longer lag phase giving rise to the same density as the drug free control.

3.5.3 What role does metabolism play?

E.coli is not a fastidious organism, and can be easily cultured on a range of media, with a wide variety of carbon sources. However, glucose is the most easily utilised carbon source for cell growth, and therefore high intracellular glucose concentrations will repress the expression of genes required for extended metabolic pathways such as the TCA cycle (Figure 3.16). When AG100 is grown using acetate as the sole carbon source, growth is far slower due to a lower Gibbs free energy, and achieves much lower densities than the stoichiometric equivalent amount of carbon coming from glucose.

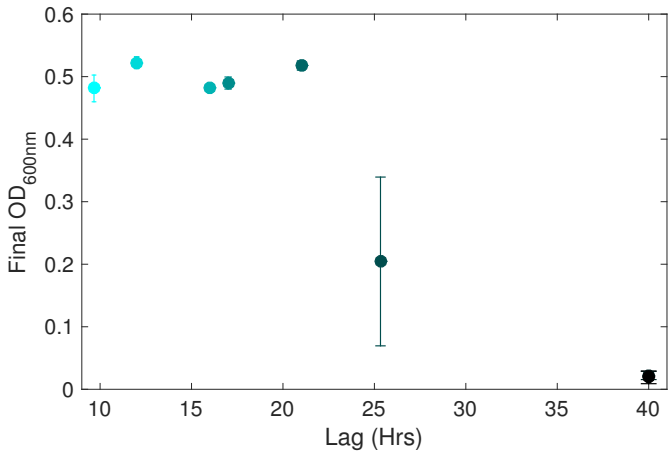


Fig. 3.15: Plot of Lag against Yield for growth curves plotted in Fig 3.13 showing that there is no increase in yield with increasing lag.

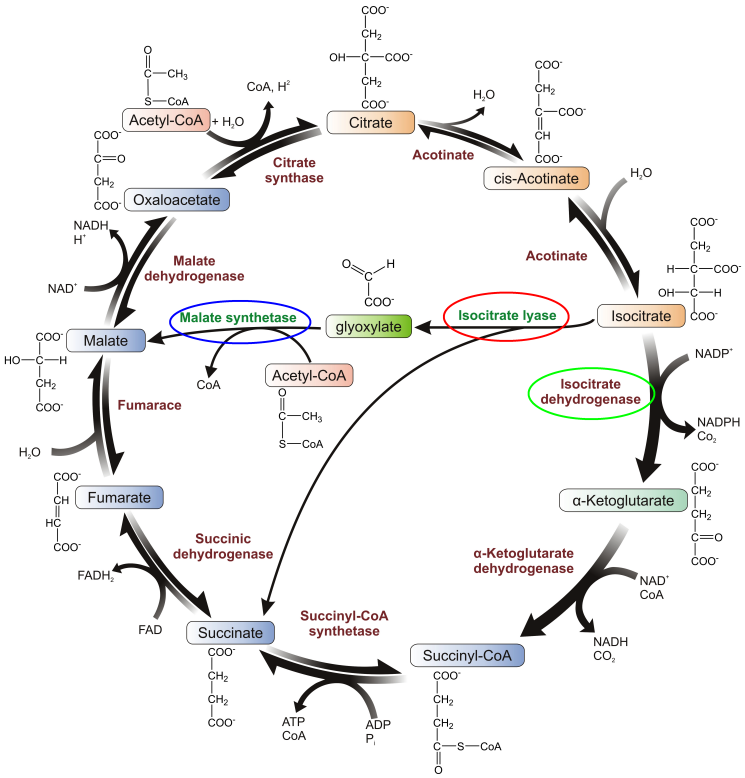


Fig. 3.16: Schematic representation of TCA cycle and glyoxylate shunt [58].

3.5.4 Dox improves growth on acetate

We now show that doxycycline can accelerate this reduced growth, and that increasing levels of the antibiotic can have a dramatic effect on improving the fitness of the cells in acetate media (Figure 3.17). Also of note is that the MIC in an acetate carbon source environment is approximately four times higher than when supplied with glucose.

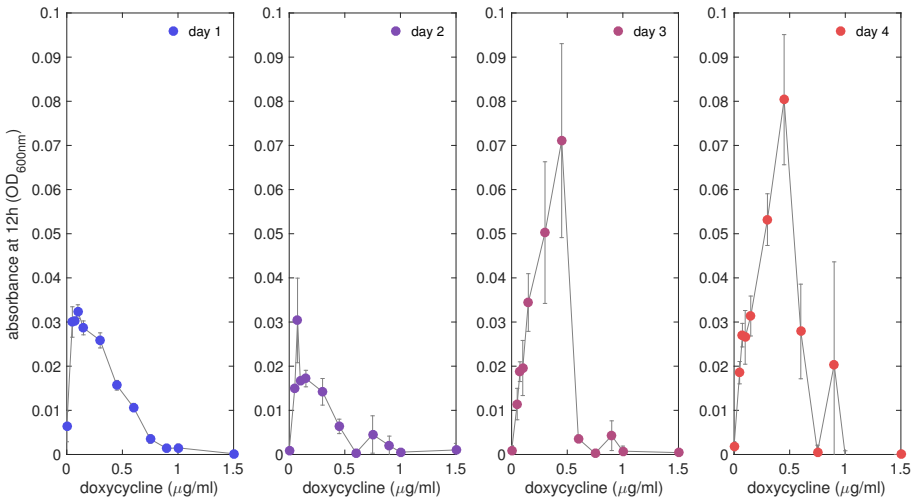


Fig. 3.17: Doxycycline seems to mediate the metabolism of acetate, allowing improved growth on this poor carbon source. This effect is visible immediately (left hand panel) and increases over 4 seasons (1 season is 24 hours). See Appendix 2.1 (pg. 181) for methods.

E.coli can grow on a number of carbon sources, utilising different metabolic pathways to maximise energy yield from the carbon resource presented to them. In this study, M9 minimal media is used to limit the availability of carbon in the media to precise amounts of glucose. However, during the metabolism of glucose, E.coli dissimilate acetate into the extra-cellular environment. Acetate can be metabolised utilising either the TCA cycle, or the modified TCA cycle, known as the glyoxylate pathway. The advantages to the organism that these pathways provide are that a complete TCA cycle will yield 36 ATP molecules, per acetate molecule metabolised, compared with only two ATPs per glucose in glycolysis. The Glyoxylate pathway does not yield ATP, but can transform acetate into sugars, such as succinate, that can more easily be transformed via gluconeogenesis into glucose. The TCA pathway, despite having a very high energy yield has an oxidative cost, as NADPH is oxidised to $NADH^{+}$, making this pathway unsustainable without reducing agents present. For this reason, the TCA cycle is normally repressed in organisms growing in abundant glucose sources [86]. Bacteria growing in acetate media can activate the glyoxylate pathway, with acetate entering the TCA cycle as acetyl-CoA. The decarboxylation steps present in the TCA cycle are skipped in this process, and the glyoxylate shunt can re-route the metabolic process to form succinate, an intermediate in gluconeogenesis. Therefore the glyoxylate shunt pathway can be used to generate ATP from acetate starting materials, by feeding succinate into gluconeogenesis [66].

Table S3.1: **Table showing the knockout functions from obtained from the Keio collection.**

Strain	Gene knockout	Function	Reference	Figure 3.16 colour
ΔAceA	Isocitrate lyase	Controls glyoxylate shunt entry	[80]	Red
ΔAceB	Malate synthase A	Determines malate or succinate fate of citrate	[28]	Blue
ΔIcd	Isocitrate dehydrogenase	Required for TCA cycle to progress instead of Glyoxylate shunt	[64]	Green
ΔManZ	Mannose phosphotransferase	Primary phosphorylation enzyme for D-glucose	[29]	Glycolysis not shown
BW25113	Wild Type strain	n/a	[11]	All enzymes present

To attempt to determine if a specific pathway was essential to stimulatory effect of Dox, we repeated the daily transfer experiments using the strains listed in Table 3.1. The Keio collection comprising single gene knockouts of every non-essential gene in *E.coli*, based on the parent strain BW215513. The strains chosen to test the importance of metabolic pathways for stimulation by Dox have gene deletions for critical enzymes. ΔaceA and ΔaceB, for example, are lacking the enzymes required for succinate to be processed through the glyoxylate overflow pathway, constraining or removing the ability of these strains to utilise acetyl-CoA to create biosynthetic intermediates. As can be seen from Appendix 2.2; Figures B.13, B.12, none of the knockout strains exhibited overgrowth, but neither did the parent strain, BW25113. This shows that the strain specificity of this phenotype must be explained by some pre-existing genetic difference between AG100 and other K12 strains that have been tested.

3.5.5 RYTO and LYTO in Dox-adapted AG100 strains

Having seen that certain Dox concentrations can improve the yield of AG100, as well is increasing the rate by which such cells adapt to their environment, we was interested in determining whether the richness of carbon resources would create a rate-yield trade off within or between strains with different evolutionary histories. Trade-off geometry is theorised to be responsible for the maintenance of biodiversity in microbes [99] and it has been shown previously that when trade-offs can be found *in vivo*, their shape can be predictive of co-existence [83]. There are numerous biophysical constraints on bacteria that can lead to trade-offs, with the 'waste from haste' rate-yield trade off being just one. Equally, there exist everyday constraints to bacterial fitness that we may not associate with trade-offs, such as the reduction in growth rate (the fitness cost) of antibiotic resistant strains; this could be viewed as rate-resistance trade-off. It is therefore necessary to consider a multitude of possible physiological factors which are interdependent when looking for trade-offs in noisy *in vivo* laboratory experiments.

AG100 and MG1655 was grown in media supplemented with glucose (at 6 different concentrations) and casamino acids only, or with Dox at MIC and 2xMIC concentrations. The serial transfer protocol was similar as previously described in this chapter, but used a 24 hours season time (see Appendix 2.2 for results). Once the populations had adapted to their new environments, they were then transferred in the same manner to each of the other environments that they had not experienced (*e.g.,* the AG100 strain adapted to MIC was transferred to drug-free and 2xMIC environments). Growth was then measured by absorbance at OD_{600nm}. In this way we hoped to determine any underlying trade-offs present that were dependent on environmental conditions, and whether these trade-offs are able to modulated by controlling glucose input. The raw data (growth curves) used to calculate RYTO

and LYTO for AG100 and MG1655 is available in Appendix 2.2 (pg. 202-210).

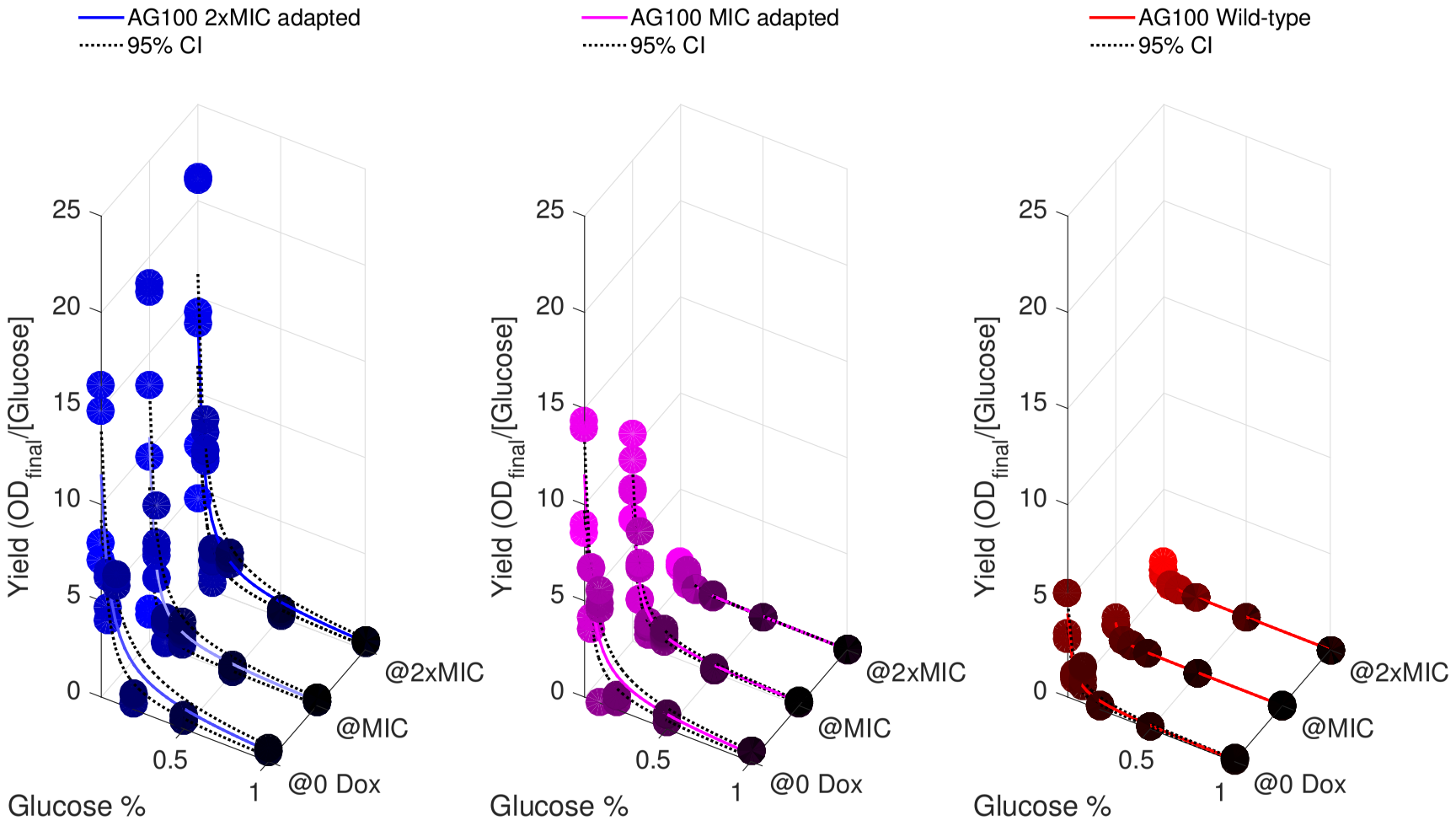


Fig. 3.18: Showing the efficiency of cells in conditions differing from their adaptive background. This plot shows AG100 that has adapted to MIC concentrations of Doxycycline, and is subsequently grown in drug free media, [MIC] and [2 x MIC] Dox. The yield increases when the drug is removed, as the cells become more efficient at converting glucose into biomass.

Figure 3.18 shows the results of growing MIC adapted AG100 in three different drug conditions ranging from drug free to twice the MIC concentration. The yield of cells in this context is defined as the final optical density divided by the amount of glucose supplied. This show the yield in terms of cells produced per unit of glucose supplied. However, this measurement does assume that glucose has been exhausted and is the limiting factor, which may not be the case when growth is inhibited by antibiotics. All three conditions show a clear within strain trade-off, that cellular efficiency at converting glucose into optical density (and therefore increasing yield) is considerably higher at low glucose. This is explained by a previously theorised trade-off between rate and yield which is due to the flux through the metabolic pathway that is used; a rapid glycolytic pathway that produces a small amount of energy (a net production of 2 ATP per glucose molecule) or a TCA cycle pathway that requires a more sizeable investment of protein, and imposes a redox stress cost on the cell, but which can produce a greater amount of energy (a net production of 36 ATP per glucose molecule) [83]. Although in this figure yield is plotted against glucose concentration, as rate increases with increasing glucose, we can make assumptions about rate-yield trade offs, based on this information. Figure 3.18 utilises the model proposed by the authors of that study, to fit a model of the following form:

$$c(S_0) = c_{hi}\frac{1}{1+pS_0} + \frac{pS_0}{1+pS_0}c_{lo} \tag{3.2}$$

This model (taken from previously published work by J.Meyer *et. al.,* [83]) has parameters of c (yield), S_0 (the sugar concentration supplied) and p (a phenotype that controls the rate of decrease of yield in increasing sugar concentration). Therefore c_{hi} represents the highest yield attainable, achieved at the lowest sugar concentrations, with c_{lo} being the inverse. It is notable that whilst it is traditionally assumed that a resistance mutation will impose a cost to fitness, often defined as a decreased growth rate, the Dox adapted strains shown in Figure 3.18 have an even greater efficiency when removed from the drug-stress. The RYTO theory would suggest that rate will decrease in such cases, and yet as as is shown in Appendix 2.2 (Fig. 27), the growth rate of is does not decrease in sub-MIC concentrations of Dox. Therefore if a trade off is responsible for maintaining the increasing yield, it must be involving a fitness parameter other than growth rate.

Previously we saw that Dox imposes a cost on cells, not in terms of growth rate as is the case with other strains such as MG1655 (Figure 3.14, but by elongating the lag phase. This lag phase elongation gave the benefit of higher cell yield, forming a trade-off between lag and yield that was visible at 40 hours. However, given that a yield increase becomes apparent within 12 or 24 hour seasons, after adaptation to the environment, it would seem that adaptation is removing some lag cost, and maintaining the yield benefit. Cells from each adapted condition were inoculated and then grown at the condition to which they were adapted. The results show that

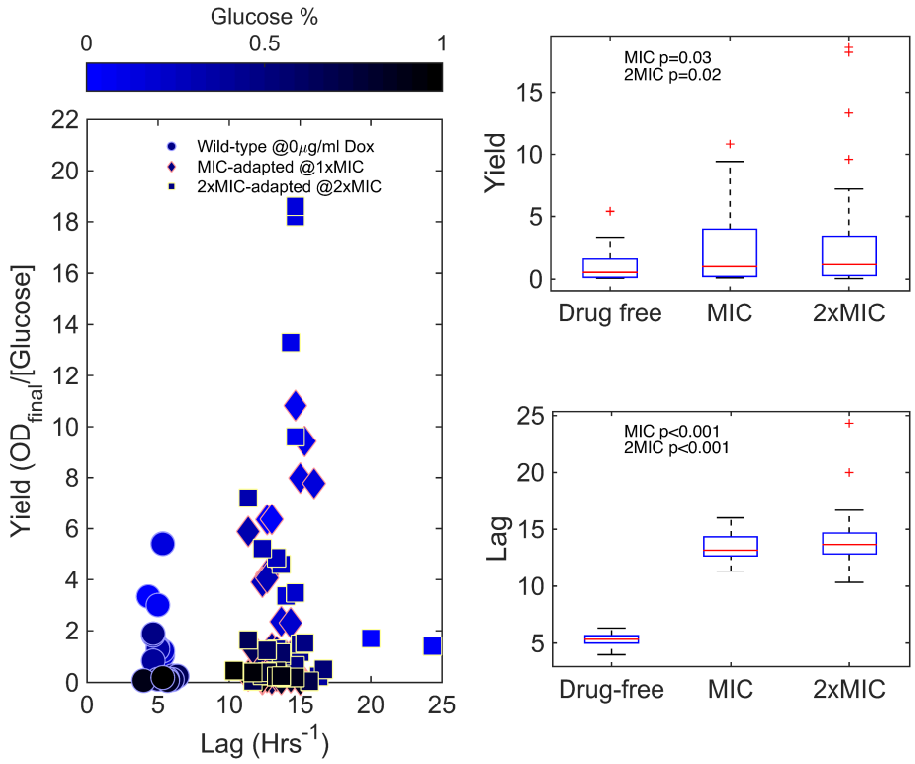


Fig. 3.19: When *E.coli* populations are grown for 5 seasons in Dox concentrations of 0/ml, 0.3/ml (MIC) and 0.6/ml (2 x MIC) it can be clearly seen that as drug concentration increases, so to does lag phase length, and yield. This shows a trade off between lag and yield.

although the drug-adapted populations have a longer lag than the wild-type, their yield is significantly increased, thus showing the LYTO is still in effect (Figure 3.19). It also appears that adaptation to the media has remediated some of the inhibitory effect of doxycycline, with a mean lag phase in the MIC-adapted population of 13 hours, compared to 21 hours when wild type cells are challenged by Dox (Appendix 2.2;. Figures B.16 B.17).

This LYTO is also maintained when the bacteria are grown in drug-free media (Figure 3.20). This provides further evidence that this adaptation is not costly to AG100, and shows that an extended lag is not merely a function of Dox concentration, but seems to be an advantageous growth feature to increase cell yield. We have already seen that at every glucose concentration tested, AG100 that has adapted to grow in inhibitory concentrations of doxycycline gains a significant increase in yield over the wild-type growth. This increase in yield is not correlated with a reduced growth rate, but appears to correlate with a increased lag-phase. As a further check, the lag and yield were assessed for the MIC adapted populations, grown in zero-drug. MIC and 2xMIC environments. Although the same LYTO is present between the drug adapted and wild type cells (when Dox is removed, the MIC adapted population still has a lag phase slightly longer than that of the wild type, with a considerably higher yield), there is no within strain trade off, as the lag cost to cells is almost entirely ameliorated by adaptation, when antibiotic pressure

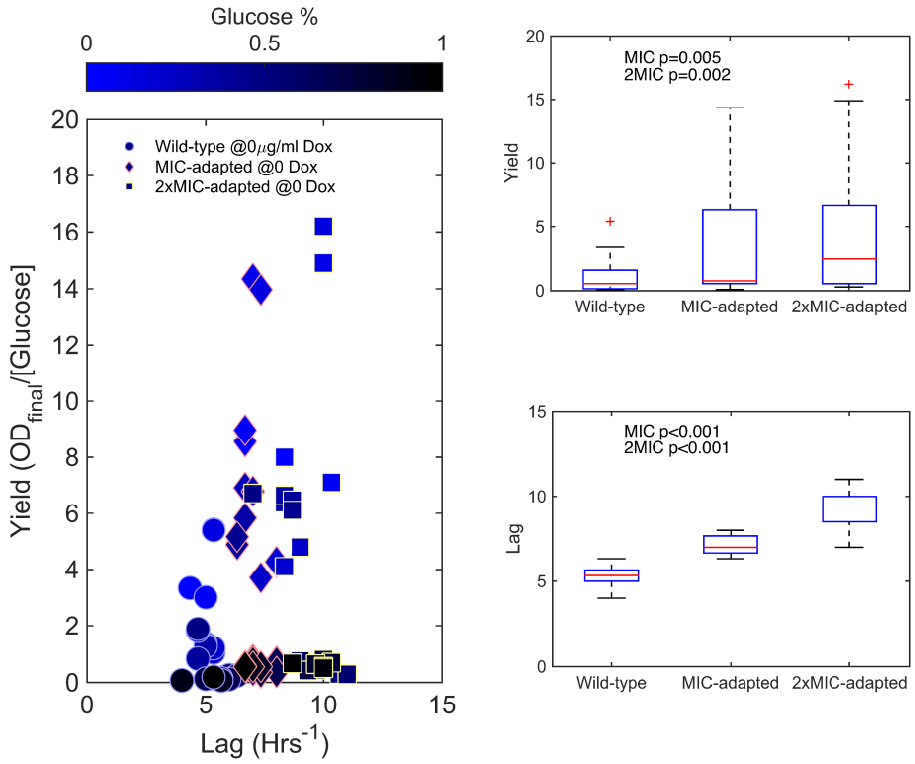


Fig. 3.20: Data showing the lag and yield of AG100 adapted to three different Dox concentrations. As cells experience a greater drug stress the time spent in lag phase increases. When these adapted strains are grown without the drug present, the lag times remain greater for those populations that have adapted to a stressful environment. However, at the same time the yield is also higher for these cells.

is removed (Appendix 2.2; Figure B.2). In MG1655 where the original overgrowth phenotype is not observed, we would not expect to see this increase in yield over wild-type, however analysing the growth kinetics at 6 glucose concentrations shows that there is indeed a transient increase in yield above wild-type levels at low glucose (Appendix 2.2;.Figure B.25). The difference in lag phase is much smaller, however, than in AG100 (Appendix 2.2; Figures B.19 B.20), meaning that there is no significant difference in lag phase between the MIC-adapted and wild-type strains, whether grown in their adaptive environments or at 0/ml Dox (Appendix 2.2; Figures B.21 B.22 B.23). Likewise, the keio mutant strains that were grown for 48 hours did not display an increased yield, and showed no LYTO (Appendix 2.2; Figure B.14). There is also however still clear evidence of the RYTO within strains (Appendix 2.2; Figure B.24).

Table S3.2: **Table showing non-carbon metabolism related mutations in K12 strains**

Mutation	Protein Function	Strains with mutation
lam−	Extracellular matrix protein	BW25113, MC4100, MG1655
rpsL	b.p substitution in Streptomycin resistance gene	AG100
rph−	Decrease in pyrimidine synthesis	BW25113, MG1655
glnX	B.p sub in tRNA suppressor	AG100
fimB	Deletion of structural gene recombinase promotor	MC4100
argE	Arginine synthesis	AG100

Table S3.3: **Table showing carbon metabolism related mutations in K12 strains**

Carbon metabolism mutation	Function	Strains with mutation
araD	Del in PPP pathway	BW25113, MC4100
galK	Galactose metabolism sub	AG100
xylA	Fermentation of xylose	AG100
mtl−	Mannitol uptake deletion	AG100
lacZ	Increases beta-galactosidase activity	BW25113
fruK	29 codon deletion to fructokinase	MC4100
rhaD	Rhamnose metabolism deletion	BW25113

3.6 Sequential treatments can create overgrowth phenotype

3.6.1 Genetic differences between strains

Why is it that *E.coli* AG100 shows such a clear overgrowth at a Doxycycline hotspot of around the MIC, when other strains that are closely related do not show the same phenotype? Table 3.2 below shows the non-metabolism related genetic differences between the strains tested. The high degree of similarity between the strains suggests that a small genetic difference, perhaps as small as a single SNP can have a dramatic effect on the response of a bacterium to a relatively short exposure to antibiotic. It is of course already well known that such small genetic changes can greatly increase resistance or tolerance, but the lack of available information about the effect of adaptation on the MIC exposes a weakness of understanding of how pre-existing genotypes can effect the resistance profile of bacteria.

Strains BW25113 and MG1655 are very closely related, sharing two mutations, *lam−* and *rph−*. The genetic similarities may provide a hypothesis as to why these two strains have the most similar dose-response and evolutionary adaptation to Doxycycline. Neither of these mutations effect carbon metabolism, whereas strain AG100 contains numerous deletions and nonsense mutations that effect several components of carbon metabolism, including mannose, galactose and xylose metabolism (Table 3.3). This could provide an explanation for the relative lack of robust growth in AG100 compared to other WT strains - in the same carbon resource environments, the AG100 strains achieves approximately three-fold lower optical density growth within 24 hours. Conversely the *araD* mutation present in BW25113 and MC4100 decreases the efficiency of the PPP (pentose-phosphate pathway), meaning that less reducing agents (NADPH) will be produced to protect the cell from ROS (Reactive Oxygen Species), and fewer phosphorylated glucose molecules will be

made available to glycolysis having been converted to fructose-6-phosphate. This is a somewhat similar effect as you would expect to see for the *xylA* mutant in AG100, however *araD* has a more critical flux control role through the pentose-phosphate pathway.

3.6.2 Forcing overgrowth phenotype through sequential treatments

AG100 has a unique genetic difference from the other strains tested, that the literature suggests may play a role in altering the metabolic regulation of *E.coli*. A mutation in *rpsL* changes the conformation of the ribosomal 16S subunit, at the binding site of Streptomycin (Strep). This conformational change gives AG100 a high innate resistance to streptomycin, and also has the affect of repressing genes required for stress response pathways, including many starvation, or stationary phase genes, whilst up-regulating genes required for unrestricted division [93]. The same authors also demonstrated that such a mutation created a trade-off between long term survival (harmed by the down-regulation of stress response pathways) and short term growth (which is maximised by rpsL mutation), causes an increased growth rate on poor carbon sources, compared to control with a wild-type *rpsL* gene. As we have seen, Dox can mediate the metabolism of acetate and other poor carbon sources, presumably up-regulating genes required to ensure higher flux through the TCA cycle for the production of energy. I hypothesise that the presence of both an *rpsL* mutation, and a Dox containing environment could lead to a boost in cell yield by de-repressing TCA cycle genes, whilst at the same time down-regulating stationary phase genes, meaning that these cells will make more efficient use of the exogenous sugars to maximise growth.

The most common resistance mechanism to Streptomycin (Strep) is a target site mutation in *rpsL*, changing the conformation of the drug binding site on the ribosome, allowing translation to proceed without being blocked by the antibiotic. As this mutation is extant in AG100, meaning that it has a very high *a priori* Strep resistance, we hypothesised that forcing a similar mutation in other strains may confer the overgrowing phenotype. A selection of strains, including MG1655, BW25513 and four single gene deletions from the Keio collection (a library of single-gene knockouts from the BW25113 parent strain) were exposed to high concentrations of Strep (see Materials and Methods), before undergoing the same adaptation to a doxycycline dose-response gradient. As mutation in *rpsL* is the most common streptomycin resistance mutation, a stringent regime of gradually subjecting strains to streptomycin (at concentrations of up to 1000 x MIC) was assumed to confer this mutation. It would be advantageous to confirm this mutation with sequencing in the future. Although this protocol was unsuccessful at conferring the overgrowth phenotype to the majority of strains tested, one particular Keio knockout strain demonstrated that

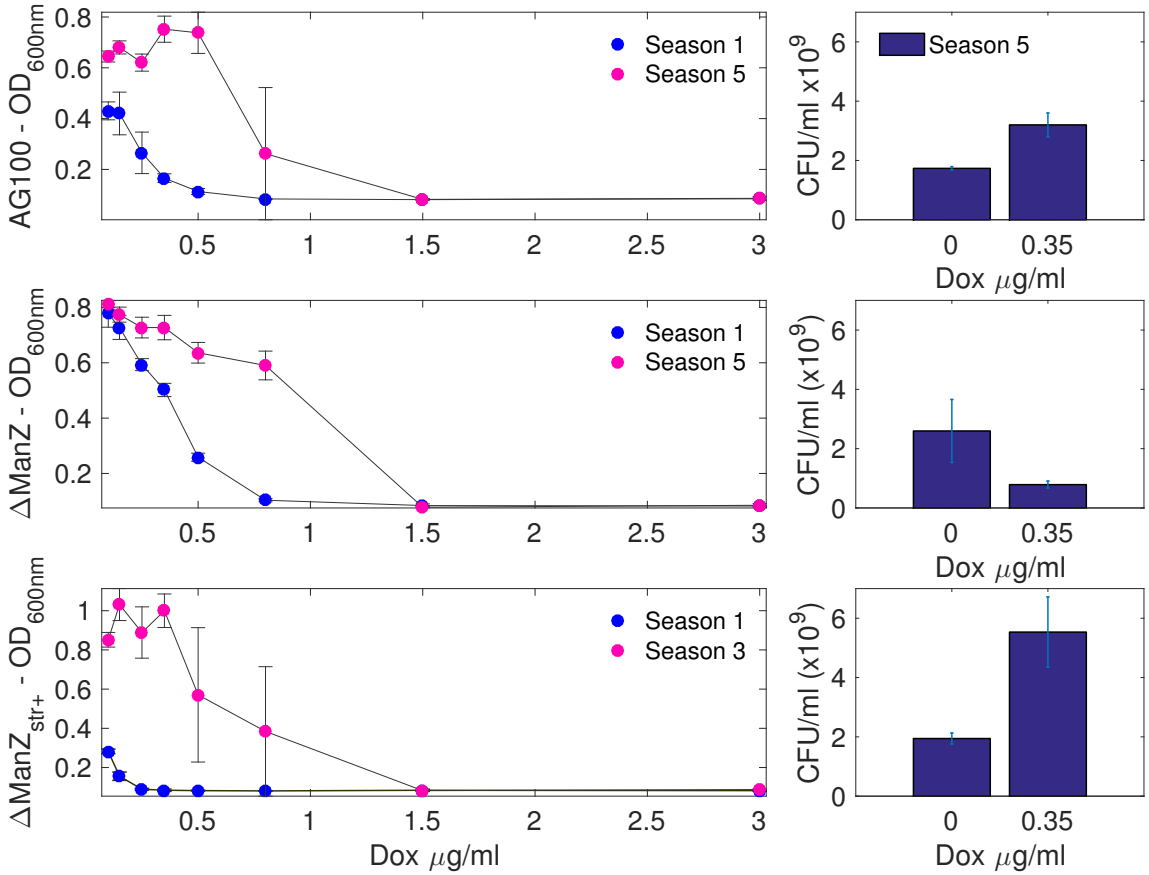


Fig. 3.21: $\Delta manZ_{strpR}$ (Strep resistant manZ knockout strain from BW25513), when compared to ΔmanZ (non-strep resistant control of the same strain) (middle panel) and AG100 (top panel) shows a distinct overgrowth, as well as a dramatic increase in fitness over the course of five seasons. Although both the top and bottom panels show overgrowth stimulated by Dox, this phenotype appeared more quickly in $ManZ_{strepR}, by season 3, whereas in AG100 the phenotype appeared after 5 seasons.$

the shape of the evolved dose response can be changed to show a yield increase at a similar hotspot region as seen in AG100 (Figure 3.21). Methodology and raw data for these experiments are found in Appendix 2.2 (pg. 198-199).

The overgrowth visible from season three onwards is not as substantial in BW25513$\Delta manZ_{strepR}$ as it is in AG100, and is more transient, with the phenotype having been lost at all but the lowest Dox concentration by season five. However, it is clear comparing the ΔmanZ and $\Delta manZ_{strep+}$ that there is a degree of non-monotonicity in the Strep resistant strain that is not seen in the un-evolved knockout (Figure 3.22). It is also evident that the rate of adaptation in $\Delta manZ_{strep+}$ is considerably higher than in any other strain; the first season growth is very low, but within two seasons the optical density is exceeding the Strep$^-$ strain. The ΔmanZ lacks the gene encoding mannosephophotransferase, and their phosphorylation of D-glucose is defective. Although there are other enzymes that can perform the same function, ΔManZ is less fit than the parental strain, and the other knockouts, suggesting that it is an important enzyme for glycolysis to occur efficiently. It therefore provides further evidence that other metabolic pathways must be important in the

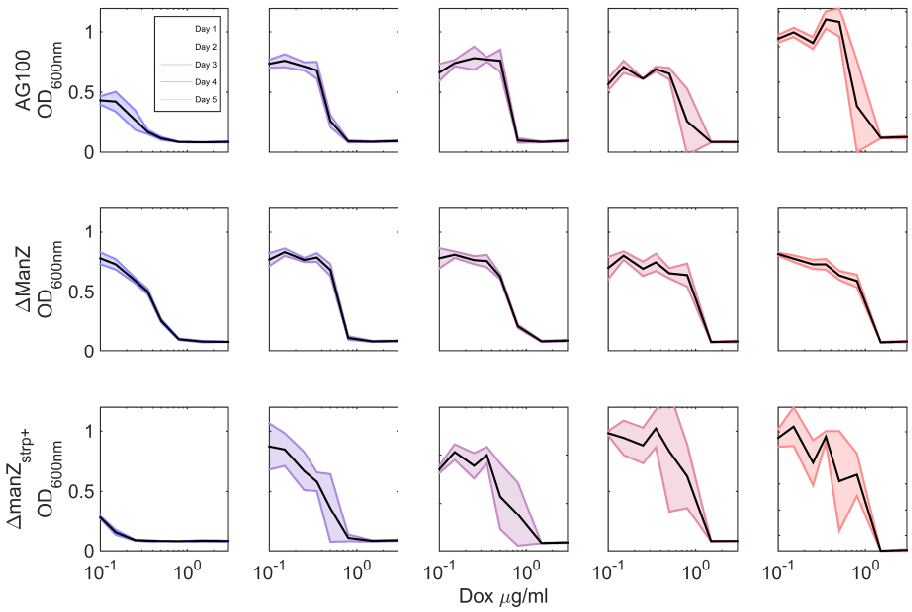


Fig. 3.22: Other than AG100, only ΔmanZ$_{strpR}$ showed any signs of overgrowth after exposure to high concentrations of streptomycin. Interestingly, although the fitness of the streptomycin resistant strain is initially very low, as expected, adaptation is extremely rapid, and overgrowth is seen at >MIC concentrations. Panels from left to right represent seasons from 1 to 5.

overgrowth phenotype.

Although a sequential treatment with Strep and Dox was able to confer a boost to cell density in ΔmanZ, the kinetics of growth are not entirely analogous with AG100. In this case, although the final density of cells increases when stimulated by Dox, the lag phase does not increase, meaning that the LYTO is not required in all circumstances for overgrowth (Appendix 2.2; Figure B.15).

3.7 Summary

We have seen in this chapter the failure of antibiotics to act as a bacterial inhibitor, and instead to stimulate growth to a level considerably higher than the wild-type. Because this trait is only exhibited in a single strain tested, it is possible to dismiss as an artefact of *in vitro* studies, and not relevant outside of the lab. This would, however, be ignoring the underlying concept of this research, that there is not sufficient understanding of the evolutionary response of bacteria to low doses of antibiotics. These sub-MIC concentrations will play an important role in the acquisition of resistance, as they allow for a dividing population to mutate, whilst under the selective pressure of an inhibitor. Furthermore, these low doses will be found in any antibiotic treatment *in vivo* regardless of the initial concentration of drug used, due to diffusion into cellular compartments, as well as drug degradation and decay. When a simple single drug treatment can lead to an inversion of the expected behaviour of a commonly used antibiotic, it demonstrates the necessity of determining

not only the inhibitory effect of a drug initially, but how robust this inhibition is over time. We also describe a previously un-reported trade off between lag phase length and yield. The 'waste from haste' efficiency trade off between rate and yield, [83] measures the fitness of cells by their growth rate. However, in this study we have shown that growth rate is not effected by Dox stress, and that the parameter effected is the length of lag phase. An increase in lag phase could be regarded as the cost of resistance in this case (as it most certainly would be in an environment with competitors), and yet if just growth rate and yield are measured it would appear that the yield increases with no commensurate reduction of rate. The lack of reporting of this lag-yield trade off may suggest that when looking for fitness costs and gains, all growth parameters should be investigated, not just rate and yield. Sequential therapy is frequently used as part of an antibiotic switching strategy designed to reduce antibiotic resistance, and yet here we have demonstrated a system whereby a specific sequential treatment can lead to the total failure of treatment, and the stimulation of growth. We have provided evidence that metabolism of poor carbon sources, such as acetate scavenging, is implicated in this overgrowth phenotype, and that a mutation in *rpsL* is responsible for the strain specific nature of the effect.

Chapter 4

Whole genome re-sequencing of Dox-adapted AG100 and MG1655

4.1 Overview

In the previous chapter we have described some unusual adaptive responses to Doxycycline, and the changes in fitness that these adaptations conferred. We performed whole genome sequencing on 18 samples of *E.coli* in order to determine any genetic variance that may help to explain the overgrowth observed. The principle findings of this sequencing were as follows:

1. Unique polymorphisms were identified in drug adapted populations, starvation metabolism genes were affected in AG100.
2. No polymorphisms were detected in MG1655 carbon-metabolism genes.
3. The acrAB/tolC efflux system was within a highly duplicated region in MG1655 adapted to Dox.
4. The acrAB/tolC efflux systems was also duplicated in AG100, although with a lower frequency.
5. There is a 30kb deletion from AG100 when adapted to antibiotic, with the loss of 49 lambda-like prophage genes, which have a poorly defined role in stress response.

4.2 Introduction

Using simple serial transfer protocols, we have highlighted the importance of understanding dose-response dynamics over the whole timescale of treatment. A concentration of Dox that inhibits 100% of visible growth during the first growth period (whether 12 or 24 hours) can be having a stimulatory effect on cell yield after just four days of adaptation. The fact that adaptation to antibiotic pressure can lead to a strongly non-monotone dose response contradicts some of the most basic assumptions of the mechanism of inhibition of antibiotics. However, non-monotone responses have been reported previously, notably being described as hormesis [75] and the Eagle effect [39]. What sets this results apart, is in addition to an non-monotonic response, the antibiotic becomes stimulatory of growth, within a clinically relevant concentration. It is important to highlight the implications for treatment success, as well as the questions raised as to out understanding of the mode of action of Doxycycline, when the efficacy of a drug can be so inverted within short timescales. Perhaps the more concerning observation made of this phenotype is that the resistance to Dox does not carry a cost in terms of growth rate as might be expected for a high yield population. The fact that the cost of resistance manifests as an increased lag phase is a novel description of the mode of action of Dox, as far as we have found. Although only one strain of *E.coli* showed this result, the fact that non-overgrowing strains can be forced into a phenotype that is stimulated by Dox has great relevancy. The method of using sequential drug treatments to minimise the spread of antibiotic resistance is used globally, and yet, without a thorough understanding of the complex interactions between stress, metabolism and replication, the interacting histories of drug exposure can lead to undesirable effects. These alterations to growth dynamics can be striking; in the case of $\Delta\text{manZ}_{strpR}$ there is a considerable gain in cell number (as measured by OD) at all concentrations of Dox, as well as the appearance of an overgrowth phenotype.

Whole genome resequencing was performed on the samples taken at various time points in the adaptive experiment showing the increase in yield in Doxycycline environments (Fig 3.2). As this phenotype appeared to be strain specific, samples of both AG100 (displaying the phenotype) and MG1655 (growing in high Dox concentrations, but showing no increase in yield) were sequenced. In this way the MG1655 strains could act as a positive control and determine if there were unique adaptations to AG100 that would explain the strain specific nature of the effect. Bacteria was sampled at the beginning of the experiment (wild-type cells) and the end of the experiment at both zero drug (media adapted controls) and in Dox (phenotype of interest). Three samples were taken of each whole population. Cells were sampled for sequencing from frozen experimental micro-plates, and DNA was extracted using a GeneJet silica micro-column method. The full methodology of the DNA extraction and sequencing pipeline is described in Appendix 3.1 materials and methods

Table S4.1: Sequenced populations

Strain	Condition	Replicates
AG100	Ancestral	3
AG100	M9-adapted	2
AG100	MIC-adapted	3
MG1655	Ancestral	3
MG1655	M9-adapted	3
MG1655	2xMIC-adpated	3

section. The strains sequenced are summarised in Table 4.1.

In Chapter 2 we conducted adaptation experiments which had the effect of increasing Dox resistance in both AG100 and MG1655 strains. However, in the AG100 strain, this resistance was accompanied by a boost to cell density, to a level greater than the drug free control. We have ascertained that this phenotype has a novel lag-yield trade off, and hypothesised that harbouring a streptomycin resistance mutation on *rpsL* may enable more efficient carbon metabolism in the presence of Dox. We also observed a greater level of resistance (to 2xMIC) in MG1655, but this was a more conventional adaptation, which conferred a cost compared to the drug free environment. By sequencing the whole genomes of these two strains, at multiple points in their evolutionary history, we can hope to determine the genetic variation contributing to increased resistance and overgrowth.

All sequencing was performed by Exeter Sequencing Service on an Illumina MiSeq v3 on paired-end Nextera libraries.

4.3 Results

4.3.1 AG100 coverage

Coverage plots of all aligned reads were generated using Blast Ring Image Generator (BRIG) [3] to easily show major structural changes to the genome. The coverage plots below (Figure 4.1, Figure 4.3) are annotated with genes related to multi-drug efflux systems, including the acrAB components of the acrAB/tolC efflux pump. There is an area of the genome with an increased coverage that is suggestive of a duplication between 308,415bp and 596,822bp. The regulatory repressor acrR is also present in the weakly duplicated region in the Dox-adapted AG100. Unlabelled red arrows on the outermost ring of the figure shows the presence of putative prophage genes, as annotated by PROKKA [111].

Dox-adapted coverage

There is an increase in read depth in the genomic region containing an important efflux system (acrAB), as would be expected from previous similar studies [13].

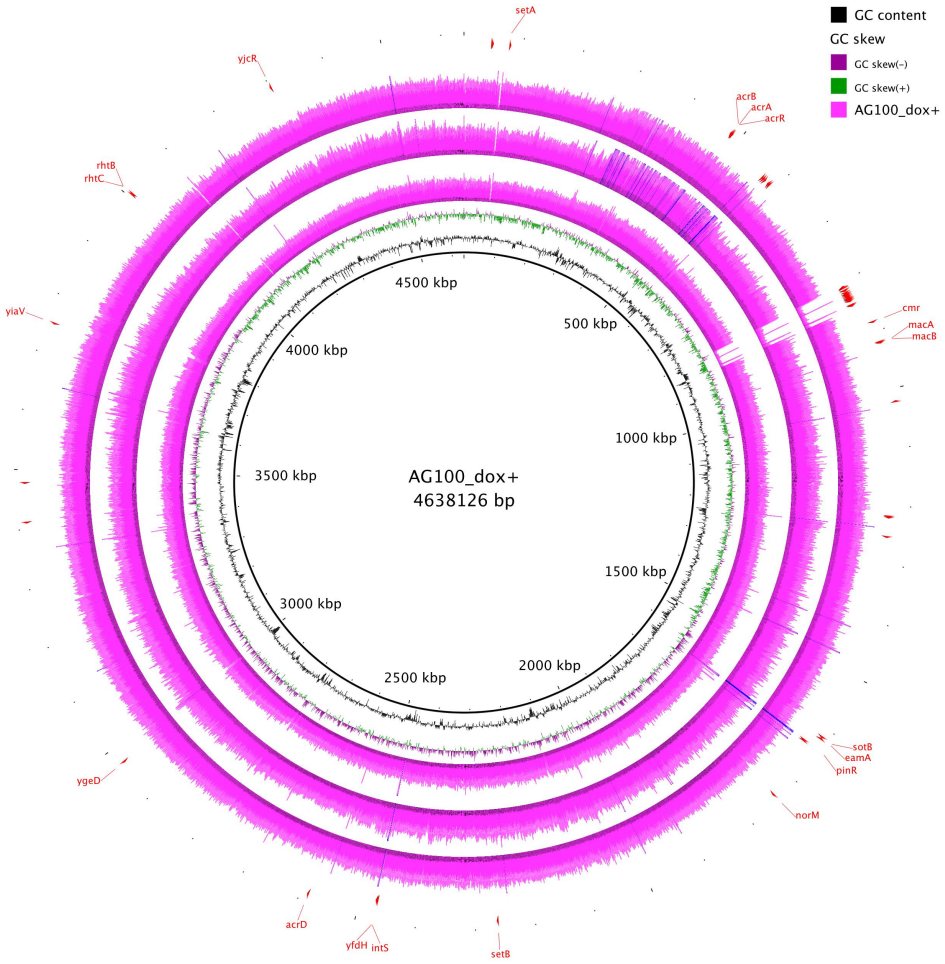


Fig. 4.1: Coverage plot for three replicates of AG100, adapted to MIC concentrations of Doxycycline. There is a deletion of a number of prophage genes, and indication of a duplication of a region containing the acrAB/tolC efflux pump in all three replicates.

In AG100, however, the frequency of the duplication is not at 100%, rather there is an increase in coverage of approximately 30% in the Dox-adapted populations. This would imply that only around a third of the population are harbouring this duplicated region, and that it is not therefore essential to the overgrowth phenotype (Figure 4.2).

Wild-type coverage

It can be seen that the wild-type strain does not have any evidence of a duplication to the region containing acrAB, and has a fully intact genome with no deleted elements. This indicates that both the loss of prophage and the duplication event are mediated by the presence of Dox.

4.3.2 Deletion in Dox-adapted AG100

A large deletion can be seen in the coverage plot of Dox adapted AG100, (around 850kbp) whereby the coverage drops from a mean of 119 to 7. Annotation with RAST and PROKKA determined this region to encode prophage genes. These prophage genes show a high homology with phage lambda, and are bookended by Galactosidase and Biotin genes as expected of lambda prophage [56]. Figure 4.4 shows the genetic differences between the Dox-adapted and wild-type populations in this prophage. A complete list of prophage genes is available in Appendix 3.1; Figure C.1.

4.3.3 AG100 Dox-adapted SNP effects

SNPs in each replicate of the Dox-adapted AG100 detected using VarScan were compared using VCFtools, and reported SNPs common to all three replicates recorded. These were then compared to the M9-adapted control SNPs to ensure that the polymorphisms were unique to the drug adapted genomes. These SNPs are shown in Table 4.2. Four non-silent polymorphisms were identified as being unique to the $AG100_{Dox}$ populations, and they were contained within two genes. One of these genes is a mobile element protein, and one is involved in carbon metabolism under starvation response. This gene, named *slp* for starvation lipoprotein, is expressed under carbon starvation, or stationary phase, and is regulated by *rpoS*. We know that *rpoS* is defective due to a mutation in *rpsL* in AG100, so this polymorphism could therefore exist as a compensatory mutation in AG100 to limit the costs of streptomycin resistance. The alternate base reading was present at a frequency of 10.61% and 10.28% for the two SNPs respectively, with full filtering and quality control data shown in Appendix 3.1 (Table C.3). Mutations in the mobile element protein had a similar frequency of between 12-14%.

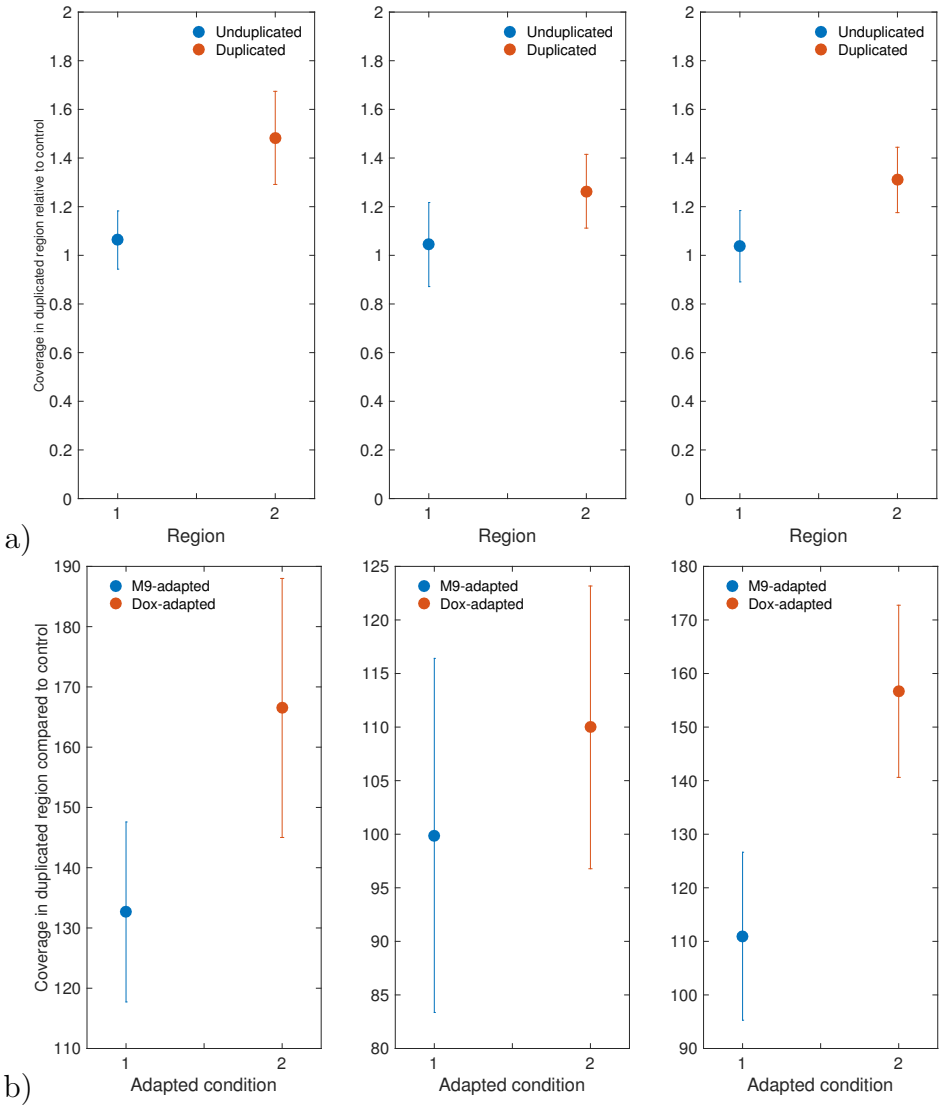


Fig. 4.2: a) shows the relative coverage of the duplicated region in the three 3 Dox-adapted replicates. The ratio does not reach two, indicating that the duplication in this region is not at a high frequency within the population, or that the duplication was not complete due to the position in the replication cycle at the time of harvest. b) shows the mean coverage of the duplicated region in the M9 and Dox-adapted populations. The higher coverage in the Dox-adapted populations shows that the duplication is an affect of the drug.

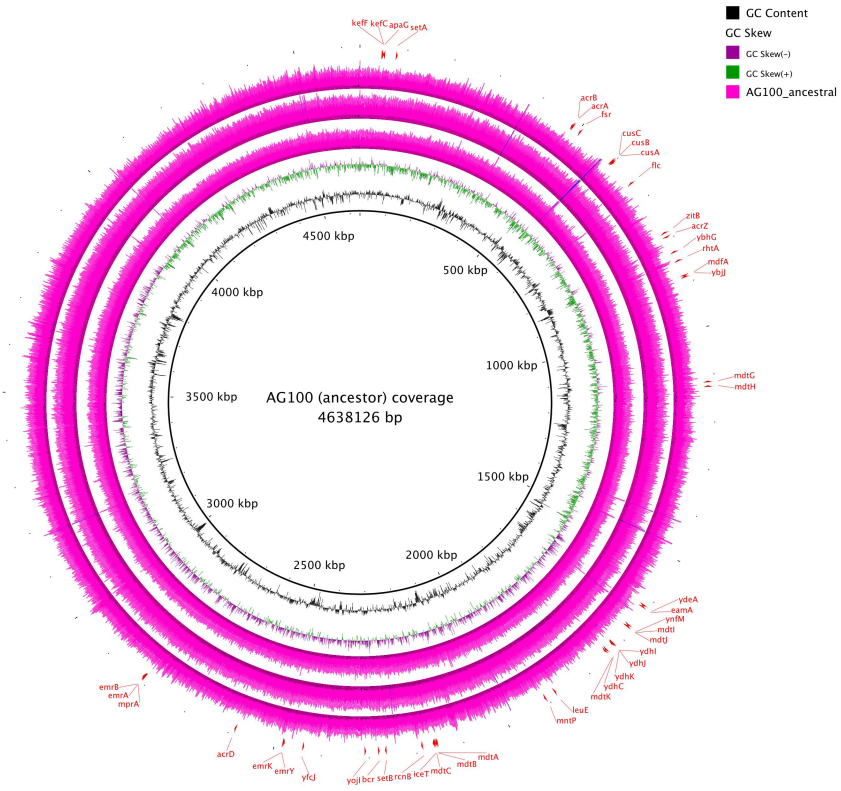


Fig. 4.3: Coverage plot for three replicates of AG100, grown in M9 for two overnight cultures, but otherwise unchanged from the wild-type ancestral strain.

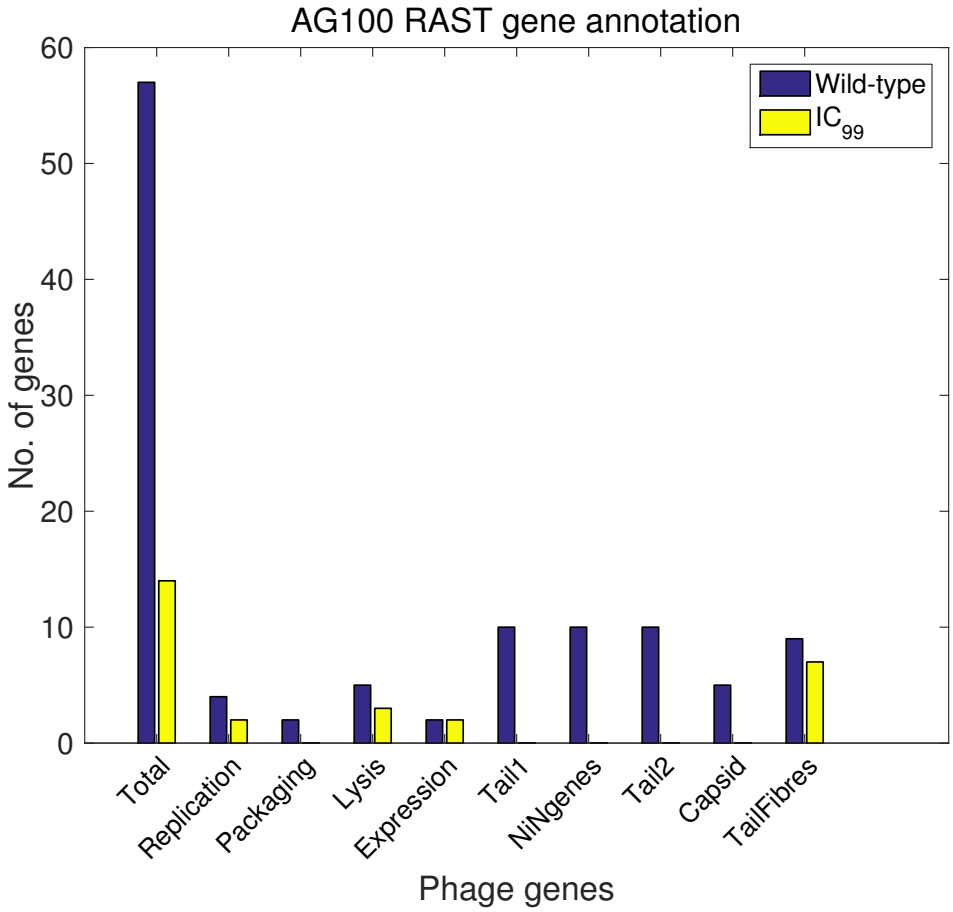


Fig. 4.4: Genes present in wild-type compared to adapted cells.

The loss of the prophage represents a change to the stress response of AG100, as does the pre-existing *rpsL* mutation. It could therefore be the case that metabolic gene function is not directly affected by mutation, but rather is controlled at the regulatory level, effected by obtuse upstream genetic variations.

Table S4.2: SNPs common to all AG100$_{Dox}$ replicates

Strain	Position	Reference	Alternate	Gene description	Status
EK12AG100	3127332	G	T	Mobile element protein	silent cgg ->cgT
EK12AG100	3127350	A	G	Mobile element protein	silent cca ->ccG
EK12AG100	3127368	A	G	Mobile element protein	silent cga ->cgG
EK12AG100	3127402	T	C	Mobile element protein	silent ttg ->Ctg
EK12AG100	3127440	G	C	Mobile element protein	silent acg ->acC
EK12AG100	3127494	C	G	Mobile element protein	non-silent cac ->caG
EK12AG100	3127460	C	T	Mobile element protein	non-silent gcg ->gTg
EK12AG100	3648916	G	T	Starvation lipoprotein Slp	non-silent atg ->atT
EK12AG100	3648944	G	A	Starvation lipoprotein Slp	non-silent gcg ->Acg

4.3.4 MG1655 coverage

Dox-adapted coverage

BRIG was used to plot the coverage of MG1655 in drug and non-drug adapted environment, in exactly the same manner as for AG100. There is a duplication in the Dox-adapted population, however in MG1655 the proportion of the genome that undergoes duplication is larger, with the region spanning from 271,063bp to 693,445bp. The duplication is also far more clearly defined, suggesting a higher frequency in the population. All genes contained in the duplicated region are recorded in Appendix 3.1; Table C.2 There is a second duplication in a region annotated with the putative transmembrane efflux protein yiaV. This duplicated region however was present in all sample of MG1655, including the wild-type. It seems that the duplication in this region is a pre-existing feature of the ancestral strain used to start this experiment. This may have occurred during a past culturing or re-culturing of the ancestral strain, but not noticed previously.

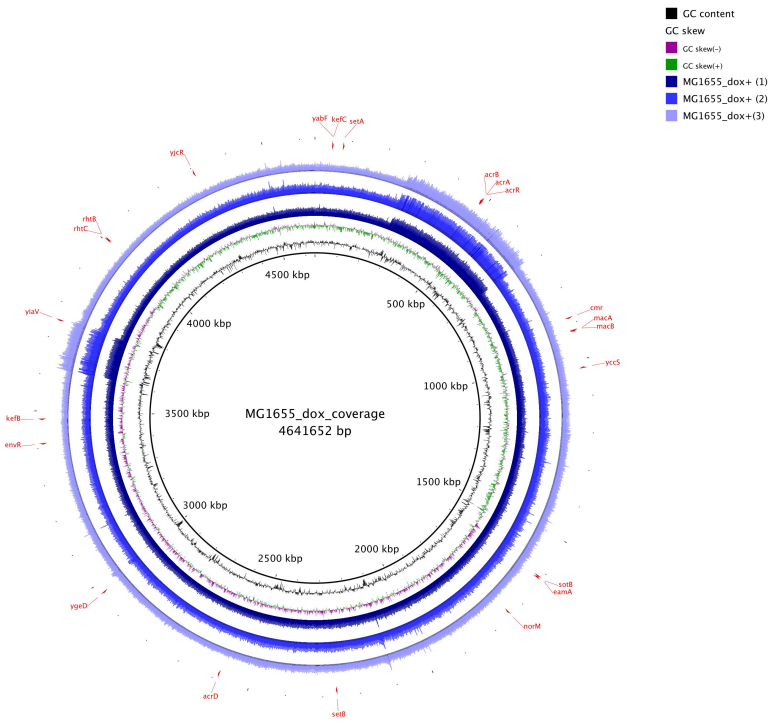


Fig. 4.5: Coverage plot for three replicates of MG1655, adapted to 2 x MIC concentrations of Doxycycline. As previously seen there is a large duplication in the region encoding the acrAB/tolC efflux pump. Interesting there is also a duplication in the putative transmembrane efflux pump of yiaV.

Wild-type coverage

The key feature shown by the coverage plot of $MG1655_{WT}$ is a duplicated region between 3,674,130bp to 3,802,567bp, containing the putative efflux system yiaV, as observed in $MG1655_{Dox}$ populations. The occurrence of this duplication in the wild type cells demonstrates a *de facto* change in genomic conformation from the assumed ancestral MG1655 strain. This could have occurred during a prior re-sampling and freezing, during transit or during the revival and growth of the samples for this experiment. Because the duplication remains present in media and Dox adapted populations, the inference must be that the cost of carrying the additional genomic material is not sufficient to create a selective pressure favouring its loss. MG1655 is commonly used as a model organism, and it has been reported that there is genetic variation between laboratory maintained strains, and the CGSC strains, for example, the isolate used by Soupene et al,. [36] contained a fnr deletion, the isolate currently being distributed by the CGSC is fnr^+ [57].

The frequency of the duplicated acrAB region is high in MG1655, with a ratio of nearly 2:1 in all replicates, indicating that the over expression of efflux is a requirement to resistance to high Dox doses, and that it has swept to fixation through the population.

4.3.5 $MG1655_{Dox}$ SNPs

Non-silent mutations were found in four genes in all three replicates of Dox-adapted MG1655, as shown in Table 4.3. These SNPs were found with low frequency (4-8%), the complete information on quality thresholds can be found in Appendix 3.1; Table C.3. There are also a number of synonymous mutations found in all replicates, including two polymorphisms to a secondary efflux system, yiaV, found in the second duplicated region of the genome. Although these polymorphisms do not change the amino acid encoded, there may be non-silent effects due, for example, tRNA specificity. The fact that they are present in all replicates of the Dox-adapted populations, but not present in the media-adapted or wild type populations would suggest that efflux is under a selective pressure.

One of the non-silent mutations appears in an intergenic region, with no annotation made by either RAST or PROKKA. The other three SNPs that could have a phenotypic effect were in diverse target systems, none with any immediate relevance to increased drug resistance. The *pnp* protein, a polyribonucleotide transferase, is a component of the RNA degradasome, and is a cold-shock protein, essential for growth at low temperatures. A mutation in lyxK, a xylulose kinase enzyme represents the only carbon metabolism mutation found in the sequencing of drug adapted MG1655. The lyxK gene is normally silent, however a mutation to the yiaJ transcriptional repressor (which occurs in the second duplicated region of these MG1655

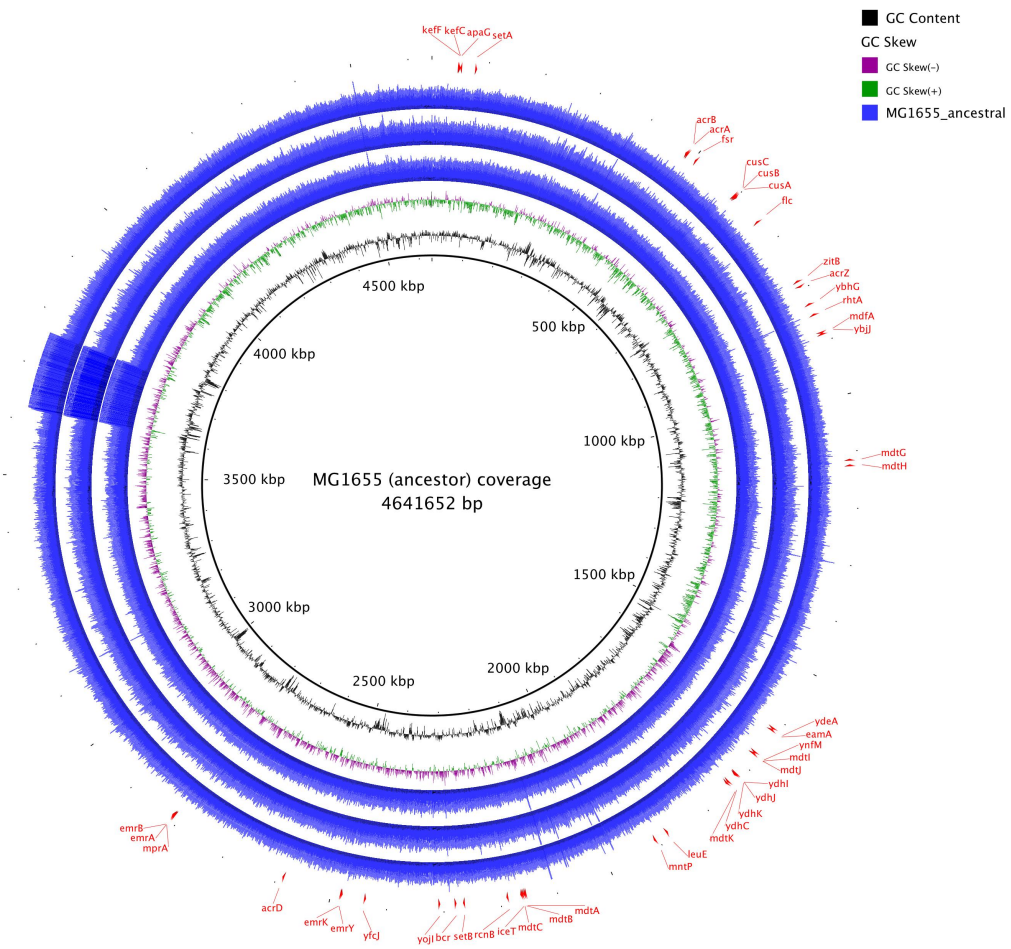


Fig. 4.6: Coverage plot for three replicates of MG1655, adapted to 2 x MIC concentrations of Doxycycline. As previously seen there is a large duplication in the region encoding the acrAB/tolC efflux pump. The putatative transmembrane efflux system yiaV (duplicated at around 3700kbp) is also present in duplicate in the wild-type, meaning that its presence in the MG1655$_{Dox}$ populations can not be attributed to adaptation to Dox.

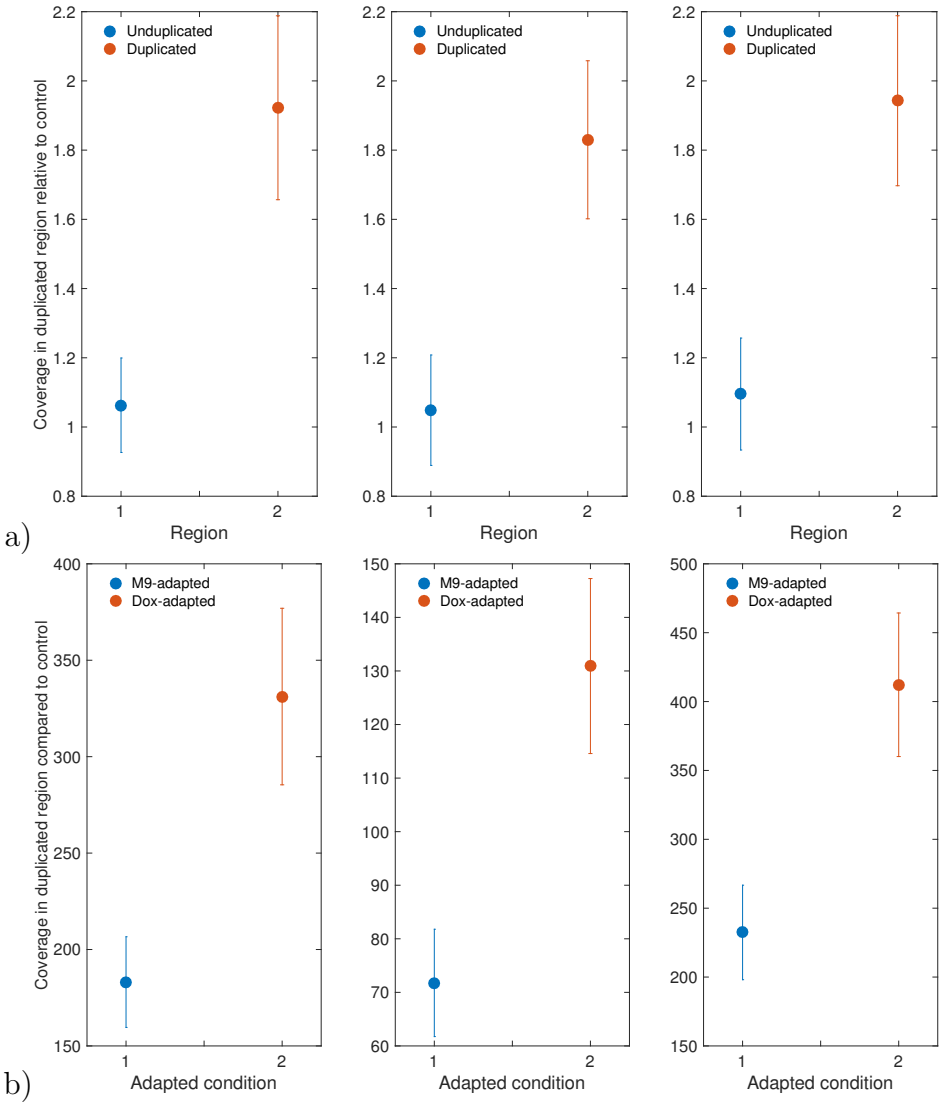


Fig. 4.7: a) shows the relative coverage of the duplicated region in the three 3 Dox-adapted replicates. A value of close to 2 shows a duplication has taken place. b) shows the mean coverage in terms of actual reads averaged over each position in the duplicated region in the M9 and Dox-adapted populations. The higher coverage in the Dox-adapted populations shows that the duplication is an affect of the drug.

Table S4.3: SNPs common to all MG1655$_{DOX}$ replicates

Strain	Position	Reference	Alternate	Gene description	Status
EK12MG1655	3608715	A	G	perM permease	silent
EK12MG1655	1466438	T	G	Phenylacetate-coenzyme A ligase: PaaF	silent
EK12MG1655	4296060	C	T	Membrane fusion component of tripartite multidrug resistance system	silent
EK12MG1655	4296190	N	A	Inner membrane component of tripartite multidrug resistance system	silent
EK12MG1655	3305354	A	T	Polyribonucleotide nucleotidyltransferase: pnp	non-silent att ->aAt;
EK12MG1655	3688287	A	C	No gene annotation	non-silent gtg ->gGg;
EK12MG1655	3723928	A	C	Xylulose kinase: lyxK	non-silent ttc ->tGc;
EK12MG1655	1468135	C	A	Phenylacetic acid degradation operon negative regulatory protein: PaaX	non-silent gga ->gTa;

populations) switches the gene on and allows for phosphorylation of L-xylose and further metabolisation of L-xylulose 5-P by the pentose phosphate pathway [59]. It can not be assumed that xylulose is being metabolised in this fashion, because a large inversion event which deactivates the repressor is required for expression of lyxK, and in this case the repressor lies within the duplicated region of the genome. The final polymorphism identified is paaX, regulatory protein of phenylacetic acid degradation [10]. This gene negatively regulates DNA transcription, as well as phenylacetate catabolism (part of the process of aromatic hydrocarbon metabolism).

Chapter 5

Trade-offs and trade-ups between bacterial traits and phage resistance: why does phage therapy fail in *in vitro* models?

Note

This document is the result of a collaboration between:

1. Robert Beardmore, Exeter University, UK (data analysis, algorithm development)
2. Mark Hewlett, Exeter University, UK (bacterial experiements)
3. Rafael Pena Miller, CCG(UNAM) Mexico (image analysis)
4. Justin Meyer, UCSD, United States (coevolution and other phage experiments)

5.1 Overview

In this chapter we observe the following key properties of data on the coevolution of a bacterial strain and a lytic phage:

1. coexistence in an *in vitro* model of failed phage treatments of bacteria between λ-phage-susceptible and apparently λ-phage-resistant *Escherichia coli*.

2. Seeking a trade-off between phage resistance and bacterial 'fitness' to explain this, in fact we observe the opposite: a trade-up whereby simultaneous increases occur in both resistance to phage and fitness through sugar uptake in certain strains that we call "superstrains".

3. This trade-up has a structural basis: mutations the cause increases in sugar uptake occur at the same locations on the protein as mutations that lead to enhanced phage resistance.

4. We then show this property of the *lamB* is sufficient to explain the failed of phage therapy that we observe.

5. En route, we elucidate a number of trade-offs whose existence has been hypothesised in the literature.

Core idea

In this experimental work we use a laboratory microcosm that contains *E.coli* which are predated on by λ-phage and the liquid medium in which the cells grow has a single limiting carbon source, maltotriose. *E.coli* can express a maltoporin on its outer membrane assembled from the product of the *lamB* gene and by increasing the concentration of the products of this gene, the cell can increase the import rate of maltotriose into its periplasm. However, we have left the bacterium with a regulatory and evolutionary decision to make because λ-phage gain access to the cell through the same maltoporin.

This experimental protocol forces selection to act upon the *lamB* gene, but how exactly? The choice is this: increase the ability of the porin, by mutation for example, to import an essential sugar into the cell, or else decrease the likelihood of the phage binding to the porin and gaining entry, thus killing the cell. Using this protocol we will be able to pinpoint a trade-off between these two traits and so, according to coevolutionary theory, understand how the coevolution of bacterium and phage proceed.

5.2 Introduction

The concept that we have insufficiently many antibiotic drug targets to support the future needs of human medicine when faced with the rapidity of bacterial evolution has been much-discussed recently, both in the scientific literature and within the wider public. However, irrespective of our own future ability, or indeed inability, to cure microbial infection using antibiotics, we can be completely confident that bacteriophage will maintain their potency against bacteria, however they evolve, because they have done so for several hundred millions years already.

By contrast, the discovery and subsequent optimisation for medical purposes by humans of small molecule antibiotics is a process that began less than a century ago and yet, without doubt as a direct result of the increase in circulation of such molecules, aided and abetted by transmissible resistance plasmids, some bacteria are now capable of exhibiting pan-antibiotic resistant phenotypes. We therefore contend that the science of microbial viruses, including bacteriophage, both natural and synthetic, will come to play an increasingly core part of our fight against microbial infection. We have much to learn from phage about how to tackle bacteria, indeed, despite their very small genomes, phage are able to overwhelm the CRISPR-mediated adaptive immunity of bacteria.

But if we are to successfully exploit phage to produce new and successful treatments for infection, topic infections in particular, the following question must be addressed. Within the context of bacterial treatments: how long do phage maintain their potency against bacteria within closed infections? Moreover, we need to know, why do phage do so badly in *in vitro* model study systems? If we are to optimise phage therapy for medical purposes, we need go *in vitro* model systems but, as we show below, phage are rarely able sterilise *in vitro* cultures.

To make this discussion a little more precise, from an *in vitro* model system for infection we can ask the following research question. When a given phage strain is known to be able to lyse a particular bacterial strain, let us suppose they are then both co-inoculated into the same culture device, like a chemostat or a microtitre plate, as if the phage were being used to treat the bacterial population as an infection. How many bacterial generations can pass before the, presumably evolved, phage has no remaining potency against the co-evolved bacterial strain? Moreover, if the phage does maintain its potency in this way, how often does the bacterial population collapse, meaning, how often is the treatment successful due to phage?

Figure 5.1 is typical of the kind of phenomenon that we have in mind as motivation for these questions. In this figure, data is shown for three different experiments whereby a wild-type *Escherichia coli* bacterium, strain REL606, has been inoculated into microtitre plate wells along with a single strain of phage λ. In none of these experiments, of 18 conducted, did the phage sterilise the device of bacteria and, moreover, in 8/18 of the experiments the phage was lost, not the bacteria. Bacterial

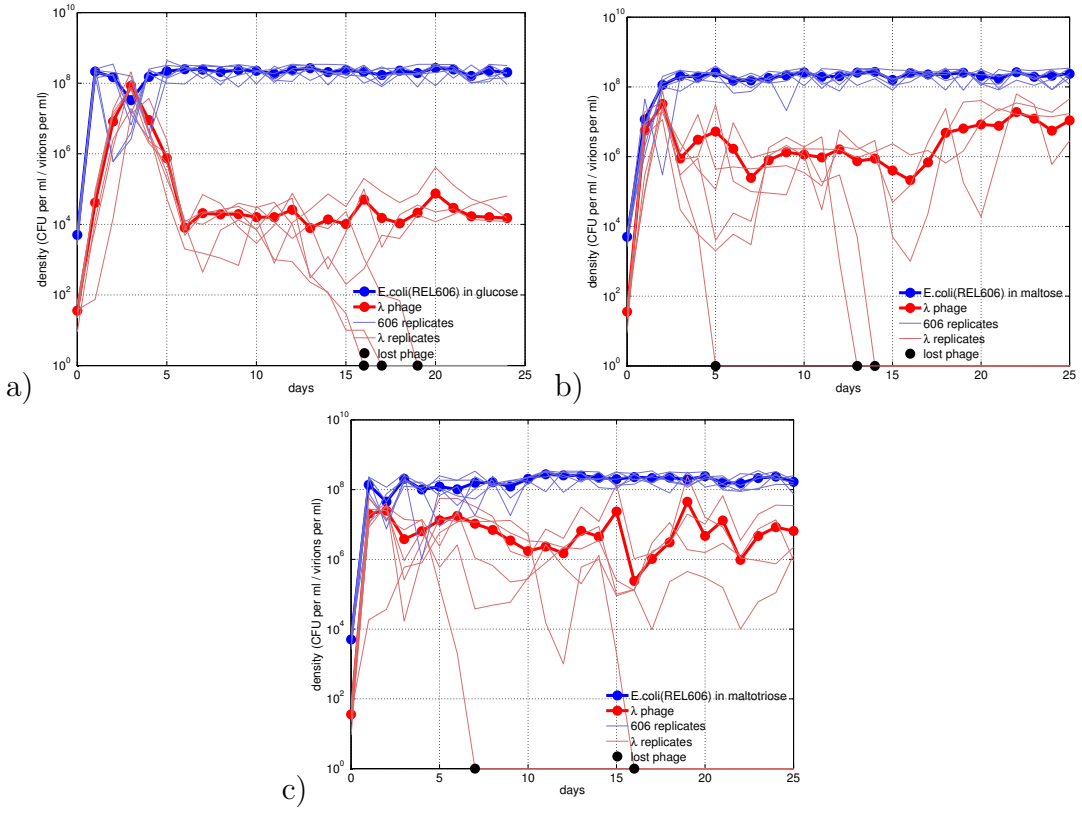


Fig. 5.1: **We are interested in the question of how often a lytic phage can sterilise an in vitro culture of bacteria**. Here, three contexts for such *in vitro* phage therapies are studied whereby 18/18 treatments failed to clear the bacterium, and 8/18 failed because the phage was not maintained in the device. The three cases have different carbon sources supplied to the bacterium, namely (a) glucose (3/6 lost the phage), (b) maltose (3/6 lost the phage) and (c) maltotriose (2/6 lost the phage).

and viral quantification was performed by removing a sample from the experiment each day and using standard colony counting to assess the number of bacteria. To determine virion number, the sample was spread onto a lawn plate of bacteria, and covered with agar to prevent indiscriminate viral spreading. Plaque counting was performed and adjusted for the dilution used to calculate virions per ml (assuming that each plaque is caused by a single virus particle).

So, resolving the above suite of questions is key to understanding the utility of phage as part of our arsenal for treating infection. For if, it turns out, that the evolutionary response of the bacteria is so rapid that *de novo* mutations quickly render wild-type phage obsolescent within a matter of hours or days from the start of treatment, then we will need new ways of using phage to treat infection that may incorporate a need for adjuvant small molecules that are able to help the phage achieve control over the bacteria.

5.2.1 Aside: the use of synthetic phage

Although there are regulatory impediments currently precluding the use of genetically modified phage for human treatments, synthetic or engineered phage are a reality that may well, in time, become a useful clinical tool and their efficacy has been demonstrated in *in vivo* murine models as adjuvants to small molecule antibiotics [79]. From the perspective of engineering principles for phage design, it is important to acknowledge that if we are to expend energy into the design and synthesis of such antibiotic agents, it is essential we undertake the analyses required to understand the longevity or 'evolutionary resilience' of phage as a class of antibiotic agents. It is also important to understand what might make for a successful phage therapy and what distinguishes those from unsuccessful ones.

5.3 Bacteria-Phage Coevolution

Phage are viruses that are pervasive predators of bacteria. There are numerous studies showing that when phage and bacteria are co-evolved, an evolutionary arms race occurs, in which the resistance on the bacteria and the infectivity of the phage both increase, seeking a replicative advantage. Although this has been demonstrated in the lab, in rich media [2], it is questioned whether this is a plausible scenario in the wild, with recent research showing that in non-sterile soil co-evolution experiments, *P.fluorescens* gained more resistance to contemporary phage than previous phage; this demonstrates a more fluctuating evolution [21]. It is frequently stated that co-evolution is an important driver of ecological diversity, as it creates a large number of niches for generalist (low infectivity, wide host range) and specialist (high infectivity, narrow host range) phage to exploit, with a similar diversifying effect on the bacterial host. The understanding of the dynamics of co-evolution have been highly influenced by the proposal of Van Valen's Red Queen Hypothesis (RQH) [115] which states that continuous evolution is driven by biotic factors, and can occur in a fluctuating, escalating or chasing manner. In the fluctuating RQH where interactions between host and predator are controlled by a few traits under tight genetic control, allele frequency will oscillate in a negative frequency dependent manner, rare host genotypes will be at an advantage and will therefore become more common - and therefore more predated [53]. The escalating RQH is perhaps the most commonly described, whereby interacting traits are selected for in a single direction, leading to successive adaptations in both parties, with the change in fitness being zero relative to the other party [87]: in evolutionary terms this is a case of running to stand still. Finally the chasing the Red Queen theory is similar to an arms race, however it occurs normally when there are numerous traits which can adapt to increase fitness; therefore the co-evolutionary game is constantly changing [68]. In this chapter, the interaction between lambda phage and REL606 will be expected to follow fluctu-

ating or escalating dynamics, as the attachment site for the phage is a maltoporin essential for uptake to the bacteria of maltotriose (the only carbon source supplied).

The success of phage therapy, defined as the ability of phage to sterilise (of some target bacterium) the environment into which it is inoculated, is likely to be contingent upon the ecological niche of the host bacterium and the availability of extracellular nutrients within that niche. For example, if we take the case of T7 phage and *Escherichia coli* co-cultured in a chemostat environment, the bacterium need not be cleared by the phage [108]. In that case it is shown that if extracellular nutrients, like carbon sources, reside at sufficiently low concentrations, the bacteria are able to persist in such a way that can also support a phage population in steady-state. It is therefore already known that phage treatment alone need not be sufficient to clear an infection and that lack of success can be due to an insufficient supply of carbon which prevents enough phage from being made to clear the bacterium.

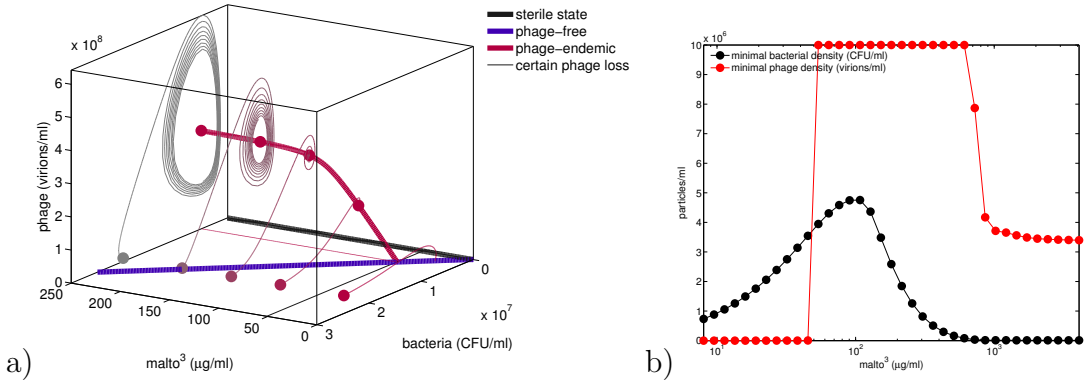


Fig. 5.2: **The predictions of a very simple mathematical model of phage therapy**. (a) Shows the locus of steady-states in (B, P)-space as a function of sugar supply, S_0, shown in black (the sterilised state where $B = P = 0$), blue (the phage-lost state where $B > 0$ and $P = 0$) and the phage-endemic state (where $B > 0$ and $P > 0$). Superimposed on this are five dynamical plots determined at different S_0 concentrations which show ever-larger Lotka-Volterra-like convergent but oscillatory dynamics that grow in amplitude as S_0 increases following the introduction of a small number of phage into a bacterial population. Plot (b) then shows how the *smallest* observed bacterial density across each oscillation (a proxy for the likelihood of phage therapy success) eventually decreases as S_0 increases.

We can see this in a very simplified mathematical of phage-bacterium dynamics in which neither are permitted to evolve with respect to any of the modelled traits. So consider this model of their respective population dynamics in a chemostat with

washout rate d and carbon supply parameter S_0:

$$\frac{d}{dt}B = \left(c\frac{VS}{K+S} - d - \phi P\right)B, \tag{5.1a}$$

$$\frac{d}{dt}P = (\beta\phi B - d)P, \tag{5.1b}$$

$$\frac{d}{dt}S = d(S_0 - S) - \frac{VS}{K+S}B. \tag{5.1c}$$

Here β is the phage burst size (number of phage released upon lysis), ϕ is the infection likelihood on contact between bacteria B and phage P, S is the carbon source concentration, K is the half-saturation constant of bacteria for the sugar, S, and V is the maximal uptake rate of sugar by bacteria. The dynamics of equation (5.1a-c) are shown in Figure 5.2. This figure shows how small S_0 values are sure to either lead to phage loss; at higher S_0 values one observes convergence towards a steady-state in which phage are endemic.

The question therefore remains, how can phage therapy work in this simple model and this is answered by Figure 5.2(b). This shows how the oscillations seen in the dynamics of Figure 5.2(a) are able to lead to the stochastic loss of the bacterium only when a bloom of bacteria subsequently creates a bloom in phage which then, in turn, massively suppresses the bacterial population later in time. This suppression, if sufficiently large, might be enough because of the high, yet momentary, phage supply to completely remove the bacterium; this argument requires a little demographic stochasticity in addition to the model in (5.1a-c) so that small bacterial populations are not able to recover. Note that the model, as presented, also includes no phage resistance on the part of the bacterium.

Thus, as Figure 5.2(b) shows, when sugar supply increases, the blooms within the boom-bust cycles in the bacterial and phage population dynamics become ever more pronounced and it is this feature that leads to bacterial densities becoming ever smaller at their minimal point. However, increasing sugar in this way is a potentially double-edge sword in terms of practical considerations because it also has the property that the bacterial blooms are become larger, commensurate with the value of S_0. Therefore, in order to have a large enough bloom to enable clearance by phage, a population of bacteria may be required that is so large it would have fatal toxic effects before the clearance could occur. So while the long-term, stable steady-state value of B is independent of S_0 when the phage are present in Figure 5.2(a), the blooms of B do depend on S_0, something that could be lethal for an infection treated *in vivo*.

5.4 Pan-phage-resistant *de novo* mutations in the bacterium

Missing from the model defined by equations (5.1a-c) is the idea of adaptive evolution. Thus, the model is too simple to allow for the bacterium to become a better competitor for resources by ingesting the sugar more quickly, or by becoming resistant to the phage. In addition, the phage are not allowed to co-evolve with the bacterium and become more infective or for its burst size to adapt to changing circumstances. We therefore consider the implications of the resistance properties of co-evolving phage and bacteria as discussed in the introduction.

So, there is another possibility that can lead to the failure of phage therapy, even within theoretical windows of extracellular resource concentrations might clear the bacterium: the bacterium evolves to become pan-phage-resistant before the phage is able to sterilise the infection. One study contains just such data on the existence of bacteria that resist attack by all 'known' phage; i.e., the bacteria is resistant to the phage that is currently trying to infect it, but remains resistant to all preceding phage genotypes - this is an example of escalatory RQH [51]. More precisely, this study investigates the coevolutionary dynamics of *Pseudomonas fluorescens* SBW25 and the phage SBW25Φ2 and bacterial mutants that were universally resistant to all isolated phage emerged during the first 20 transfers in five out of six replicate batch transfer experiments (transferred every 48h, 60 transfers in total).

It is noteworthy that in none of those replicates did the completely resistant bacterial types fix in the bacterial population. In two replicates the resistant mutants fluctuated at frequencies between 0.8 and 1.0 and the authors note however that the presence of phage at times when frequency of resistance reached 1.0 indicates that sensitive bacteria where still present in the community, although they failed to sample them. In one replicate complete resistance reached a frequency of only 0.4 by the end of the experiment. In two other replicates the universally resistant types reached frequencies of 0.4 and 0.8 over the next 20 to 30 transfers and were subsequently lost from the population.

This study therefore points to an important hypothesis: complete resistance to phage can occur rapidly during phage therapy via *de novo* genetic change and, moreover, those mutations need not fix in the population. This may because these mutations are associated with metabolic, or 'competitive', costs. We are therefore interested in measuring costs of resistance in the context of the λ-phage treatments of *E. coli* in Figure 5.1 and in asking whether there are mutations that induce phage-resistance, and if so, is there any evidence of pan phage resistance, and, if so, does that resistance also come with associated metabolic costs?

The remaining data in this document will show that phage resistance can evolve *de novo* and that it need not be associated with bacterial fitness costs because the

adapted bacterial proteins can be mutated in such a way that dual adaptive benefits can accrue from those mutations. Moreover, there can even be regions, so-called adaptive hotspots, on the bacterial proteins that can provide the dual benefits.

5.5 A 3-day Bacteria-Phage Coevolution Experiment and Two Libraries

Each well of a 96-well microtitre plate containing liquid minimal growth medium with $125\mu g/ml$ maltotriose was inoculated with wild-type *E.coli* B(REL606) and λ-phage. REL606 was used as it shows good growth on maltotriose, and has a fully sequenced genome. The two populations were coevolved for three days by transferring around 1% of the contents of each well to a new plate every 24h in which fresh growth medium had been prepared. At the end of the three day period, a number of bacteria and phage were isolated and the *lamB* gene of each isolated bacterium was sequenced. The *lamB* gene was sequenced in the ancestral strain, and compared to sequenced genes from the co-evolved strains. All sequencing was done by a collaborator, Justin Meyer. The *lamB* gene is a maltoporin responsible for the transport of maltotriose into the cell. However it is also the attachment site for lambda-phage, and therefore is a suitable target for sequencing, as it is essential to both the bacteria in a maltotriose carbon environment, and to the phage as an attack site. A unique identifier was associated to each isolated genotype of bacteria and phage and the *lamB* amino acid sequence folded using an approximate free-energy-minimisation algorithm. This resulted in a library of 93 phage and 49 bacteria (Table 5.1). These co-evolution experiments were performed by Justin Meyer (unpublished results), who supplied the phage library for use in this study.

An infectivity/resistance matrix can be derived from the analysis of phage infectivity as following the methodology in Appendix 5.1 (Material and Methods). When each phage type and each bacterial type are assayed and processed in this way, the result is depicted in Appendix 5.1 (Figure 5.3). The same data is illustrated using a different algorithm in the same appendix (Figure 5.4) that is more sensitive to minor differences in the infectivity phenotype the image processing algorithm determines. Perhaps the most prominent region in Figure 5.3 is the white area showing bacteria in the library that cannot be infected by any of the isolated phage.

5.5.1 Is There A Phage Host-Range Trade-Off? No!

We can use the infectivity data in Figure 5.3 to determine whether the phage that are most infective suffer any costs of infection by having a reduced bacterial host-range, in other words, do phage suffer a host-range trade-off. Thus, to each bacterial genotype and to each phage we can associate the following phenotypes: the first is

Table S 5.1: *E.coli* **B(REL606) mutants derived from a wild-type strain following coevolution with λ-phage, showing their library identifier and the mutations found in** *lamB*.

label	**mutation found** (SNP unless stated otherwise)	**location on** *lamB* (b.p.)
wt ($j = 0$)	N/A	
2b	deletion: -GAC	292-294
4a	C → A	854
7a	insertion: +TCGCAGGGTAAAGGGCTG	795-796
8a	-GATTTCAA	1264-1271
9a	-AAC	871-873
11a	T → G	556
13a	+TACGGTCGTGCCAACTTGCGTGATAACT	652-653
13b	-GCCGTTCCTGCTGATTTCAA	1252-1271
14a	T→ G	835
17a	-AAC	871-873
18a	T → G	538
19a	+T	610-611
20b	-A	349
21b	+TCGCAGGGTAAAGGGCTG	795-796
22a	G → T	1210
23b	A → C	793
26a	-AAC	871-873
27a	+T	597-598
28b	-T	726
29a	A → G	518
30a	G → A	571
41a	A → C	284
47a	T → A	850
50b	G → T	874
51a	C → A	824
51b	C → A	843
52a	G → T	809
56a	-AAAGGGCTGTCGCAGGGT	805-822
57b	C → T	788
59a	G → C	1285
60a	G → T	709
61a	A → C	518
61b	G → T	559
62b	A → G	1211
65b	C → G	1059
66a	+ACTTCG	1240-1241
67a	C → T	136
68a	C → T	815
69a	-CAA	864-866
70a	A → C	793
70b	-TCGCAGGGTAAAGGGCTG	796-813
71a	A → C	793
71b	G → T	809
94a	-ACCGATCCGGCC	304-315
95a	G → A	821
96a	A → T	760
96b	A → C	581
97b	A → C	580
98a	deletion & insertion	704-719
99a	A → G	560

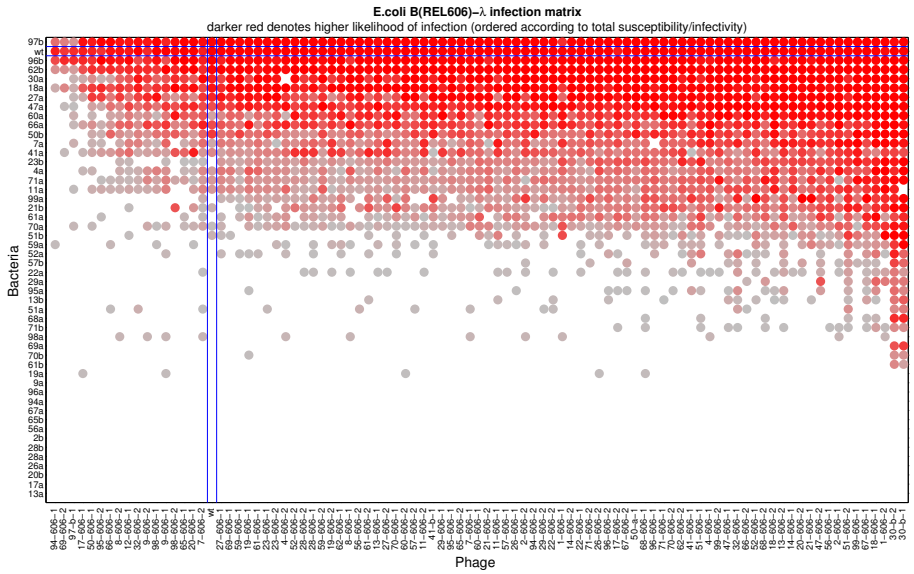


Fig. 5.3: A representation of the genotype-by-genotype infection matrix showing infectivity on a white-grey-red scale of increasing infectivity; the wild-type phage and bacteria are indicated by the blue lines. The value of the red dot at the intersection of these two lines, the wildtype-to-wildtype interaction, is unity. The main feature of interest is the large white region representing a cluster of completely resistant bacterial genotypes.

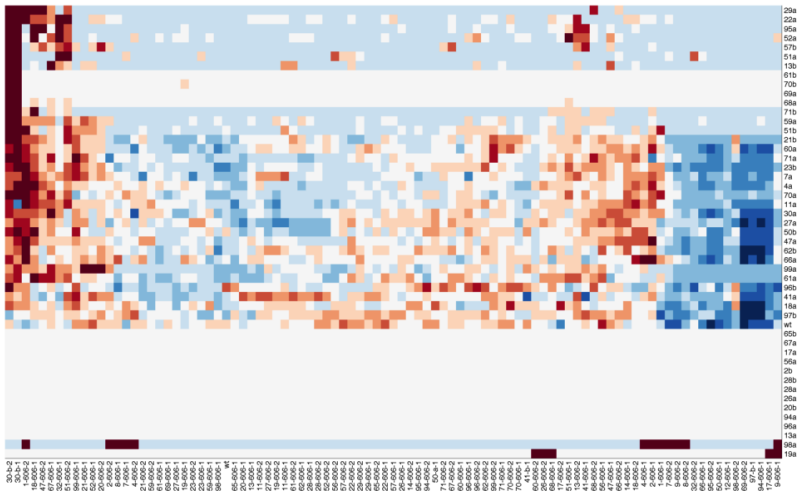


Fig. 5.4: A genotype-by-genotype infection matrix displayed using Matlab's `dendrogram` routine, bacteria are represented vertically and phage are horizontal. The main feature of interest is the large white region representing a cluster of 'pan-library-phage-resistant' bacterial genotypes.

a vector of dimensionless values between 0 and 1 representing the infectivities of each phage genotype to every bacterium, this is a sum of the matrix columns in Figure 5.3. This sum is then normalised such that the maximal observed phenotype of all genotypes in the library is unity. The second phenotype is then the number of non-zero entires that appear in this column, in other words, the size of the host range.

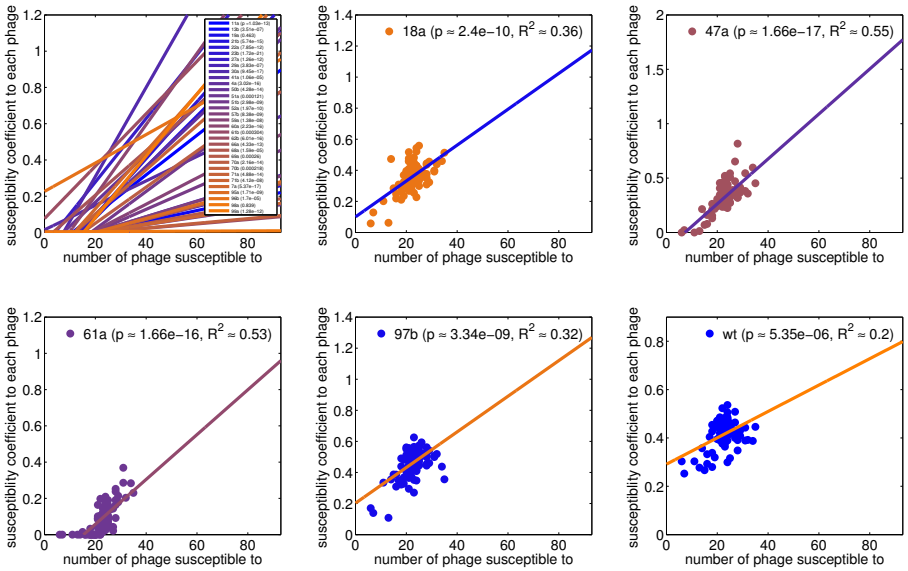


Fig. 5.5: **The smaller the set of phage that infect an E. coli the more resistant it is to the phage that can infect it**. We can also say it this way: the more susceptible an *E. coli* strain is to each phage, the more phage there are that can infect it. The top-left plot shows all the regressions for all bacterial genotypes with the significance of the non-constant nature of the linear regression given in the figure legend; all but two regressions are significant. The remaining five plots are example regressions for five genotypes, including the wild-type strain, REL606, denoted 'wt' in the figure legend.

Figure 5.5 shows the results of analysing the resulting pairs of phenotype for all phage strains using a series of robust linear regressions. If there were a host-range trade-off, one should observe that phage which are infective to more bacterial hosts should also pay the penalty, or cost, of being less infective of the hosts they can infect. Thus, the linear regressions in Figure 5.5 would gave negative slope in the presence of such a trade-off, however, this is clearly not the case. We can reject the hypothesis that the regression is negative in all but two of the regressions because each such slope is, in fact, positive and therefore consistent with a 'host-range trade-up'.

Informally, this can be seen as the idea that Figure 5.3 has a nested structure because the bacteria can be classed, or ordered, according to how susceptible or 'sticky' to phage they are. Simply put, some have phage receptors that are more sticky than others and thus bind both more genotypes of phage and, of those that can bind, they bind better, at least according to the assay as implemented.

5.6 Bacterial Growth Trade-Offs

Resistance is usually assumed to carry a fitness costs to the bacterium which resists the phage, this fitness cost is assumed to form a trade off: the greater the phage resistance phenotype, the lower bacterial fitness will be, where the term 'fitness'

can be interpreted in a variety of measures of vegetative growth. However, trade-offs often prove elusive in practise and so we sought data on a wide variety of different trade offs in our library of *E.coli* and λ-phage mutants. To study the existence of trade-offs we focus on the interactions between four bacterial traits: exponential growth rate, duration of lag phage, cell yield per carbon supplied and phage infectivity using the above infectivity matrices.

To summarise our findings, we observe that there may be no trade-offs, trade-offs that are present may be environmentally mediated, and there can even be trade-ups. A *trade-up* arises when a bacterial strain can simultaneously increase two traits through a single mutation, for example when complete resistance to phage predation arises alongside an exponential growth rate increase due to that one mutation. Indeed, it is worth stressing again that we will show that there are mutations that confer complete resistance to all of the phage in our library and yet those mutations also appear to make bacteria fitter than the wild-type bacterium from which the mutant was derived. We discuss the potential evolutionary implications of such 'superstrains' in the remainder (Table 5.2).

Table S5.2: Six possible trade-offs from four bacterial traits where $\star$ can be either 'within strain', 'between strain', and 'trade-off' or 'trade-up'.

four traits	cell yield	lag time	phage resistance
growth rate	$\star$	$\star$	$\star$
cell yield		$\star$	$\star$
lag time			$\star$

5.6.1 Bacterial library: growth kinetic parameters

Each bacterial genotype was grown in a single flask for 24h in the absence of any competitors in order to measure the following metabolic parameters. If bacterial density is given by $b(t)$, we suppose that b follows a model of the following logistic form

$$b(t) = b_0 + \frac{k}{1 + a\exp(-rt)}.$$

Using this simple sigmoidal growth model, we can fit to growth data taken from a growth assay to produce a phenotype (a, k, r) for each bacterial genotype. One exemplar fit between data and model is shown in Figure 5.6. More fits are shown in Figure 5.7

We can robustly determine the blank parameter, b_0, the maximal population density parameter, k, and the exponential growth rate r. The length of lag phase is determined in a different way. For this we perform a series of regressions on data starting at the time when there are three density datapoints available, this is

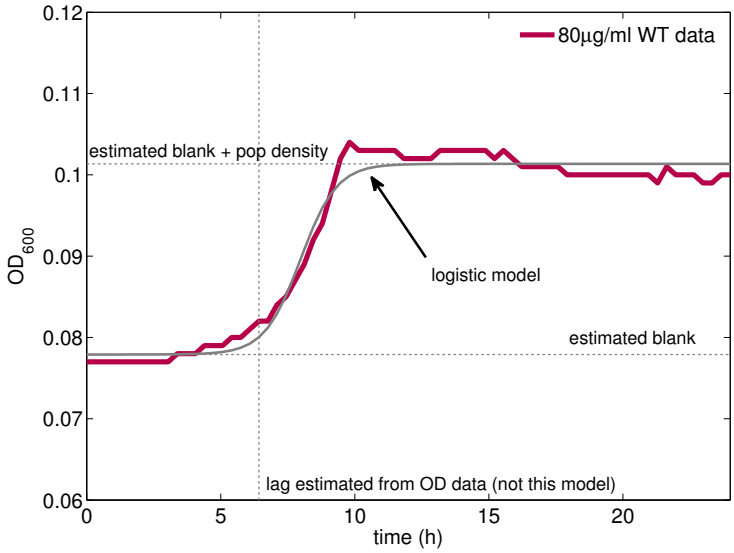


Fig. 5.6: A visual comparison of short-term bacterial growth using a mathematical model and experimental growth data determined in a microtiter plate reader and shown for the *E.coli* B(606) wildtype. The residual used for the fit only uses data up to the end of exponential phase, the x-axis denotes the length of the experiment in hours.

following one hour of growth. We then perform a linear regression on that data, once it has been filtered to remove high-frequency noise. Then, we decide if that regression is significantly non-constant at the $p = 0.01$ level, if so, we decide that the lag phase has lasted 1h. If not, we continue to add in density datapoints from later times until such time as we do achieve significance at the $p = 0.01$ level using the same regression test. If so, the earliest time point for which this test is successful is deemed to be the lag time.

The entirety of all the raw observed data is presented in Figure 5.8 where the bacterium has been cultured in 2.5, 5, 10, 20, 40, 80, 160 and 250$\mu g/ml$ of maltotriose in the Davis Minimal (DM) culture medium. DM media is recommended for the isolation of nutriotional mutants from *E.coli*, additionally the minimal nature of the media ensures that maltotriose is the only carbon source available to the bacteria from the media (although obviously there will be cross-feeding and scavenging from other carbon sources within the duration of the experiment). The figure shows the opportunity for the data to probe a number of common hypothesis relating to the nature of trade-offs. For example:

1. **RYTO** The rate-yield trade-off is a trade-off whereby higher growth rates comes at a 'cost' of lower cell yield per mole of carbon supplied. In other words, faster growth is necessarily less efficient. One can see apparent evidence for the RYTO in Figure 5.8(c) where the collation of all the raw metabolic data does indeed appear to show a negative and highly nonlinear correlation between rate (the parameter r from the datafits) and yield (total population density

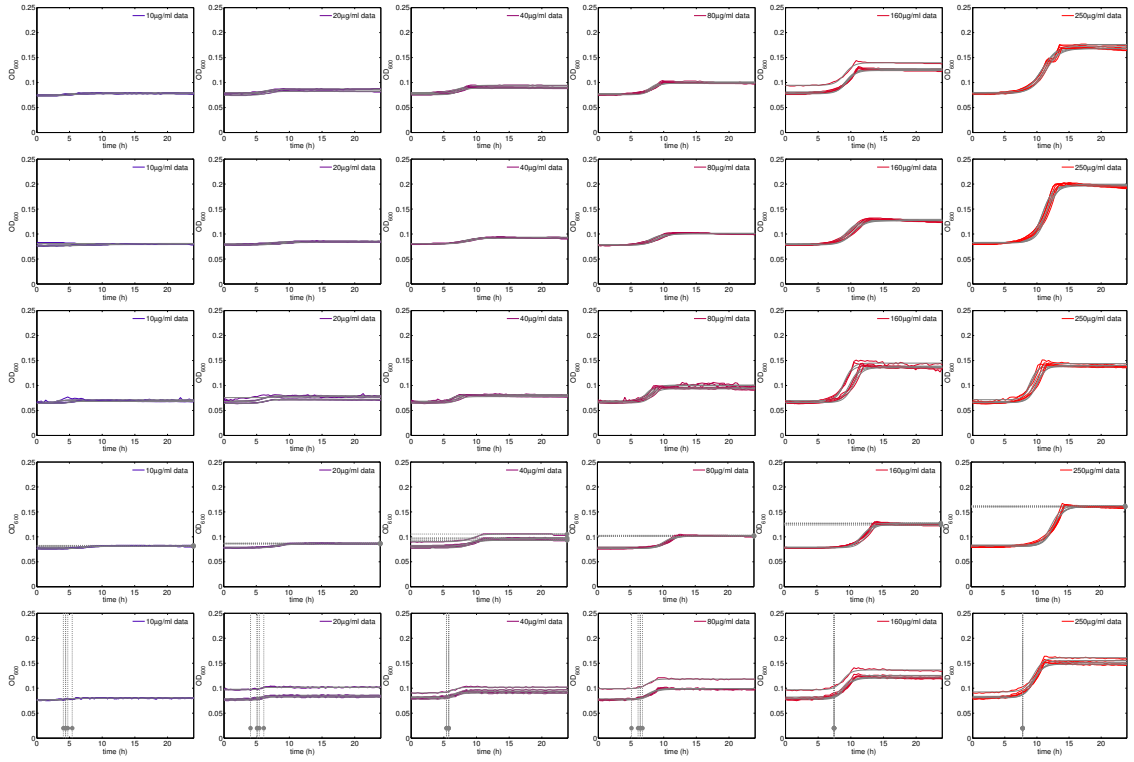


Fig. 5.7: A visual comparison of short-term bacterial growth using a mathematical model and experimental growth data determined in a microtiter plate reader and shown for the *E.coli* B(606) wildtype for six resource concentrations. The residual used for the fit only uses data up to the end of exponential phase, the x-axis denotes the length of the experiment in hours. An approximate 95% confidence interval about each prediction is shown, reflecting the uncertainty in the estimated parameters.

divided by sugar supplied).

2. **LRTO** The lag-rate trade-off is the idea, apparently consistent with the raw data in Figure 5.8(a) that an increased growth rate comes at the cost of a longer lag phase. This is not a commonly-studied trade-off in the literature.

3. **LYTO** As rate is negatively correlated with yield (RYTO) and rate is positively correlated with lag (LYTO) it follows that lag and yield must be negatively correlated, as can be seen in Figure 5.8(b).

The nature of the dependence of the duration of lag phase on the supply of sugar (maltotriose) to the cells is shown in Figure 5.9 for four library strains that illustrate the four different classes of behaviour in the entire library. If y denotes the lag phase data and x is the set of maltotriose concentrations used then the model $y = a + bx^p$ groups the dataset into four clusters (not shown) whereby $p = 0, 1/2, 1, 2$ denoting constant lag, linear increasing lag, or nonlinearly-increasing lag. The mechanistic basis of the observed behaviour is unknown at the present time.

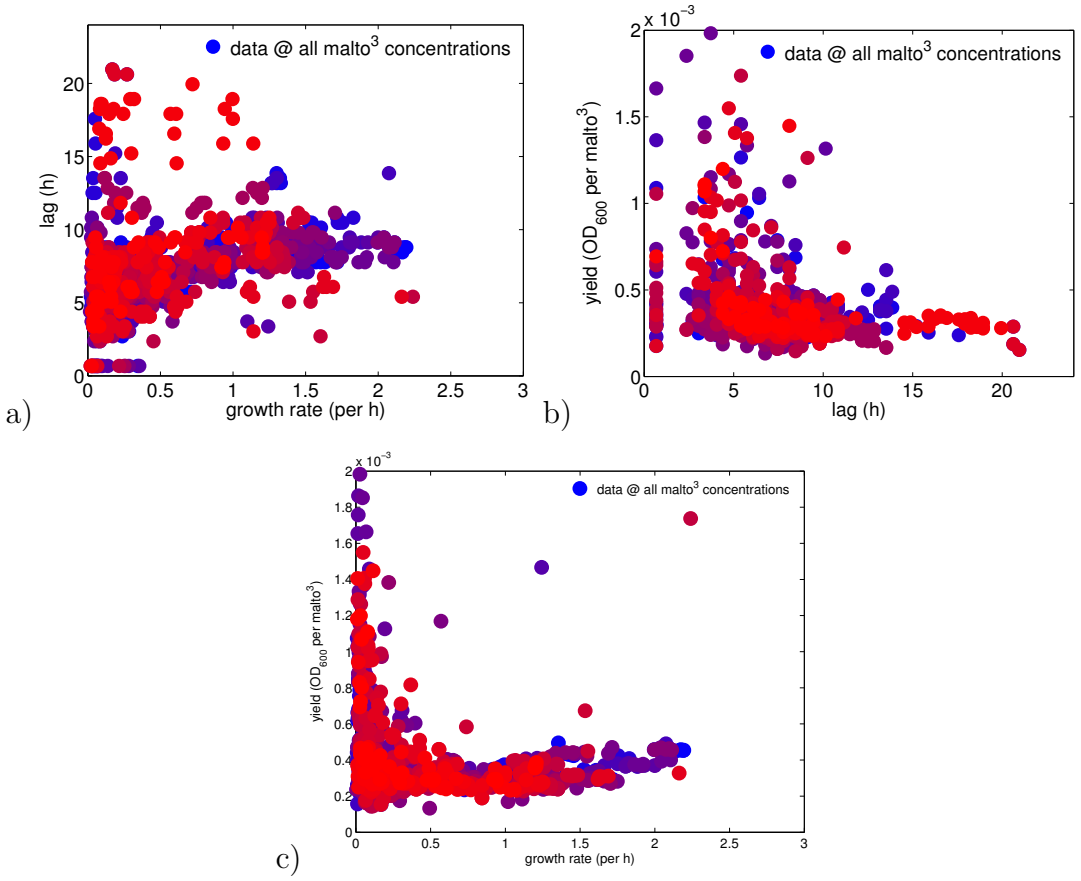


Fig. 5.8: **Three vegetative growth traits measured in laboratory conditions indicate three trade-off hypotheses**: (a) a longer lag phase yields a greater growth rate; (b) longer lag may correlate with lower yield; (c) greater growth rate may correlate with lower yield. The latter is the classical rate-yield trade off (RYTO). The colour code indicates cultures in low maltotriose conditions for all strains (the bluest data) to cultures in high maltotriose conditions (red data).

5.6.2 Is there a relationship (potential trade-off) between lag phase and phage resistance? No!

We sought a relationship, or potential trade-off, between the duration of lag phase and phage resistance, the hypothesis being that longer lag would correlate with greater resistance, and we have the datasets to probe this hypothesis in the above. When we therefore performed two different tests, one appropriate for continuous data and one appropriate for categorical data, both tests are consistent that there is no evidence for the hypothesised trade-off. This can be seen in Figure 5.10.

In Figure 5.10 we see a Kruskal-Wallis one-way analysis of variance performed on all bacteria from the library that exhibit what are defined to be two classes of resistance phenotype, these are either complete library-pan-resistance or else being susceptible to at least one of the phages. None of these tests show any significant evidence of different lag times between those two classes, although there are some maltotriose conditions (2.5, 5, 80$\mu g/ml$) that are consistent with longer lag times

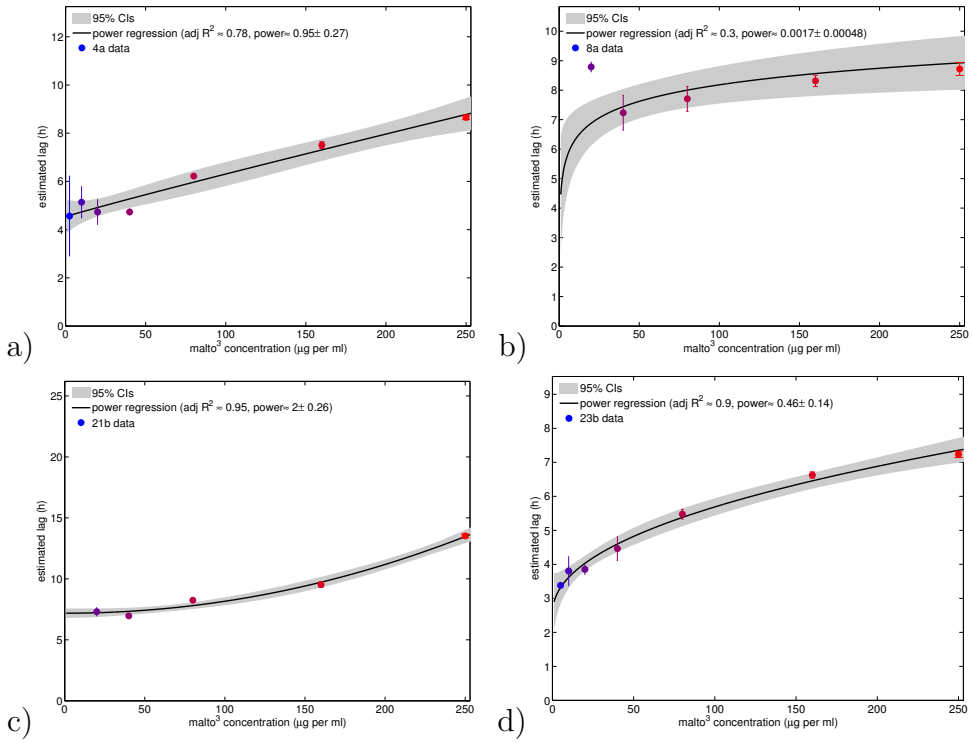


Fig. 5.9: The metabolic dataset contains four different classes of sugar versus lag behaviour: (a) shows a linear increase in lag as maltotriose increases in concentration, (b) shows constant, or near-constant, lag as sugar increases, (c) shows a quadratic increase in lag with sugar increase and (d) shows a square-root increase in lag with sugar. These have been determined by fitting a power law of the form $y = a + bx^p$ to the data, where the value of p is reported in each of the four plots where it is called the 'power'.

being associated with greater phage resistance. But the evidence for this claim across the entire dataset is not replicated between the different conditions tested.

Indeed, the totality of evidence for trade-offs in our dataset that involve the lag parameter is weak. Take Figures 5.11 and 5.12 which show attempts to correlate lag with yield and lag with growth rate, respectively. There is evidence at some resource concentrations of a negative correlation, and it is statistically significant at those particular resource concentrations, but the correlations are weak and they are not consistent between different resource conditions. We therefore refrain from declaring that there is any evidence for a lag-rate or lag-yield correlation, despite the evidence for it in some of the experiments we performed.

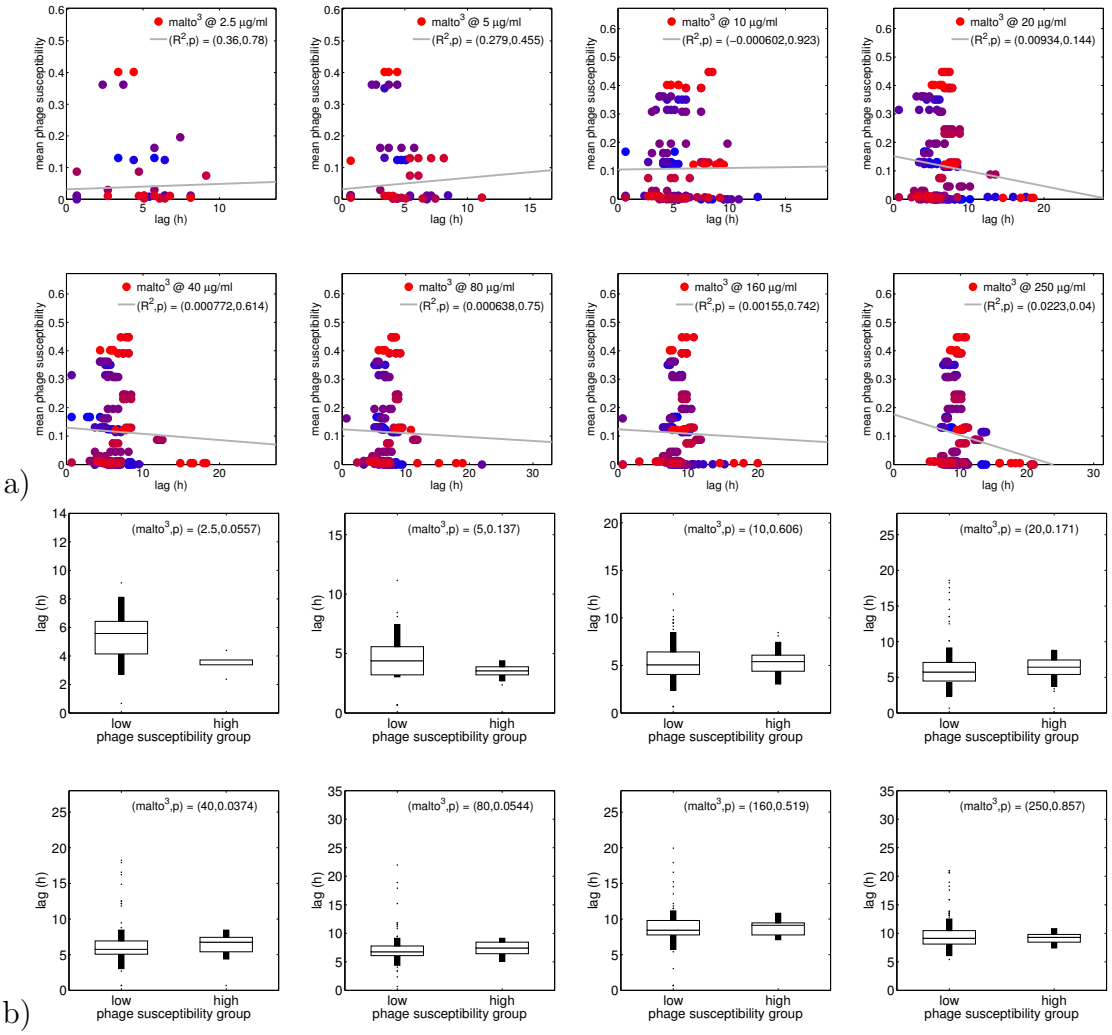


Fig. 5.10: **Two tests seeking a correlation between lag duration and phage resistance**. (a) For each maltotrise concentration the duration of lag is plotted versus the total infectivity phenotype for that bacterium and then a robust linear regression is performed. None of the regressions are significantly non-constant, apart, perhaps, the regression at $250\mu g/ml$ but even here any possible correlation is very weak. Note: different colours denote data for different bacterial strains. (b) A Kruskal–Wallis one-way analysis of variance performed on bacteria exhibiting two classes of resistance phenotype (either complete library-pan-resistance or else being susceptible to at least one phage) shows no evidence of different lag times between those two classes.

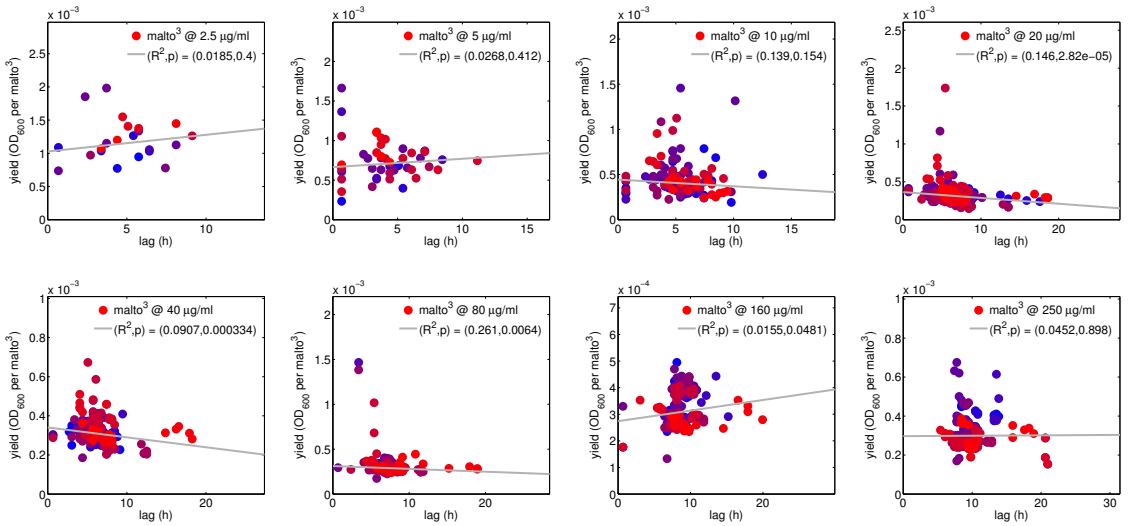


Fig. 5.11: Metabolic data seeking correlations between yield and lag time at different resource concentrations. Note: different colours denote data for different bacterial strains.

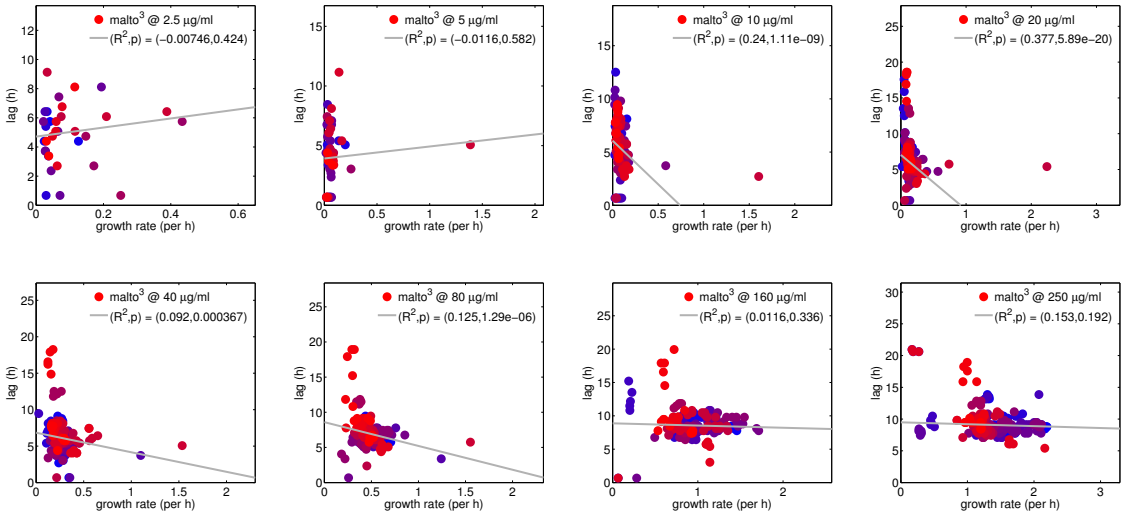


Fig. 5.12: Metabolic data seeking correlations between yield and lag time at different resource concentrations. Note: different colours denote data for different bacterial strains.

5.7 The Rate-Yield Trade-Off (RYTO)

Supporting the idea of a rate-yield trade off is the notion of a branched metabolic pathway, such as that found in glycolytic-TCA structure found in almost all living cells. A schematic of this structure that is sufficient to describe a RYTO is shown in Figure 5.13 where the legend describes the verbal rationale behind how a RYTO with respect to ATP production is derived.

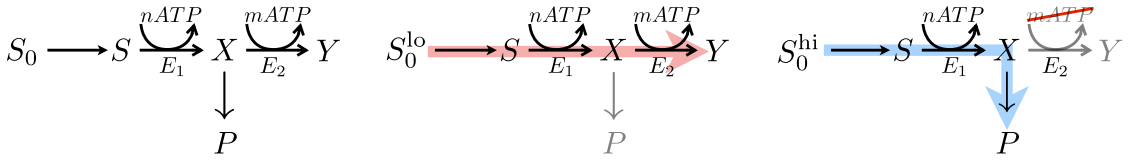


Fig. 5.13: A RYTO in terms of ATP yield as a function of the rate at which substrates are processed can be the result of a branched metabolic structure. The middle plot shows that when resource supply is low, hence so is the rate of ATP production, all metabolic intermediates, X, are processed to form the product Y because enzyme E_2 is not saturated. The right-most plot shows that when substate supply is high, E_2 is saturated and so X is completely lost to an intermediate, P, rather than via the ATP-producing step to Y, P then diffuses away and possibly out of the cell itself, whereupon the substrate produce n units of ATP very quickly, because of the high substrate concentration, rather than the full quota of n+m units; this is therefore a fast but inefficient pathway from the perspective of ATP production.This image and following model are used from a previous publication with permission [83].

It can be shown [13] that if $c(S_0)$ is the ATP yield of the pathway schematic in Figure 5.13 then its branched structure gives the form of a RYTO, with substrate supply concentration S_0, whereby

$$c(S_0) = c_{\mathrm{hi}} \frac{1}{1 + pS_0} + \frac{pS_0}{1 + pS_0} c_{\mathrm{lo}}. \tag{5.2}$$

Here c_{hi} represents the highest possible yield (obtained at the lowest sugar) and c_{lo} is the inverse, with p being the phenotype that controls the rate of yield decrease with increasing sugar. Although this has been derived from the perspective of an ATP-RYTO, we will apply this model to whole-population estimates that are pertinent to a growth-efficient RYTO, assuming that ATP yield and ATP production rate are correlated with cell yield and the rate of cell growth.

5.7.1 A within-strain rate-yield trade off

Figure 5.14 shows that the bacterial strains exhibit growth rate-yield data that is consistent with the mathematical model in equation (5.2), note how yield declines to a plateau as the growth rate increases because of increasing sugar concentration. Figure 5.14 also shows that some of the strains, according to a nonlinear regression

test, can have constant yields and the relatively likelihoods of the fit of a constant model versus equation (5.2) are shown in the legend of the figures. Thus, while the RYTO can be present in particular strains, it need not be in all cases. The growth rate and yield relationship is explicitly shown in Fig. 5.16.

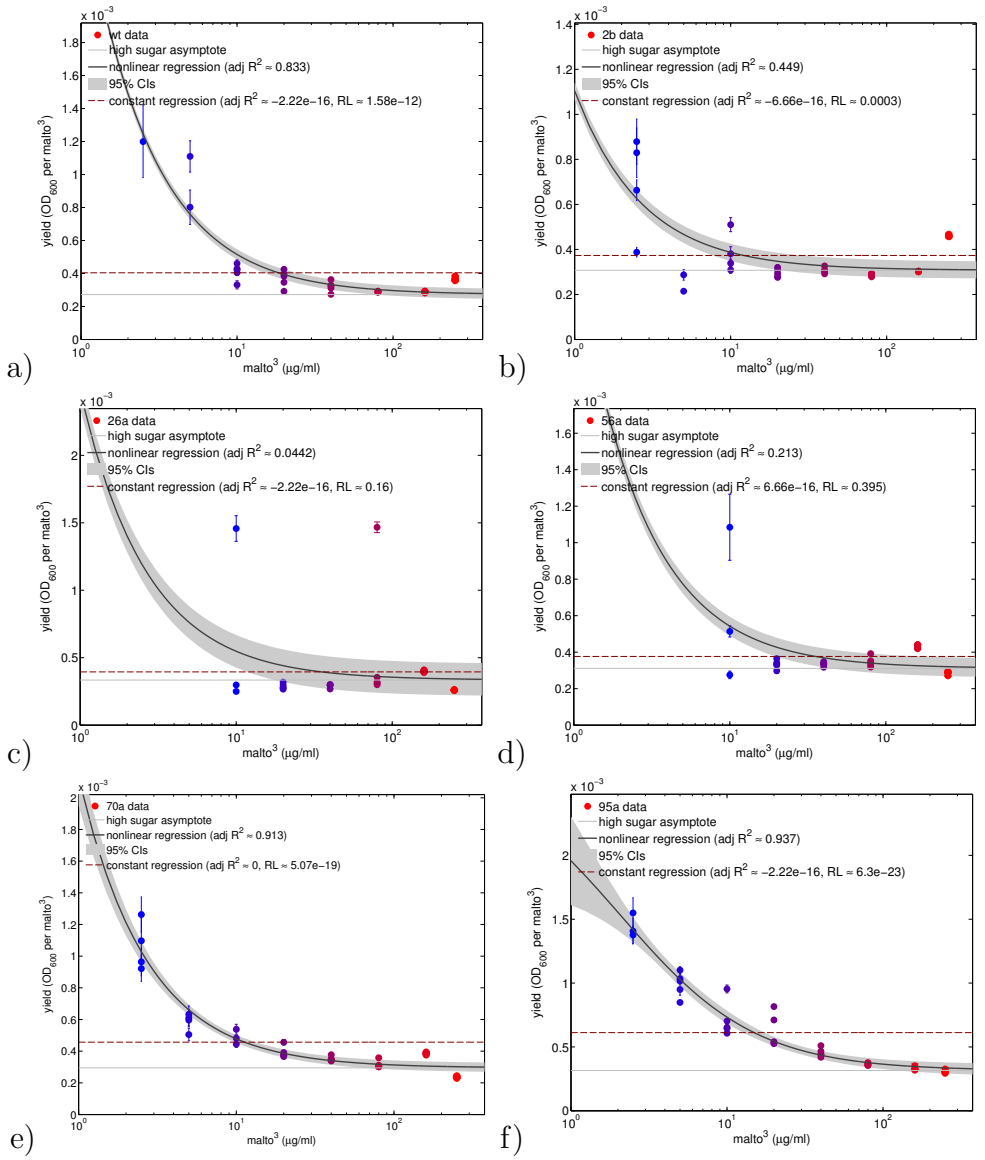


Fig. 5.14: The within-strain RYTO for six different strains from the bacterial library on top of which is superimposed datafits of the theoretical model specified by equation (5.2). The data on the x-axis is sugar concentration, note, but we still call this a RYTO because growth rate is positively correlated with sugar supply at these concentrations. 'RL' denotes the relative likelihood of a nonlinear regression versus a constant (no yield) model.

5.7.2 RYTO, non-Monod, growth kinetics

Theoretical approaches to representing growth rate as a function of extracellular resource concentration have taken the following approach for many decades. The

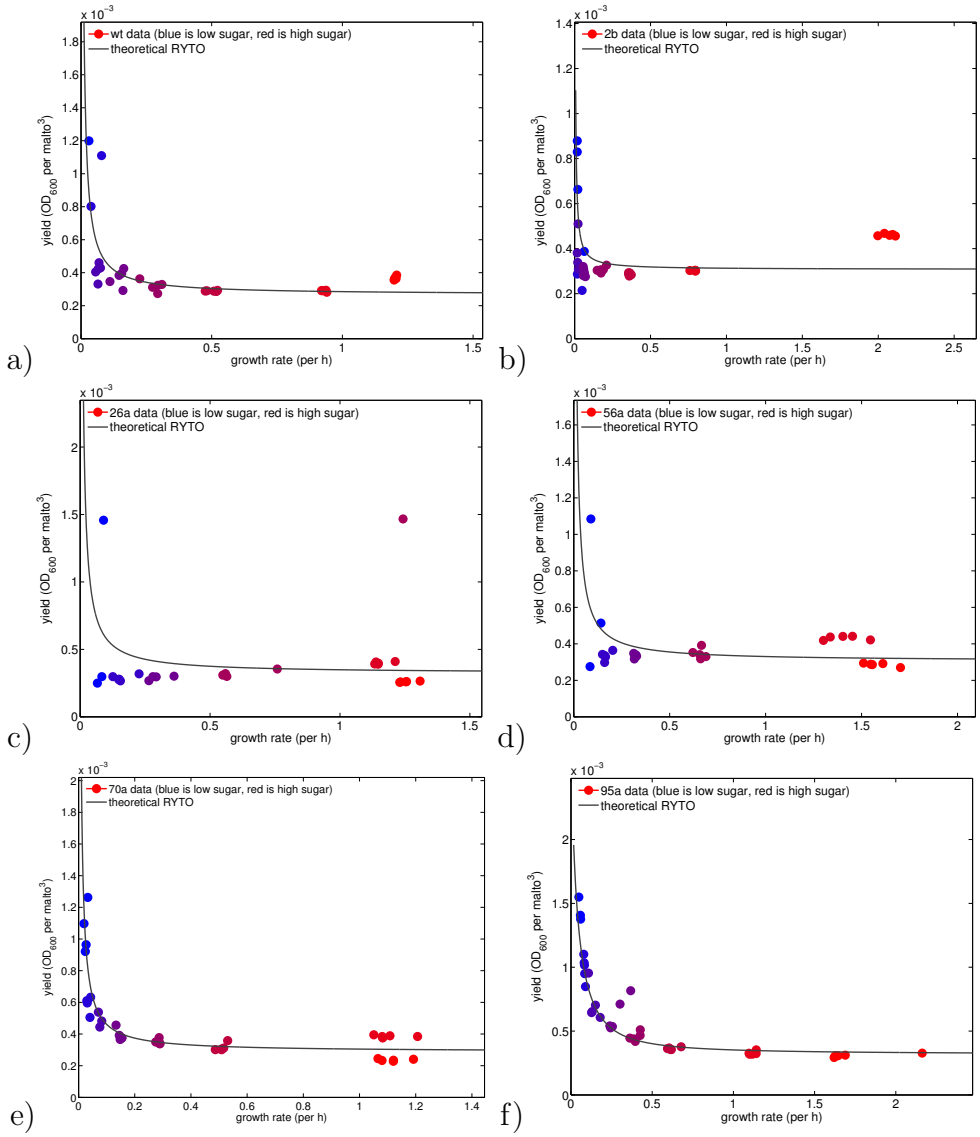


Fig. 5.15: The within-strain RYTO for six different strains from the bacterial library on top of which is superimposed datafits of the theoretical model specified by equation (5.2). The data on the x-axis is growth rate, note.

idea is that the extracellular resource diffuses, randomly attaches to a transporter, with so-called affinity $1/K$, which then translocates the molecule into the periplasm of the bacterium, or perhaps its cytoplasm, and this happens at a maximal rate per unit time, V. Various simplifying assumptions of this nature are then used to propose that the rate of change of the concentration of the resource in the periplasm (or cytoplasm) of the cell is given by

$$U(S) = \frac{VS}{K+S},$$

the 'U' denoting the fact that this is an uptake rate of a substrate, S. In order to turn this quantity, an uptake rate, into a growth rate of a cell, one typically assumes that the intracellular resource is metabolised into biomass with a given efficiency, or

conversion constant, c, so that the growth rate of the cell is

$$G(S) = c \times \frac{VS}{K+S}. \tag{5.3}$$

However, if we acknowledge that c can also depend on S when the latter is a sugar that is metabolised by the cell, because of the RYTO is may be necessary to use the trade-off form for c described above in equation (5.2), and so

$$G(S) = c(S) \cdot \frac{VS}{K+S} = \left(c_{\mathrm{hi}} \frac{1}{1+pS} + \frac{pS}{1+pS} c_{\mathrm{lo}} \right) \cdot \frac{VS}{K+S}.$$

We therefore sought strains in the library for which this composite Monod-RYTO model is a better descriptor of growth rate data than the pure Monod model in (5.3).

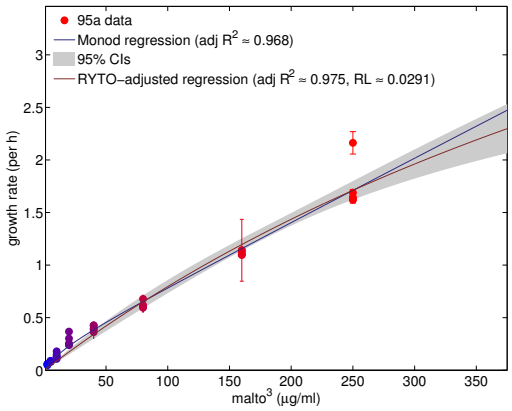


Fig. 5.16: An example of a bacterial strain where a Monod description may need to be modified to include the RYTO in order to get an appropriate description of growth rate data as a function of sugar supplied.

Although a RYTO was present in all strains measured, we found only one strain that showed significant evidence of requiring a RYTO modification of the standard Monod model in order to accurately model the effect of sugar concentration on growth rate. Goodness of model fit was according to a relative likelihood criterion based on AIC values of the respective datafits. It is shown in Figure 5.16. (Akaike informational criteria, a measure of the relative quality of statistical models for a given dataset).

5.7.3 A within-strain RYTO, but a between-strain RY Trade-Up!

Figure 5.17 shows that when one plots scattergrams at all the different maltotriose concentrations tested of growth rate versus yield data, one obtains a positive correlation between these traits. One might have anticipated that the RYTO observed

above within-strain also extended to the between-strain dataset, but it does not. The mechanistic and genetic basis for observing this behaviour in the data are not clear.

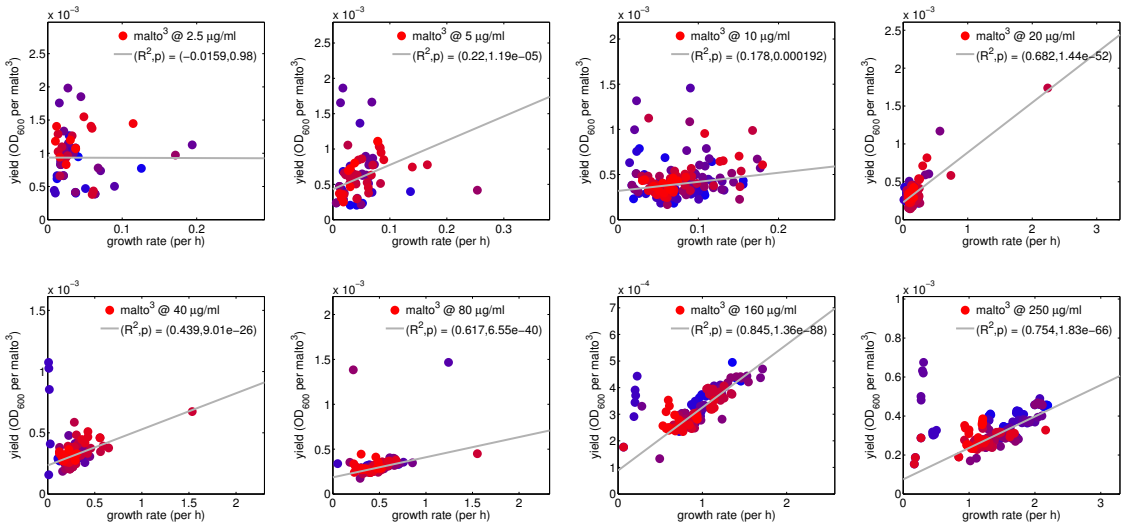


Fig. 5.17: A series of scatter plots of rate and yield data show the existence of a between-strain rate-yield trade-up (RYTU) that is observed at all but one of the maltotriose concentrations tested (namely above $2.5\mu g/ml$ where the data is noisy). Note: different colours denote data for different bacterial strains.

The evolutionary consequences for the absence of a RYTO and instead a RYTU are the following. It forces one to make the prediction that when this strain of *E.coli* evolve in the presence of phage, there may be strains who can increase their growth rates through adaptive evolution and, moreover, those strains are simultaneously capable of increasing their metabolic efficiency. It provides the possibility that there therefore exist super-mutants who can optimise their 'fitness', however the latter term is defined in relation to some combination of growth rate and yield. Thus a RYTU would appear to preclude an evolutionary trajectory towards the coexistence of two specialised mutants, one that specialises in terms of growth rate and the other than has a slower rate of growth but gains by being more efficient; this appears to not be possible among the strains we have isolated in the presence of phage.

5.8 Growth-Resistance Trade-Offs and Two 'superstrains'

It is often assumed, when faced with a stressor like antibiotics or phage, that different bacterial genotypes expressing resistant and susceptible phenotypes coexist because there are fitness costs of resistance. This is based on the idea that there is a truism of living systems whereby entities are either generalists ('jacks of all trades'), or specialists (masters of one particular trade). Thus, based on this idea, a phage-resistant genotype should pay a fitness cost in terms of its intrinsic ability to divide, for example the bacteria better able to resist phage are less able to metabolise or transport a sugar needed for growth. This particular case is the subject of our study, we go in search of such fitness costs and ask whether resistance to phage does result in impaired metabolism. We shall find that it does not and, in fact, we see features in our data of quite the opposite behaviour: masters of all trades can exist.

5.8.1 Costs of resistance: no evidence for growth-resistance trade-offs

We have a sufficiently well-structured dataset to probe hypotheses on the relationship between phage resistance and metabolic traits so we are now interested in whether, or not, there are costs of phage resistance exhibited by the bacteria. To summarise the totality of findings in this direction, we will show below that these questions can be subtle. While there may be costs in the mean, when taken over all sub-groups from the library, there are also reasons why certain strains do not exhibit costs at all. These are the 'superstrains' we described at the start of the chapter.

We first turn to Figure 5.18 which shows that there is little clear and consistent evidence of a negative, or indeed positive, correlation between growth rate and phage resistance. Thus, we have no evidence to claim the existence of costs of resistance in our library based on this data. Moreover, Figure 5.19 also shows there is no basis on which we can claim a negative, or positive, correlation between yield and phage resistance. So, again, there is no evidence here of metabolic costs of phage resistance.

Figure 5.20 provides further evidence that there are no costs of phage resistance, according to the data we have gathered. In Figure 5.20(a) a scatterplot shows the change in growth rate of each bacterial genotype from the wild-type ($\rho(\phi) = \phi_{mut}/\phi_{wt}$ where ϕ denotes the phage susceptibility phenotype) strain versus the change in infectivity from the wild-type strain ($\rho(r) = r_{mut}/r_{wt}$). It shows that while the latter decreases with respect to wild type, as indicated by the values of $\rho(r)$ being less than unity for all genotypes, meaning that the mutants are more phage resistant on average than the wild-type, the former, $\rho(\phi)$, shows no evidence

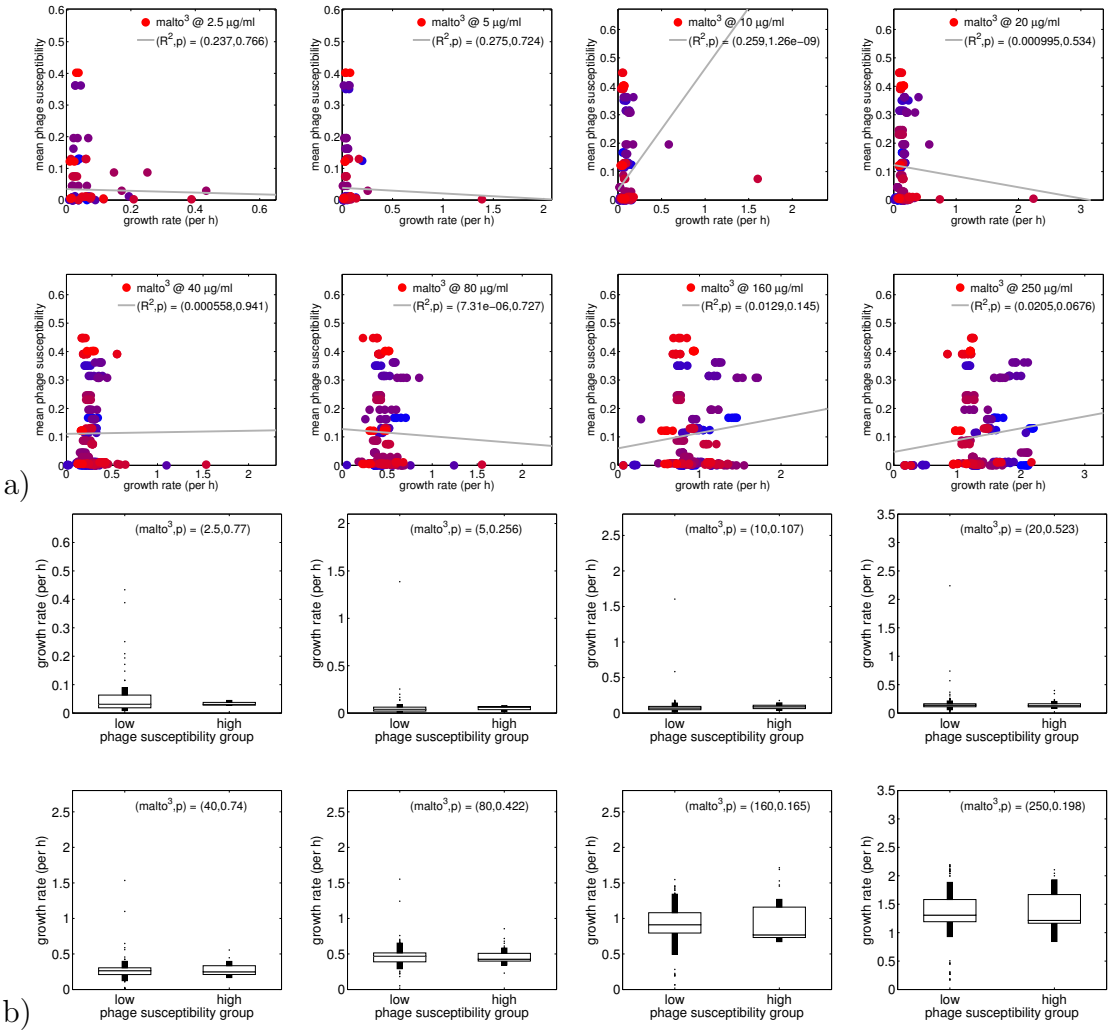


Fig. 5.18: **Two tests seeking a correlation between growth rate and phage resistance**. (a) For each maltotriose concentration growth rate is plotted versus the total infectivity phenotype for that bacterium and then a robust linear regression is performed. Although just one of the regressions is significantly non-constant, namely the regression at $10\mu g/ml$, even here any possible correlation is very weak. Note: different colours denote data for different bacterial strains. (b) A Kruskal–Wallis one-way analysis of variance performed on bacteria exhibiting two classes of phage resistance phenotype (either above or below half of the maximal susceptibility observed to all phage) shows no evidence of differential growth rates between those two classes.

at all of decreasing. Quite the opposite is, in fact, observed.

Importantly, Figure 5.20(b) shows three 'outliers' (56a, 2b, 26a) that are highlighted. These are bacteria that are both pan-resistant with respect to the library of isolated phage but which also exhibit growth rates that are greater than the median and mean of the phage-susceptible class of bacteria. These are therefore candidates for masters of all trades, the 'superstrains' whose existence we alluded to in earlier discussions.

One the left of Figure 5.21(a) we see mathematical model fits of a Monod-only

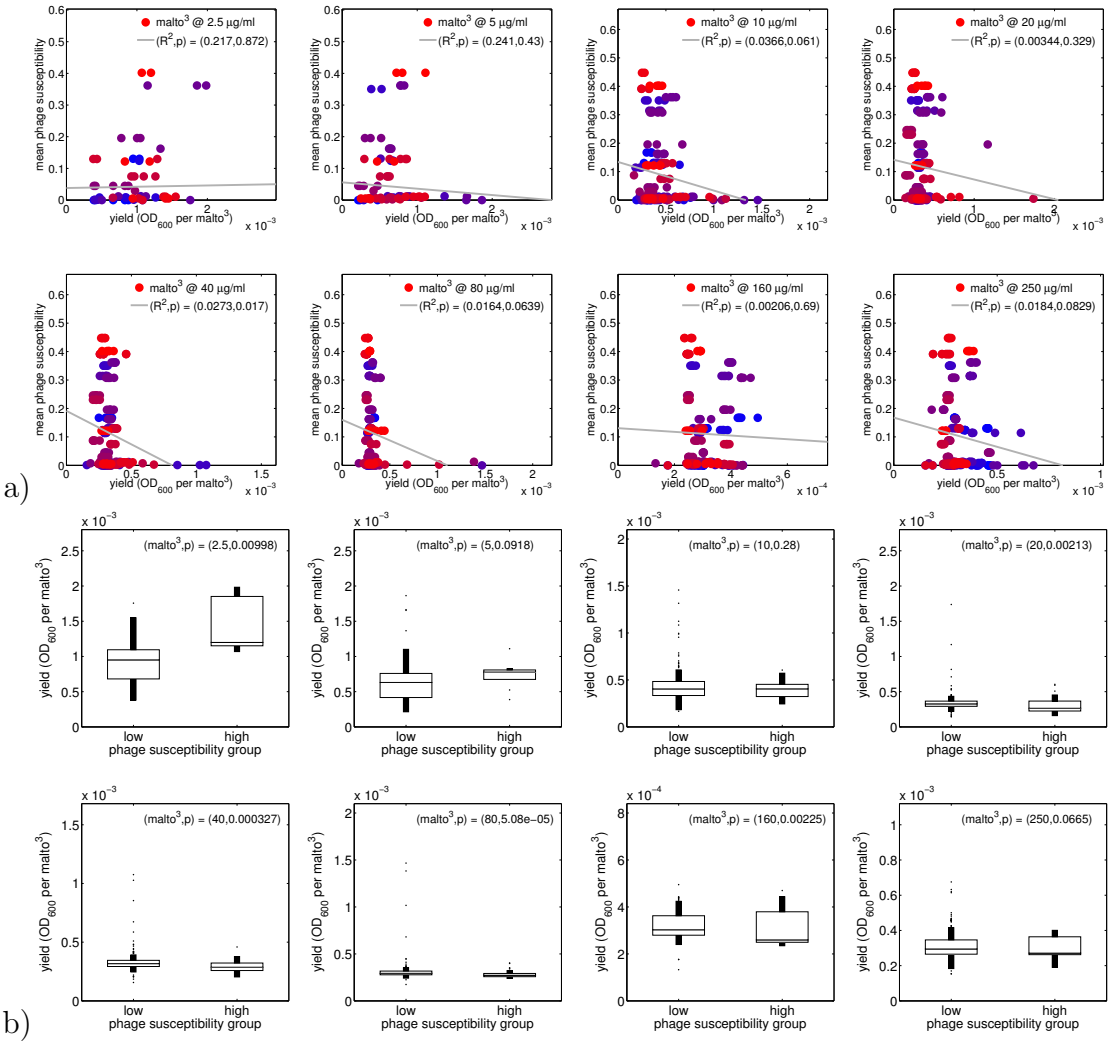


Fig. 5.19: **Two tests seeking a correlation between yield and phage resistance**. (a) For each maltotriose concentration yield is plotted versus the total infectivity phenotype for that bacterium and then a robust linear regression is performed. Although just one of the regressions is significantly non-constant, namely the regression at $40\mu g/ml$, even here any possible correlation is very weak. Note: different colours denote data for different bacterial strains. (b) A Kruskal–Wallis one-way analysis of variance performed on bacteria exhibiting two classes of phage resistance phenotype (either above or below half of the maximal susceptibility observed to all phage) shows variable evidence of differential yields between those two classes that could be interpreted as both a trade-up and a trade-off in different cases. In short, the data has variable characteristics between conditions.

form (the RYTO information is not included as part of the fits) to the growth rates of all the bacteria in the library as a function of maltotriose. The wild-type fit in Figure 5.21(a) is shown in blue. Now, pan-phage-resistant bacterial genotypes are shown in black and phage-susceptible genotypes are shown in grey and the question we have is whether any of these datafits lies above the wild-type datafit for all modelled sugar concentrations? Therefore, to examine this more closely, on the right of Figure 5.21(a) we have the same information as on the left, except that

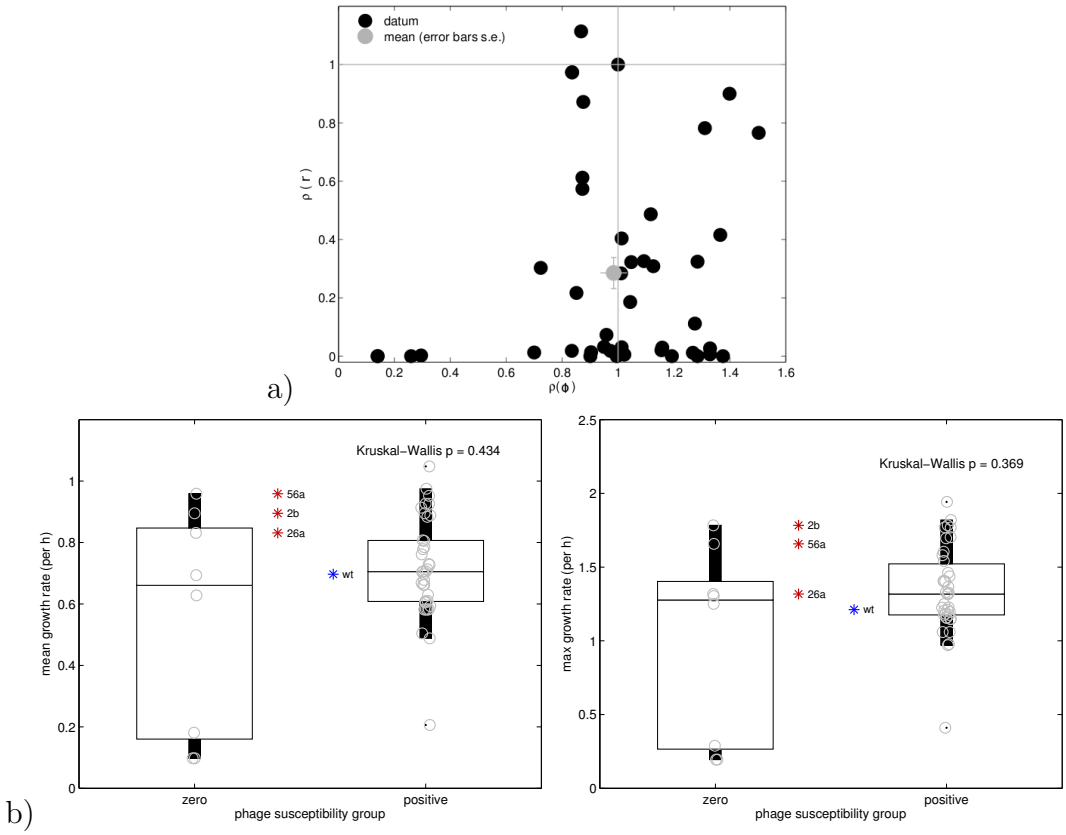


Fig. 5.20: **Further evidence that there are no costs of phage resistance**. In (a) we see a scatterplot of the entire dataset which shows the change in growth rate of each bacterial genotype from the wild-type ($\rho(\phi)$) strain versus the change in infectivity from the wild-type strain ($\rho(r)$). It shows that while the latter decreases substantially with respect to wild type, as indicated by the values less than unity, meaning that the mutants are more phage resistant on average than the wild-type, the former shows no evidence of decreasing. This is true both in the mean and in the sense that some strains are clearly both more resistant than the wild type strain and have greater growth rates. (b) This is pair of ANOVAs in which the mean growth rate (left) and the maximal observed growth rate (right) are analysed for two classes of bacteria: one that is pan-phage-resistant with respect to this library of and another which is susceptible to at least one phage. Neither are able to reject the hypothesis that pan-phage-resistant bacteria have lower growth rates, moreover three 'outliers' (56a, 2b, 26a) are shown. These are bacteria that are both pan-resistant and also exhibit growth rates that are greater than the median (and mean) of the phage-susceptible class of bacteria.

now only some of the genotype fits are plotted and the colour coding is different to the left plot. In red are datafits for genotypes that are pan-phage-resistant and highlighted in bold red are two of these which also possess the property of having faster growth than the wild-type at *all maltotriose concentrations tested*. In other words, we have identified two 'superstrains': 56a and 26a.

We now wish to address the question of how these two bacterial genotypes might have achieved the feat of resistant phage much better than the wild-type bacterium

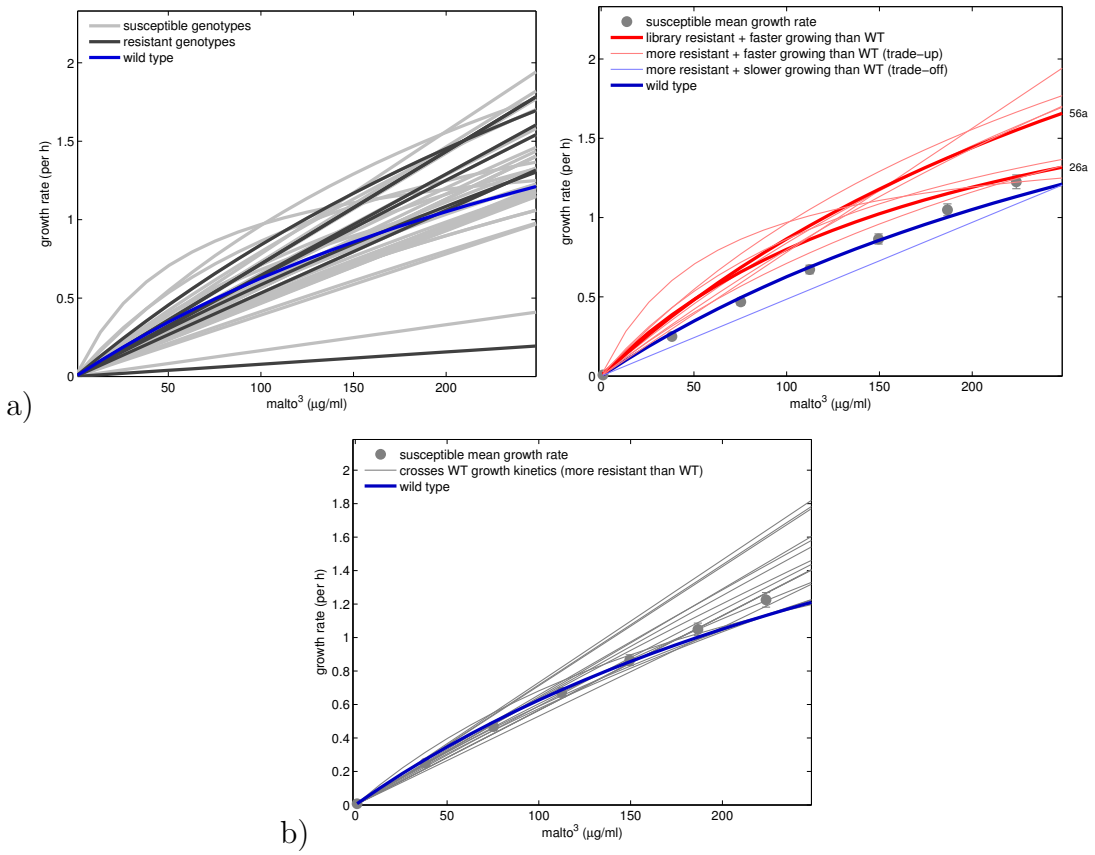


Fig. 5.21: **An initial indication as to why there are no costs of phage resistance**. One the left of plot (a) we see mathematical model fits (of a Monod form) to the growth rates of all the bacteria as a function of maltotriose supplied, the wild-type fit is shown in blue. Pan-phage-resistant bacterial genotypes are shown in black and phage-susceptible genotypes are shown in grey. On the right we have the same information except only some of the genotype fits are plotted and the colour coding is different. In red are datafits for genotypes that are pan-phage-resistant and highlighted in bold red are two of these which also possess the property of having faster growth than the wild-type at all maltotriose concentrations tested. For completeness, plot (b) shows that there are many bacterial genotypes which are fitter (in the sense of exhibiting faster than wild-type growth) at some maltotriose concentrations but which are less 'fit' at other maltotriose concentrations. Thus, these genotypes exhibit a mixture of costs of resistance and sometimes they do not have any costs at all.

and also in having much better metabolic phenotypes like growth rate across the entire spectrum of maltotriose concentrations.

Just for reasons of completeness we present Figure 5.22 which shows a range of data pertinent to demonstrating the existence of 'superstrains' 56a and 26a, moreover strain 2b displays some of the features but from the quality of the data associated with this strain we refrain from adding it to the list of monsters.

Each subplot of Figure 5.22 shows three strains of bacteria and the goodness of fit of the mathematical models used to determine the list of 'superstrains'. It compares these models to the analogous model fitted to the wild-type growth rate

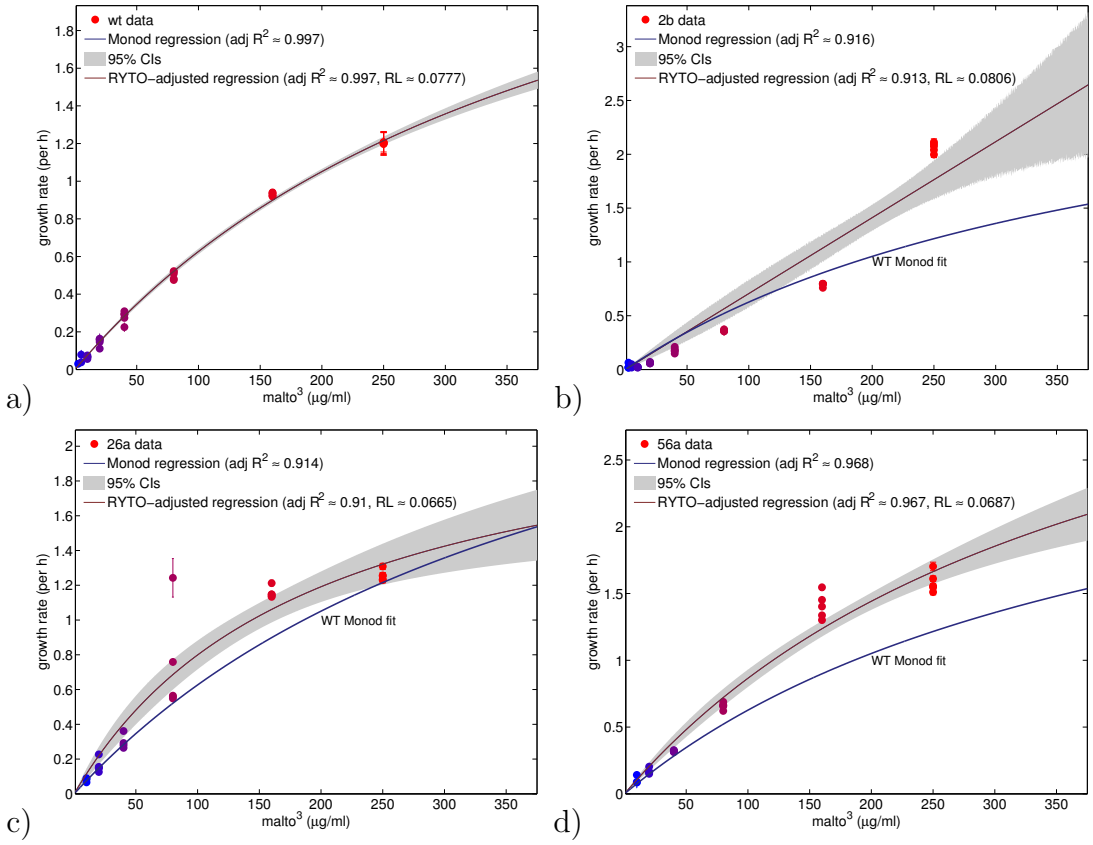


Fig. 5.22: Comparisons of growth rate model datafits (of a Monod form) to the analogous model fitted to the wild-type growth rate data, the fits for the 'superstrain' genotypes (26a and 56a) are above that of the wild-type.

data, showing that the fits for the 'superstrain' genotypes are significantly above that of the wild-type.

5.9 The protein homotrimer of interest: LamB

LamB is a protein encoded by the *mal* operon which is responsible both for the uptake of extracellular maltotriose, and other maltodextrins, but to which the phage λ is able to bind. For this reason, LamB is often known as the λ receptor and the entire porin is formed from a homotrimer of three LamB proteins, as shown in Figure 5.23.

After the co-evolution experiment had been performed, bacteria were isolated from mono-culture, and *lamB* was sequenced using Sanger sequencing. In this way, all unique mutations were detected, and each strain caryring a unique mutation was assigned a code. The 54 unique mutations comprise the library of mutants used in this study. Co-evolution and sequencing work was carried out by Justin Meyer. In order to compare two protein structures of two different LamB mutants we perform a so-called morphometric analysis allied to a range of different clustering techniques. These are applied to the α-carbon structure in each folded LamB monomer in an

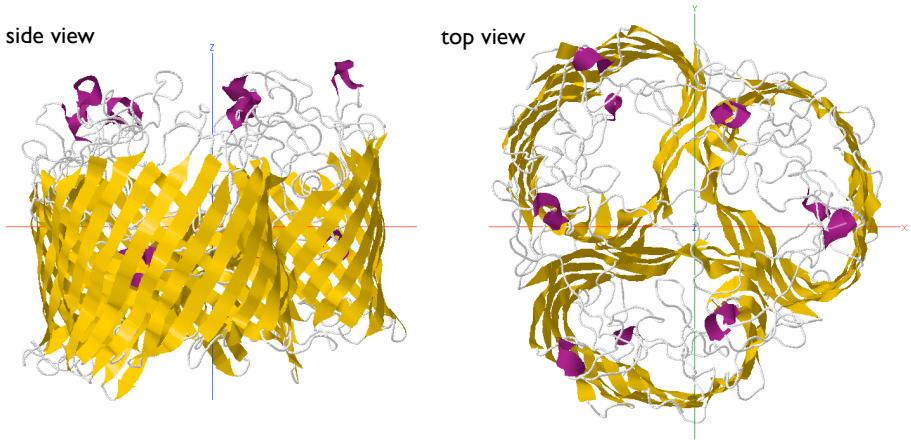


Fig. 5.23: Two views of the wild-type *E.coli* B(REL606) maltoporin formed from a trimer of the products of the *lamB* gene (produced using JMol).

attempt to determine where across LamB the greatest changes are observed, this is typically done with respect to the wild-type protein structure because we are interested in evolutionary changes in the protein's structure. In order to perform this morphometric analysis a series of algorithms was implemented by Prof. Beardmore, following the methodology set out in Appendix 5.1 (Methods and Material). Briefly, alpha-carbon data was extracted from a protein folding program (I-TASSER) and proteins were scaled to be of an equal size. Then a 3d point-matching alogorithm from [88] was implemented to determine the distance between matched points, for example between the ancestral protein type, and an evolved maltoporin containing SNPs or Indels.

5.9.1 'Superstrain' pleiotropy: spatially-localised protein changes with a dual function

Using the algorithms presented in the previous section we can begin to understand where on the LamB protein the geometrically largest changes have occurred because of the selection for phage resistance in the presence of maltotriose. This information is summarise in Figure 5.24 that is based on the determination of cluster centres and clusters using the mean shift algorithm (with a fixed bandwidth for all subfigures) applied to a kernel density estimate of all the distances from mutant proteins to the wild-type.

This procedure first calculated all pair-wise geometric distances using the $d^{\text{unmatched}}$ function defined in the above matching algorithm, call this matrix (D_{ij}) with the wild-type LamB protein situated at $i = 0$ for convenience. The largest α-carbon changes for those proteins that cluster at certain distances from the wild-type are then plotted on the wild-type version of the protein. Figure 5.24 shows that different protein subsets of the entire library have changes that cluster at particular locations on the wild-type LamB protein, not randomly across the protein. These

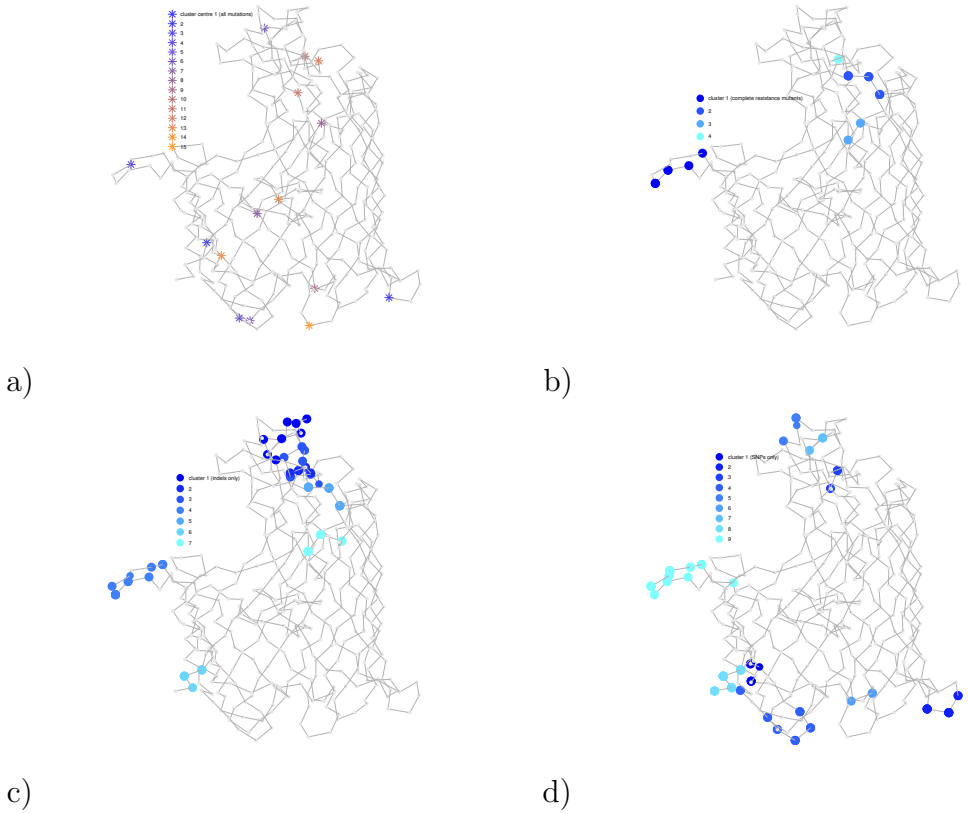


Fig. 5.24: LamB 'movement clusters' of different genotypes and phenotypes obtained using the mean shift algorithm. (a) Shows cluster centres associated with changes assessed across the entire library of LamB genotypes. (b) Shows clusters (not only the centres) determined by the mean shift algorithm associated with the pan-phage-resistant subset of the library. (c) Shows clusters determined by the mean shift algorithm associated with the insertion and deletion mutations subset of the library. (d) Shows clusters determined by the mean shift algorithm associated with SNP-only mutations in the library.

locations are shown for mutations that are either SNPs, indels or associated with pan-phage-resistant genotypes. Note that there are two particular regions on LamB that have the greatest geometric movement following the selection procedure that created the library, these are in 'outer loop' regions and also in the 'greasy slide'. The latter is a region of aromatic residues via which maltotriose gains entry into the porin and thereafter into the periplasm of the cell. The outer loops reside in the extracellular part of the protein that is responsible for the docking of the phage's J protein on its tail fibre.

If Figure 5.24 represents a summary of the totality of evolutionary changes across the LamB protein, analogous structural changes exhibited by the 'superstrain' genotypes will show how LamB can be altered by specific mutations that have a dual function of both providing phage resistance and increasing the uptake of maltotriose.

Thus, in Figure 5.25 we see how the LamB protein of strains 26a and 56a have moved to accommodate the mutations and the largest movements are highly lo-

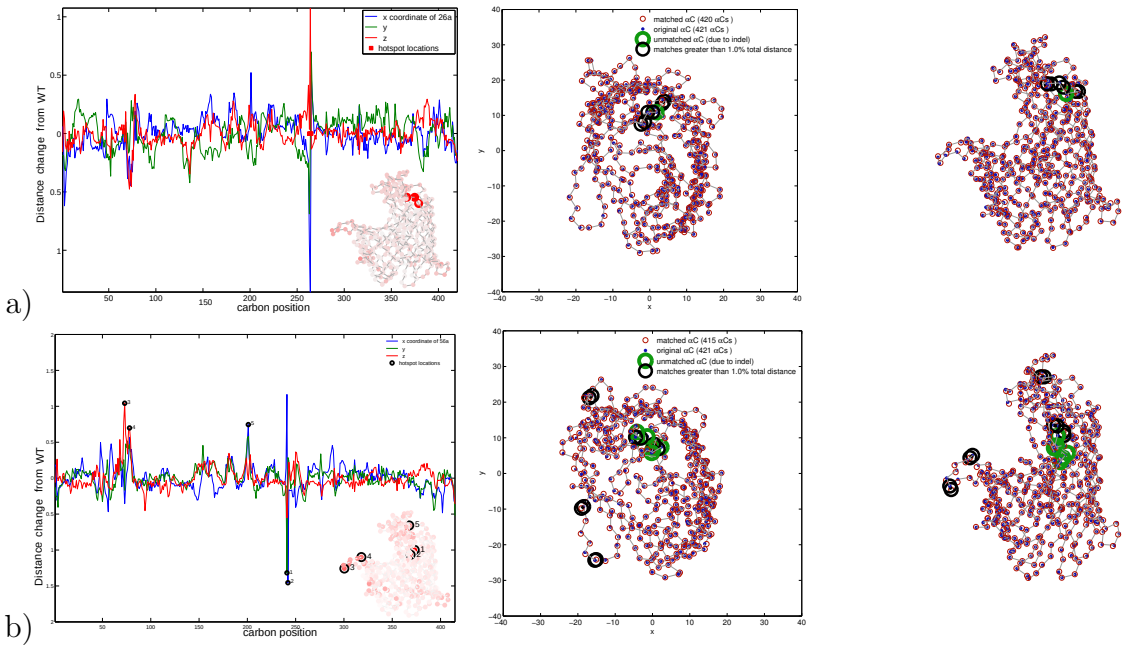


Fig. 5.25: Localised LamB 'movement clusters' of two the 'superstrain' genotypes. The left-most plot shows how each α-carbon changes in x, y and z-space and these changes are highly localised to one region of the protein which are induced in the centre and right plots: the data is relevant to strains (a) 26a and (b) 56a.

calised in the case of 26a, slightly less so in the case of 56a because, in the latter case, the regions that are deemed to have moved from the wild-type protein by the matching algorithm are described by multi-bump structures (see leftmost plot of Figure 5.25(a) and (b)).

5.9.2 The lock and key phase transition hypothesis: distance in protein space correlates with phage resistance

Phage-bacterial binding is often said to follow a lock and key property whereby the bacterial transporter in the outer membrane has a lock structure which is fitted by the key of the phage tail fibre protein [35]. With this in mind, we hypothesised that bacterial proteins that are further from the wild-type in protein space are also less susceptible to the wild-type phage. We tested this using the data shown in Figure 5.26. This shows how deviation from the wild-type protein in Euclidean distance measure correlates in a nonlinear manner with deviation from the wild-type susceptibility vector to each phage.

However, the data is quite sparse because of a cluster in the protein data at mid-range protein distances which, nevertheless, seems to be consistent with a potential for a 'phase transition'. Namely, there is a key distance from the wild-type protein of around 100 units above which it appears to become difficult for the wild-type phage to attach to the evolved LamB. This can be seen in Figure 5.26(a) which shows a sudden rise in Euclidean deviation from the infectivity (aka susceptibility) of the

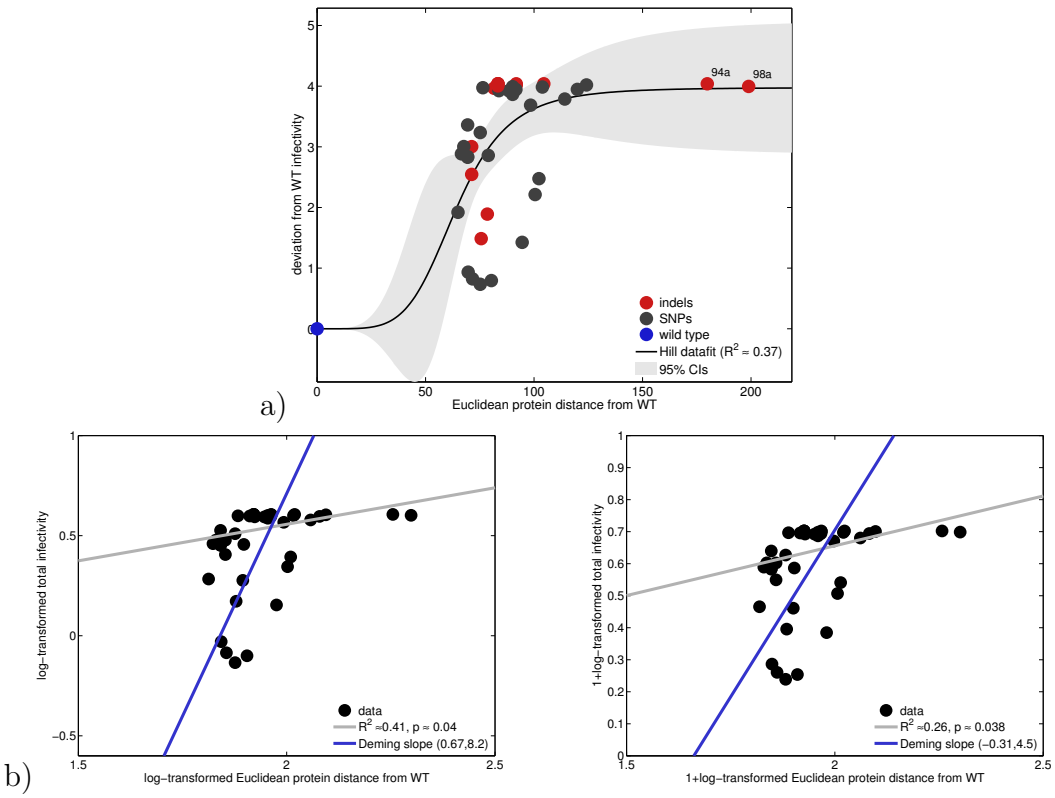


Fig. 5.26: (a) A sigmoidal model fit is shown on data that indicates a potential relationship between the distance (d^{matched}) from wild type LamB in protein space with the distance from wild type in the susceptibility to wild type phage phenotype. A greater distance from WT indicates reduced susceptibility to phage. (b) A log-log transform of the data in (a) is shown and two different linear fits are presented, one accounting for noise in the y direction, and one accounting for noise in both x and y directions (a Deming regression). Both are significantly non-constant, indicating a positive correlation in the data. The 95% confidence interval is indicated for the slope of the Deming regression (blue).

wild-type bacterial strain. 5.26(b) then uses two different measures of correlation to assess whether protein distance to wild-type LamB is correlated with this infectivity phenotype by taking logs of the data in two different senses and then applying two different linear regressions. The first (in grey) is a standard, robust linear regression whereas the one presented in blue is a Deming regression which accounts for potential errors in both the x and y data. Both of these indicate significant positive correlations between protein distance and infectivity distance to the wild-type.

We tested if this property remained when we added in not just the wild-type phage but all the evolved phage too. Indeed, the property remains in terms of the correlation between protein distance and susceptibility to phage; this can be seen in Figure 5.27 where the same features can be noted there as in Figure 5.26. The same is true of Figure Figure 5.28 which uses the unmatched α-carbon distance measure as part of the data analysis.

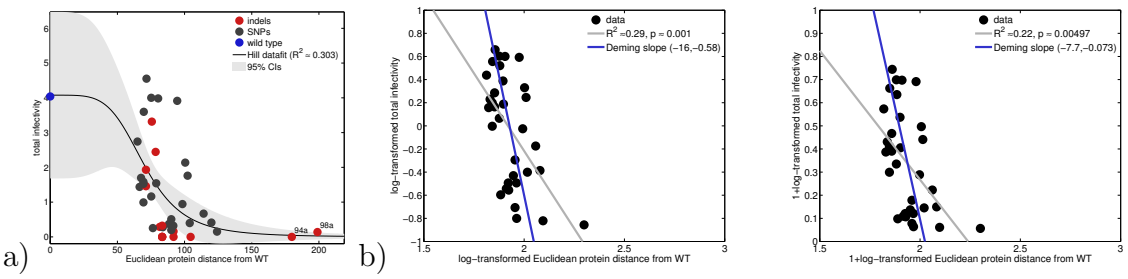


Fig. 5.27: (a) A sigmoidal model fit is shown on data that indicates a potential relationship between the distance (d^{matched}) from wild type LamB in protein space with the total susceptibility to all phage phenotype. (b) A log-log transform of the data in (a) is shown and two different linear fits are presented, one accounting for noise in the y direction, and one accounting for noise in both x and y directions (a Deming regression). Both are significantly non-constant, indicating a positive correlation in the data. The 95% confidence interval is indicated for the slope of the Deming regression (blue).

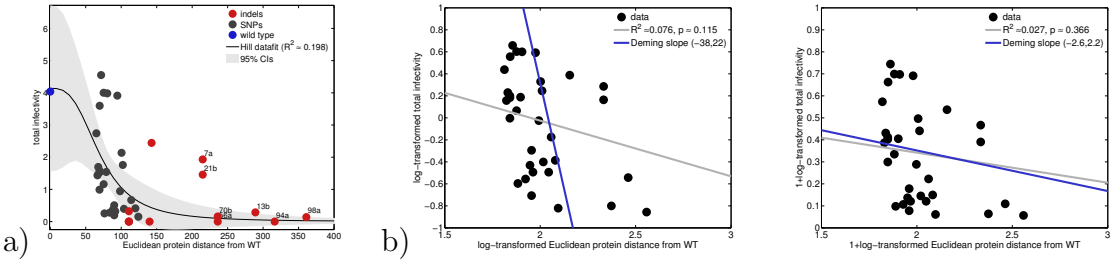


Fig. 5.28: (a) A sigmoidal model fit is shown on data that indicates a potential relationship between the total protein distance ($d^{\text{unmatched}}$) from wild type LamB in protein space with the total susceptibility to all phage phenotype. (b) A log-log transform of the data in (a) is shown and two different linear fits are presented, one accounting for noise in the y direction, and one accounting for noise in both x and y directions (a Deming regression). Both are significantly non-constant, indicating a positive correlation in the data. The 95% confidence interval is indicated for the slope of the Deming regression (blue).

Note that Figures 5.26, 5.27 and 5.28 use '$\log - \log$' and '$\log(1+\cdot) - \log(1+\cdot)$' transforms of the data. This is because the data contains zeros where the bacteria are resistant to all known phage and so a $\log - \log$ transform forces some datapoints to become Inf in Matlab which are then not plotted. Thus a similar transform was used that does not have this property to ensure this procedure did not introduce any artifactual results because of the loss of data.

From the above analysis that protein structure could correlate with resistance to phage properties, we then asked whether clustering strains according to proteins using the distance matrix (D_{ij}) defined above would also cluster like phage-susceptibility phenotypes. The result of this analysis is represented in Figure 5.29.

Figure 5.29 is based on the matched distance metric, $d^{matched}$, shows how like proteins cluster within the dendrogram with like resistance phenotypes that are

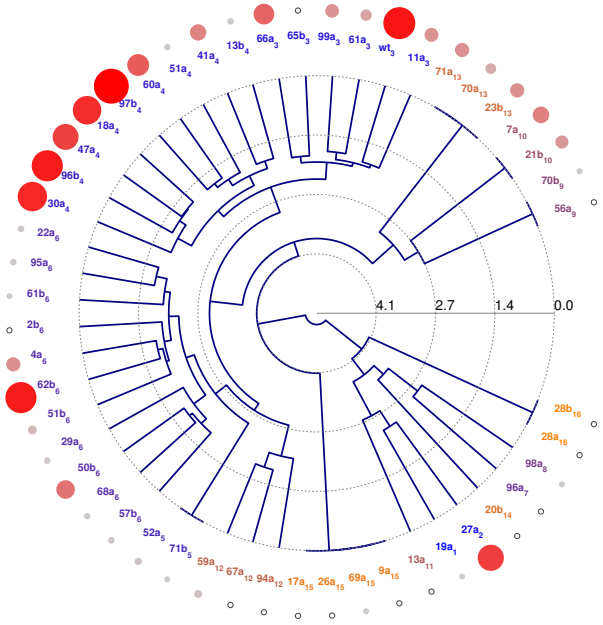


Fig. 5.29: A dendrogram created using the protein distance matrix (D_{ij}) to which a hierarchical spectral clustering algorithm has been applied. The analysis is based on a choice of fifteen clusters which are indicated as a numerical label on each bacterial genotype. The circular blobs are sized and colour coded according to the phage-susceptibility phenotype determined from the prior image analysis. Open white circles with black edges are pan-resistant bacteria.

marked with a coloured blob. There are some visible exceptions to this, so for example strain 27a has been clustered into a region containing a large number of phage that are highly resistant, despite 27a being highly phage susceptible. Similarly, 65b has been placed into a region containing many phage-susceptible strains despite it being completely phage resistant (with respect to the genotypes in the library).

As the clustering process using $d^{matched}$ contained these anomalies, we then also used $d^{unmatched}$ and followed the same hierarchical cluttering procedure to see if it would provide a better discrimination of phage resistance properties based on LamB structures. So, in a similar vein we present Figure 5.29 that is based on the unmatched distance metric, $d^{unmatched}$. It also shows how like proteins cluster within the dendrogram with like resistance phenotypes that are marked with a coloured blob. Again, there are visible exceptions to this, so strains 27a and 47a have now been clustered into region containing large numbers of phage that are highly resistant, despite them being highly phage susceptible. Similarly, 56a has been placed into a region containing many phage-susceptible strains despite it being phage resistant.

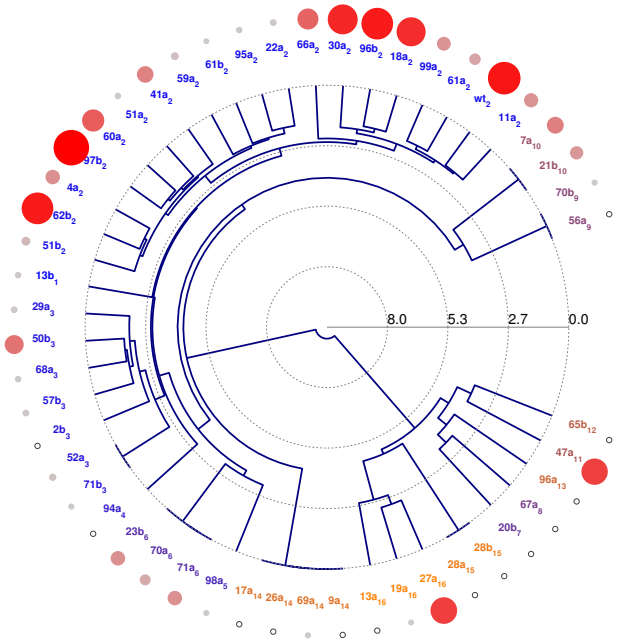


Fig. 5.30: A dendrogram created using an 'unmatched' protein-to-protein distance matrix to which a hierarchical spectral clustering algorithm has been applied. The analysis is based on a choice of fifteen clusters which are indicated as a numerical label on each bacterial genotype. The circular blobs are sized and colour coded according to the phage-susceptibility phenotype determined from the prior image analysis. Open white circles with black edges are pan-resistant bacteria.

5.9.3 Summary

We utilise a library of REL606 *E.coli* that have been co-evolved with lambda phage to test the effect on co-evolution on fitness, and whether predictions can be made as to the resistance of bacteria to phage by the change in the maltoporin protein. We also demonstrate that there are trade-offs and trade-ups present within the library, with some superstrains showing an increase in resistance, growth rate and yield. Figure 5.28 shows that as euclidean protein distance increases from the wild-type, the susceptibility to phage decreases. However, this does not occur in a linear fashion, but rather demonstrates a switch-like response, that a certain minimum change in protein distance is required to confer resistance, and this change can happen in a narrow range of protein distance. We therefore conclude that while LamB is a potential predictor of phage resistance, it is not a perfect predictor. This is not surprising, there are many ways a bacterium can resistant exploitation by phage that involve the targeting of phage DNA once it has entered the cell as part of the CRISPR-Cas mediated DNA editing system that provides bacteria with an adaptive immune response. Thus it is possible to have a mutant phage with a wild-type LamB that could, in principle, be partially phage resistant.

Interestingly we again see a trade-off between lag phase length and a growth parameter, in this case, growth rate. Because the length of the lag phase is not

frequently quantified in studies searching for trade-offs, there is little discussion on this effect in literature, and yet it is an important growth parameter to assess in order to understand the growth kinetics of the bacterial population. We also show that within each strain there is a consistent RYTO, i.e., as sugar concentration is increased, rate increases and yield decreases. However, when looking at trade-offs between strains, we observe a trade-up between rate and yield; that is to say that at any given maltotriose concentration, a high yield will be correlated with a high rate.

Chapter 6

Selection for efflux: Non-Monotone Selection for Resistance in a Spatial Antibiotic Gradient

6.1 Overview

By implementing an analogy of the laboratory protocol used to determine antibiotic susceptibility, the so-called E-test, we are able to quantify aspects of selection for antibiotic resistance in a spatial drug gradient. The principle findings of this chapter are:

1. Diffusion of antibiotic in this experiment is described by a power law.
2. One might expect greater drug dose to correlate with lower population densities, we show that this expectation is not met in practice.
3. Reasons for this are both ecological and evolutionary.
4. Using a combination of theoretical modelling and spatially-extended laboratory microbial growth experiments, we show that the growth rate of a bacterial population with genetically distinct subpopulations can be maximised at intermediate distances from the antibiotic drug.
5. Each subpopulation is distinguishable by the number of resistance genes they carry and can have a maximal grown rate at a different spatial locations.

   One prediction from this observation is that microbes growing in an antibiotic gradient can exhibit, at least temporarily, a certain 'bullseye' pattern formed from concentric rings. Experiments using *Escherichia coli* provide evidence for the existence of these rings and we show, using genetic manipulation of the resistance genes in question, that such a genetic manipulation can move the location of this pattern.

   This chapter represented the work of a within-lab collaboration between the following people who each contributed to different aspects of the project. Their contributions are detailed as follows.

   R. C. Reding Roman: implemented experimental protocols, analysed data, designed experiments and mathematical models;

   M. Hewlett: designed and implemented experimental protocols, analysed data, aided development of custom imaging system;

   F. Gori: implemented computer codes to simulate mathematical models;

   S. Gardner: developed a multi-fluorescence video capture device based on an Arduino controller and a commercial camera;

   I. Gudelj: designed experiments, provided funding for the project from a NERC grant;

   R. Beardmore: provided funding from an EPSRC grant, wrote and analysed mathematical and computer models.

6.2 Introduction: selection for resistance at different dosages

The purpose of this study is to test the following hypothesis: is it true that a greater antibiotic dose necessarily yields fewer bacterial cells? The Eagle effect of decreased penicillin effect at high doses, observed many years ago [39], provides one of the earliest answers to this question, and that answer is negative for it can be the case that bacterial densities increase with increasing antibiotic dose. However, how Darwinian selection for resistance might also produce datasets like those reported by Eagle, and others, have received little attention in the literature. We therefore study this question from an evolutionary and ecological perspective, focussing on how resistance by efflux can contribute to datasets that also lead to the rejection of the above hypothesis. In short, we will conclude that a phenomenon known as *competitive release* can combine with selection on gene duplication mutations that mediate drug resistance by efflux to produce spatiogenomic patterns in which more drug need not correlate with fewer bacteria.

One of the most common assays performed in hospitals is the antibiogram [106, 91], it is an *in vitro* test for susceptibility to a range of antibiotics. At the core of antibiotic susceptibility testing is a simple laboratory test that determine zones of inhibition, these are often performed using manufacturers' so-called E-strips [117]. The E-strip is impregnated with a gradient of antibiotics and it is placed onto a bacterial lawn whereupon killing and growth inhibition of some of the bacteria can result. From this killing, a clearance appears on the lawn up to a certain drug concentration from which one can read concentrations at which the microbe is sensitive to the drug used.

As this protocol by its very design creates a spatial gradient of antibiotic, we are interested in ecological and evolutionary features that might result from the implementation of such a test. The bespoke implementation in our laboratory of an analogue of the test is illustrated in Figure 6.1.

This drug susceptibility protocol is performed on an agar plate containing minimal growth medium (M9), a carbon source (glucose was used in all cases, at a concentration of 0.2%), salts and amino acids. At the centre of the plate a circular region of agar is excised by hand and replaced with agar that has been impregnated with an antibiotic at a controlled dose, an antibiotic gradient therefore ensues across the plate. Soft agar is used, as this allows bacteria to be inoculated throughout the agar, not just on the surface. This was done so that a homogeneous inocula could be achieved throughout the whole plate. Plates were incubated at 30°C in static incubator, or in our custom image system.

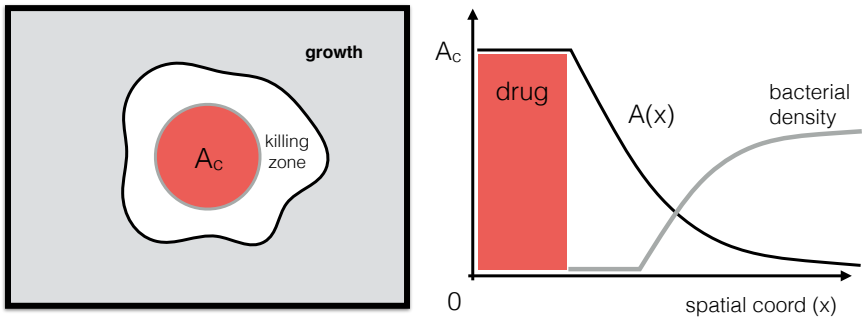


Fig. 6.1: A schematic of our implementation of the antibiotic susceptibility test: an antibiotic drug held at high dosage at the centre of an agar plate diffuses out into a bacterial lawn, producing a killing zone or zone of inhibition where by the higher the dose, the larger the zone.

Photographs were taken every 12 hours for 3 days, or every 20 minutes in cases where a time-lapse video was taken.

It is thought that antibiotic gradients are associated with rapid selection for resistance, the idea being that a microbial population is exposed to small increments of drug which can be easily dealt with. A little like the tale in which a frog that can be boiled slowly by steadily increasing the temperature of the water of the pan in which it sits, so a microbe is thought to readily survive small increments in antibiotic dosage when an abrupt change would be lethal [114].

We therefore use a laboratory model system in which changes in antibiotic resistance can be measured over short timescales. This system sees the bacterium *Escherichia coli* K12, strains TB108, MG1655 and AG100, treated with tetracycline and macrolide antibiotics, we use this combination because K12 has a clinically important efflux pump [18] formed from the products of the operon *acr* and the protein TolC [74]. Moreover, *acr* resides in a genomic region that can be recombined to produce duplications and triplications very rapidly under antibiotic selection pressure involving macrolides and tetracyclines [70].

6.2.1 Increases in the zone of inhibition of MG1655 and AG100 with dose are consistent with linear diffusion theory

The fundamental solution of the linear diffusion equation

$$\mathsf{A}_t = \sigma \mathsf{A}_{xx},$$

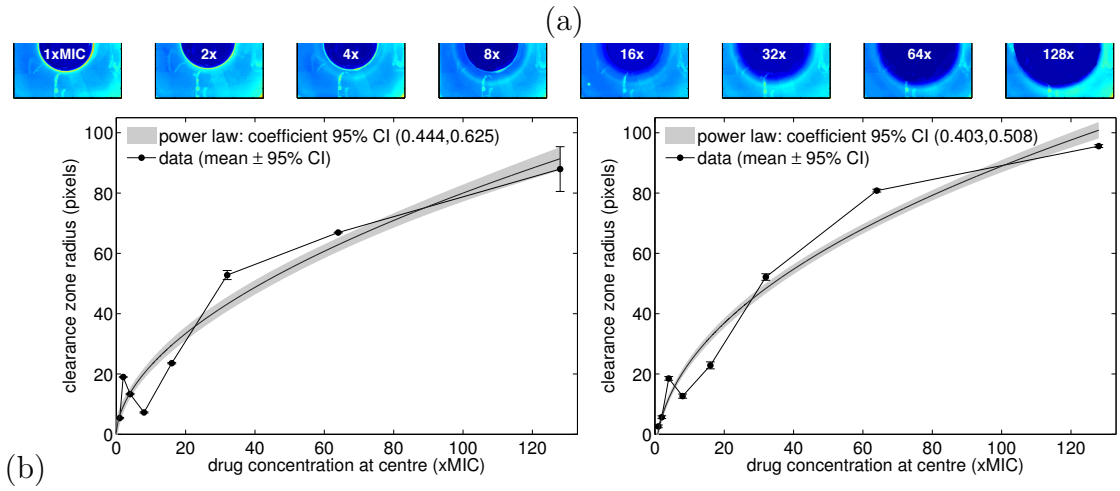


Fig. 6.2: (a) Raw data: an image of a bacterial lawn (in false colour) showing how the zone of inhibition on that lawn increases in area with increasing dose, here at 1,2,4,....,128 times the MIC dose determined in planktonic culture conditions. (b) The increase in area for the tetracycline drug for strains MG1655 and AG100 follow a power law with coefficient close to a value of two: a quadratic. This is consistent with increases in zone of inhibition being described by a threshold killing model whereby escape of the drug fro the centre following a diffusion equation, as shown in the text (Statistical note: correlation coefficients are $R^2 = \{0.898, 0.959\}$ respectively, F-statistics versus constant models are $F = \{728, 1930\}$ and corresponding p-values are given by: $p = \{1.43 \cdot 10^{-82}, 3.45 \cdot 10^{-115}\}$.)

on an infinite two-dimensional domain with Dirac delta initial condition of mass A_c is given by the expression

$$\mathsf{A}(t) = \frac{\mathsf{A}_c}{4\pi\sigma t} \exp\left(-(x^2+y^2)/4t\sigma\right)$$

which represents a temporally-modulated normal distribution, in which σ represents the diffusion rate. . If we assume, as is natural, that extracellular and intracellular drug concentrations are highly correlated and, also, that the concentration of A leads to bacterial killing provided it is of a sufficiently high value, A_d say, a value known as the minimal bactericidal concentration (MBC), then the zone of inhibition, or killing, is given by those pairs (x, y) for which

$$\mathsf{A}_d < \max_{t>0} \frac{\mathsf{A}_c}{4\pi\sigma t} \cdot \exp\left(-(x^2+y^2)/4t\sigma\right).$$

Let us write $r = \sqrt{x^2+y^2}$. By taking the derivative with respect to t of the function $\mathsf{A}(t)$ defined within the latter inequality and setting this derivative to zero, we find this maximum occurs when $t = r^2/4\sigma$. At this value of t there results $\mathsf{A}(t) = \mathsf{A}_c/(\pi r^2 e)$. Hence, equating $\mathsf{A}_c/(\pi r^2 e)$ with A_d, we deduce that killing occurs up to a certain spatial radius, $r < r_d$ say, where

$$r_d = \left(\mathsf{A}_c/(\mathsf{A}_d \pi e)\right)^{1/2}.$$

If we therefore examine plot of empirically-determined zones of inhibition, like those in Figure 6.2, we ought to observe that the increase in radius of that zone with the dosage applied scales like the square root of the drug dose. From a nonlinear regression $p_1 + p_2 x^{p_3}$, we observe that the value $p_3 = 1/2$ lies within the 95% confidence interval for the estimate of this power coefficient. We therefore cannot reject the hypothesis that some value of the power coefficient provides a better fit than the theoretically-predicted square root law. We shall modify this simple diffusion-based theory later in order to provide more refined predictions of the nature of population dynamics in the growth region outside these zones of inhibition.

6.2.2 The AcrAB-tolC efflux system, the *acr* operon and its duplications

Outside the zones of inhibition where drug dosage is so high that bacterial killing is assured, we anticipate a complex set of population dynamics between different tetracycline and macrolide resistance mechanisms available to the *E.coli* cell. In particular, the *E.coli* chromosome carries genes that encode proteins AcrA, AcrB and TolC which form a membrane-spanning efflux pump that includes tetracyclines and macrolides in the family of small molecule substrates that it can pump. Moreover, structure-altering mutations in AcrB are known to mediate clinical resistance to a range of antibiotics [18].

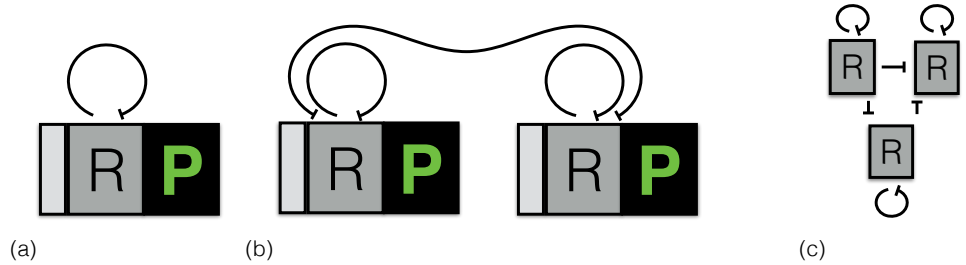


Fig. 6.3: (a) The basic structure of an operon: a promotor region in light grey, followed by a gene coding a protein that represses transcription of the same operon, followed by a second protein that encodes part of an efflux pump. We use the green font colour to highlight the fact that we have a strain which has GFP fused to the protein P, we also have strains without GFP fused to P. We have in mind that P represents the A and B proteins of the *acrRAB* operon. (b) *E.coli* can duplicate the number of copies of the *acr* operon in its genome which leads to a novel network structure following duplication in which the two copies repress each other. (c) Following a further duplication of one of the operons, a three-node network results with all nodes repressing each other.

The efflux proteins AcrA and AcrB are encoded within an operon *acr*, that we shall also write *acrRAB* to highlight the fact that the operon contains a triple

of proteins, two of which form the pump but the first of which to be transcribed encodes a repressor of the same operon. Now, *acr* is contained within a genomic region in the *E.coli* chromosome that can be recombined into the chromosome to form a large mutation with a high mutation rate consisting of about 8% of the entire bacterial genome. Under strong selection for resistance as encoded by this pump, this region has been observed to be duplicated in doxycycline monotherapy and triplicated in a doxycycline-erythromycin cocktail treatment [70] within five days of exposure to these antibiotic conditions at low population densities.

When these operons are duplicated, triplicated, or held at an even greater copy number, this creates a growing network of ever stronger co-repression, as depicted in Figure 6.3. Moreover, the *acr* operon is regulated by a complex stress network that includes the *mar* regulon and regulatory mutations in this operon, through *marR* can mediate expression of the AcrAB-TolC efflux pump [70].

In order to illustrate how this pump is regulated during different phases of bacterial growth and in different antibiotic concentrations, we present the data in Figure 6.4. Using an altered MG1655 strain in which GFP is physically fused to AcrB, a strain denoted TB108. The latter strain, which allows us to infer data on efflux pump numbers per cell, shows that the pump decreases in concentration in lag and stationary phase, during the period up to 12h in Figure 6.4(a). In the absence of drug as stationary phase is entered, the pump is up-regulated to a value deemed to be unity for this discussion. A region of stasis is reached from 18h onwards in which the pump concentration neither declines nor increases. When the drug erythromycin is used in the growth medium, a similar dynamic of regulation of the pump is observed except that the final concentration per cell is much higher and it continues to increase in stationary phase, to a value about 40% higher than that observed in the absence of drug. Now, Figure 6.4(a) and (b) both show that an increasing down-regulation of the *acrB* protein is observed as more drug is used in the growth medium. Thus, while it might be hypothesised that more drug necessarily leads to an increase in pumps expressed per cell, this statement is only true for certain dosages. The opposite is also observed in a single season of bacterial growth.

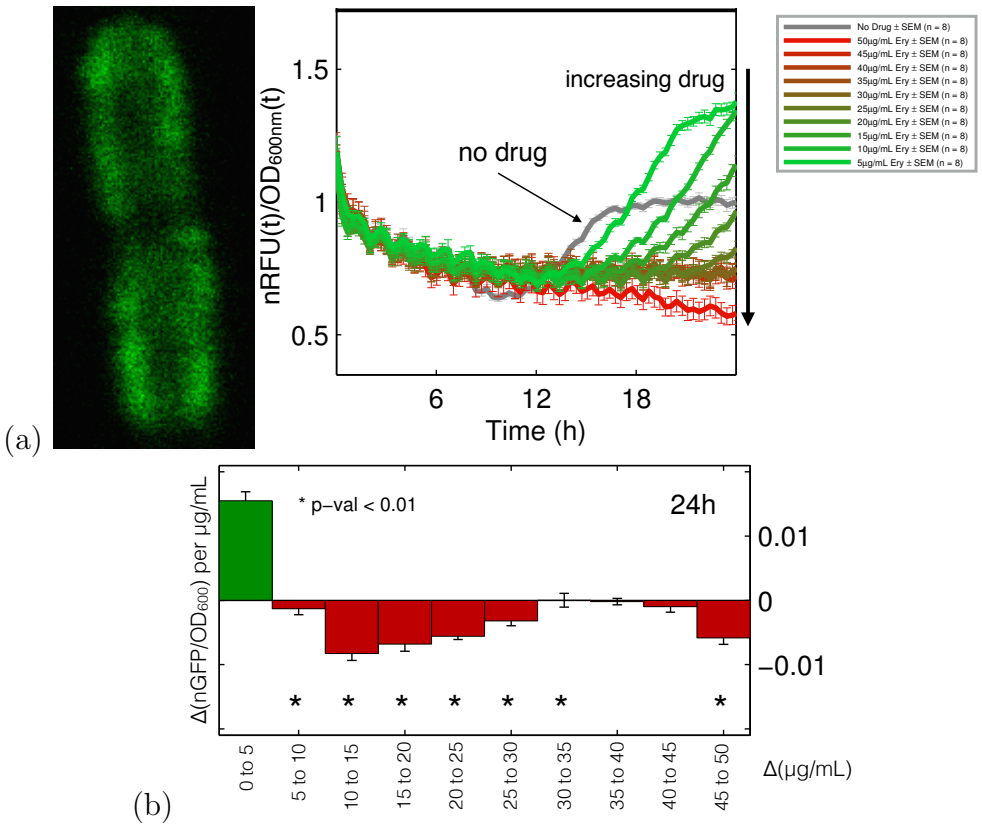


Fig. 6.4: (a) A fluorescence microscope image showing the localisation of GFP within a section of an *E. coli* cell close to the moment of division. Using this labelled strain of bacterium we are able to deduce how the dynamics of regulation of the protein AcrB correlates with the use of the macrolide, ERY. In the absence of the drug (grey curve) the protein is down-regulated through lag and exponential phase (less than 12h) before being up-regulated and then stabilising in stationary phase (12h and beyond). When drug dosage is applied, first at a low dosage of $5\mu g/ml$, the concentration of AcrB per cell increases substantially to level about 40% higher at its maximum in the absence of drug. However, as the dose is further increased we observed a negative correlation between drug and AcrB concentration per cell. (b) The negative correlation so-described is significant across a wide range of ERY dosages, as shown by the results of determining a t-statistic (for the derivative of GFP per cell with respect to dose) following a linear regression that is testing for the increase or decrease in AcrB concentration as dose changes.

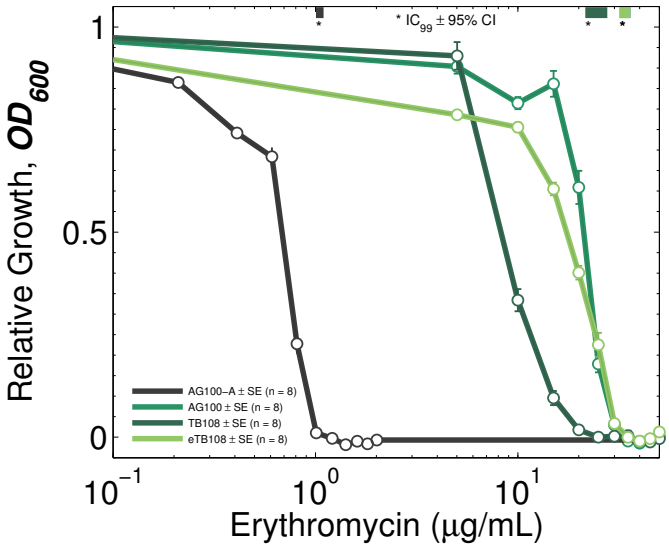


Fig. 6.5: A (relative) dose-response of three strains of *E.coli* with respect to the drug erythromycin using units of optical density measured at 600nm. The *acr* knockout strain AG100-A is most sensitive to the drug, followed by TB108 which has a GFP physically fused to AcrB, followed by the AG100 and eTB108 strains, which have very similar dose-response profiles. TB108 is a derivative of *E.coli* K-12 MG1655, with GFP fused to the acrB protein.

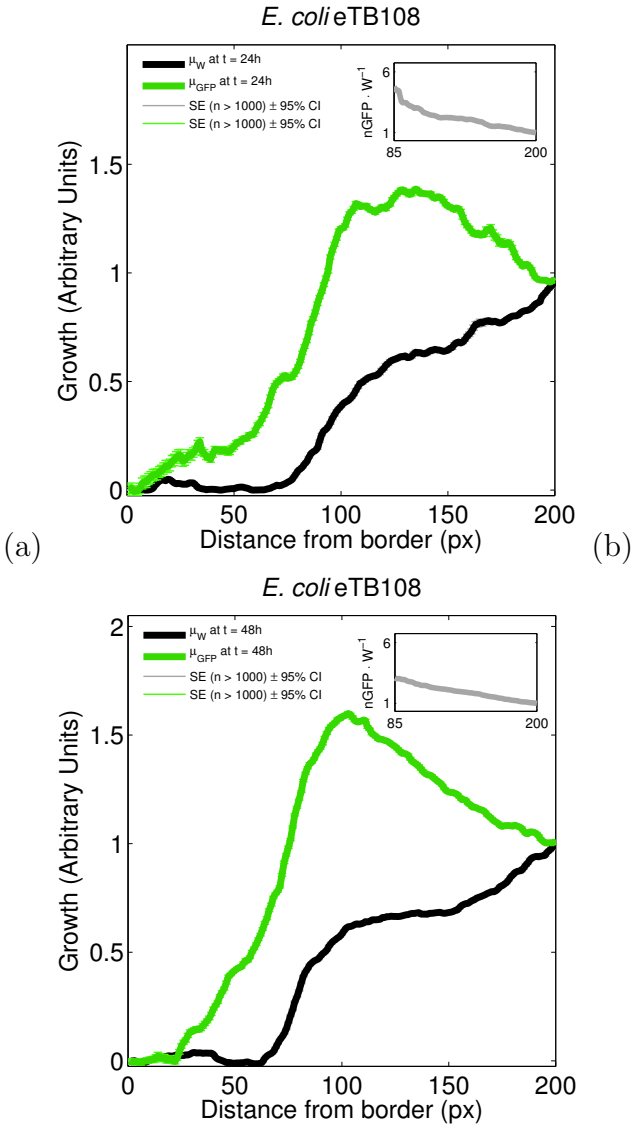


Fig. 6.6: A spatial dose-response assay using strain eTB108 showing the population-wide expression of AcrB using the proxy of green fluorescence levels (shown next to optical density (i.e. white light)) at 24h (left) and 48h (right). eTB108 refers to the strain TB108 which has been evolved for 8 days in Erythromycin, and has a higher degree of resistance. This yields a per cell measure of AcrB concentration which is done by calculating fluorescence observed per optical density. This measure is shown in the inset of both plots which indicate a linear relationship between the spatial coordinate and AcrB per cell, at least for this strain of bacterium, which indicates a positive correlation between drug dose and pumps per cell. Moreover, the levels of AcrB expression are higher at 24h than at 48h during conditions when extracellular resources, like glucose, are at lower concentrations.

6.2.3 A duplication of the *acr* operon does not double protein *AcrB* concentration

Consider the following differential equation model of the *acr* operon shown in Figure 6.3. The auto-repressive nature of structure, whereby the repressor R is transcribed, followed by the efflux protein, P, and then R represses the further transcription of the operon leads to one possible model of the following form:

$$\frac{d}{dt}\mathsf{P} = -d_0\mathsf{P} + \rho\mathsf{M}, \tag{6.1a}$$

$$\frac{d}{dt}\mathsf{M} = -d_1\mathsf{M} + g\frac{V}{1+k\mathsf{P}} + M_0. \tag{6.1b}$$

Here d_0 and d_1 are degradation rates of the protein, P, and of mRNA associated with the transcription of the operon. We assume that the concentration of the repressor protein, R is proportional to that of P, which is reasonable given they are encoded by the same operon. The parameter M_0 is a basis level of transition of the operon, $gV + M_0$ is the maximal rate of transcription, where the parameter g is a proxy for the number of copies of the operon held in the genome.

In steady-state, equation (6.1a-b) satisfies the following. First, $\mathsf{P} = \rho\mathsf{M}/d_0$, and therefore $M_0 + gV/(1 + k\mathsf{P}) = d_1 m = d_1 d_0 \mathsf{P}/\rho$. This is a quadratic in P which, when solved, gives

$$\mathsf{P} = \mathsf{P}^*(g) := \alpha + \sqrt{1 + \beta \cdot g},$$

where $\alpha = (k^{-1} - \rho M_0/(d_0 d_1))/(k^{-1} + \rho M_0/(d_0 d_1))$ and $\beta = \frac{4V\rho}{d_0 d_1 k}(k^{-1} + \rho M_0/(d_0 d_1))^{-2}$.

Let us now assume that $\mathsf{A}(t)$ is the internal concentration of antibiotic in the cell and that A_0 is extracellular concentration of the drug, therefore,

$$\frac{d}{dt}\mathsf{A} = \varphi(A_0 - \mathsf{A}) - \rho\mathsf{A}\cdot\mathsf{P}. \tag{6.1c}$$

where ϕ is the diffusion rate of A in or out of the cell, and ρ is the association rate between A and P.
Then, in steady-state,

$$\mathsf{A} = \mathsf{A}^*(g) := \frac{\varphi A_0}{\varphi + \rho\mathsf{P}} = \frac{A_0}{1 + \frac{\varphi}{\rho}(\alpha + \sqrt{1+\beta\cdot g})}.$$

We therefore deduce, because of the auto-repressive nature of the efflux operon, that internal drug concentration and concentration of the efflux protein itself, scale as $\sqrt{g}$, the square root of the number of copies of the operon in the

genome.

At this point we make the following remark. In the above, the value of M_0 is assumed to be small and should in fact be $g \cdot M_0$ in this analysis. This reflects the idea that if a gene is copied in the chromosome, then the basal transcription rate could well, indeed should, increase approximately linearly with the number of copies of the transcribed gene. This modification has the effect of changing the terms α and β above so that they also depend on g. However, if M_0 is small enough, then the above predicted square root dependence of $\mathsf{A}^*(g)$ on g will be largely unaffected, although there clearly do exist parameter regions where that dependence will be lost. In short, the replacement $M_0 \to g \cdot M_0$ in the above manipulations makes clear in which parameter regimes that can happen. Throughout the remainder, for simplicity, we assume that α and β do not depend on g.

6.3 Theoretical predictions: the bullseye pattern

6.3.1 The bullseye pattern: a heuristic derivation

Assume for now that environmental, meaning extracellular, antibiotics degrade sufficiently during the time of an experimental protocol in order for the diffusion equation with decay to be a reasonable description of the dynamics of the antibiotic molecules through space, ignoring uptake and efflux for the moment. Doxycylcline degrades to methacycline and 6-epidoxycycline when subjected to elevated temperatures and UV exposure. At 40°C around 2% of Dox has degraded to these products. Our results agree with these published values, with Dox stored under our experimental conditions retaining over 95% efficacy after 2 days. Thus we suppose that the parameter that was denoted A_0 above now depends on space and time in such a way that

$$\mathsf{A}_t = \sigma \mathsf{A}_{xx} - d\mathsf{A},$$

where $\mathsf{A}(x = 0, t) = A_c$ and $\mathsf{A}(x \to \infty, t) = 0$ hold for all $t > 0$. In steady-state, this equation has solution

$$\mathsf{A} = \mathsf{A}(x) = A_c \cdot e^{-(\sigma/d)^{1/2} x}.$$

Consistent with the experimental protocol depicted in Figure 6.1, if we assume the existence of a diffusing extracellular nutrient, call it S, then we may also introduce two further diffusion and decay parameters, this time for the carbon

source glucose, and write $\mathsf{S} = \mathsf{S}(x) = S_c \cdot e^{-(\Sigma/D)^{1/2}x}$. In this case it should be clarified that the decay term for glucose actually refers to sugar lost to the environment due to consumption by bacteria, not physical degradation. We use the latter form as a coarse approximation of the true nature of the nutrient concentration, S, because no cells can be found within the red, high-drug region in the centre of Figure 6.1. However, cells are found inside the grey region and they will begin to ingest those nutrients and grow, this will create a nutrient gradient whereby more nutrients can be found inside the red region than outside it. Thus, we assume for now that the nutrient has a similar form to the drug profile. As this is a heuristic derivation, to makes matters simpler still we will assume

$$\mathsf{S} = \mathsf{S}(x) \approx S_c \cdot e^{-(\sigma/d)^{1/2}x}.$$

Now, suppose that bacteria grow through time according to Michaelis-Menten-Monod kinetics in the sense that their growth rate, G, can be written

$$G = G(\mathsf{S}, \mathsf{A}) = c \cdot \frac{v\mathsf{S}}{\kappa + \mathsf{S}} \cdot f(\mathsf{A}),$$

where $f(\mathsf{A})$ is an A-dependent inhibition function which slows growth as a function of intracellular drug concentration, A. Thus, following the dependencies in this simple model, G is a function of x. We are interested in where the maxima of this function, $G(x)$, might occur and so we now compute $\frac{\partial}{\partial x}G(\mathsf{S}, \mathsf{A})$. So, we first note that

$$\frac{\partial}{\partial x}G(\mathsf{S}, \mathsf{A}) = f'(\mathsf{A})\mathsf{A}'\frac{cv\mathsf{S}}{\kappa + \mathsf{S}} + f(\mathsf{A})\frac{cv\mathsf{S}'}{\kappa + \mathsf{S}} - f(\mathsf{A})\mathsf{S}'\frac{cv\mathsf{S}}{(\kappa + \mathsf{S})^2}$$

and as we are seeking maxima with respect to x of this quantity, we now determine whether, or not, the latter expression has zeros. This occurs if and only if

$$0 \;=\; f'(\mathsf{A})\mathsf{A}'\mathsf{S} + f(\mathsf{A})\mathsf{S}' - f(\mathsf{A})\mathsf{S}'\frac{\mathsf{S}}{\kappa + \mathsf{S}}. \tag{6.2}$$

Now, collating the information we have so far and using the fact that A represents intracellular drug concentration, by relating this to extracellular drug concentration, we have $A_0 = A_c \cdot e^{-(\sigma/d)^{1/2}x}$ and, therefore,

$$\mathsf{A} = \frac{A_c \cdot e^{-(\sigma/d)^{1/2}x}}{1 + \frac{\varphi}{\rho}(\alpha + \sqrt{1 + \beta \cdot g})} = \mathsf{A}^*(g)\frac{\mathsf{S}}{S_c}.$$

Thus, equation (6.2) has a solution if and only if the following quadratic in S does too:

$$0 \;=\; \frac{\mathsf{A}^*(g)}{S_c} \cdot \frac{f'(\mathsf{A})}{f(\mathsf{A})} \cdot \mathsf{S}(\kappa + \mathsf{S}) + \kappa. \tag{6.3}$$

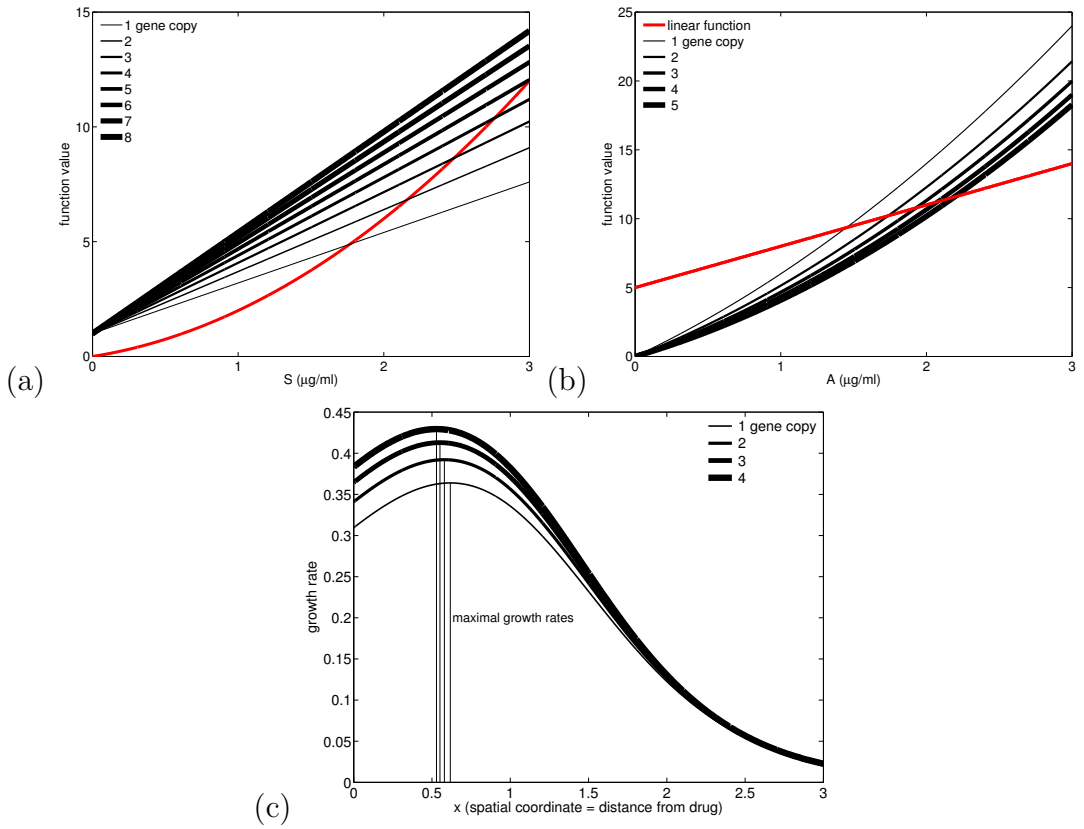


Fig. 6.7: **Illustrating the theory behind a spatiogenomic bullseye pattern.** (a) The intersection of the line and the quadratic functions give the values of the environmental sugar concentration, S, for which growth rate has a local extremum (a maximum). In this illustrate example, it is only in those regions for which S lies between 2 and $3\mu g/ml$ (at the intersection of the red and black lines) for which growth rate can have a local maximum and, then, it is only for a certain range of the number of additional gene copies. (b) The analogous plot to that given in (a), but showing dependence on A rather than S. (c) From (a) and (b) we deduce that, for the correct values of A and S, there can be local spatial maxima with respect to growth rates whose location can change with the number of copies of the resistance operon, *acr*.

We have reached a point beyond which we cannot easily continue without specifying a form for the function $f(\mathsf{A})$ which denotes the rate of decrease in growth rate as a function of antibiotic. These are not generally well-understood and, in the absence of any better model, we assume that we are working with a protein synthesis inhibitor that binds to a pocket on the ribosome and that growth rate is proportional to the number of drug-free ribosomes. If we impose mass-action kinetics onto the following standard schema whereby antibiotic, A, binds to the ribosome, R, to form the inhibited complex, $[AR]$,

$$A + R[k_{-1}]k_1[AR]$$

then one can readily derive

$$\frac{R}{R+[AR]} = \frac{1}{1+qA}$$

as being the fraction of antibiotic-free ribosomes. We therefore set $f(\mathsf{A}) = 1/(1+q\mathsf{A})$ where q is a parameter. Given this set of assumptions, we deduce that equation (6.3) is equivalent to

$$\mathsf{S}\cdot(\kappa+\mathsf{S}) = q\kappa\left(q\mathsf{S}+\frac{S^c}{\mathsf{A}^*(g)}\right). \tag{6.4}$$

Equation (6.4) can be re-written with respect to antibiotic too, whereupon it takes the form

$$\mathsf{A}\cdot\left(\kappa S_c^{-1}\mathsf{A}^*(g)+\mathsf{A}\right) = q\kappa\left(1+\mathsf{A}\right). \tag{6.5}$$

It is clear from the linearity and quadratic nature of their respective LHS and RHS that equations (6.4) and (6.5) have solutions that depend on g and, as illustrated in Figure 6.7, the growth rate function G, defined above, can have local maxima in space that are located in different places for different values of g. In order words, an antibiotic can create a spatio-genomic pattern with a multi-bump structure with respect to the drug gradient.

Whether, or not, the potential for such a multi-bump structure is realised in practise depends on many parameters within the model. For example, not only is an antibiotic gradient required in the theory, a resource gradient has to be created by virtue of the fact that the use of high antibiotic dosages kills so many cells that it the resources are 'released'. This also requires the presence of cells some distance from the drug source in exactly the manner depicted in Figure 6.1. In that figure, the indicated 'killing zone' contains no bacterial cells but it does contain antibiotic and glucose, and other nutrients, in the agar that will both diffuse outwards to create the requisite gradients. The glucose, or any other sugar, gradient is created by virtue of the greatest population densities being present furthest from the drug at the start of the experiment.

6.3.2 A theoretical genetics model of drug efflux

Having used a highly stylised form of mathematical modelling to infer that a spatio-genomic pattern, in the form of a bullseye, can appear in an antibiotic gradient because of a spatial nutrient gradient, as illustrated in Figure 6.7, we now turn to a mathematical that is more explicit in its ability to capture both the diffusing nature of different bacterial and chemical species and the manner in which ecological dynamics mediate the potential existence of the bullseye pattern.

To begin the specification of the model, we first define the radial diffusion operator, written L, of a function $f(r)$, by

$$L(f) = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r}.$$

This is defined for sufficiently smooth functions, f that satisfy $f'(0) = 0$ and $f'(r_0) = 0$ where $r_0 > 0$ is a positive constant that defines the radius of the circular domain in which the diffusion processes are assumed to take place.

The following partial differential equations are now used to describe the spatial dynamics of antibiotic inhibition to which we allude above:

$$\begin{aligned}
\frac{\partial}{\partial t}B_0 &= G_0(S, a_0)B_0 + \delta(1+\Delta)B_1 - \delta B_0 + \rho_b L(B_0), && \text{(6.6a)}\\
\frac{\partial}{\partial t}B_1 &= G_1(S, a_1)B_1 - \delta(1+\Delta)B_1 + \delta B_0 + \rho_b L(B_1), && \text{(6.6b)}\\
\frac{\partial}{\partial t}a_0 &= -\eta a_0 + B_0\varphi(a_{\text{ext}} - a_0) - B_0\frac{vp_0}{k+p_0}a_0 + \delta(1+\Delta)a_1 - \delta a_0 + \rho_a L(a_0), && \text{(6.6c)}\\
\frac{\partial}{\partial t}a_1 &= -\eta a_1 + B_1\varphi(a_{\text{ext}} - a_1) - B_1\frac{vp_1}{k+p_1}a_1 - \delta(1+\Delta)a_1 + \delta a_0 + \rho_a L(a_1), && \text{(6.6d)}\\
\frac{\partial}{\partial t}a_{\text{ext}} &= -\eta a_{\text{ext}} - \sum_{i=0}^{1} B_i\varphi(a_{\text{ext}} - a_i) - \sum_{i=0}^{1} B_i\frac{vp_i}{k+p_i}a_i + \rho_e L(a_{\text{ext}}), && \text{(6.6e)}\\
\frac{\partial}{\partial t}S &= -\frac{vS}{\kappa+S}\sum_{i=0}^{1} B_i + \rho_u L(S), && \text{(6.6f)}
\end{aligned}$$

subject to initial conditions as depicted in the schema of the protocol in Figure 6.1 and using no-flux (Neumann) boundary conditions so that no mass can spill out of the simulated experimental agar plate at any time.

The model (6.6) captures the densities of two bacterial subpopulations having density $B_i(t)$ at time t and expressing different numbers of copies of an efflux gene or operon, where the $i = 1$ genotype has one more than the $i = 0$ genotype, where this gene codes a protein, or protein complex, that transports drug from the cell. The intracellular drug concentrations for each subpopulation B_i is denoted by a_i and is measured per unit volume. The variable a_{ext} is the extracellular concentration of drug, S denotes the concentration of a limiting carbon source, glucose in our experimental systems.

Note that mutational changes in the genome are encoded in such a way that passage from B_1 to B_0 is more frequent than from B_0 to B_1. We have done this because the genomic region in which *acr* resides is unstable in the sense that replication of the chromosome usually leads to just one copy of the operon, and mother cells with a duplication are more likely to have two 'daughter' cells where one such daughter is without that duplication than to have two daughters where both daughters carry it. In other words, there is strong

physiological purifying selection acting against the duplication mutation [85].

Equations (6.6a) and (6.6b) describe the variation of bacterial density with respect to time. The model assumes that this variation is caused by three phenomena, one for each term of the equation. The first part of these equations describes the temporal variation of bacterial density B_i due to the growth and division of bacteria with intracellular drug concentration a_i and in an environment with glucose concentration S. This phenomenon is modelled as an exponential growth of the bacterial population, with cellular growth rate

$$G_j(S, A) := c_j \cdot \frac{1}{1 + pA} \cdot \frac{vS}{k + S},$$

where c is cell yield per glucose supplied, v and k are the maximal uptake rate and half-saturation constants associated with Michaelis–Menten uptake of the single carbon source S, and p is the reciprocal of the half–saturation constant due to the inhibition from ribosome-antibiotic binding.

To represent the cost of expressing the efflux gene, we chose the yield of c_i of each sub-population, B_i, to satisfy the condition that their respective yields are ordered such that the greater resistance genotype has a lower per-glucose cell yield, in other words $c_1 < c_0$. We have assumed this because the chromosomal recombination events that produce increases in the copy of the *acr* operon carry with them 8% of the entire set of genes held in the chromosome. It is highly likely that the duplication will therefore come at a large 'protein cost' whereby many proteins will be over-expressed following the duplication that have no selective advantage and yet energy in the form of carbon and ATP is used in order to express those genes.

The second terms of equations (6.6a) and (6.6b) describe the variation of B_i caused by down/up–regulation of the efflux gene in some of the bacteria. The down–regulation of the efflux gene in some of the B_1 cells make them become B_0, while B_0 that up–regulate that gene become B_1. It should be noted, however, that if *acr* is downregulated in B_0, the cell may still pay a cost due to the increased production of proteins on the operon that may not be downregulated. The process of up–regulation of efflux genes is assumed to occur randomly as a Poisson process at a certain rate δ per cell per hour (the probability of expression and amplification of the gene per cell per unit time are assumed, for simplicity, to be the same); similarly, the down–regulation is a Poisson process with rate $\delta(1 + \Delta)$, a value necessarily greater than δ. The final terms of (6.6a) and (6.6b) give the variation of $b_i(t)$ due the radial diffusion of b_i with rate ρ_b.

The process of up-regulation of efflux genes and both increases and decreases in the copy number of the efflux genes are assumed to occur randomly as

a Poisson process at a certain rate δ per cell per hour (the probability of expression and amplification of the gene per cell per unit time are assumed, for simplicity, to be the same).

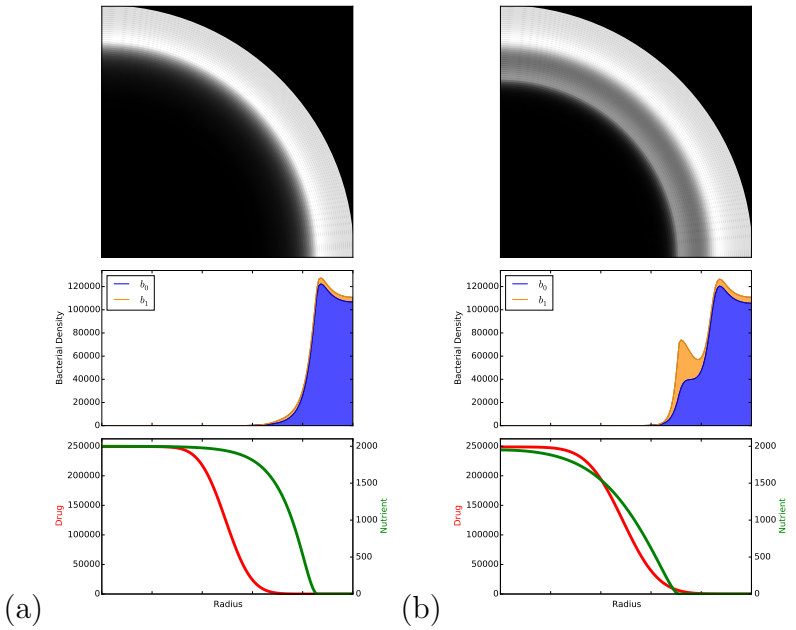


Fig. 6.8: **Predictions of two versions of the model (6.6) when efflux operon duplication is impossible (in case (a)) and then possible (in case (b))**: (a) this shows a result of simulating (6.6) when it is impossible for a second operon to appear in the chromosome, using parameter values that therefore simulate an 'operon knockout' experiment. In this case the spatio-genomic density pattern is formed from a single bright ring. (b) This is analogous to (a), except that the model parameters are now chosen so that duplication of the efflux operon is possible; note how an 'inner ring' of bacteria develops which is formed from b_1, the bacterium with that second operon copy (shown in yellow in the middle plot).

We assume a functional form for p_j (the number of pumps produced by j copies of *acr*) that is monotone increasing and bounded in j, controlled by a dimensionless constant g (the Michaelis-Menten function $p_j = (j-1)/(1+g(j-1))$). Thus p_j is also dimensionless and the quantity $p_j/(k+p_j)$ is the probability of finding a pump in the state where it momentarily is bound to drug. The rationale for this is that the polymerase transcription complex, assumed limited in number, has to compete for each gene copy, thus providing a limit on the number of efflux genes that can be simultaneously expressed.

The remaining variables in (6.6) have the following meaning: φ is the antibiotic diffusion rates across the cell membrane, v is the maximal antibiotic efflux rates and k is half-saturation constants associated with efflux pump-antibiotic molecule binding; V and K are the maximal uptake rate and half-saturation constant associated with Michaelis-Menten uptake of the single carbon source, glucose S; $G_j(S,A)$ is the per hour growth rate of each cell detailed above; δ is the rate of amplification of the efflux gene and $\delta(1+\Delta)$, a value necessarily

greater than δ, is the rate of decay of the efflux protein expressed by this gene. Finally, therefore, $n-1$ is the maximum copy number of the gene.

The main result of using equation (6.6) is shown in Figure 6.8 where it can be seen that there is a clear difference in the nature of the spatio-genomic pattern depending on whether the antibiotic efflux operon present in the chromosome can be duplicated, or not. In the case where it cannot be duplicated, the spatial pattern is simple and consists of a single 'growth ring' where bacterial densities have achieved a local optimum. This is consistent with the analysis illustrated in Figure 6.7, but it indicates that only one of the predicted rings is realised using the diffusion model (6.6). Although nutrient release from dead cells may will play a role in nutrient availability, this is not modelled here and is not necessary in order to form multiple rings. The duplication of the efflux operon allows a bacterial sub-population to approach closer to the centre of the plate, i.e., to a region with a higher concentration of glucose. Figure 6.8 shows that in order to achieve something like the bullseye pattern predicted by Figure 6.7, we must use a strain of bacterium that is able to duplicate the efflux operon, acr. These provide testable experimental predictions that we address in the following section.

6.4 Experimental Data

We now provide an analysis of the result of some bacterial growth experiments that, we believe, show evidence of the bullseye pattern predicted in theory. Before we are able to do this, we first present some mathematical measures of monotonicity that we can apply to dose-response curves in order to provide a quantitative description of how many rings a spatial dose-response pattern possesses.

6.4.1 Mathematical measures of (non-)monotonicity

A First Numerical Measure of Non-Monotonicity: the Hill function

We explore three methods of measuring non-monotonicity to determine which is sensitive to changes in density, without identifying a large number of false positives from experimental noise. The benefits of each method will be briefly discussed now. The basic idea of an empirical antibiotic dose-response is encapsulated in Figure 6.5. In it, an exponential scale of increasing dosages is presented on the x-axis and the y-axis contains data relating to the growth of the microorganism in question, in that case optical density although other measures, such as growth rate, are also used in the antibiotic literature. It is

common, as we have also done using the dose-response data in Figure 6.9(a), to assume that the data decreases monotonically with dose and to numerically fit a Hill function to it. A Hill function takes the form

$$H(A) = H_0 \cdot \frac{K^n}{K^n + A^n}$$

where H_0, K and n are three parameters to be determined as part of the data fit. Using this fit, we can then estimate IC_x values which are the dosages that give an $x\%$ reduction in growth. In other words, IC_x satisfies the relationship $H(IC_x) = \frac{x}{100} \cdot H_0$, thus the value of K is the IC_{50} of the data.

However, as Figure 6.9(b) shows, Hill functions are not accurate descriptors of dose-response data when one is interested in studying, as we are here, the effect selection for resistance has at different dosages. For the latter figure shows the effect on the Hill function-like data, gathered at 24h, after that *E.coli* data has been evolved in the presence of erythromycin for a further 72h. The details of how this is done are contained in the methods section, but it is clear from the figure that the monotone nature of the data has changed; while there is short-term monotonicty of dose-response the strength of selection at different dosages differs in such a way that monotonicity is not preserved through adaptive and evolutionary changes. This observation is consistent with the theory of section 6.3.1 which shows that the maximal growth rates for different genotypes can occur at different dosages and it is difference of these growth rates between genotypes that determine the strength of selection for resistance.

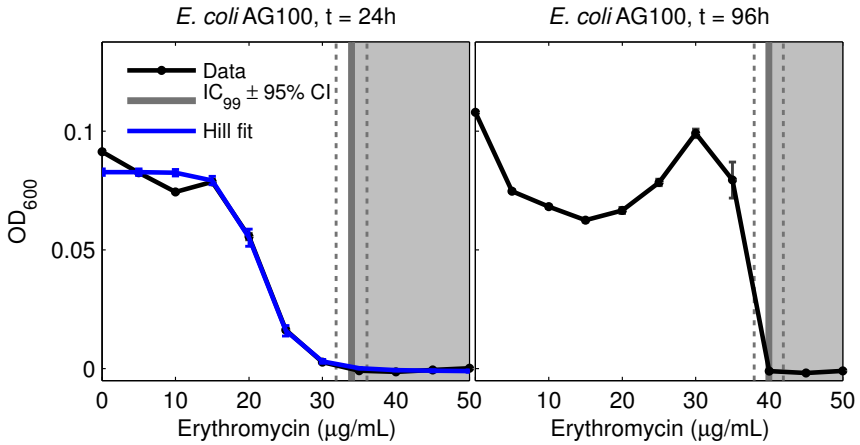


Fig. 6.9: An erythromycin dose-response curve of the *E.coli* AG100 strain with a Hill function fit super-imposed onto the data. The latter is used to determine an IC_{99} within a certain confidence. Drug dose is shown on a linear scale on the x-axis, optical density at 24h is shown on the y-axis. As can be seen in the left-hand plot, at 96 hours a hill-function can no longer be fit to the data due to the non-monotonicity observed.

We therefore require statistical methods for quantifying the the non-monotonicity properties of dose-responses. One could, of course, use the goodness of fit,

or poorness of fit, of a Hill curve to dose-response data but this would not capture information about the oscillatory properties of the dose response. Another approach is to perform a best-monotone function fit to data, this could be achieved as follows.

A Second Numerical Measure of Non-Monotonicity: monotone datafitting

Suppose that $(u_0, u_1) \in H := \mathbb{R} \times L^2((0,1), \mathbb{R})$, we can define a monotonic function $M(u) \in W^{1,1}((0,1), \mathbb{R})$ by

$$M(u_0, u_1)(x) = u_0 + \int_0^x u_1(y)^2 dy.$$

Note that $\frac{d}{dx} M(u_0, u_1) = u_1^2 \geq 0$ almost everywhere and u_1^2 is an element of L^1 by definition, justifying our claims on the properties of M.

Now, any dataset resulting from the experimental construction of a dose-response curve gives rise to a function in $W^{1,1}$ through linear interpolation: if (x_i, y_i) defines a discrete set of points with $0 = x_1 < x_2 < ... < x_N = 1$ representing a series of antibiotic dosages, we take the maximal dose to be $x_N = 1$ which should be thought of simply as a normalisation to unity, then the linear interpolant of the data points y_i is not only a continuous function, but it is in fact Lipschitz. It is therefore a member of the space $W^{1,\infty}$ which is contained within $W^{1,1}$. We can therefore apply the operator M to experimental data.

Given a function $d(\cdot) \in W^{1,1}$, supposed to represent an empirical dose response dataset, we then define the best monotone approximation of d to be the element $(v_0, v_1) \in H$ that achieves

$$\min \left\{ \|M(u_0, u_1) - d\|_{L^2} : (u_0, u_1) \in H, \|M(u_0, u_1)\|_\infty \leq \|d\|_\infty \right\}.$$

If we define $r(u_0, u_1) := \|M(u_0, u_1) - d\|_{L^2}$ then r is a sum and composition of convex and linear functionals and operators and H is a Hilbert space on which infimising sequences of M are bounded and from these observations one can prove that there is a unique minimiser of r. Moreover, approximations to this can be readily computed using a direct optimisation algorithm if d is the piecewise linear interpolant of a discrete dataset. We call the value of r so obtained the measure of monotonicity of the function or dataset d. It is not a fine measure of how the data oscillates, but it is a measure of how 'not monotone' the data is.

In order to illustrate that the appearance of non-monotone dose-response profiles with mid-dose local maxima can appear for a range of antibiotic drug

molecules, we applied this numerical measure of non-monotonicity to a video of *E.coli* AG100 growth in the presence of kanamycin, see Figure 6.11. The raw data is contained within a video, a few frames of which are illustrated in Figure 6.10. Consistent with the theories developed earlier, this data exhibits a transition from monotone to non-monotone dose response.

A Third Numerical Measure of Non-Monotonicity: the oscillation profile

Although a metric with respect to monotonicity allows one to quantify the loss of monotonicity in the dose-response profile through time, the above theoretical modelling arguments have indicated that there is the opportunity, given the right environmental conditions, for the presence of subtle structures within those nonlinear profiles. For example, the non-monotonicity may be due to selection for certain copies of specific drug resistance operons occurring at different drug dosages. In order to discern those effects, which are predicted to be associated with multi-bump dose-response patterns, we need a finer measure of these oscillation that we can apply to experimental data.

We therefore turn to the winding number of a given function, call the function $f(t)$ and let $W(f)$ denote its winding number about the number α. We consider f to be normalised in two ways such that its domain is $[0,1]$ and, over this domain, $\inf(f) = 0$ and $\sup(f) = 1$. Its winding number is the value $W(f,\alpha) = w(f-\alpha, f')$ where

$$w(x,y) = \int \frac{xdy - ydx}{x^2+y^2}.$$

We say that the oscillation profile of $f(t)$, is the function of α defined by

$$\mathcal{O}(f)(\alpha) := W(f,\alpha)$$

defined for all values of α, but potentially non-zero only for $\alpha \in \mathrm{ran}(f)$ where $\mathrm{ran}(f) := (\inf(f), \sup(f)) = (0,1)$. Note that $\mathcal{O}(f)(\alpha) = 0$ for values of α outside that range, inside that range $\mathcal{O}(f)$ only takes integer values. If we now define a sequence $O_n := \int_{\mathcal{O}(f)(\alpha)<n} \mathcal{O}(f)(\alpha)d\alpha$, with $O_0 = 0$, and then set

$$B_n = O_{n+1} - O_n,$$

we call the latter the *bumpiness spectrum* of the function f.

In order to illustrate how the oscillation profile and the bumpiness spectrum might be used to provide insight into the non-montone and multi-bump nature of empirical data, we turn to Figure 6.12. It shows synthetic examples of both monotone dose-repsonse profiles and also non-monotone profiles. The figure

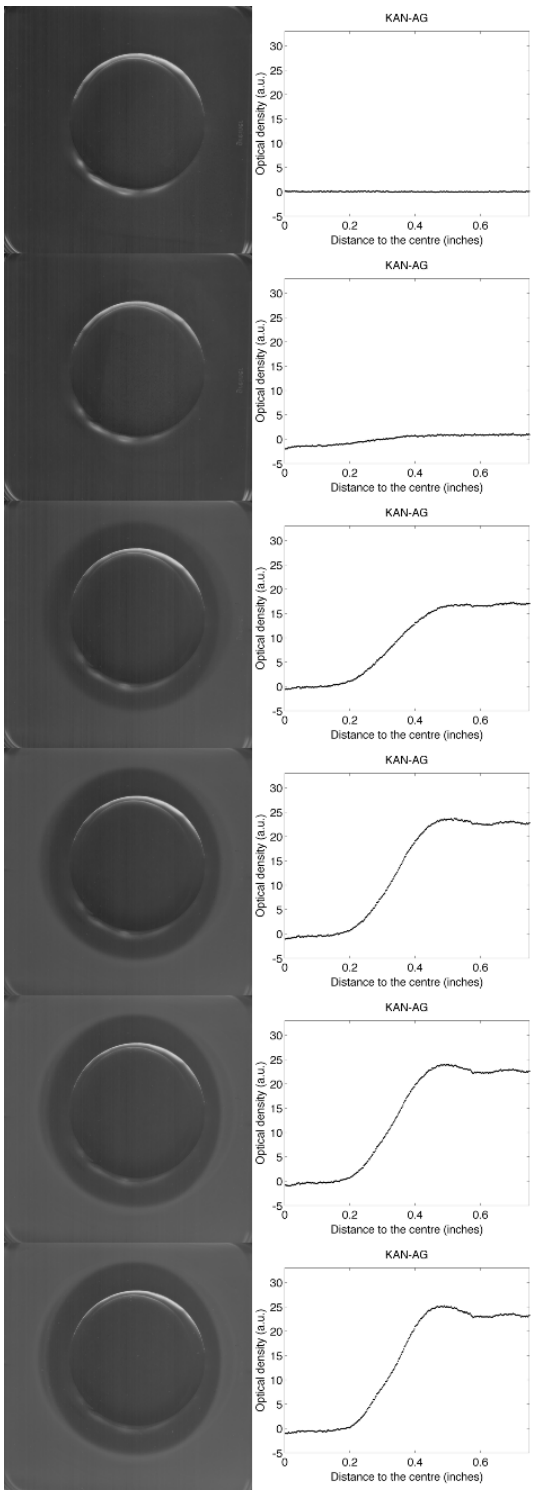


Fig. 6.10: Frames from a video showing the transition from monotone to non-monotone dose response in the inhibition of *E.coli* growth by kanamycin held in the central circular region. The top-most image was taken at time, $t = 0h$, the bottom-most at $t = 24h$. The right column shows the mean dose-response determined from each image in the left-hand column.

also shows how the bumpiness spectrum encapsulates the oscillatory character of those profiles in a manner that is reminiscent of Fourier series coefficients.

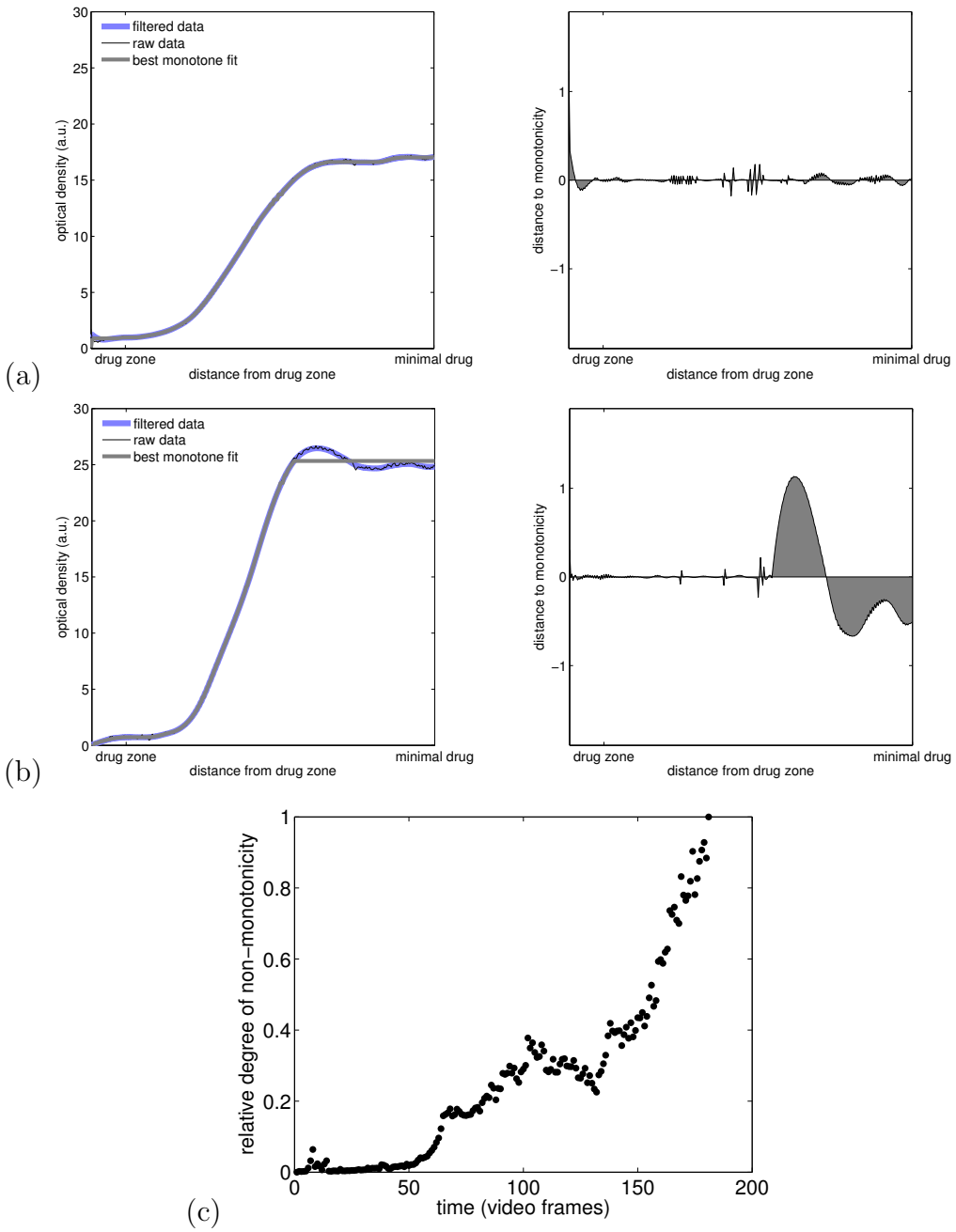


Fig. 6.11: The result of quantifying the transition in the dose-response data of Figure 6.10. Plot (a) shows how the data exhibits a low degree of non-monotonicity at early on the in the experiment because the best monotone fit to data is a good descriptor of the dose-response. Plot (b) shows, at a later time, that the best monotone fit to data gives a relatively poor fit. The right-hand images in (a) and (b) show the difference between filtered data and the best monotone fit. (c) Tracing the goodness of fit of the best monotone descriptor of the dose-response data shows deterioration through time, indicative of a transition from monotone to non-monotone dose response.

6.4.2 Applying the bumpiness spectrum to laboratory spatial dose-response data

We will now claim that the bumpiness spectrum so-defined can be used to corroborate visual evidence for an emergent pattern exhibiting concentric rings

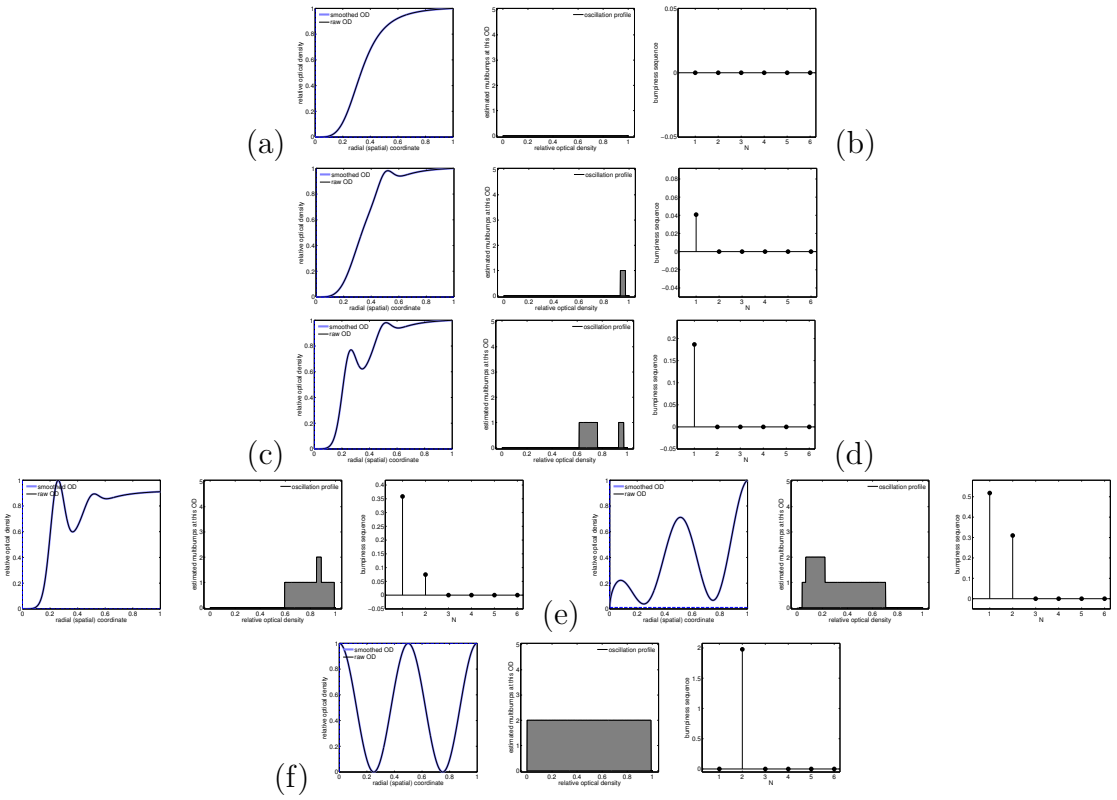


Fig. 6.12: Measures of non-monotonicity (the oscillation profile and bumpiness measure) using the winding number of some synthetic data. In each of (a-f) one can see, in the left column of the two images, a function representing a particular dose-response pattern with zero drug at the $x = 0$ position and higher drug where $x \gg 0$. Plot (a) shows a monotone dose-response which is reflected in the oscillation profile equation, shown in the middle figure of plot (a), to the zero function. Plots (b-f) have increasing degrees of non-monotonicity which is reflected in the oscillation profile in the right-hand figure of each plot whereby the grey regions have increasing area. In each window, the right-most plot corresponds to the bumpiness spectrum. Plot (f) accords with what one might expect from intuition: the bumpiness spectrum of a $\cos(2\pi \cdot x)$ function, suitably normalised, is $(0, 2, 0, ...)$ as it exhibits exactly two bumps. Plot (a) shows that the bumpiness spectrum of a monotone function is the zero sequence $(0, 0, ...)$.

of increasing radius when *E.coli* grows in a spatial antibiotic gradient, as predicted by the above theory.

First, the experimental data for this is a pair of images that we present in Figure 6.13 (a) and (b). The left-hand image shows the optical density profile as a function of the distance from the source of the drug at the centre of the plate whereby drug dosage necessarily decreases towards the edge of the plate. Images taken from these agar plates are shown in the right-hand image of the same figure whereby the white semi-circular region is the drug source. This figure illustrates bacterial population density at 24h and, to the eye, appears to exhibit bullseye ring patterns. This appears consistent with the density images on the left, and although the *acr* knockout strain has an essentially increasing density profile with decreasing drug, as may be expected, the AG100 wild-type

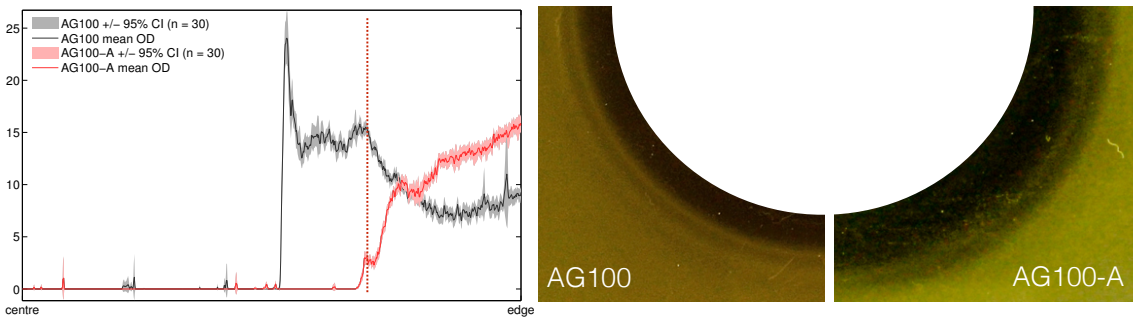


Fig. 6.13: The left-hand image shows the mean optical density of two bacterial strains at 24h, AG100 and AG100-A, the latter does not have a functional *acr* operon whereas the former does. The two right-hand images are false colour representations of two different agar plates on which the respective strains were cultured. The centre of each plate contained 1/ml of Doxycycline.

strain has an altogether different, and clearly non-monotone, character.

To quantify this apparent non-monotonicity, effectively in order to produce numerical measures of what the human eye perceives in Figure 6.13(b), we present Figure 6.14. It shows the oscillation profile of the population density data as a function of the spatial coordinate for the two *E.coli* strains AG100 and AG100A, Figure 6.14 is an analysis of data taken from Figure 6.13(a).

The population density data relating to AG100A in the left-hand image of Figure 6.14 looks almost monotone to the human eye and, indeed, Figure 6.14(a) shows little quantitative evidence of non-monotonicity. There is some evidence of a mutli-bump profile to be seen in the bumpiness spectrum, but certainly not to the same extent as that seen in the analogous data for strain AG100. This data can be seen in Figure 6.14(b). The latter data alluded to, for AG100, shows in quantitative terms what appears to the human eye in Figure 6.13 whereby distinct rings highlighting regions of growth can be seen for this strain, consistent with predictions made earlier using mathematical models.

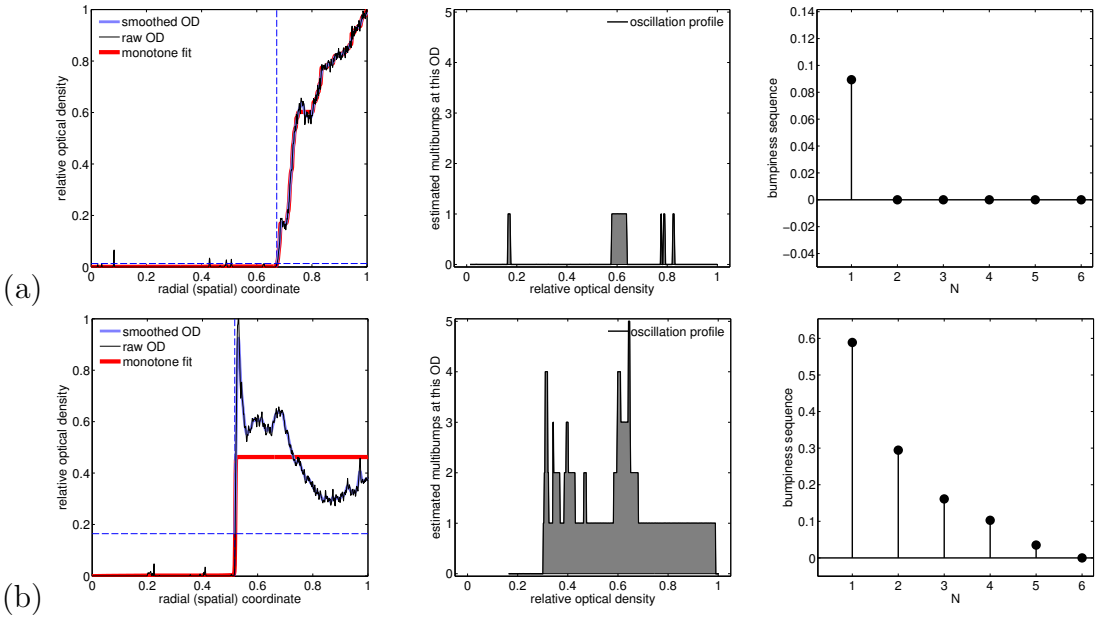


Fig. 6.14: (a) The left-most plot shows an analysis of the dose-response profile which constitutes the observed mean optical density as a function of the spatial distance from the source of the drug (at position zero), and in this case it is derived from strain AG100A that lacks the efflux pump operon, *acr*. The left-hand plot shows (in red) the best monotone function fit to data, filtered data (in blue) and the dash blue lines indicate the leftmost point where the optical density has been established to be significantly above zero using a t-test with significance level $p = 0.05$. To the right of this point we assume that optical density is significantly positive. The middle plot is the oscillation profile $\mathcal{O}(f)(\alpha)$ for this AG100A data and the rightmost plot is the bumpiness spectrum that is derived from the oscillation profile. Neither of the latter two measures are consistent with the presence of oscillations. (b) This is the analogous analysis to (a) but now for the strain AG100 that possesses the *acr* operon.

6.5 Conclusions

We have used a prediction from the results of several theoretical models to show that bacterial growth in an antibiotic gradient can result in a visible spatial pattern, moreover this pattern can be used to reject the hypothesis that more antibiotic necessarily equates to less bacterial population growth.

The laboratory model system we have used is not based on the occurrence of single nucleotide polymorphisms (SNPs) as the drivers of the evolutionary processes that produce this spatial pattern, but rather a drug efflux operon that is rapidly duplicated by recombination in the face of antibiotic stress is the mechanism of resistance selected for by evolution. It is selection on this operon that creates the requisite spatial pattern. Moreover, we have introduced a series of quantitative measures of monotonicity to help discern in quantitative terms what is apparent to the human eye in images of bacterial growth on agar plates in antibiotic gradients.

Our results are not directly relevant to the interpretation of antibiograms in a clinical context, although we have shown that an increase in drug dosage increases the size of the zone of bacterial growth inhibition according to a square root law, consistent with diffusion theoretic models of bacterial killing. However, our observations are key to understanding conditions for which an increase in drug concentration will not necessarily reduce bacterial densities. After all, all practical uses of antibiotics create decaying dose profiles as one moves further way from the point source of an antibiotic drug. So, while it may be thought that one can mitigate, or even obviate, the process of selection for drug resistance by always passing to ever greater dosages [**?**, 90], those dosages must eventually decay to zero in regions of treatment located not too far from the source whereby one might be able to observe the effects we allude to in our study.

Chapter 7

Conclusions

7.1 Antibiotic combinations

Much faith has been placed in the idea of utilising cocktails and combinations of antibiotics, either in mixture, sequence or cycling, in order to minimise the problems caused by resistance. The logic behind such a hypothesis is that a cell that gains resistance to a single agent will then be confronted with a second (or third, etc) agent, to which it has no resistance. For this reason combinations of drugs are often in different classes, to minimise the chance of cross-resistance. However, the literature shows that determination of drug interactions as synergistic or antagonistic are complicated measurements that are confounded by a number of factors, including the time of measurement, resource background, drug dose and the method of measurement. It is for these reasons that the drug pair used in Chapter 1, Rifampicin and Doxycycline, have been assessed as both synergistic and antagonistic in the treatment of *Brucellosis melentiasis*. What is more, my research shows that an interaction that appears to be strongly indicating a favourable interaction (normally synergy) can be misleading, depending on the growth stage of the cells; therefore a difference in the time of measurement of just 12 hours can show a complete inversion of interaction. Initially an experiment was devised to test the hypothesis that antagonism can slow the emergence of antibiotic resistance, but in order to test this hypothesis accurately we must be assured that the interaction is stable and robust - if the original interaction is mis-assigned then the results of such a test can not be relied upon.

7.2 Antibiotics as growth stimulants

The first three chapters of this work are predominantly concerned with non-monotone dose response, and an adaptive response to Dox that leads to a

stimulation of bacterial growth. Initially this seems to be of very great concern - although antibiotic resistance is a widely recognised grave threat, the idea of "super-resistance" that allows cells to experience a growth boost in the presence of antibiotics has not before been shown. We show that in *E.coli* AG100 several days of adaptation to Dox is sufficient to confer this overgrowth phenotype. However, recapitulations of this experiment using different but closely related strains, failed to replicate the result. The fact that this effect appears to be strain specific allows for an exploration of the mechanism behind such an adaptive effect. To this end, the genetic differences between AG100 and MG1655 were compared. Despite a very high degree of homology between the two strains, a mutation in *rpsL* in AG100 is not present in MG1655, This mutation represses the function of *rpsL*, which in turn leads to a downregulation of the production of sigma-factor; *rpoS*. This protein is responsible for many of the physiological changes that a cell undergoes when resources are limited, and the growth enters the stationary phase; the absence of *rpoS* is reported to lead unrestricted division, with the cell failing to enter stationary phase. Equally, this mutation is known to improve growth on poor carbon sources, such as acetate, which our results also demonstrate. For this reason a range of strains (including metabolic mutants from the Keio collection) were subjected to high doses of Streptomycin in order to attempt to force a mutation in *rpsL* and recreate the overgrowth phenotype in a new strain. This experiment was partially successful, with one of the strains tested showing an overgrowth in the presence of Dox, after having become resistant to Strep (around 1000xMIC). The strain that showed this overgrowth was $\Delta \text{manZ}_{strepR}$. Although it is interesting that a strain that is already glycolysis limited, through a knockout of mannosephosphotransferase, is the strain to show an increase in yield when treated with Dox, an important future step would be to sequence this strain, ideally with whole genome sequencing, although sequencing the *rpsL* gene would also be sufficient to determine if the strep resistance mutation present in AG100 is also present in this strain. Likewise, sequencing of other strains that did not develop an overgrowth phenotype would enable us to determine if the hypothesis is correct, or if that lack of overgrowth is due to those strains not developing the strep resistance mutation in the same location.

In addition to confirming the hypothesis through sequencing, it would be advantageous to repeat this protocol with a far greater number of replicates. Our results showed that although $\Delta \text{manZ}_{strepR}$ did gain a stimulatory effect from Dox, this effect was transitory, lasting only 24-48 hours, and there was a high level of variability between replicates. An experiment is currently being performed using automated liquid handling technology to increase the throughput of this test, with 96 replicates being tested, after anywhere from 1-12 days of

exposure to increasing concentration of strep. This protocol will allow us to determine the conditions that maximise bacterial yield, and gain more robust results. Furthermore, in order to definitively ascertain the specificity of this effect, a high throughput screening of bacterial species and strains to Dox should be performed. Using automated liquid handling and OD reading, it would be possible to screen dozens of common pathogenic and non-pathogenic bacteria and assess the inhibition over evolutionary timescales.

7.3 Phage co-evolved metabolism

We show, through a carefully controlled model system using co-evolved REL606 *E.coli* and lambda phage, that metabolic trade-offs are present within each strain when grown in differing concentrations of maltotriose as the sole carbon source. However, we also show that when compared between strains these growth parameter correlations may manifest as trade-ups, whereby at a given carbon resource concentration, both yield and rate are increased. We do not observe a noticeable trade-off or trade-up between the level of resistance to phage and the growth rate or length of lag phase of bacteria, but we do notice that yield and resistance are positively correlated at high maltotriose concentrations and negatively correlated at high concentrations, thereby displaying both trade-off and trade-up behaviour dependent on environment. We finally show that the change in the protein shape of the maltoporin is a reasonable predictor of phage resistance, with a switch to phage resistance when the shape of the protein changes beyond a certain threshold, with similarly shaped maltoporin mutants tending to cluster with similarly resistant mutants.

7.4 Non-monotone spatial dose-response

The greatest challenge in collecting experimental data of this type was ensuring that image capture was precise enough to detect small changes in density, and convert that data to a pixel intensity score. The experimental setup created some image artefacts, for example condensation or reflection on the plate confounded accurate readings. Additionally, small imperfections in the agar plate construction can again lead to non-symmetrical diffusion, which impacts image analysis. Since the completion of this study, the image platform has undergone several design iterations, including the addition of superior insulation and a air flow (filtered) system in order to reduce condensation. Likewise, the position of the camera and light sources has been modified to remove reflections. These changes have improved image quality, which further reduces

noise and means that oscillatory profiles are cleaner. Repeating the bullseye pattern experiments in this new system design may somewhat improve results. This imaging system allows for exciting studies to be made in real time, in the future similar experiments could be performed using swarming strains of bacteria on plates with various drug 'pathways' allowing visualisation of the pathway that most rapidly selects for resistance. Plate construction could also be potentially improved; the lab has acquired a 3D printer, allowing custom design of agar templates, rather than improvising using lab equipment (for example, using a 40mm petri dishes as the template for the central drug containing agar section); this would remove the non-symmetry previously discussed. Bullseye patterns are frequently seen (although often ignored) when performing E-tests on a large variety of bacteria. It would be an interesting extension to this study to repeat the setup using a wider selection of species and drugs: if the bullseye pattern is replicated, sequencing could be used to determine if duplication of efflux is a necessary feature in each case.

Appendix A

Chapter 2: Appendix

A.1 Materials and Methods

Reagents

Bacterial cultures were grown from the same colony for 16 hours at 30^{o}C and 180rpm in M9 minimal broth (Part A: 40mM K_2HPO_4, 15mM KH_2HPO_4; Part B: 2.3mM Trisodium citrate, 7.5mM $(NH_4)_2SO_4$, 1mM $MgSO_4$) supplemented with 0.2% glucose and 0.1% Casamino Acids (Fischer Scientific and Duchefa).

Strains

The AG100 strain of *E.coli* was chosen as the model organism of this study, the genome is well-understood and the Keio library of single-gene knockouts is available. We consider AG100 a wild-type strain, although it was first isolated as a seperate K-12 strain in 1983. The AG100A(Δacr) mutant lacking the full operon *acrAB* required for producing essential efflux pump components was also used. All strains used were provided by Dr Laura McMurry (Tufts University). Strains were stored frozen in glycerol (30%, -80° with μl samples of AG100 thawed from frozen aliquots and grown in overnight cultures in M9 media (37°C, 180 r.p.m shaking) for use. Colonies were visually checked after 24 hours culture on LB agar and subsequently re-cultured in M9 overnight, as above, for use.

Dose-response assay

Rif and Dox powder (Sigma Aldrich) were prepared and stored at 4° in the dark, as per manufacturers instructions. Stock solutions were made by dissolving the powder in methanol and mqH_2O respectively (50μg/ml and 5μg/ml). Row-wise increasing concentrations of antibiotic were inoculated with AG100

in a 96 well microplate (150μl total volume per well) using a 96 pin replicator. Optical density was measured every twenty minutes (Biotek ElX808, 600nm). Control growth was established in M9 media (glucose 0.2% + casamine 0.1%) as described in [119].

Rifampicin: static/cidal test

The growth curves for a range of concentrations of rifampicin are shown for AG100 and AG100A . The relevant optical density measures are given in Figure A.1 whereas Figure A.2 below shows CFU data at three time points at zero drug and 2.2μg/ml Rif. The CFU data is consistent with a reduction in the number of live cells and, therefore, bactericidal effect of the antibiotic. (mean CFU based on three replicates and 24h total incubation; 37° used for CFU data) beyond 12h in Figure A.1. Rifampicin has a bacteriostatic effect on wild-type AG100.

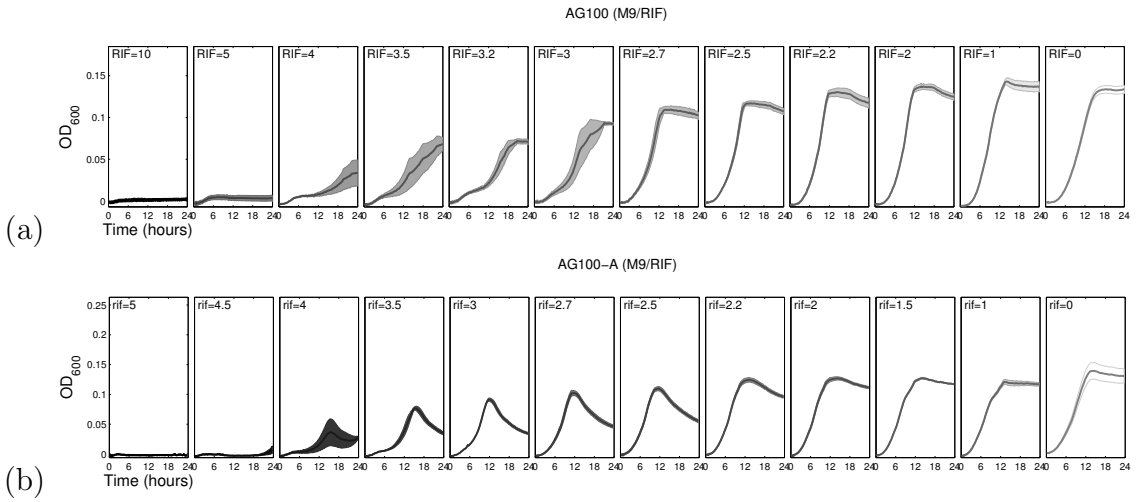


Fig. A.1: (a) Optical density profiles of AG100 through time cultured for 24h at different concentrations of rifampicin. The highest dosage used leads to no growth at all, moreover growth is slower at increased dose but there is no evidence of cell death or lysis (see appendix for controls). (b) Optical density profiles of the AG100A(Δacr) strain through time cultured for 24h at different concentrations of rifampicin. The loss of population density in stationary phase is consistent with bactericidal activity and the live cell count (see appendix) shows a reduction in CFUs after the 12h timepoint is reached.

Drug-Interaction assay

Single-drug basal concentrations for both Rif and Dox given by the $IC_{50,60,70,80}$ concentrations were obtained from dose-response curves and used as the basis of a combination treatment of in a synergy assay, as follows. Combinations of Rif-Dox with equal inhibitory effect were then combined in the following

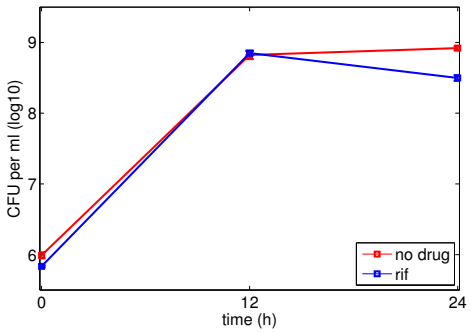


Fig. A.2: A CFU/ml count for strain AG100A at three different time points both with (red; at 1.8μg/ml) and without rif (blue). Three readings were taken at 0h, 12h and 24h showing the reduction in viable cells between the latter two time points in the presence of rif (error bars are shown (s.e.); $p < 0.01$).

proportions; $(R, D) = (1, 0), (\frac{3}{4}, \frac{1}{4}), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{4}, \frac{3}{4}), (0, 1)$ relative to those basal dosages. The microtitre plate was inoculated and automatically read as described above. The most challenging aspect of this experiment was ensuring that the IC_x levels for both drugs were calibrated correctly, so that for example, an IC_{50} dose of dox gave rise to the same cell density as the IC_{50} concentration of Rif. The dose response curves for all the drugs and strains used is below.

Rifampicin and doxycycline dose-response curves

Dose-response curves for both these drugs and for *E.coli* K12 strain AG100 are shown in figure A.3. These figures are used to calculate the different concentrations of both drugs are required to achieve, for example, the IC_{90}. Dose-response curves for both these drugs and for *E.coli* K12 strain AG100A(Δacr) are shown in figure A.4. Dose-response curves for Rifampicin and Erythromycin are shown in figure A.5.

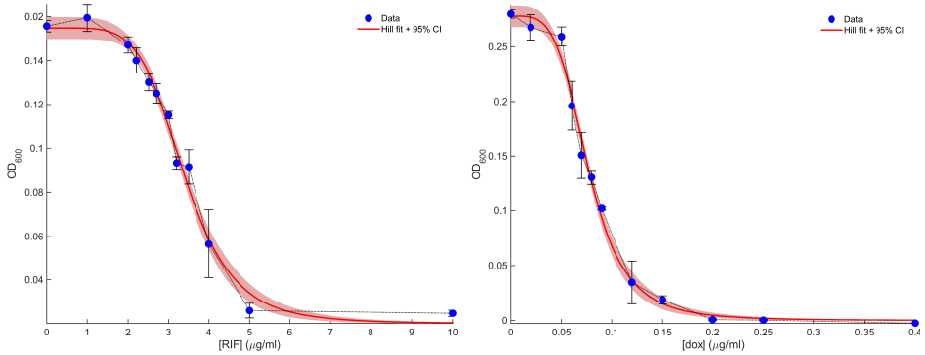


Fig. A.3: (a) Dose-response of AG100 for rifampicin at 12h. (b) Dose-response of AG100 for doxycycline at 12h. Basal dosages used to test for dynamic drug interaction profiles are IC_{60}, IC_{70}, IC_{80} and IC_{90}.

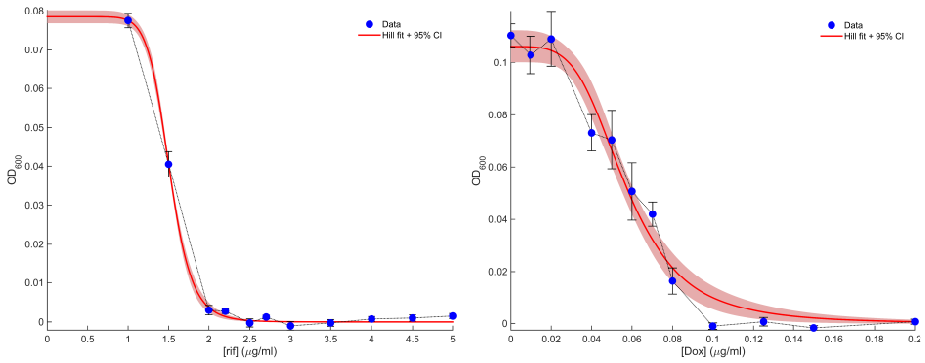


Fig. A.4: (a) Dose-response of K12 strain AG100(Δacr) for rifampicin at 12h. (b) Dose-response of AG100(Δacr) for doxycycline at 12h. Basal dosages used to test for dynamic drug interaction profiles are IC_{50}, IC_{60}, IC_{70} and IC_{80}.

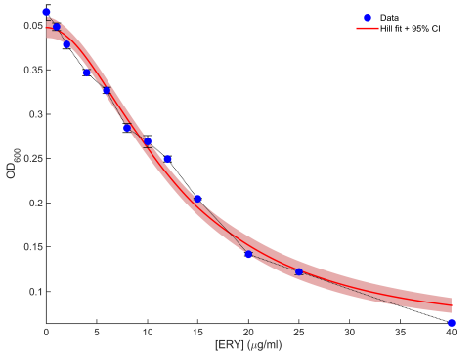


Fig. A.5: Dose-response of AG100 for Erythromycin at 12h. Illustrates the IC50 basal dosages used to test for dynamic drug interaction profiles in M9 media supplemented with 1% glucose and 0.5% casamino acids .

CFU to OD conversion

Throughout this document, cell population density has been measured using OD at 600nm as a proxy for that density. This can be confounded by cell size changes which can vary through growth phases and at different sugar concentrations. In order to ensure stability of the relationship between different sugar supply concentrations and cell numbers in CFU per ml, we performed an experiment where cells were grown in different maltotrise concentrations and both OD and CFU/ml were correlated. This is shown in Figure A.6.

In order to ensure that optical density measurements used in this document correspond to live cells at all glucose concentrations, we additionally performed a calibration of the plate reader by comparing OD measurements to CFU (colony-forming units) at multiple glucose concentrations, comparable to the range used throughout this research. The data produced using this experiment across a range of glucose concentrations in M9 minimal media are presented in figure A.6 which shows an approximately linear relationship holds between the two measures.

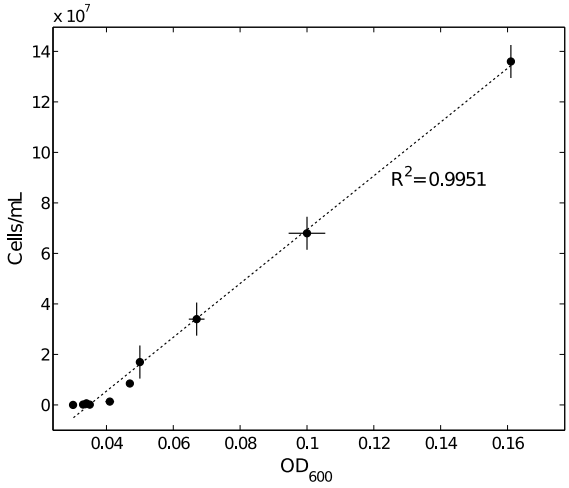


Fig. A.6: **CFU and OD are equivalent measures of *E.coli* population densities, provided that density is high enough.** Validation that *E.coli* cell densities as measured in units of OD_{600} or CFU (colony forming units) are approximately equivalent: from this data we deduce that 10^8CFU per ml $\approx$ 0.128OD.

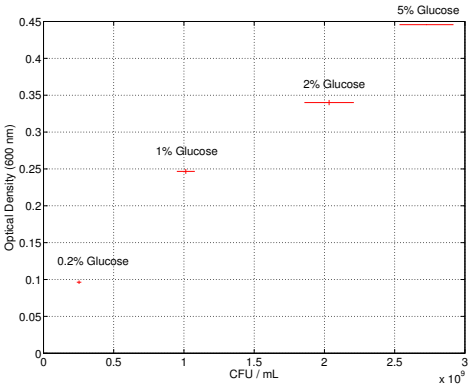


Fig. A.7: Illustration of the linear correspondence between cell densities at different glucose concentrations, measured using optical density at 600nm light wavelength and a live cell count.

If we accept the datapoint at 1% glucose in figure A.7, from this we deduce that 10^9 CFU/mL corresponds to 0.25 OD units. As the culture devices we use contained a liquid volume of $150\mu L$, we deduce that 0.25 OD correspond to 1.5×10^8 live cells, 0.1 OD therefore corresponds to 6×10^7 live cells.

Serial-transfer protocol

A 96 well micro-titre plate was set up as follows: A dose response assay was performed to determine the concentrations of Rif and Dox required to achieve $IC_{50,60,70,80}$ at twelve hours. These concentrations were used to prepare combination treatments. The five Rif:Dox treatments and controls (combinations proportions used were; 100:0, 99:1, 95:5, 90:10, 0:100, Positive control, Negative control, Empty well) filled the wells from top to bottom of the plate. 3

replicates were used for each basal inhibitory concentration. Master solutions of all well environments were made in sufficient quantities to prepare 9 identical plates. Plates 1-8 are used for the 8 serial 12h transfers in the experiment. Plate 9 is re-inoculated with fresh *E.coli* at the end of the experiment as a drug degradation control. Transfers are carried out using a 96 pin replicator, and optical density is measured every twenty minutes (600nm).

A.2 Supplementary data

Rif-adjuvant rate of adaptation calculation

Rate of adaptation calculated as per Hegreness *et. al.,* [54], the methodology of which is discussed in more detail in chapter 2. Increasing amount of Rifampicin in the combination treatment lead to higher levels of antagonism (and hyper-antagonism, with the growth of bacteria being greater than that of control), but also lead to an increase in adaptation rate. The seasonal change in growth rate, and the time taken for half maximal growth rate change are shown in Figure A.8, with the correlation between increasing OD and rate of adaptation shown in Figure A.10.

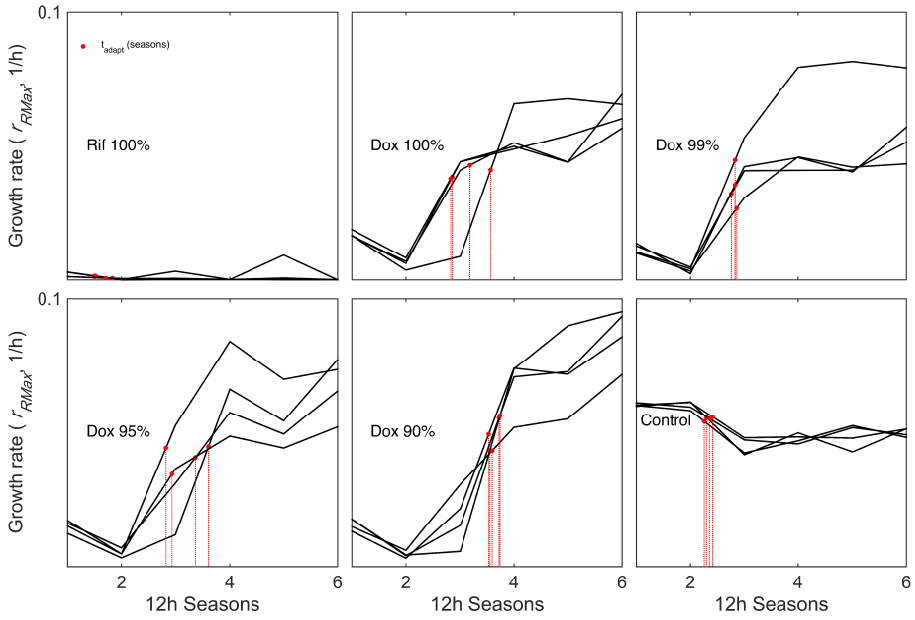


Fig. A.8: Change in growth rate shown for all environments. Only the rifampicin mono-therapy shows a complete inhibition of bacteria across the course of the experiment. Doxycycline quickly becomes useless as an inhibitor and starts to act as a stimulant to growth. Surprisingly however, the addition of small proportions of Rif (up to 10%) increase this stimulatory behaviour.

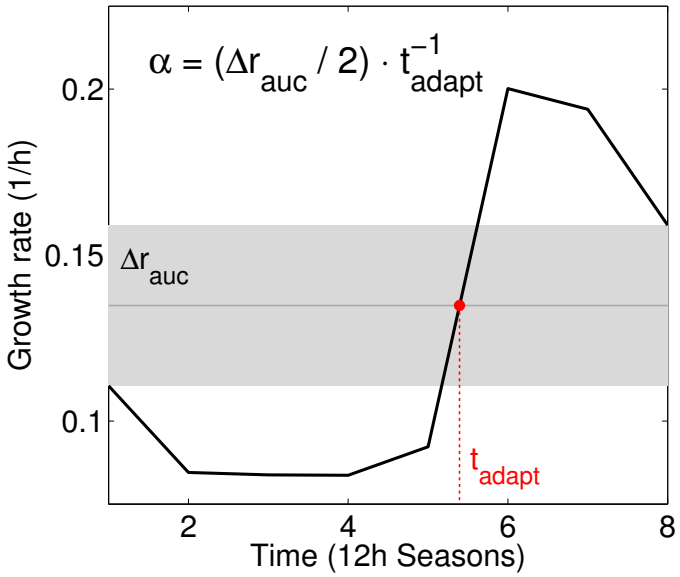


Fig. A.9: Representative data plot showing the cumulative change in growth rate over the 8 seasons of the experiment. The rate of adaptation (α) is calculated using the formula in the plot. The grey shaded area shows the change in growth rate (rate as measured by area under the curve (AUC).

The rate of adaptation is measured as shown on the above plot (Figure A.9), and is a second derivative of the growth growth rate. In this way, the time taken for half the maximal change in rate is multiplied by half maximal change in rate to derive the rate of adaptation.

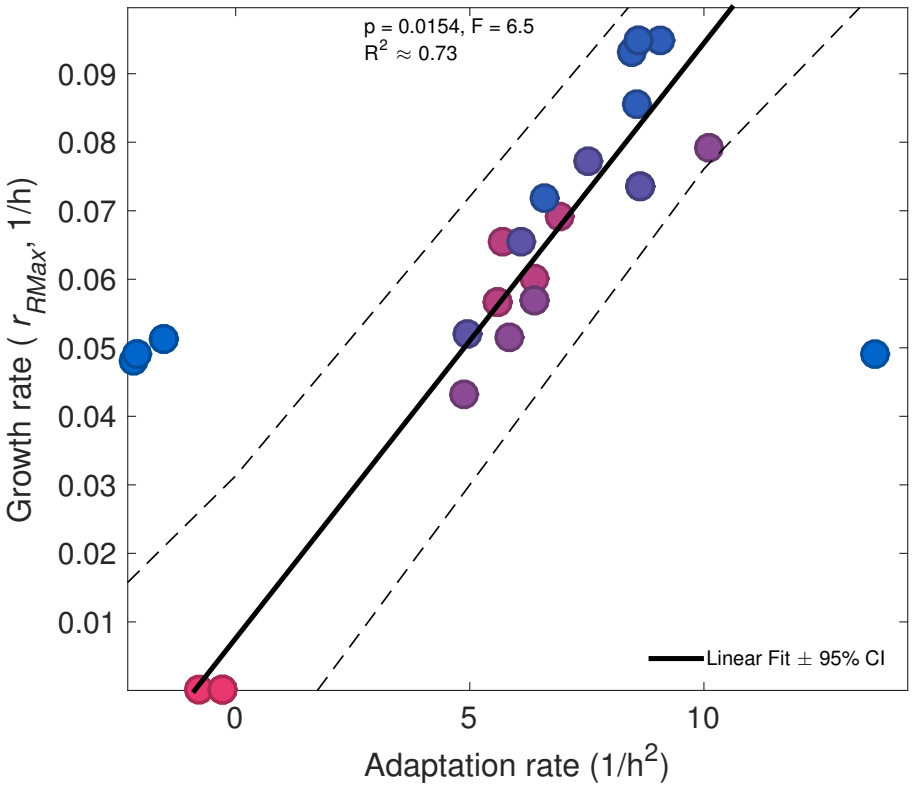


Fig. A.10: As expected, when optical density increases over the course of the experiment, it denotes adaptation the environment, and therefore rate of adaptation increases as well. The outlier show Rif monotherapy (Red) with zero growth and the control with growth, but negative adaptation. Dox mono-therapy is shown in blue, with purple data points representing increasing Rif proportions.

Appendix B

Chapter 3: Appendix

B.1 Methods and Materials

Dose-response

Concentrations of drug from twice MIC to 0 were selected. Unlike in NCLSI guidelines, drug concentrations did not conform to a two-fold dilution method. Instead concentrations were chosen to fall within the range of 'sub-inhibitory' doses. 96 well microtitre plates (*Greiner*) were loaded column-wise with 11 drug concentration and 1 control. Replicates were loaded row-wise, at 150μl per well. Inoculation was performed using a 96 pin replicator (*Sigma*), and Growth was recorded by *BioTek eLx808 UV/ABS reader, 600nm* every 20 minutes. All drug solutions were made in media containing the requisite carbon source at the desired concentration.

Evolution of Dose-Response assay

A concentration gradient of eleven drug environments was inoculated with E.coli (AG100). Drug concentrations ranged from twice MIC (from prior experiments) to zero. Growth was recorded by *BioTek eLx808 UV/ABS reader, 600nm* every 20 minutes. Approximately 1% inocula was transferred from each well to an identical plate (using 96 pin replicator from *sigma*) to refresh the carbon source in the media. Transfers were carried out every twelve hours in the initial experiment, and every 24 hours in the subsequent experiment. Experiments were continued for four days.

CFU counts

Colony forming unit counts were used to confirm optical density reading as a proxy for cell number. CFU counts were performed for all principle ex-

periments, and calibration curves comparing OD to CFU in differing glucose environments were used to corroborate using OD as a proxy for density. Unless otherwise stated in the text, samples were taken from 3 biological replicates, and underwent sequential 10-fold diluted in nutrient free M9. 100μl was added to the surface of a solid LB agar petri disk and spread to cover the whole surface using a disposable L-shaped spreader (*Sigma*). Three technical replicates were made for each sample. Plates were inverted and incubated for 24 hours at 30°C and then counted manually.

Acetate dose-response

As above, however 1.37% acetate was used, to provide a stoichiometrically identical number of carbon atoms to the 1% glucose used in the prior experiments.

Rate-yield trade off in Doxycycline resistant strains

To investigate whether the increased rate of adaptation seen at several concentrations of Doxycycline, and the commensurate resistance gained had any effect on the metabolic efficiency of these strains, growth was recorded at six glucose concentrations. The strains that were used were all derived from AG100 and MG1655, shown in Table B.1, and had been frozen (-80°) following the 10 day adaptation protocol.

All strains were grown for 24 hours, *OD* measured at 600_{nm} every 20 minutes. Strains were grown in six glucose concentrations (0.01%, 0.05%, 0.1%, 0.2%, 0.5% and 1%) and in three different environments: M9 (supplemented with casamino acid, 0.1%), the IC_{99} of the wild-type and the IC_{198} of the wild type. This will allow a comparison of growth rates and yields for these strains and sub-strains, both in the environment that they are adapted to, and in environments that are more or less challenging than they are used to. In this way we can detect any costs of resistance to cellular efficiency or metabolism.

Table SB.1: Strains used to measure rate and yield in Doxycycline adapted cells.

Parent Strain	Adapted conditions	Name
AG100	/	$AG100_{wt}$
AG100	M9	$AG100_{m9}$
AG100	MIC	$AG100_{MIC}$
AG100	2xMIC	$AG100_{2MIC}$
MG1655	/	$MG1655_{wt}$
MG1655	M9	$MG1655_{m9}$
MG1655	MIC	$MG1655_{MIC}$
MG1655	2xMIC	$MG1655_{2MIC}$

Rate was calculated in three ways; by fitting an exponential model, a logistic model and by calculating the area under the curve (AUC). The goodness of the model fits were compared by R^2 number. However, the AUC was the measure adopted in all further calculations, as rate estimates from this method gave the most consistency for all growth curves, whereas model fit methods tended to fail in cases where growth was very low, or stationary phase was not reached.

Measuring costs of resistance

To determine whether Doxycycline resistant sub-strains are carrying a fitness cost due to their resistance, when the selective pressure of the drug is removed, the growth kinetics must be compared to the wild-type growth. All the strains from the adaptive experiment were revived from -80° stocks and revived as previously outlined. Strains were revived in the environment that they had previously been exposed to (thus $AG100_{MIC}$ is revived overnight in M9 supplemented with 1% glucose, 0.1% casamino acid and 0.3μg/ml Dox) before being grown in drug-containing and drug-free environments. Serial transfers (1%, 96-pin replicator) were made every 24 hours, for 10 seasons. Thus the cells were removed from there adaptive environment for the same length of time as the initial adaptation took to occur. By using the OD_{24hr} as a proxy for fitness (with CFU counts carried out on season 1 and season 10 to confirm cell numbers), comparisons can be made across strains and across environments.

Developing Streptomycin resistance

Strains were exposed to increasing concentrations of Strep in order to increase resistance through the acquisition of a mutation to *rpsL*. Exposure was performed on solid LB agar plates, at concentrations of 2xMIC, 8xMIC, 32xMIC and 128xMIC. The MIC was determined with a standard dose response curve, as described previously. Large inoculate (250l) from dense overnight cultures in LB media were spread on the initial 2xMIC Strep-containing plates. Between each subsequent plate, 5 colonies were sampled from the plate and grown overnight in LB media supplemented with the same concentration of Strep as they had been grown on (to ensure that resistance was not lost), before replating onto the next Strep concentration. If less than 5 colonies were successful in growing at any concentration, all viable colonies were sampled. There were no occasions with no viable colonies for sampling. The use of large inocula and colony plating allowed for a rapid increase in antibiotic concentration.

B.2 Supplementary Data

Dose-response adaptation

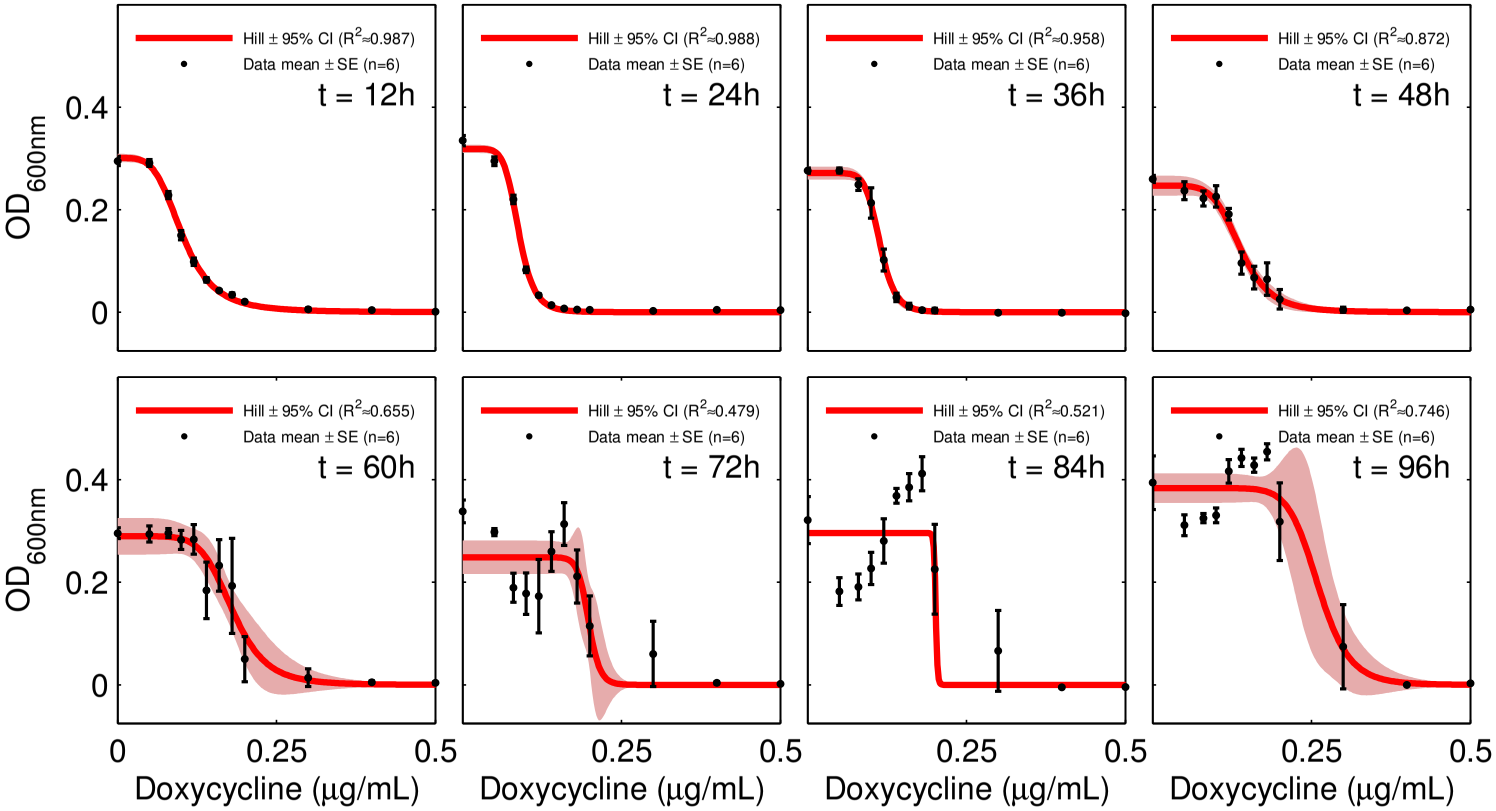


Fig. B.1: Fitting a Hill function to the dose response curves obtained from 12hr serial transfers for 8 seasons shows that the non-monotonicity of the dose response is sufficient by season 6 for the fit to fail.

Growth rate in 12 hour transfers

Evolution of dose response was carried out using 24 hour seasons to confirm that overgrowth is not an artefact caused by selecting for increased growth rate in exponential phase (figure B.4).

Dose response transfer experiments were assessed for non-monotonicity by fitting a Hill-function to the data set, and subtracting the true from predicted values. Therefore in Figure B.5, a deviation from the 0 (red horizontal line) represents an overgrowth (positive) or undergrowth (negative) compared to the predicted monotonically decreasing values. In Figure B.5 the fit is not exact at any season. This is due to a significant hormesis effect at low drug concentration, even apparent on the first season. However, the overgrowth phenotype is still seen clearly in later seasons, as the MIC increases (to the point where the highest Dox concentration used becomes sub-lethal), and the final cell density in the overgrowth concentrations reaches approximately twice the cell density of the drug-free control.

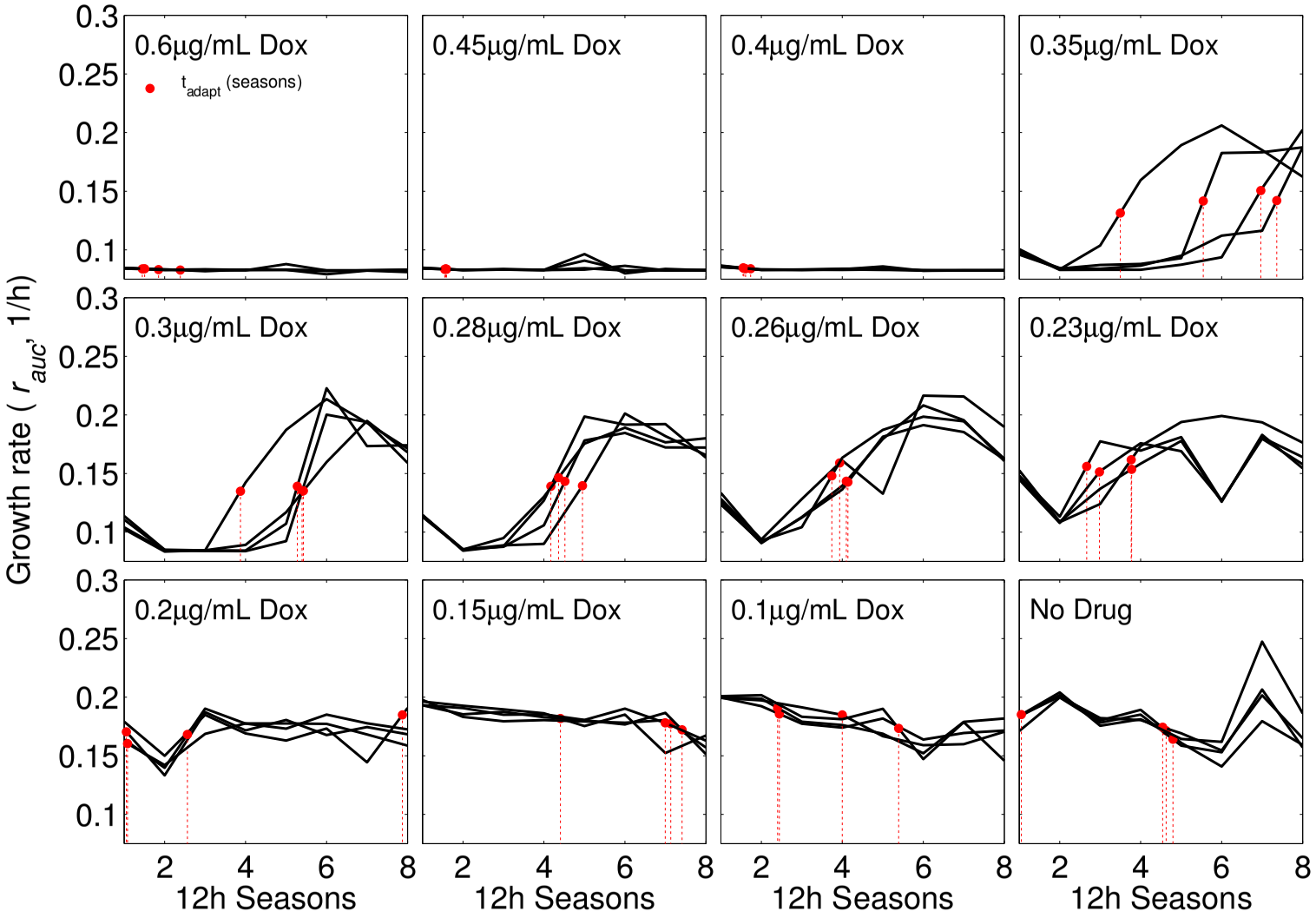


Fig. B.2: The cumulative growth rate (1/hr, defined as $\mathrm{OD}_{(final)}$)/AUC) for all replicates in all drug concentrations is shown with a dashed red line denoting the ΔR/2 ($\mathrm{Rate}_{(final)} - Rate_{(initial)}/2$). The rate of adaptation is defined as the $(\Delta \mathrm{R}/2)/T_{(\Delta}\mathrm{R}/2)$ Data for 12 hour seasons.

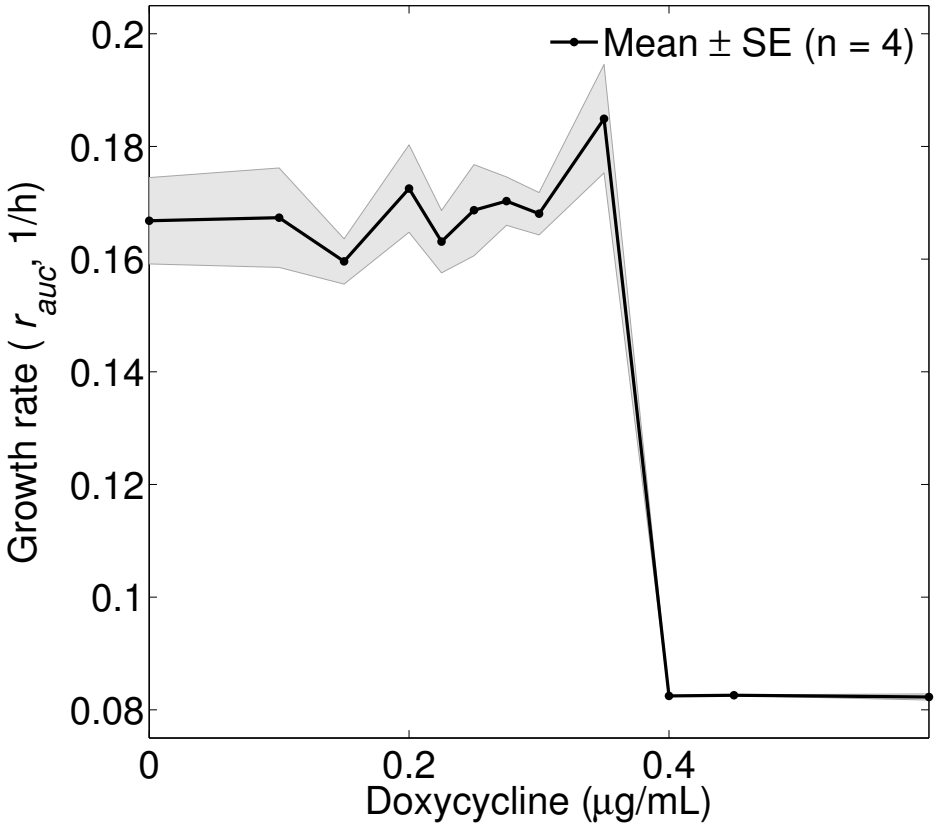


Fig. B.3: Rate decreases as doxycycline concentration increases, however at the dosages where the overgrowth is noticed, the rate improves to a level not significantly above the wild-type. As the drug becomes entirely inhibitory, the rate decreases to zero as expected.

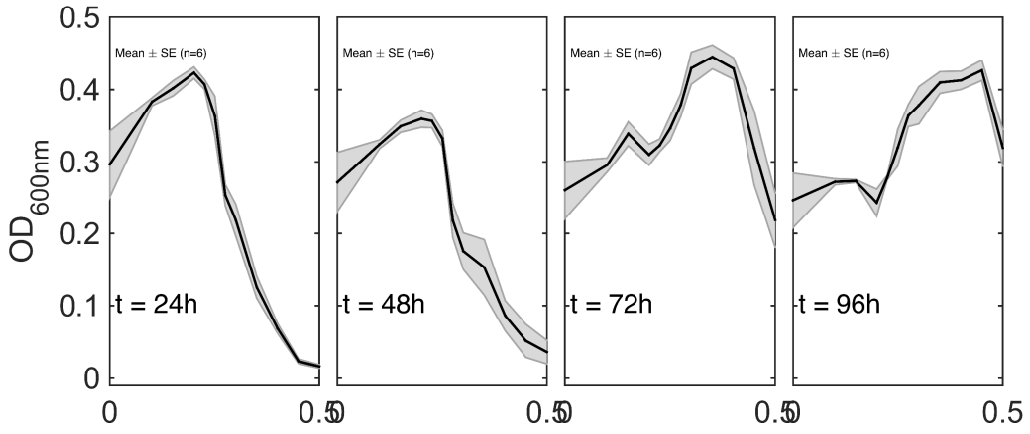


Fig. B.4: Evolution of dose response with 24 hour seasons.

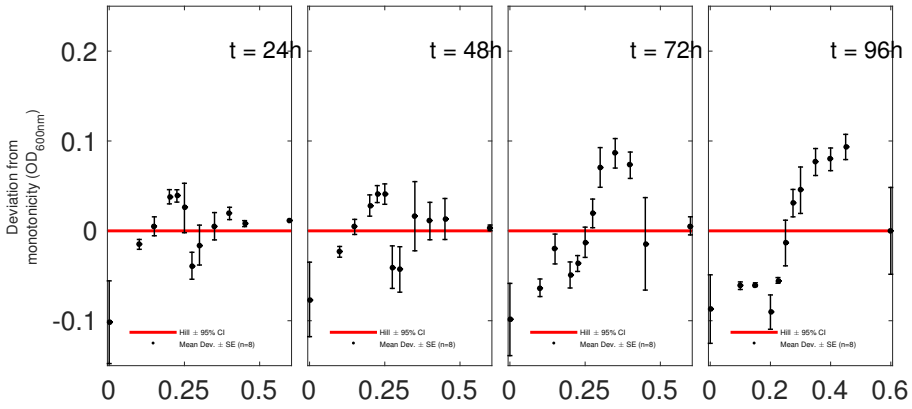


Fig. B.5: Hill fit to evolution of dose response with 24 hour seasons.

No overgrowth in MG1655

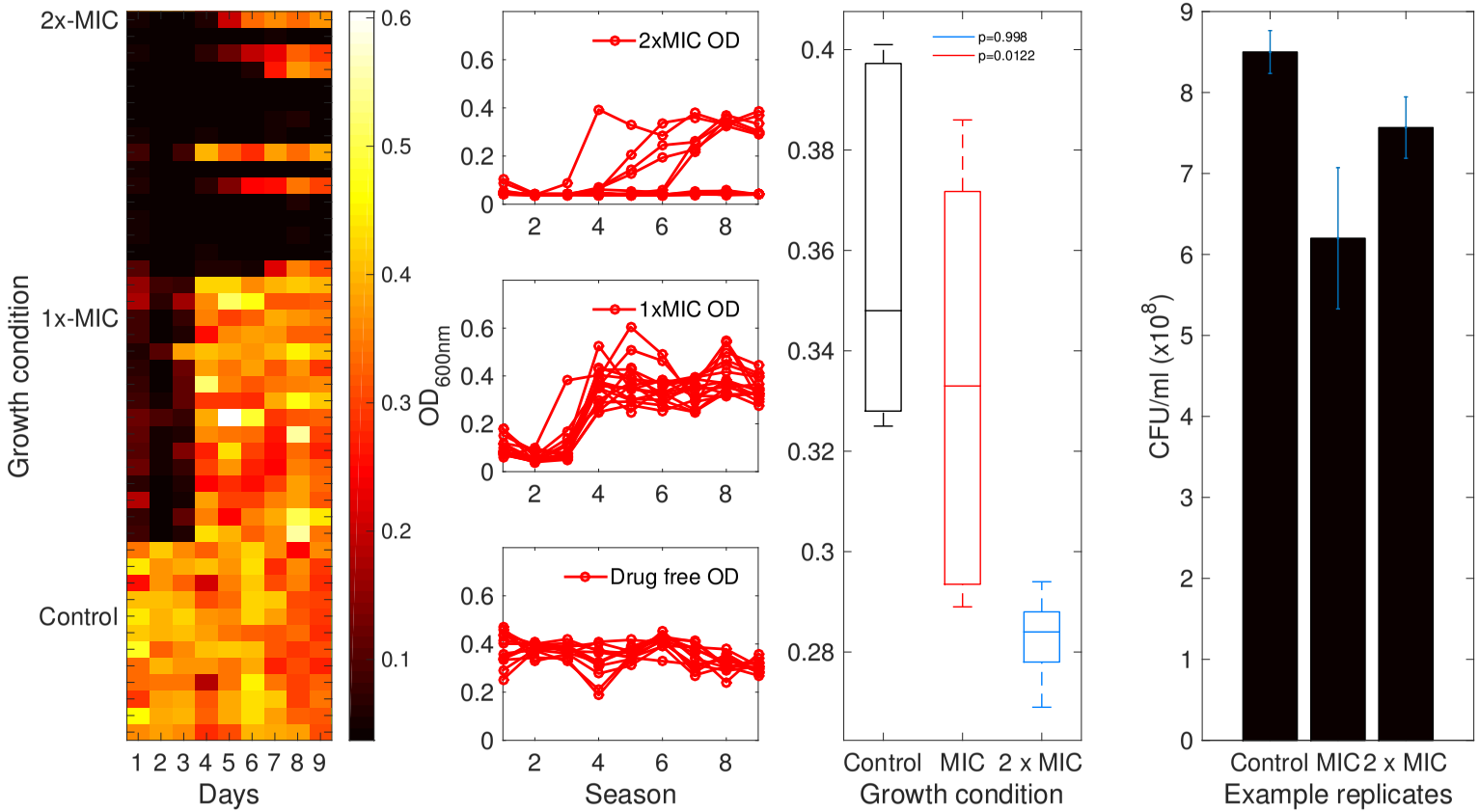


Fig. B.6: Although MG1655 adapts to high dose, up to 2 x MIC in a large number of replicates, the growth is at a level below that of the drug free control.

MG1655 is very closely related to AG100, however the same experimental protocol does not produce an overgrowth of cell density in MG1655. This may provide insight as to the mechanism of overgrowth, by comparing the known genetic differences between these two strains.

Change in rate of adaptation for 24 hour transfers

The same rate of adaptation measurement from [119] is used to confirm that adaptation rate increases in AG100 with 24 hour seasons in the same manner as the transfers that were undertaken with 12 hour seasons. As the shorter seasons should select for rate, with the longer seasons selecting on k, it may be expected that the phenotype that demonstrates a stimulatory effect on cell density will be more apparent in the longer season transfer. Indeed, the overgrowth seen is more substantial in the 24 hour seasons, and appears more quickly. In addition, the growth rate also increases to a greater degree than in the 12 hour transfers, therefore selecting more strongly for k also seems to select an increase in growth rate as well. This is seen later in the rate-yield trade off described, and yet is insufficient to explain the ability of Dox to enable a higher cell density.

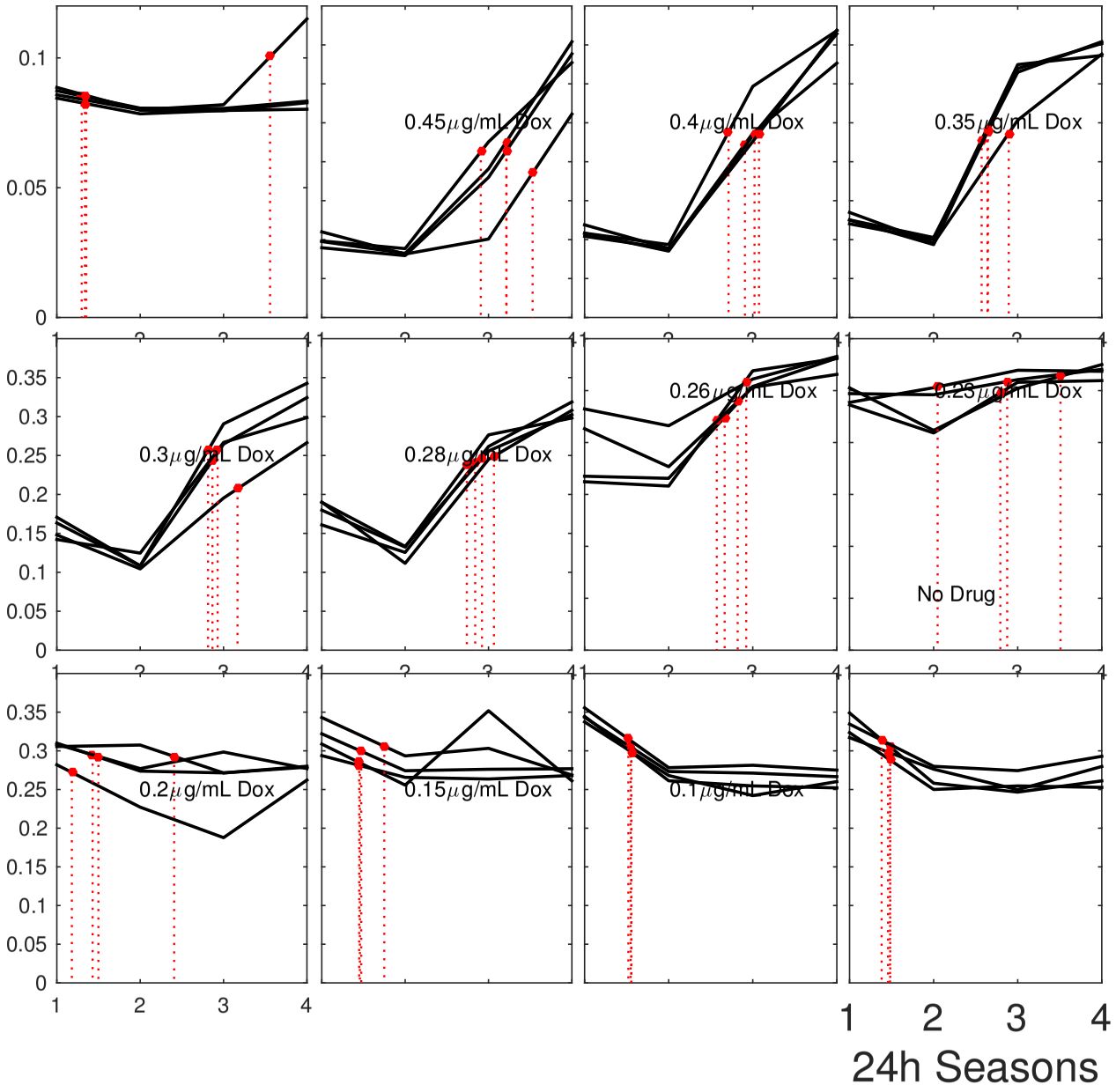
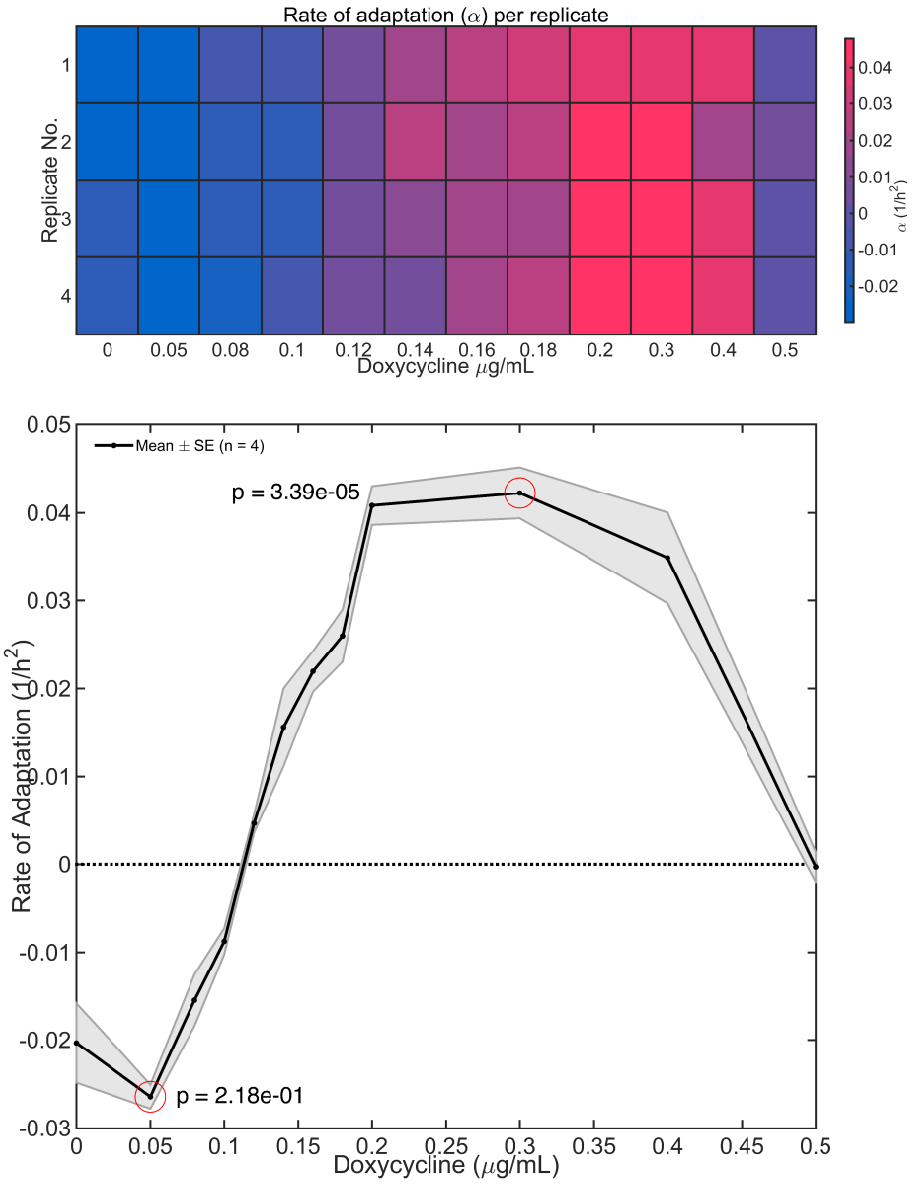


Fig. B.7: All replicates showing the growth rate for each 24 hour season, and the point at which half the maximum change in rate is achieved (red marker). In drug free conditions the adaptation rate is negative as the growth rate slightly declines over the course of the experiment



for 24 hour seasons]The rate of adaptation is again fastest in the 'hotspot' region of Dox. There is no concentration of Dox used that led to a clearing of bacteria.

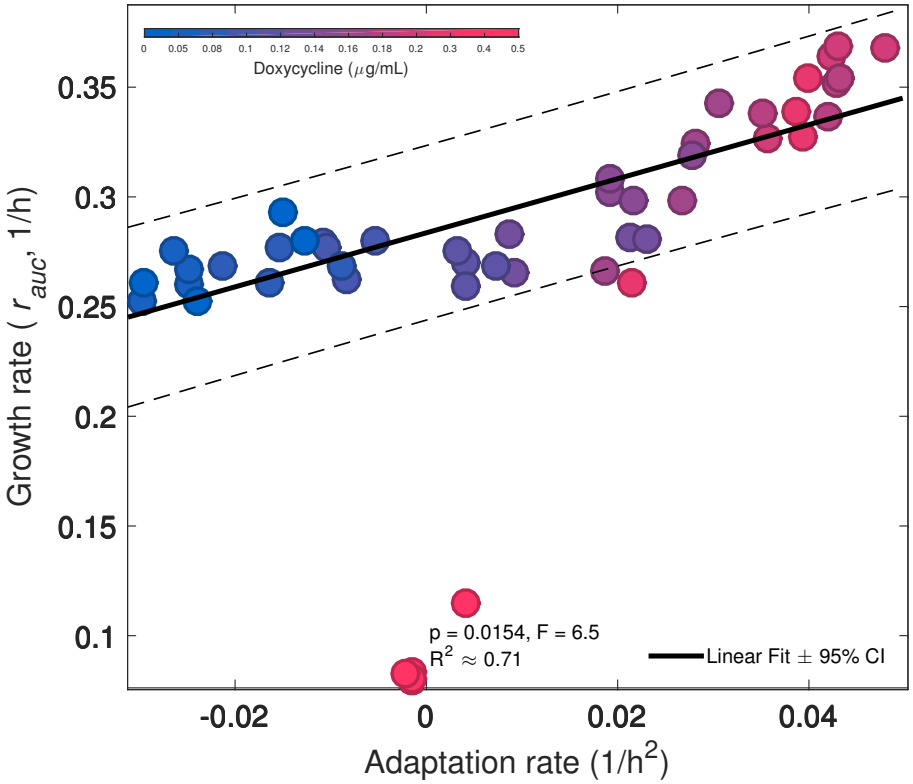


Fig. B.9: As with the twelve hour transfers, there is a weak positive correlation between the slightly increasing growth rate and the rapidly increasing adaptation rate as the [Dox] increases. The exception is the 0.5/ml concentration, whereby growth was still inhibited by $\tilde{8}0\%$, and therefore does not fit the pattern.

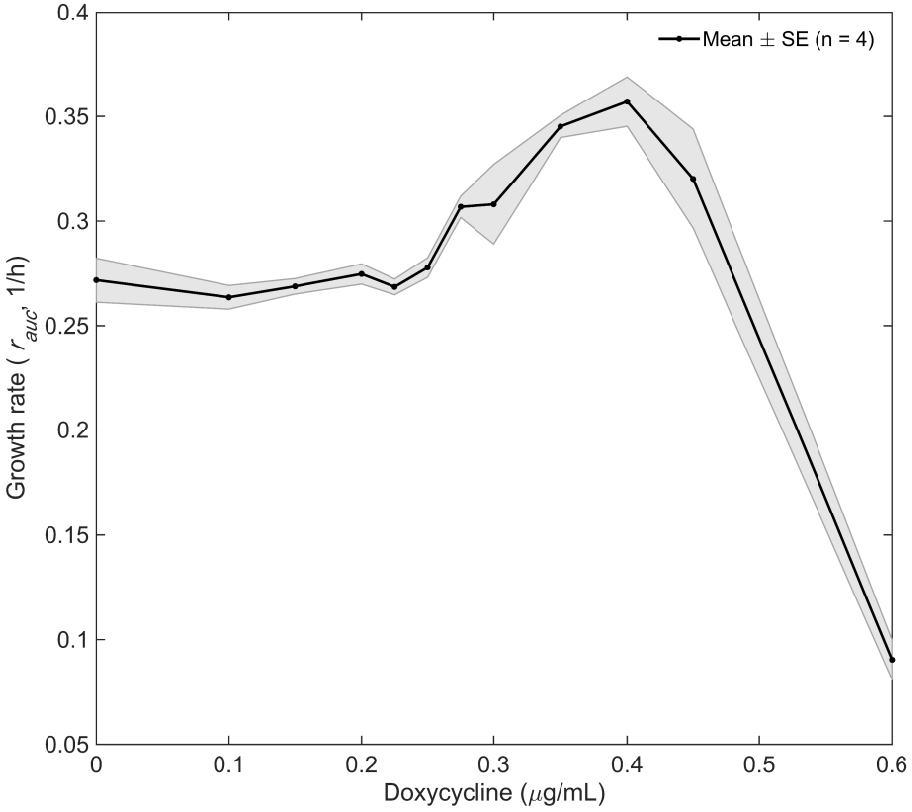


Fig. B.10: Growth rate vs Dox

Efflux is necessary for overgrowth

The *acrAB/tolC* efflux system is a multi-drug efflux pump, implicated in drug resistance in E.coli for a large number of antibiotics, as well as antiseptics and other toxic compounds. This is an innate system, and therefore resistance can be controlled by regulation, without the requirement of a *de novo* protection mutation. The consequence of efflux can be imagined to be effectively decreasing the amount of drug experienced intra-cellularly, therefore it would be expected that the *acrAB* knockout strain AG100 would have a decreased MIC compared to the wild-type. This is indeed the case, with the MIC of 0.15/ml around threefold lower than the MIC for the wild-type (both strains grown for 24 hours seasons). Figure B.11 also demonstrates that as well as the expected decrease in MIC, there is no evidence of overgrowth at any concentration, therefore implying that this efflux system is required in order for cell density to be stimulated by Dox.

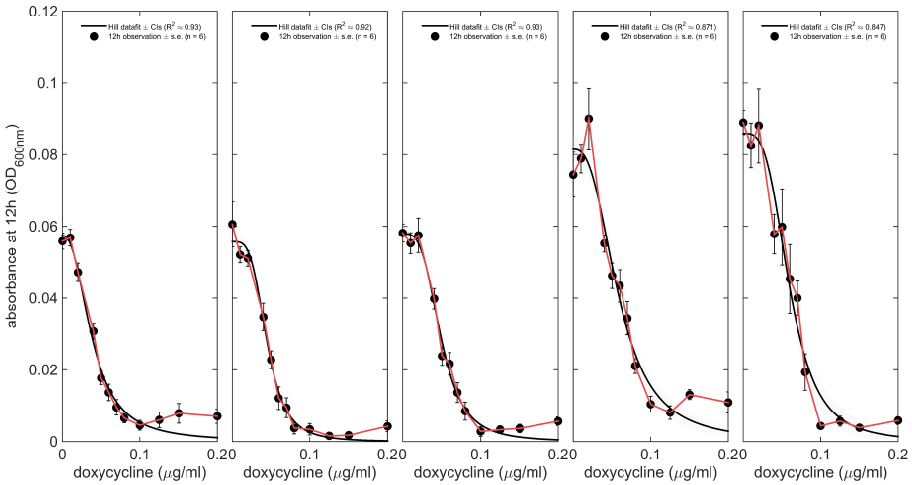


Fig. B.11: Dose-response curves showing the evolution of the *E.coli* Ag100a strain to Doxycycline. Hill function fits every curve, monotonicity of antibiotic inhibition is maintained in the absence of *acrAB/tolC* efflux

No overgrowth in other strains

Figures B.12 and B.13 provide additional evidence that this stimulatory effect of Dox is a strain specific phenomena, with closely related K12 strains not exhibiting the same phenotype. Four strains (MC4100, MG1655, BW25113 and AG100) are compared, with only AG100 showing stimulation, and only MG1655 showing high levels of resistance (at 2xMIC). Figure B.13 shows genetically engineered strains of BW25113, from the Keio collection (provided by the CGSC at Yale). Again, none of these strains demonstrate the phenotype of interest, and therefore unfortunately could not be used to test the hypoth-

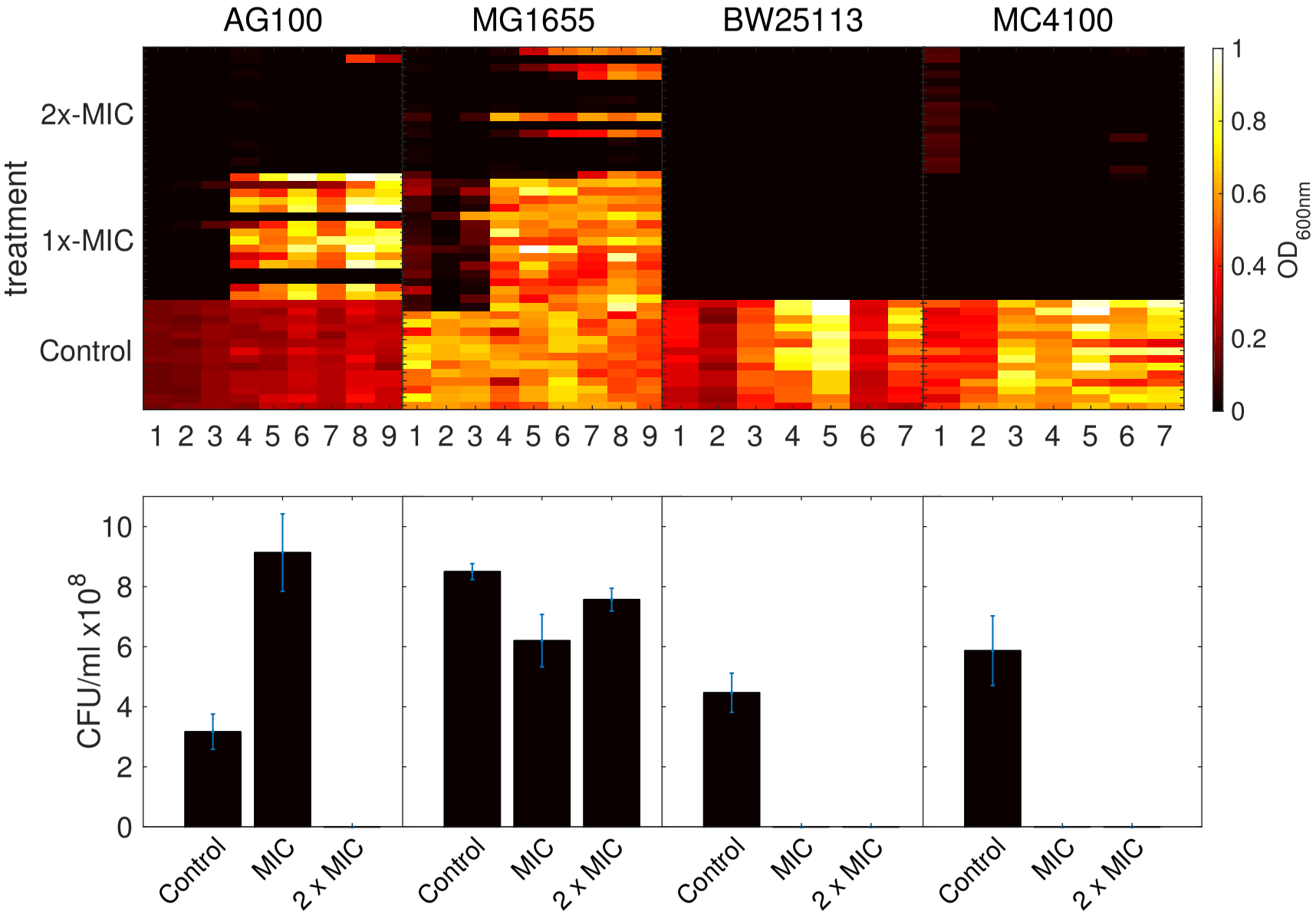


Fig. B.12: Other strains, including Keio parent strain BW25113, do not show overgrowth, or significant adaptation to high doxycyline concentrations

esis that metabolic competency in different pathways could be necessary for stimulation.

No overgrowth in Keio knockouts

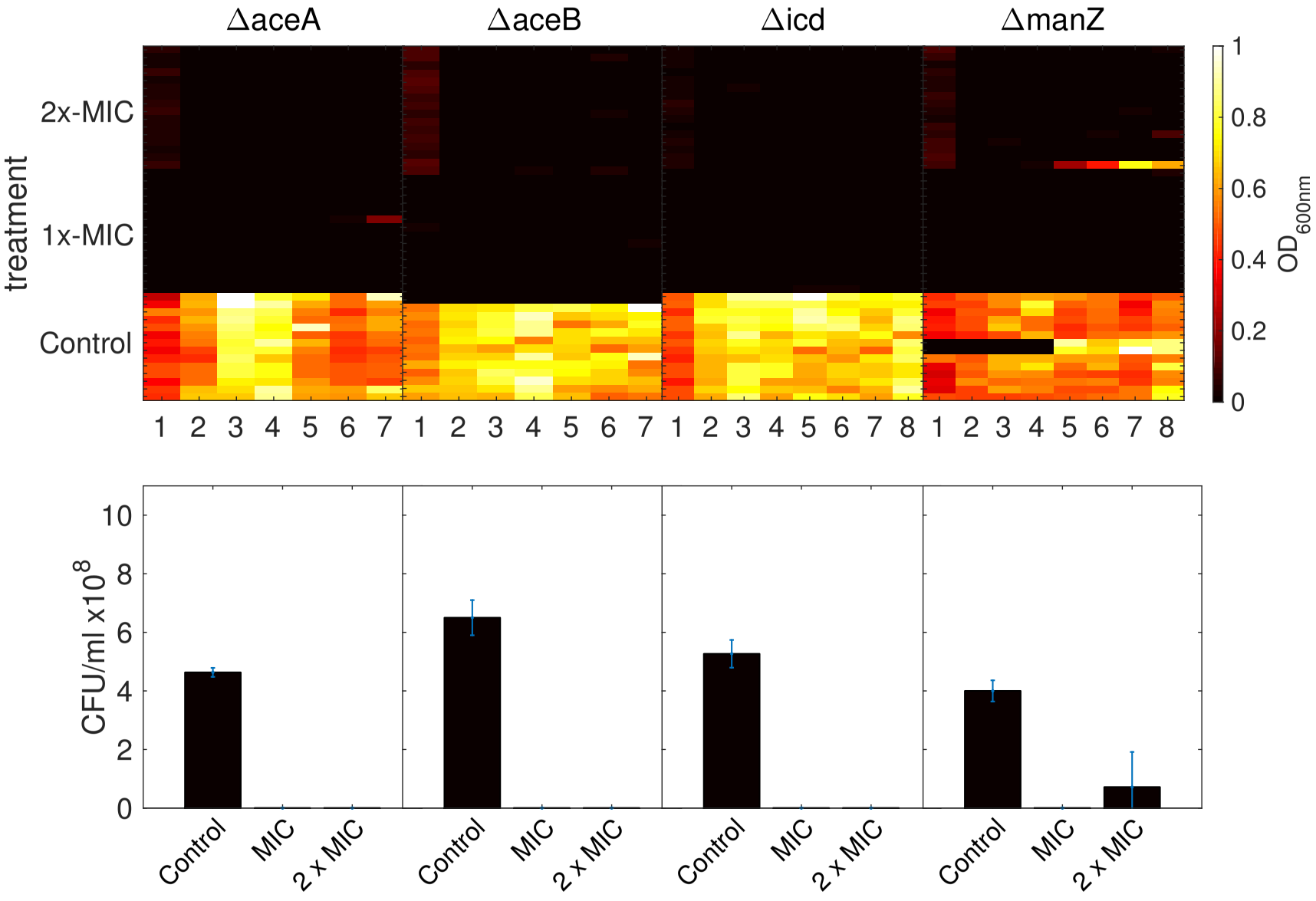


Fig. B.13: No Keio library mutants show overgrowth or high dose adaptation.

Streptomycin resistance does not cause overgrowth in most strains

In order to determine if *rpsL* mutation (the predominant mechanism of Streptomycin resistance) could enable the lag-yield trade off seen in AG100, each strain from the Keio collection was subject to four days of increasing Strep pressure, from 2xMIC to 128xMIC. Rather than adaptation in liquid media, cells were grown on solid agar impregnated with the appropriate concentration of Streptomycin. By inoculating the agar with cells at a high density, the large inoculum size allowed a higher probability of inoculating a resistant cell line. This enabled a high level of Strep resistance to emerge rapidly. None of the strains so treated exhibited the increased lag leading to a higher yield over 48 hours of growth in Dox (Figure B.14), however one strain (ΔmanZ) did develop such an overgrowth within 3 seasons, when adapted to Dox.

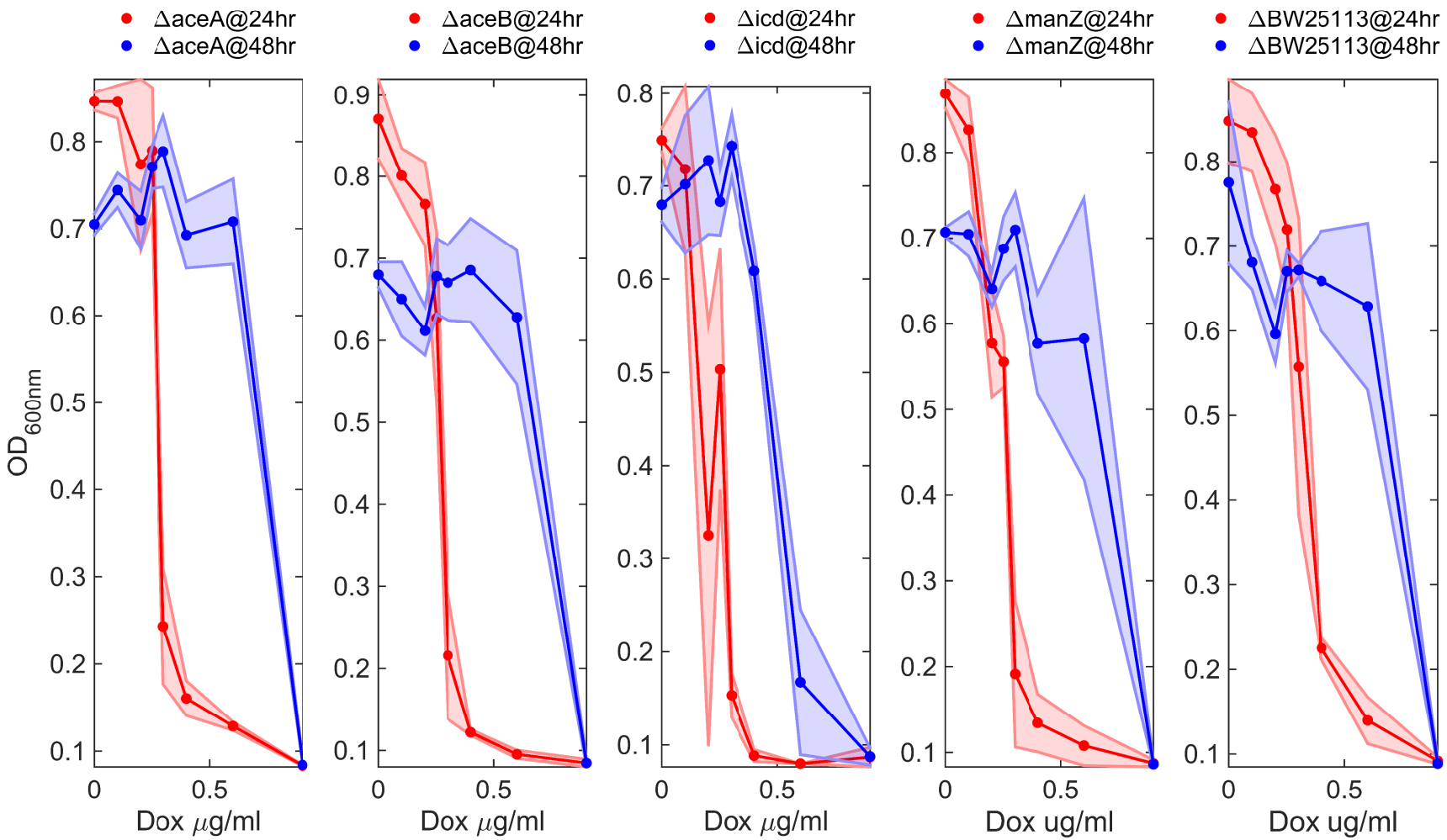


Fig. B.14: No other strain shows evidence of a trade of between lag and yield, although it can be seen that the dose response profile changes significantly if a 48 hour response period is used, allowing some drug containing environments to overgrow the drug free control, in contrast to what is seen at 24 hours.

ΔmanZ$_{strpR}$ does not exhibit a lag yield trade-off

Although ΔmanZ$_{strp+}$ does exhibit an increase in cell density comparable to that seen in AG100, as Figure 3.21 shows, there is not a significant increase in lag time. Potentially the longer lag phase is not required for the cells to up-regulate TCA cycle genes, as ΔmanZ is already glycolysis limited, and therefore is more likely to be utilising alternative metabolic pathways already.

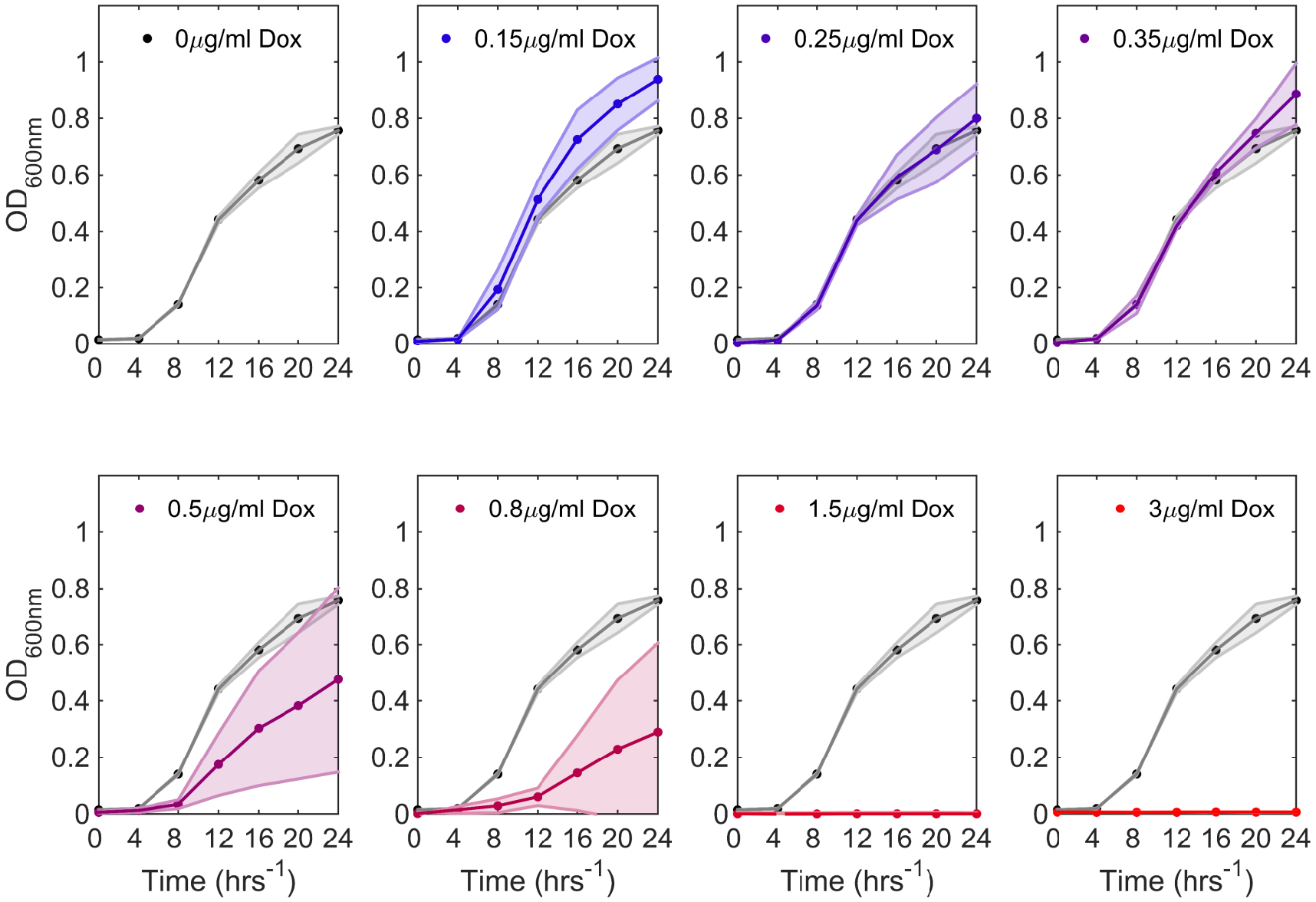


Fig. B.15: No other strain shows evidence of a trade of between lag and yield, although it can be seen that the dose response profile changes significantly if a 48 hour response period is used, allowing some drug containing environments to overgrow the drug free control, in contrast to what is seen at 24 hours.

Rate-yield and Lag-yield trade off raw data

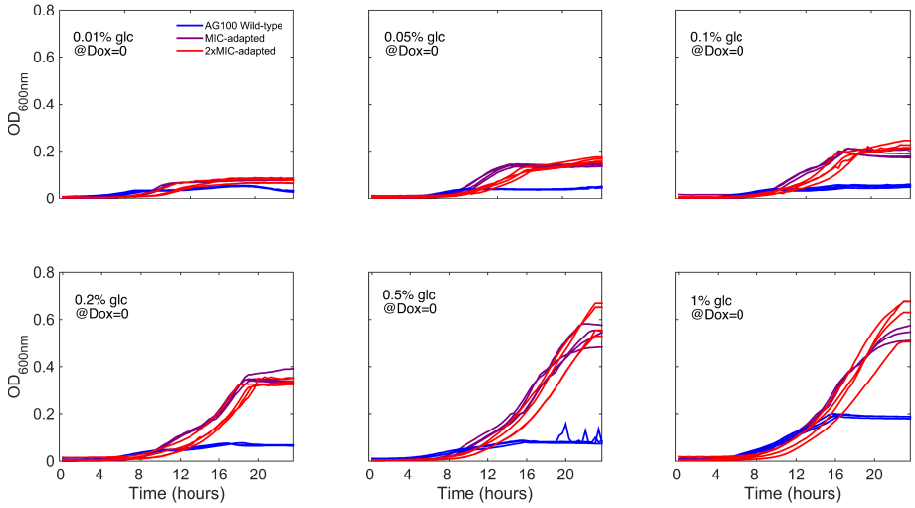


Fig. B.16: ⌈AG100 displays LYTO]E.coli AG100 strains adapted for one week in conditions of either zero drug, MIC or twice MIC were grown in drug free media. Both antibiotic-adapted strains outperform the wild-type bacteria, the growth curves showing a clear lag-yield trade off, as well suggesting that there is not a growth rate cost for resistance in this case.

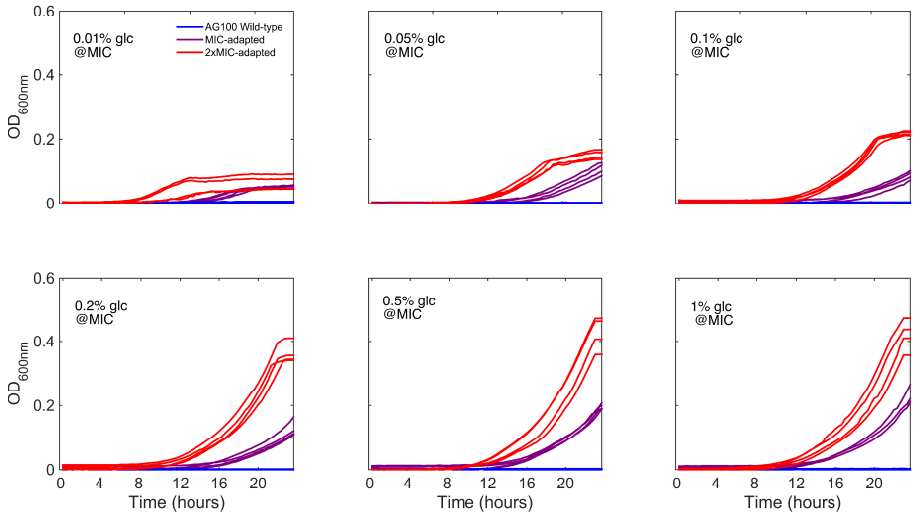


Fig. B.17: Strains adapted to MIC and 2xMIC show growth when in media supplemented with MIC concentration of Dox. Although growth occurs, and occurs to a higher level than wild-type bacteria in drug-free media, growth is not as high as when the drug stress is removed.

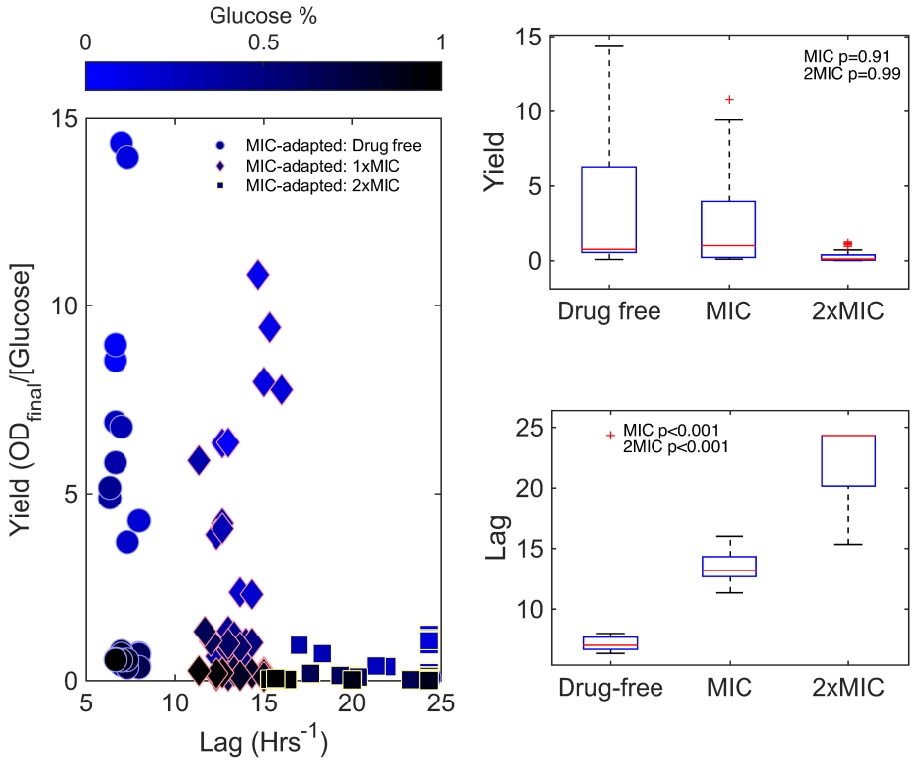


Fig. B.18: This plot shows the comparison of yield and lag for AG100 populations adapted to 1xMIC [Dox] and subsequently grown at three concentrations of the drug. As expected there is minimal growth at the Dox concentration that is higher than the adaptive background of the population, but what is very clear is the dramatic increase of yield that the population shows when the drug is removed. Although the lag decreases when drug is removed, it does not decrease to wild-type levels, so although adaption can mitigate the strength of the lag-yield trade off, it can not remove it or invert it to a trade up.

Lag-yield data for MG1655

Where can trade-off's be found?

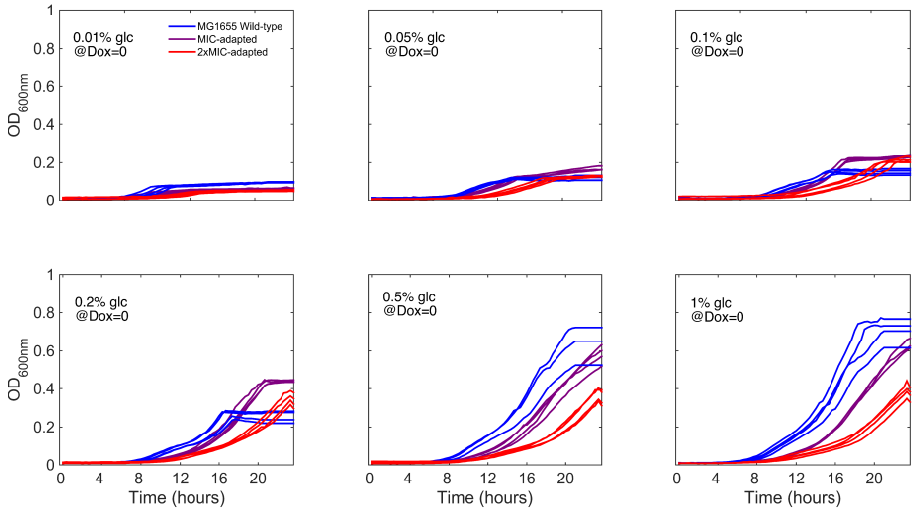


Fig. B.19: Using the same system as above, MG1655 wild-type and adapted strains were grown in drug free media. The drug adapted strains show signs of outcompeting the wild-type only in glucose concentrations of 0.1% and 0.2%, elsewhere the wild-type strain has the highest fitness in this media.

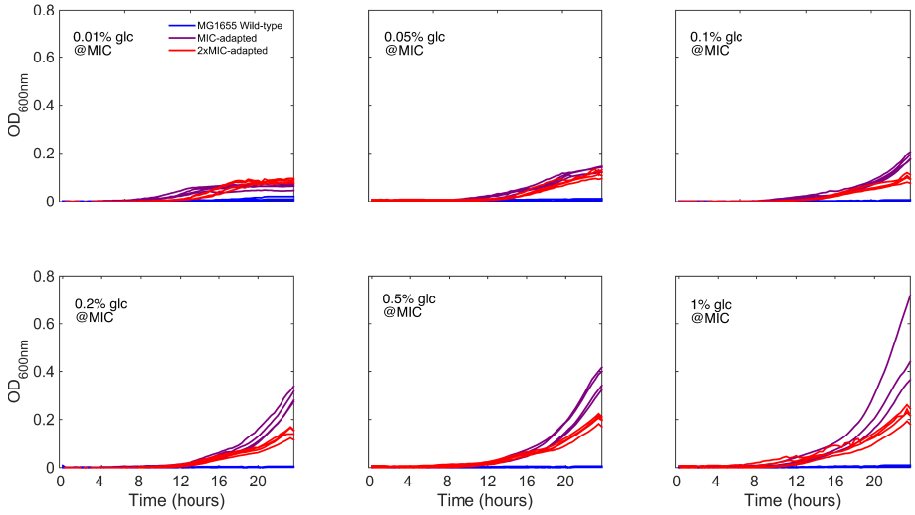


Fig. B.20: As expected, in media containing Dox at the level of the MIC, the wild-type strain fails to grow, and the drug adapted strains show levels of growth, with the MIC-adapted strain being most fit. OD does not reach the same density as wild-type growth in zero drug.

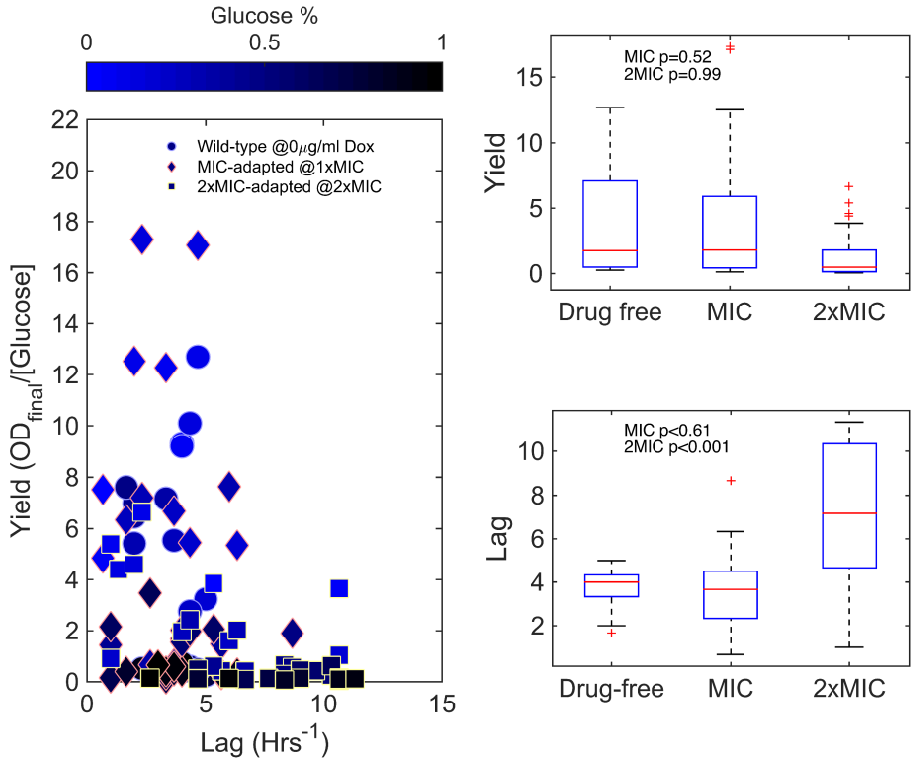


Fig. B.21: The yield is lower for all drug adapted MG1655 populations than the control

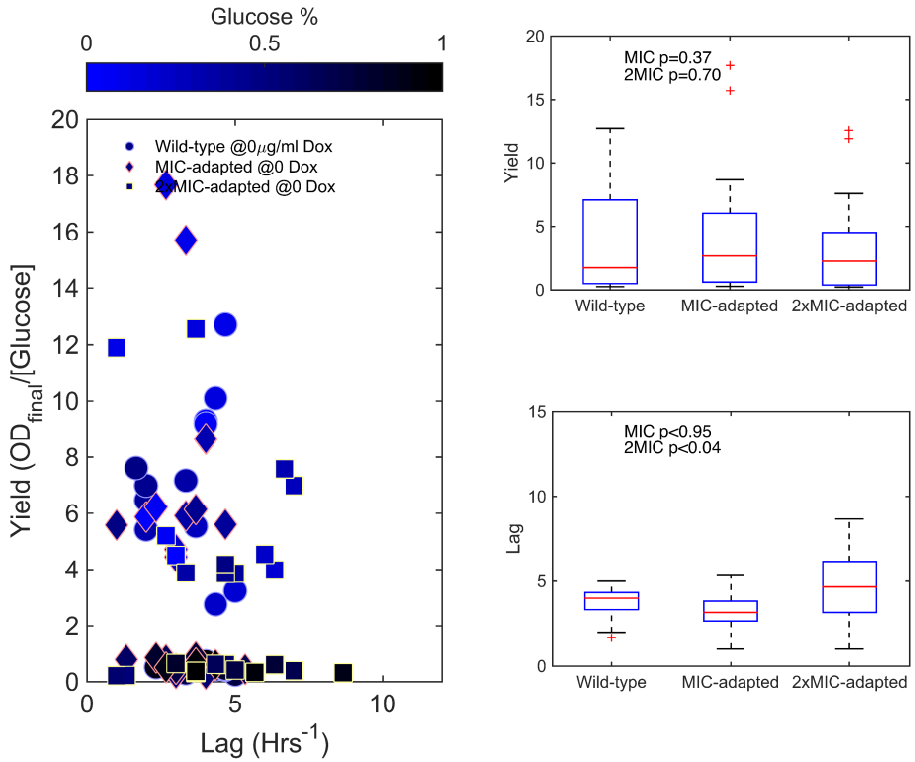


Fig. B.22: There is no LYTO present in MG1655 when drug challenge is removed

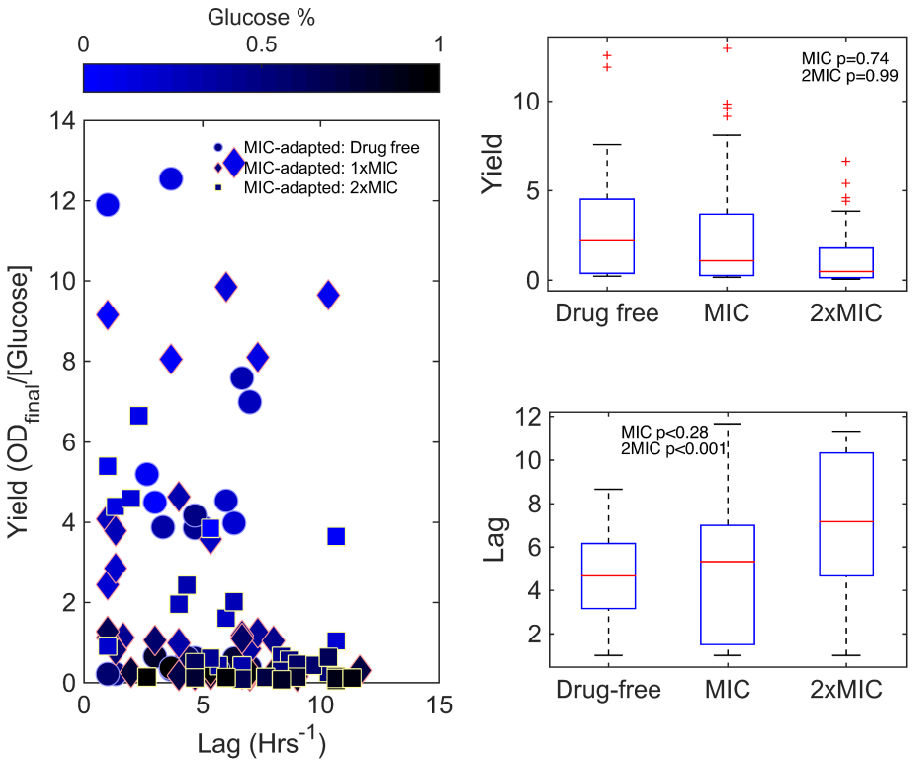


Fig. B.23: No LYTO present in MG1655 in Dox

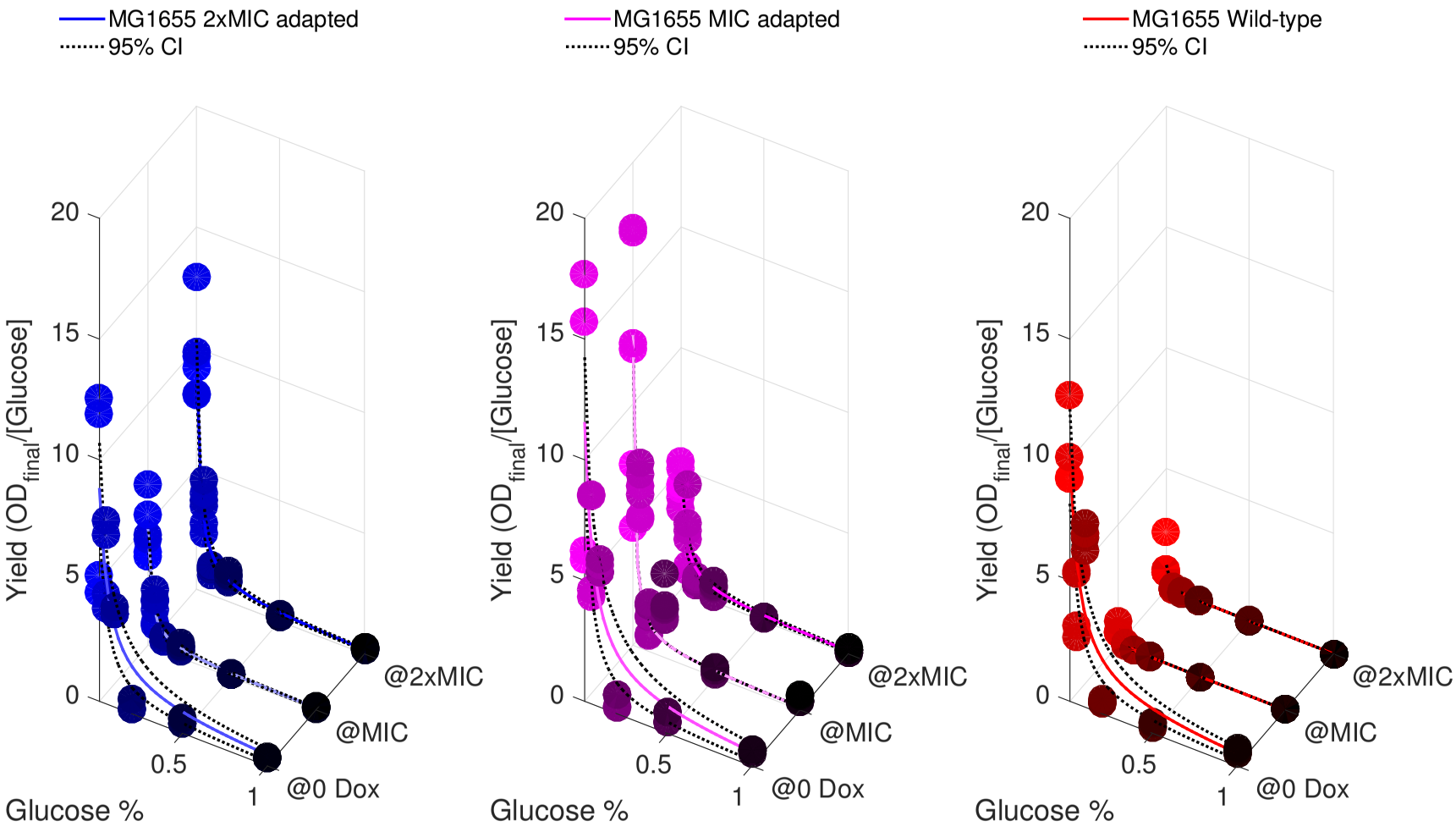


Fig. B.24: RYTO also present in MG1655

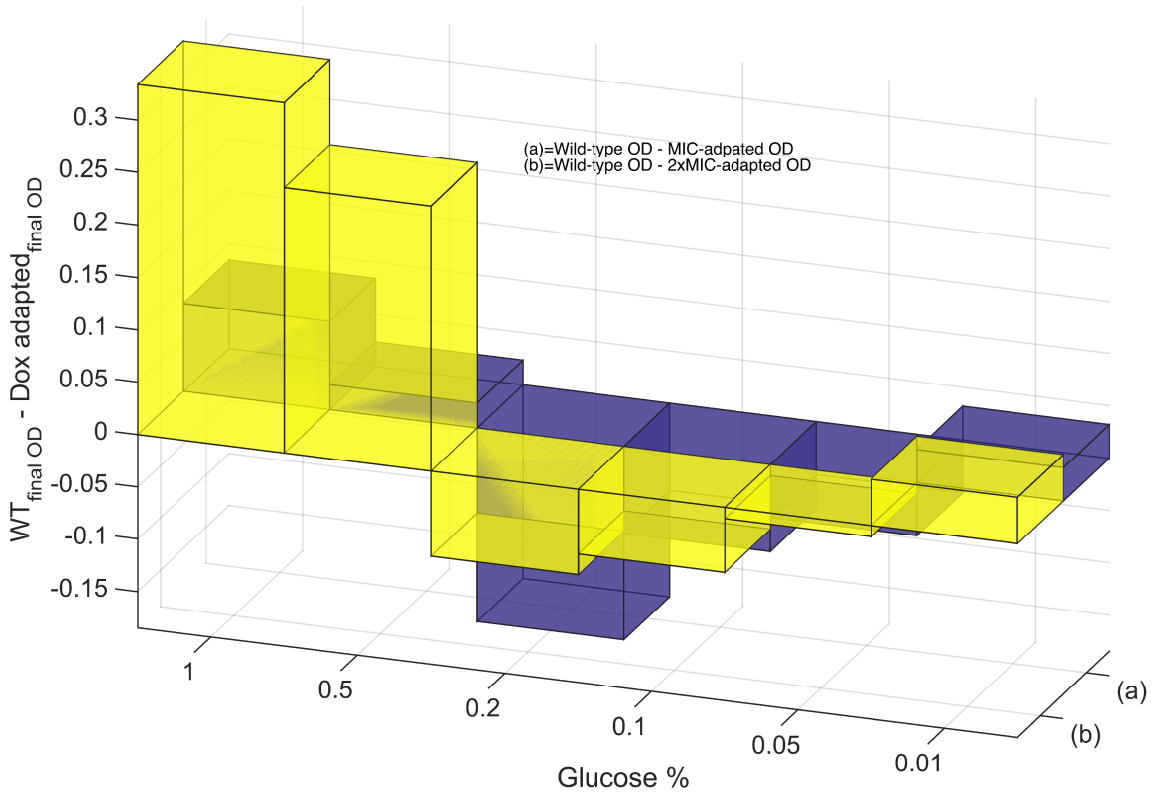


Fig. B.25: Although MG1655 did not show overgrowth during the initial dose response testing, the adapted strains do have a higher than wild-type growth (as measured by OD) at certain low glucose concentrations

Appendix C

Chapter 4: Appendix

C.1 Material and Methods

DNA Extraction: DNA was extracted from frozen samples (150μl culture + 75μ/l 80% glycerol) using an Ambion gene jet DNA extraction kit. Sample was extracted from populations taken from season 8 of the adaptive experiment, results shown in Chapter 2, Fig 3.8. A small quantity (1μl) was removed from the frozen population using a sterile tip, and inoculated into M9 media containing the same concentration of nutrients and antibiotic as the adaptive condition, therefore recapitulating the conditions of season 9. The only difference between the experimental condition and the condition used for growing an overnight culture for DNA extraction was that a larger volume (5ml) was used for DNA extraction (in 15ml Falcon tubes). As well a season 8 populations, a sample was taken from the ancestral AG100 and MG1655 frozen populations to act as a control. 18 samples were extracted in total, as shown in Table 4.1. The kit extraction protocol was slightly modified to increase yield, with a larger sample (5ml) being used, and centrifuged at 4°C for 10 minutes. The supernatant was discarded and the pellet was resuspended in the lysis buffer as per the instructions. After DNA binding to the silica column, an additional wash step was used, with a 2 minute rinse in 80% ethanol, in order to remove all guanidine salts. The binding tubes were centrifuged whilst open for 1 minute following this wash to ensure that all ethanol was removed, as this would stop the elution of DNA during the final step of the extraction. To further maximise yield, the wash buffer was heated to 45°C and left to soak the column for 5 minutes before elution. DNA quantitation was performed initially on a NanoDrop spectrophotometer, which also gave an indication of contamination with protein or organic salts. DNA was also run on a 1% agarose gel to ensure that DNA size was in excess of 10,000bp and to act as a secondary check to ensure that there was no protein contamination (which would show

up as smearing on the gel). Finally DNA was accurately quantitated using the Qubit system, ensuring yields of over 25ng/μl in 50μl. Samples with sufficient DNA in excess washing buffer were spun on a vacuum centrifuge at 45°C until volume was reduced by half, to increase the concentration of DNA.

Paired-end DNA libraries were prepared by Exeter Sequencing Service, using the Nextera (Illumina) library preparation protocol [14]. Although three biological replicates were sequenced for each condition, only 2 replicates for the M9-adapted AG100 population were used due to an error in the library preparation stage. The purpose of sequencing both ancestral and M9-adapted cell lines was to ensure that any adaptation observed by whole genome re-sequencing was due to adaptation to Doxycycline, and not adaptation to minimal media, or random genetic drift.

DNA Quality Control

After extraction of genomic DNA, quality control checks were carried out. Bioinformatics strongly follows the principle of 'Rubbish in, rubbish out', in terms of the quality of genetic material will have a large effect on the reliability of the final DNA sequence. Therefore quality control steps are vital at every point of the process, from the DNA extraction, library prep and through to bioinformatics. The first checks which were carried out was quantification of DNA volume, using NanoDrop (Figure C.1(b)) and Qubit systems. The NanoDrop can also give information about possible contaminants, but can be unreliable at low concentrations of DNA. For this reason, the preferred pre-submission quality control step of the Exeter Sequencing Centre is to run the DNA on a 1% agarose gel to check for contamination, degradation and sufficient fragment length (Figure C.1(a)).

DNA QC

Sequence data quality control

Once the sequence data was returned from ESS, some quality control steps are also necessary to remove bias inherent in the sequencing system. Reads are scored for quality, with quality tending to diminish towards the end of the read.

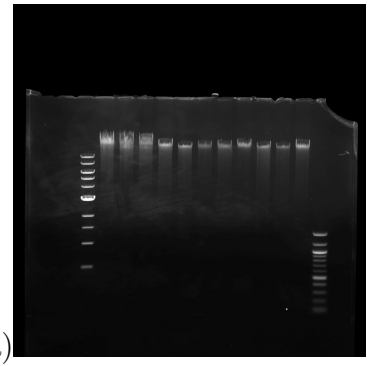
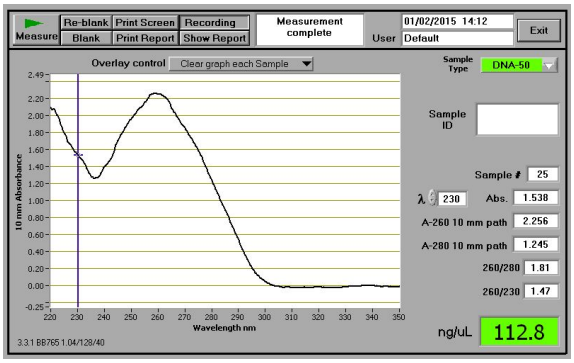
a)  b) 

Fig. C.1: a) Minimal smearing and bright bands suggest an un-degraded sample free from contaminants. On the left hand side is shown at 10kb ladder, so fragment size is confirmed as greater than 10kb. b) NanoDrop spectrophotometry shows absorbance at a range of wavelengths.

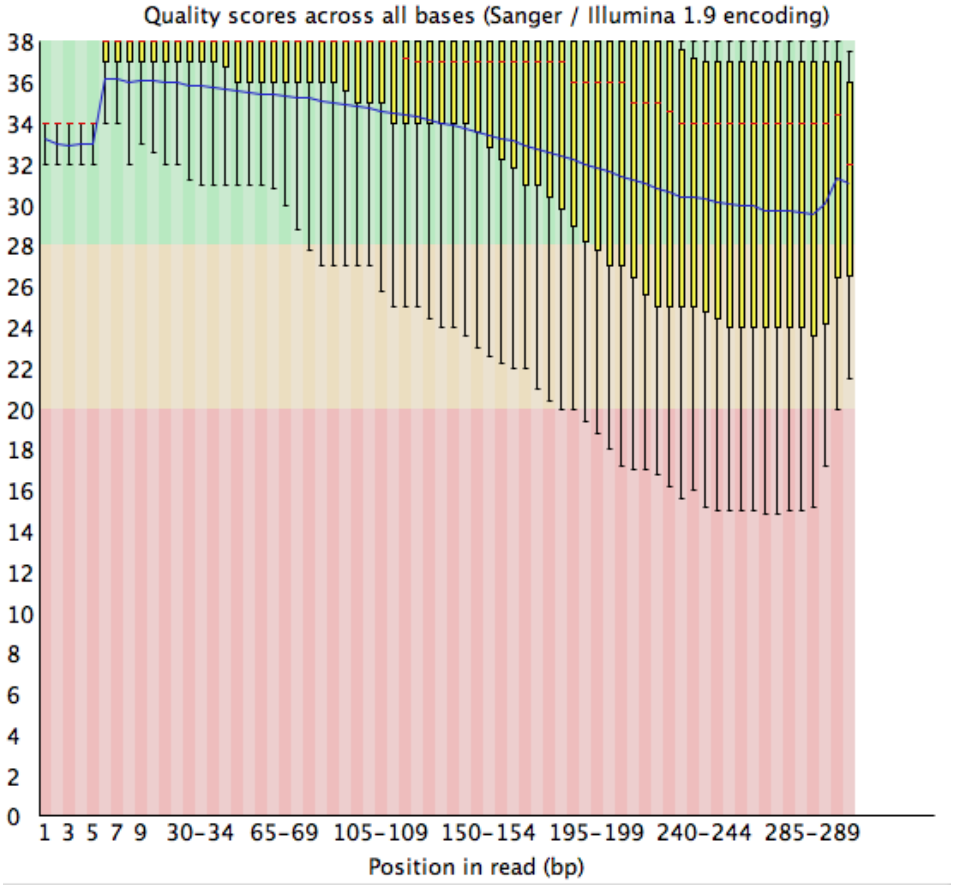


Fig. C.2: FASTQC trimmed read file, showing Q score of ¿25 for all position of the read. Representative image of data quality.

Bioinformatic pipeline

[Bioinformatic pipeline]Schematic representation of the final pipeline used to analyse AG100 and MG1655 whole genome data. Image adapted from www.ebi.ac.uk

Quality Control

Exeter sequencing service performed some initial quality assessment, as well as adapter trimming, but otherwise the sequence data was raw when delivered. All reads were quality controlled before further analysis. FastQC was used to perform read trimming, removing read ends to ensure that quality across the whole read length was in excess of a Phred quality score of 20 (although the end of some reads was as low as 20, the mean across the reads was much higher, with most read positions with a quality above 30.

Scaffold plots

Contigs were assembled and aligned against the previously published genomes for AG100 and MG1655 respectively. Each sequenced population mapped well against the references, with a maximum of 8 contigs required for assembly (Figure C.3).

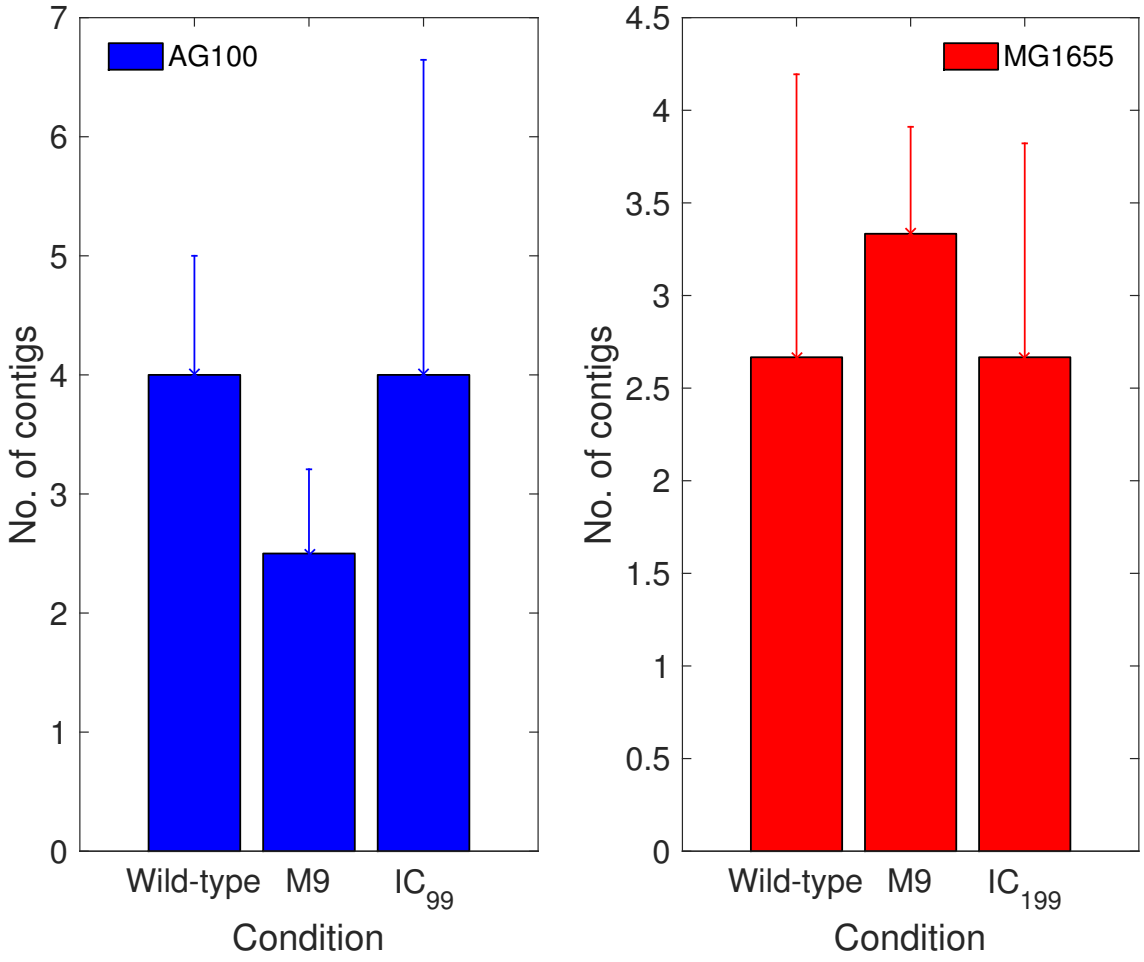


Fig. C.3: Contig numbers

Alignment

All sample reads were mapped to the previously published E. coli K12 (AG100) reference genome [95]. Each sequence was indexed to the reference in fasta format, using the BWA (Burrows-Wheeler alignment) index subroutine (BWA 0.7.4). Alignment was also performed using BWA, and aligned files were subsequently sorted to genomic position and indexed using Samtools and filtered to remove any non-paired end matches [73]. Coverage was analysed using genomeCoverageBed in Bedtools [98], and data was further analysed in Matlab.

Variant calling

Single Nucleotide Polymorphisms (SNPs) were called using VarScan 2.3.8. VarScan uses a heuristic/statistic method to based on read depth, base quality and variant frequency in order to call SNPs. Filters were added to SNP calling to ensure that high quality calls were made. These quality filters were; minimum read depth of 60, with at least 10 reads supporting the alternative variant, a P-value of 0.05 or less and a minimum average quality of 30. Additionally VarScan calls SNPs as homologous if over 90% of the reads support one or the other variant. VarScan also detected short indels in the same manner. Pindel was used to detect structural variations, including inversions, rearrangements and indels. Pindel works by detecting breakpoints using a pattern-growth algorithm. This identities paired reads whereby only one of the reads is able to map to the reference. By breaking the unmapped read into shorter fragments, and re-mapping, the breakpoint of insertions or deletions can be detected. Duplication events were determined using CNVnator v0.3 [20], which analyses mapping density statistically to discover copy number variants. The genome must be split into 'bins' for CNVnator to compare, the bin size was left as the default of 100 after calibration. Although the high mean coverage meant that a smaller bin size (and therefore a more accurate copy number variance for a given genome position) could be used, CNVnator documentation suggests that the bin size not be smaller than the read length.

Annotation

Both genomes sequenced have been annotated previously and these annotation files were used in annotation of wild-type populations of AG100 and MG165. These files were accessed from EBI and ENSEMBLE respectively [40], [41]. Annotation of the evolved genomes was also performed using RAST for initial web-based viewing, and PROKKA to annotate the consensus fasta files of all experimental replicates [9], [89], [19], [111].

Table SC.1: Phage deletion genes - $AG100_{Dox}$

Start	Stop	Length (bp)	Function	Subsystems
806940	806392	549	Phage protein	- none -
807439	807257	183	DNA-damage-inducible protein we	DNA repair
807663	807550	114	Phage protein	- none -
808573	807893	681	Phage exonuclease (EC 3.1.11.3);	- none -
809355	808570	786	DNA repair protein RecT (prophage associated)	DNA repair
809657	809361	297	Mobile element protein	- none -
809875	809732	144	Kil protein	- none -
810449	810081	369	Single stranded DNA-binding protein	phage-associated
810754	810632	123	antirestriction protein	- none -
811905	811582	324	Phage antitermination protein N	Phage nin genes
813659	812820	840	hypothetical protein	- none -
814485	813772	714	Phage repressor	- none -
814586	814786	201	Regulatory protein cro	- none -
815061	815198	138	Phage repressor	- none -
815231	816130	900	Origin specific replication initiation factor	Phage replication
816127	816828	702	Replication protein P	- none -
816825	817115	291	Ren protein	- none -
817189	817629	441	Phage NinB DNA recombination	Phage nin genes
817608	818498	891	Phage NinC	Phage nin genes
818814	818984	171	Phage NinF	Phage nin genes
819391	819588	198	Phage NinG rap recombination	Phage nin genes
819585	819791	207	Phage NinH	Phage nin genes
819769	820434	666	Phage NinI serine/threonine phosphatase	- none -
820431	821054	624	Phage antitermination protein Q	Phage nin genes
821731	822054	324	Phage holin	Phage lysis modules
822038	822514	477	Phage lysin	1
822511	824610	2100	Phage terminase	large subunit
824607	824813	207	Phage head-tail joining protein	Phage capsid proteins
824810	826411	1602	Phage portal protein	Phage packaging machinery
826392	827711	1320	Phage capsid and scaffold	Phage capsid proteins
827721	828053	333	Head decoration protein	- none -
828109	829134	1026	Phage major capsid protein	Phage capsid proteins
829176	829574	399	Phage capsid and scaffold	Phage capsid proteins
829586	829939	354	Phage capsid and scaffold	Phage capsid proteins
829951	830529	579	Phage tail completion protein	Phage tail proteins
830526	830921	396	Phage minor tail protein	Phage tail proteins
830950	831669	720	Phage tail assembly	Phage tail proteins
831685	832107	423	Phage minor tail protein	Phage tail proteins
832089	832523	435	Phage minor tail protein	Phage tail proteins
832516	835077	2562	Phage tail length tape-measure protein 1	Phage tail proteins
835074	835403	330	Phage minor tail protein #tail protein M	Phage tail proteins
835403	836101	699	Phage minor tail protein	Phage tail proteins
836251	836850	600	Phage tail assembly protein	Phage tail proteins
836847	837419	573	Phage tail assembly protein we	Phage tail proteins

C.2 Supplementary Data

Prophage genes in deleted region

Genes in duplicated region

Table SC.2: Genes in Duplicated region

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
553910	555295	1386	Cysteinyl-tRNA synthetase	Zinc regulated enzymes
284787	286169	1383	Xyloside transporter XynT	Xylose utilization
286181	287791	1611	Beta-xylosidase	Xylose utilization
511875	513167	1293	Inner membrane transport protein YbaT	Uptake of selenate and selenite
638808	638380	429	Universal stress protein G	Universal stress protein family
557511	558053	543	type 1 fimbriae major subunit FimA	Type 1 pili (mannose-sensitive fimbriae
655888	656091	204	Twin-arginine translocation protein TatE	Twin-arginine translocation system
506381	505902	480	Cys-tRNA(Pro) deacylase YbaK	tRNA aminoacylation
640908	640498	411	Regulator of nucleoside diphosphate kinase	Transcription factors bacterial
507517	507858	342	HigA protein (antitoxin to HigB)	Toxin-antitoxin replicon stabilization systems
463701	464099	399	4-hydroxybenzoyl-CoA thioesterase family active site	Ton and Tol transport systems
635886	636449	564	Alkyl hydroperoxide reductase protein C	Thioredoxin-disulfide reductase
636694	638259	1566	Alkyl hydroperoxide reductase protein F	Thioredoxin-disulfide reductase
558273	558965	693	Fimbrial periplasmic chaperone SfmC	The fimbrial Sfm cluster
558996	561599	2604	Outer membrane usher protein SfmD	The fimbrial Sfm cluster
561641	562618	978	Fimbriae-like adhesin SfmH	The fimbrial Sfm cluster
562947	563144	198	Fimbriae-like periplasmic protein SfmF	The fimbrial Sfm cluster
563779	563147	633	Transcriptional regulator of fimbriae expression FimZ (LuxR/UhpA family)	The fimbrial Sfm cluster
447291	447016	276	Cytochrome O ubiquinol oxidase subunit IV	Terminal cytochrome O ubiquinol oxidase;

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
447959	447345	615	Cytochrome O ubiquinol oxidase subunit III	Terminal cytochrome O ubiquinol oxidase;
449940	447949	1992	Cytochrome O ubiquinol oxidase subunit we	Terminal cytochrome O ubiquinol oxidase;
450849	449962	888	Cytochrome O ubiquinol oxidase subunit II	Terminal cytochrome O ubiquinol oxidase;
384531	385493	963	Taurine-binding periplasmic protein TauA	Taurine Utilization
385506	386273	768	Taurine transport ATP-binding protein TauB	Taurine Utilization
386270	387097	828	Taurine transport system permease protein TauC	Taurine Utilization
609435	607195	2241	TonB-dependent receptor; Outer membrane receptor for ferric enterobactin and colicins B	D
607029	606400	630	Phosphopantetheinyl transferase	Siderophore Enterobactin
609678	610880	1203	Enterobactin esterase	Siderophore Enterobactin
611098	614979	3882	Enterobactin synthetase component F	serine activating enzyme
615195	616328	1134	Ferric enterobactin uptake protein FepE	Siderophore Enterobactin
617140	616325	816	Ferric enterobactin transport ATP-binding protein FepC	Siderophore Enterobactin
619241	620491	1251	Enterobactin exporter EntS	Siderophore Enterobactin
621451	620495	957	Ferric enterobactin-binding periplasmic protein FepB	Siderophore Enterobactin
623011	624621	1611	2	3-dihydroxybenzoate-AMP ligase: siderophore
626241	626654	414	Proofreading thioesterase in enterobactin biosynthesis EntH	Siderophore Enterobactin
530525	529431	1095	Selenophosphate-dependent tRNA 2-selenouridine synthase	Selenocysteine metabolism; tRNA modification
641944	641138	807	Ribonuclease we precursor	RNA processing and degradation

Table SC.2 – *Continued from previous page*

Start	Stop	Length (bp)	Annotation	Subsystem
665656	665189	468	LSU m3Psi1915 methyltransferase RlmH	RNA methylation; Ribosome biogenesis; tRNA modification
312172	311906	267	LSU ribosomal protein L31p @ LSU ribosomal protein L31p	zinc-independent
311906	311766	141	LSU ribosomal protein L36p	Ribosome LSU bacterial
432355	432750	396	Ribonucleotide reductase transcriptional regulator NrdR	Ribonucleotide reduction
432754	433857	1104	Diaminohydroxyphosphoribosylaminopyrimidine deaminase	Riboflavin
433946	434416	471	6	7-dimethyl-8-ribityllumazine synthase
434436	434855	420	Transcription termination protein NusB	Riboflavin synthesis cluster; Transcription factors
380586	380143	444	Glucose-1-phosphate thymidylyltransferase	Rhamnose containing glycans; dTDP-rhamnose synthesis
383234	382038	1197	Glucose-1-phosphate thymidylyltransferase	Rhamnose containing glycans; dTDP-rhamnose synthesis
630418	629330	1089	Glycerol dehydrogenase	Respiratory dehydrogenases 1
452844	451369	1476	AmpG permease	Recycling of Peptidoglycan Amino Acids
464846	464151	696	Queuosine Biosynthesis QueC ATPase	Queuosine-Archaeosine Biosynthesis; tRNA modification Bacteria
539864	541165	1302	Xanthine permease	Purine Utilization; Xanthine Metabolism in Bacteria
301279	300323	957	Periplasmic aromatic aldehyde oxidoreductase	FAD binding subunit YagS
681353	680418	936	Inosine-uridine preferring nucleoside hydrolase	Purine conversions;

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
504213	505865	1653	UDP-sugar hydrolase	Purine conversions;
490711	491262	552	Adenine phosphoribosyltransferase	Purine conversions;
496474	497118	645	Adenylate kinase	Purine conversions
499424	500728	1305	Inosine-guanosine kinase	Purine conversions
494419	496293	1875	Chaperone protein HtpG	Protein chaperones
455976	456599	624	ATP-dependent Clp protease proteolytic subunit	Proteasome; Proteolysis
456725	457999	1275	ATP-dependent Clp protease ATP-binding subunit ClpX	Proteasome; Proteolysis
458187	460541	2355	ATP-dependent protease La	Proteasome; Proteolysis
420339	421658	1320	Proline-specific permease proY	Proline
485835	489197	3363	Potassium efflux system KefA protein	Potassium homeostasis
502537	500861	1677	POTASSIUM/PROTON ANTIPORTER ROSB	Potassium homeostasis
401046	402461	1416	Alkaline phosphatase	Phosphate metabolism
578484	579038	555	Phage tail fiber assembly protein	Phage tail fiber proteins
579816	580001	186	Phage tail fiber assembly protein	Phage tail fiber proteins
565774	565986	213	Origin specific replication binding factor	Phage replication
573255	573638	384	Phage antitermination protein Q	Phage nin genes - N-independent survival
575208	575510	303	Phage outer membrane lytic protein Rz	Phage lysis modules
635514	634768	747	Thiol:disulfide interchange protein DsbG precursor	Periplasmic disulfide interchange
553736	553242	495	Peptidyl-prolyl cis-trans isomerase PpiB	Peptidyl-prolyl cis-trans isomerase;
461214	463085	1872	Peptidyl-prolyl cis-trans isomerase PpiD	Peptidyl-prolyl cis-trans isomerase
395586	394429	1158	Penicillin-binding protein AmpH	Peptidoglycan Crosslinking of Peptide Stems

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
400222	399128	1095	D-alanine–D-alanine ligase A	Peptidoglycan Biosynthesis;
662131	661043	1089	Rare lipoprotein A precursor	Peptidoglycan Biosynthesis
667513	666872	642	Nicotinate-nucleotide adenylyltransferase	NAD and NADP cofactor biosynthesis global
483702	480553	3150	RND efflux system	inner membrane transporter CmeB
484810	483725	1086	Membrane fusion protein of RND family multidrug efflux pump	Multidrug Resistance Efflux Pumps
485060	485707	648	Transcription repressor of multidrug efflux pump acrAB operon	TetR (AcrR) family
347981	348871	891	Methylisocitrate lyase	Methylcitrate cycle
349311	350480	1170	2-methylcitrate synthase	Methylcitrate cycle
350514	351965	1452	2-methylcitrate dehydratase	Methylcitrate cycle
352005	353891	1887	Propionate–CoA ligase	Methylcitrate cycle
347742	346156	1587	Propionate catabolism operon regulatory protein PrpR	Methylcitrate cycle
660976	659693	1284	D-alanyl-D-alanine carboxypeptidase	Metallocarboxypeptidases; Peptidoglycan Biosynthesis
319393	318068	1326	Putative Dihydrolipoamide dehydrogenase	FAD-dependent NAD(P)-disulphide oxidoreductase
421814	423631	1818	Maltodextrin glucosidase	Maltose and Maltodextrin Utilization
479217	478666	552	Maltose O-acetyltransferase	Maltose and Maltodextrin Utilization
634559	633657	903	LysR-family transcriptional regulator YbeF	LysR-family proteins
658319	657366	954	LysR-family transcriptional regulator YbeF	LysR-family proteins
281670	282578	909	4-hydroxy-tetrahydrodipicolinate synthase	Lysine Biosynthesis DAP Pathway

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
657157	656192	966	Lipoate synthase	Lipoic acid metabolism
659219	658578	642	Octanoate-[acyl-carrier-protein]-protein-N-octanoyltransferase	Lipoic acid metabolism
659583	659320	264	Proposed lipoate regulatory protein YbeD	Lipoic acid metabolism
653498	654058	561	Lipid A acylation protein PagP	palmitoyltransferase
361159	360548	612	Galactoside O-acetyltransferase	Lactose utilization
362478	361225	1254	Lactose permease	Lactose and Galactose Uptake and Utilization
366809	365727	1083	Transcriptional repressor of the lac operon	Lactose and Galactose Uptake and Utilization
321000	321719	720	Predicted L-lactate dehydrogenase	Fe-S oxidoreductase subunit YkgE
323150	323845	696	Predicted L-lactate dehydrogenase	hypothetical protein subunit YkgG
321730	323157	1428	Predicted L-lactate dehydrogenase	Iron-sulfur cluster-binding subunit YkgF
411864	410596	1269	Protein AraJ precursor	L-Arabinose utilization
553239	552517	723	UDP-2	3-diacylglucosamine diphosphatase
439476	437614	1863	1-deoxy-D-xylulose 5-phosphate synthase	Isoprenoid Biosynthesis
440400	439501	900	Octaprenyl diphosphate synthase	6E)-farnesyl diphosphate synthase
632305	631688	618	Co-activator of prophage gene expression IbrB	IbrA and IbrB: co-activators of prophage gene expression
633510	632290	1221	Co-activator of prophage gene expression IbrA	IbrA and IbrB: co-activators of prophage gene expression
466711	467529	819	Cof protein	HD superfamily hydrolase
416441	417130	690	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	High affinity phosphate transporter and control of PHO regulon

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
417188	418483	1296	Phosphate regulon sensor protein PhoR (SphS)	High affinity phosphate transporter and control of PHO regulonregulatory system
389026	388052	975	Porphobilinogen synthase	Heme and Siroheme Biosynthesis
497354	498316	963	Ferrochelatase	protoheme ferro-lyase
680334	678664	1671	Chaperone protein DnaK	GroEL GroES; Protein chaperones
435897	436406	510	Phosphatidylglycerophosphatase A	Glycerolipid and Glycerophospholipid Metabolism
535009	535785	777	Hydroxypyruvate isomerase	Glycerate metabolism
377667	376834	834	S-formylglutathione hydrolase	Glutathione-dependent pathway of formaldehyde detoxification
378870	377761	1110	S-(hydroxymethyl)glutathione dehydrogenase	Glutathione-dependent pathway of formaldehyde detoxification
379180	378905	276	FrmR: Negative transcriptional regulator of formaldehyde detoxification operon	Glutathione-dependent pathway of formaldehyde detoxification
510940	511872	933	Glutaminase	Glutamine
682196	681471	726	Glutamate Aspartate transport ATP-binding protein GltL	Glutamate and Aspartate uptake in Bacteria
682870	682196	675	Glutamate Aspartate transport system permease protein GltK	Glutamate and Aspartate uptake in Bacteria
683610	682870	741	Glutamate Aspartate transport system permease protein GltJ	Glutamate and Aspartate uptake in Bacteria
684688	683780	909	Glutamate Aspartate periplasmic binding protein precursor GltI	Glutamate and Aspartate uptake in Bacteria

Table SC.2 – *Continued from previous page*

Start	**Stop**	**Length (bp)**	**Annotation**	**Subsystem**
365604	362530	3075	Beta-galactosidase	Galactosylceramide and Sulfatide metabolism
618129	617137	993	ABC-type Fe3+-siderophore transport system	permease 2 component
424217	423636	582	Acyl carrier protein phosphodiesterase	Fatty Acid Biosynthesis FASII
460750	461022	273	DNA-binding protein HU-beta	DNA structural proteins
556172	555960	213	FIG002958: hypothetical protein	DNA replication cluster 1
415052	411906	3147	Exonuclease SbcC	DNA repair
416251	415049	1203	Exonuclease SbcD	DNA repair
440642	440400	243	Exodeoxyribonuclease VII small subunit	DNA repair
577326	577192	135	DNA-damage-inducible protein we	DNA repair
493704	494309	606	Recombination protein RecR	DNA processing cluster
491460	493322	1863	DNA polymerase III subunits gamma and tau	DNA processing cluster
493375	493704	330	FIG000557: hypothetical protein co-occurring with RecR	DNA processing cluster
551893	550826	1068	Phosphoribosylaminoimidazole carboxylase ATPase subunit	De Novo Purine Biosynthesis
552399	551890	510	Phosphoribosylaminoimidazole carboxylase catalytic subunit	De Novo Purine Biosynthesis
358098	358757	660	Carbonic anhydrase	Cyanate hydrolysis
357989	357090	900	Cyn operon transcriptional activator	Cyanate hydrolysis
358788	359258	471	Cyanate hydratase	Cyanate hydrolysis
359291	360445	1155	Cyanate transport protein CynX	Cyanate hydrolysis
354221	355480	1260	Cytosine permease	Creatine and Creatinine Degradation

Table SC.2 – *Continued from previous page*

Start	Stop	Length (bp)	Annotation	Subsystem
510678	508174	2505	Lead	cadmium
687822	686284	1539	Apolipoprotein N-acyltransferase	Copper homeostasis: copper tolerance
316312	315842	471	Transcriptional regulator YkgA	Commensurate regulon activation
304597	303887	711	CFA/I fimbrial chaperone	Colonization factor antigen we fimbriae
306137	304566	1572	CFA/I fimbrial minor adhesin	Colonization factor antigen we fimbriae
308724	306199	2526	CFA/I fimbrial subunit C usher protein	Colonization factor antigen we fimbriae
309418	308750	669	CFA/I fimbrial auxiliary subunit	Colonization factor antigen we fimbriae
310063	309476	588	CFA/I fimbrial major subunit	Colonization factor antigen we fimbriae
310341	310138	204	FIGfam014588: Predicted regulator of CFA/I fimbriae	Colonization factor antigen we fimbriae
654233	654442	210	Cold shock protein CspA	Cold shock
591711	590269	1443	Copper sensory histidine kinase CusS	Cobalt-zinc-cadmium resistance
595655	598798	3144	Cobalt-zinc-cadmium resistance protein CzcA; Cation efflux system protein CusA	Cobalt-zinc-cadmium resistance
592384	591701	684	DNA-binding heavy metal response regulator	Cobalt-zinc-cadmium resistance
592541	593914	1374	Cation efflux system protein CusC precursor	Cobalt-zinc-cadmium resistance
594072	594404	333	Cation efflux system protein CusF precursor	Cobalt-zinc-cadmium resistance
594420	595643	1224	Cobalt/zinc/cadmium efflux RND transporter	membrane fusion protein
647724	647428	297	Citrate lyase gamma chain	acyl carrier protein
643488	642058	1431	Citrate Succinate antiporter	Citrate Metabolism
644444	643572	873	2-(5&39;&39;-triphosphoribosyl)-3&39;-dephosphocoenzyme-A synthase	Citrate Metabolism
644976	644425	552	Apo-citrate lyase	Citrate Metabolism

Table SC.3: Quality filters for all reported polymorphisms

Strain	Position	Reference	Alternate	Read Depth	High quality reads	Reference depth	Alternate depth	Frequency	P-value
MG1655	1468135	C	A	104	104	99	5	4.81	$2.97x10^{-2}$
MG1655	3305354	A	T	52	52	0	52	100	$6.31x10^{-3}$
MG1655	3608715	A	G	46	46	41	5	10.87	$2.78x10^{-2}$
MG1655	3688287	A	C	75	75	68	7	9.33	$6.74x10^{-3}$
MG1655	3723928	A	C	124	124	119	5	4.03	$2.99x10^{-2}$
MG1655	4296060	C	T	60	60	53	7	67	$6.49x10^{-3}$
AG100	227791	T	C	105	93	67	26	27.96	$1.95x10^{-9}$
AG100	3127332	G	T	85	77	65	12	15.58	$1.53x10^{-4}$
AG100	3127350	A	G	81	73	60	13	17.81	$6.79x10^{-5}$
AG100	3127368	A	G	83	77	65	12	15.58	$1.53x10^{-4}$
AG100	3127402	T	C	86	79	65	14	17.72	$3.24x10^{-5}$
AG100	3127440	G	C	83	80	66	14	17.5	$3.27x10^{-5}$
AG100	3127460	C	T	82	76	61	15	19.74	$1.42x10^{-5}$
AG100	3127494	C	G	71	67	54	13	19.4	$6.41x10^{-5}$
AG100	3421431	A	G	105	93	73	20	20.21	$3.03x10^{-7}$
AG100	3648916	G	T	109	95	83	12	12.63	$1.68x10^{-4}$
AG100	3648944	G	A	103	96	85	11	11.46	$3.6x10^{-4}$
AG100	4613143	T	G	96	91	81	10	10.99	$7.52x10^{-4}$

Appendix D

Chapter 5: Appendix

D.1 Materials and Methods

Bacteria and Phage Library: a genotype × genotype Infectivity matrix from imaging algorithms

In order to determine an infectivity phenotype for each bacteria-phage pair, a bacterial lawn on an agar plate was inoculated with each of the phage to produce a plaque the size of which determines the relative infectivity of each phage for that bacterium. A typical bacterial lawn resulting from this assay is depicted in Figure D.1(a). A photograph of each plate was taken with a digital camera at a resolution of 72dpi (see Figure D.1(a)) and processed using the public-domain software `ImageJ` [1]. After acquisition, the picture was converted into an 8-bit image and filtered to reduce the effect of non-uniform illumination. The local normalisation algorithm used is based on local mean and variance values computed using local spatial smoothing [30].

The obtained image was then smoothed using a Gaussian blur (with $\sigma = 4$) and enhanced using a contrast stretching algorithm (normalized, with 3 saturated pixels). The resulting image was then thresholded using the iterative isodata algorithm [102], inverted and converted to a mask. By combining the resulting image with the original image we get the pre-processed image shown in Figure D.1(b). Then the region of interest (ROI) was established manually for each picture in order to determine the 8×12 grid where each phage was inoculated. From each cell in the grid a sub-image was automatically cropped and the area of each blob measured using a growing region algorithm with optimised thresholds. Both the grid and the obtained blobs are illustrated in Figure D.1(c).

The blob-detection algorithm can be summarised as follows: 1) first determine a seed pixel, for instance by choosing a local maximum near the centre of the image; 2) then recursively add in neighbouring pixels that are above the optimised threshold, and thus increasing the size of the region; 3) finally convert the resulting blob into a polygonal ROI and compute its area. After repeating this process for each of the 96

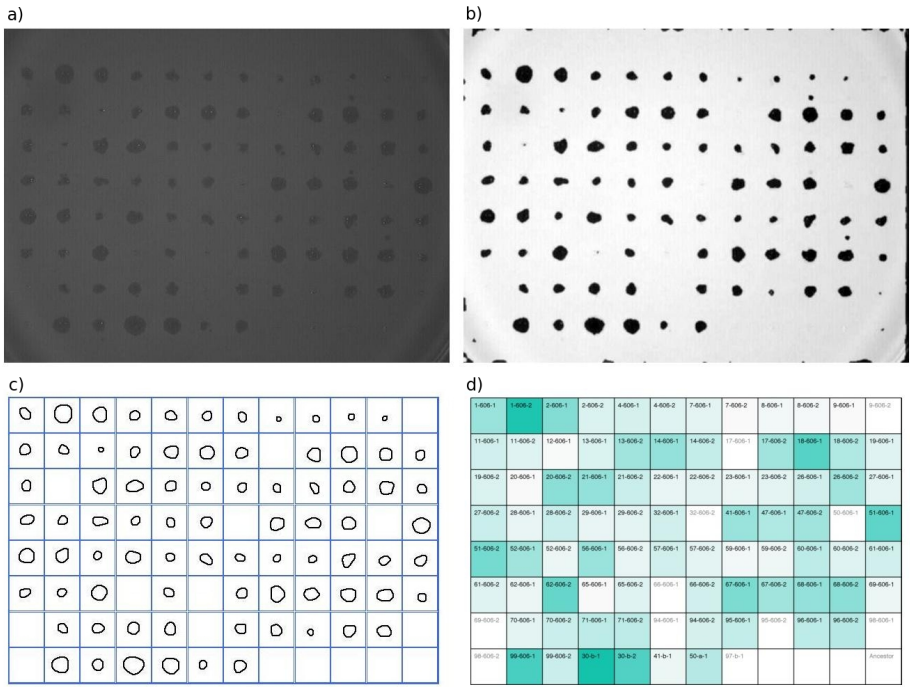


Fig. D.1: An example (bacterium 23b) illustrating the image processing steps taken to quantify phage infectivity patterns from the picture of a plate. a) Photograph at 72dpi. b) Pre-processed image. c) Blob detection algorithm. d) Resulting normalised infectivity data (darker green denotes higher infectivity).

cells in the grid, the obtained values are then normalised to provide a proxy for the likelihood of infection of each phage type, a quantity that can be visualised using a heat map, as in Figure D.1(d).

Growth Kinetics

A library of 56 strains that had been co-evolved with λ phage was obtained from Justin Meyer (San Diego) with thanks. Image analysis was carried out by Rafael Pena-Miller (UNAM, Cuernavaca) to determine plaque size and depth to give a quantitative measure of infectivity fot every strain. In order to assess any potential change in growth kinetics as a result of mutated maltoporin transport protein, each strain was grown for 24 hours at 8 concentrations of maltotriose as the sole carbon source. All strains had the same REL606 ancestor, with the genetic differences between all strains had been elucidated using illumina sequencing of the lamB gene. In order to determine the trade off between growth rate and yield, as well as any correlation with resistance, all strains were grown for 24 hours in 8 concentrations of maltotriose, ranging from 2.5μg/ml to 250μg/ml. Davis minimal media was used throughout to ensure that maltotriose was the only carbon source supplied to the bacteria. Experiments were carried out with five replicates for each strain (plus positive and negative controls) using two strains per 96 well microplane (150μl well volume). Optical density was read using a Biotek Elx808 uv/vis absorbance reader

every twenty minutes. To ensure measurable growth levels at the lowest maltotriose concentrations, cells were revived in LB media for 12 hours after sampling from the freezer. 100μL of this overnight culture was then added to Davis media supplemented with the appropriate level of sugar. After a second overnight growth, cells were used to inoculate the micro-plate using a 96-pin replicator, transferring approximately 1μL per well.

Calculating growth rate

Growth rate is a key calculation in deterring rate-yield trade off dynamics, and yet is not as trivial a calculation as may be assumed. The standard method of calculating a rate value is to differentiate the exponential growth curve, and derive a doubling time from this information. However, as previously mentioned, such as approach not only takes no account of the length of the lag phase, or the growth arrest phase, it can also be difficult to arrive at a realistic number, especially in cases where growth is very low and fails to follow a sigmoidal growth pattern, making curve fitting more difficult.

For this reason, other methods of calculating rate were attempted and compared. The stepwise difference was measured, and the largest stepwise difference (or R_{max}, as it will be described hereafter) used as a rate measurement. Alternatively the area under the curve for the whole growth curve was calculated (in this case, using the trapezoid approximation method) from the raw data. Both of these techniques had the advantage of making calculations based on the raw data, and not requiring curve fitting. Although using raw data without model fitting has advantages, in this case the standard sigmoidal model was used to determine rate. This decisions was made for a number of factors; firstly to ensure consistency with other studies of growth rate. As opposed to experiments in Chapter 2, there was no antibiotic to act as a growth limitation in this study, therefore all strains reached stationary phase within the 24 hours of the experiment. For this reason, a sigmoidal fit could be achieved more easily, and was therefore adopted to find the rate.

A Protein-Matching Heuristic

A protein matching algorithm begins by extracting the α-carbon data from the approximate folding package (I-TASSER). We then perform a Procrustes-like analysis, extended to include the matching of proteins with SNPs, insertions and deletions. In order to match two proteins of different sizes we used an iterative and heuristic extension of the rectangular munkres algorithm to extend the 3-d point-matching algorithm given in [88], as follows. This algorithm also defines a distance between two proteins' α-carbon structures, $\mathcal{P}$ and $\mathcal{Q}$, the algorithm is defined as follows.

Suppose $\{p_j\}_{j=1}^{N} = \mathcal{P}$ and $\{q_j\}_{j=1}^{M} = \mathcal{Q}$ are given.

0. If $N = M$, the Procrustes algorithm (using only direct isometries) followed by the munkres linear matching algorithm defines an optimal pairing of minimal Euclidean distance such that $j = J(i)$ for $i = 1, .2, ..., N$.

   Otherwise, the following iteration is implemented when $N > M$ with the goal of locating a subset of $\mathcal{P}$ that matches $\mathcal{Q}$ optimally, or close to optimally:

1. Let $j := 0, \mathcal{P}_0 := \mathcal{P}, \mathcal{K}_{\text{best}} := \{p_1, ..., p_M\}, \mathcal{L}_{\text{best}} := \{p_{M+1}, ..., p_N\}$ and define $d_{\text{best}}^{\text{matched}} := \infty$. Also define an iteration limit, j_{lim}.
2. Use munkres to find a set of indicies that provide the best match between $\mathcal{P}_j$ and $\mathcal{Q}$ and so write $\mathcal{P}_j = \mathcal{K}_j \cup \mathcal{L}_j$ where $\mathcal{K}_j$ is a set of N points that matched best and $\mathcal{L}_j$ is the set of $M - N$ that match worst.
3. Define a flag, `stop` := is($\mathcal{L}_{\text{best}} = \mathcal{L}_j$ or $j \geq j_{\text{lim}}$)?
4. Define a distance from the unmatched points to the target set and a distance from the matched points to the target set, respectively:

$$d_j^{\text{unmatched}} := \sum_{l \in \mathcal{L}_j, p \in \mathcal{Q}} d(l, q) \qquad \text{and} \qquad d_j^{\text{matched}} := \sum_{k \in \mathcal{K}_j, p \in \mathcal{Q}} d(k, q).$$

   If

$$d_j^{\text{matched}} < d_{\text{best}}^{\text{matched}}$$

   then

$$d_{\text{best}}^{\text{matched}} := d_j^{\text{matched}}, d_{\text{best}}^{\text{unmatched}} := d_j^{\text{unmatched}}, \mathcal{K}_{\text{best}} := \mathcal{K}_j, \mathcal{L}_{\text{best}} := \mathcal{L}_j.$$

5. Find the direct Procrustes optimal isometry given by a pair denoting the linear isometry and a translation, (R_j, τ_j), between $\mathcal{K}_j$ and $\mathcal{Q}$ and set $\mathcal{P}_{j+1} := R_j[\mathcal{P}_j] + \tau_j$.
6. If $j \geq 1$ and `stop` is false, set $j := j + 1$ and return to (2).

At this point $\mathcal{K}_{\text{best}}$ is the set of best approximating points derived from $\mathcal{Q}$ through a linear, direct isometry that and $\mathcal{L}_{\text{best}}$ are the image of the points in $\mathcal{P}$ under this isometry that we have rejected. The algorithm has also produced two notations of distance, one taken over the points that match and one that adds to this number the distance taken over all the remaining points that do not match.

This algorithm is applied to an artificial dataset shown in Figure D.2(a) and the results presented in Figure D.2(b): a discrete 'figure 8' is matched to a larger and randomly perturbed version of the same shape (larger by two α-carbons, analogous to having two single base-pair insertion mutations). By the green circles Figure D.2(b) illustrates that the algorithm correctly identifies the two additional carbons. The larger dataset is shown as a set of blue dots and black lines are used to indicate the largest matches (here greater than 3.5% of the total matched protein distance) to the smaller dataset, the latter shown as red circles.

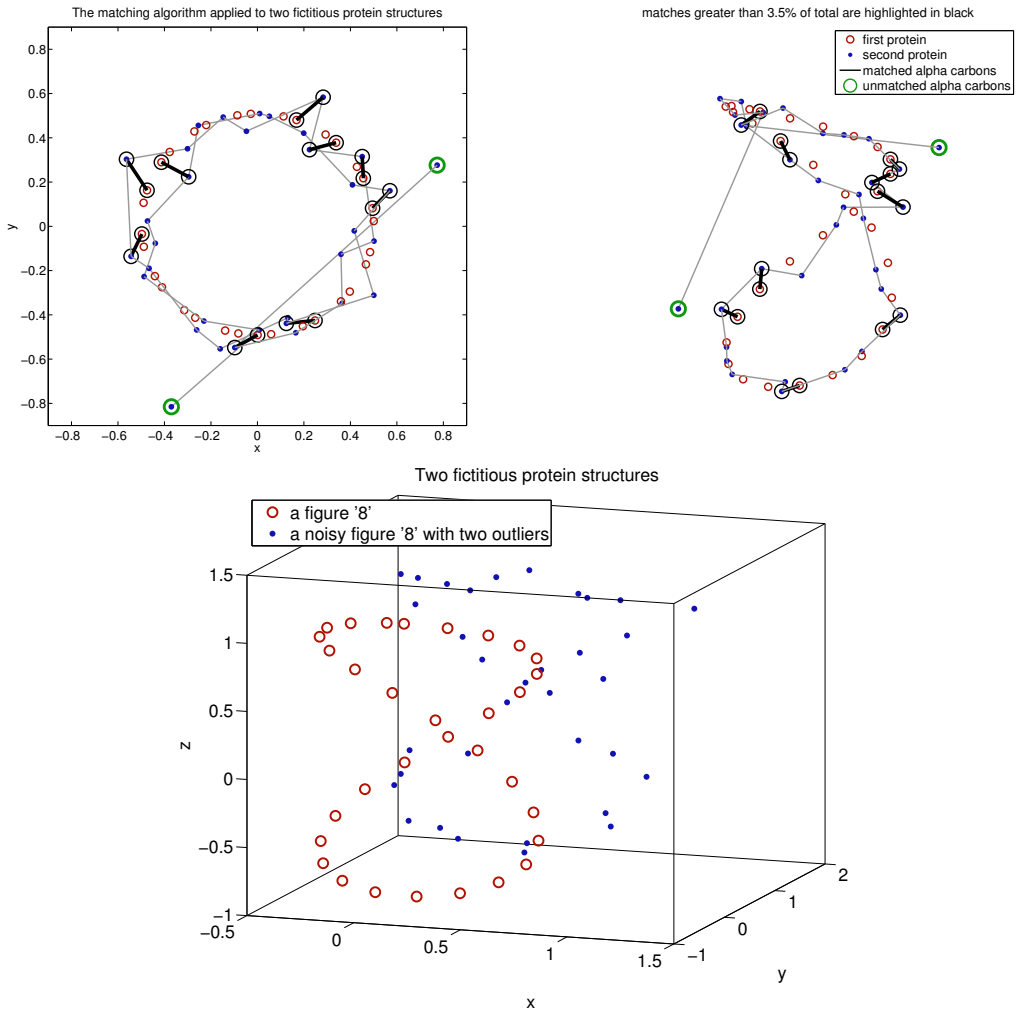


Fig. D.2: (a) This shows two different views of the matched data, one from above on the left, and one from the side on the right. The algorithm described in the text determines matched points and identifies two 'homologous' subsets in each dataset, it identifies the remaining two points in the larger protein that cannot be matched and the worst matched pairs are then indicated with a black line. (b) Two 3-d datasets are shown, one is a figure '8' comprised of 32, 3-vectors representing α-carbon positions, the other is a similar figure '8' to which noise has been added, a rigid motion applied and two additional points added to make a new dataset of 34, 3-vectors. We seek the best match between these two fictitious 'proteins' with respect to a set (ie. Hausdorff) metric based on the Euclidean norm.

D.2 Supplementary Data

For completeness we sought a relationship between each strain's competitive ability measured using a standard relative fitness phenotype and phage resistance, but we saw no correlation, just as there is no correlation between resistance and lag, rate or yield. Data supporting this absence of correlation is shown in Figure D.3.

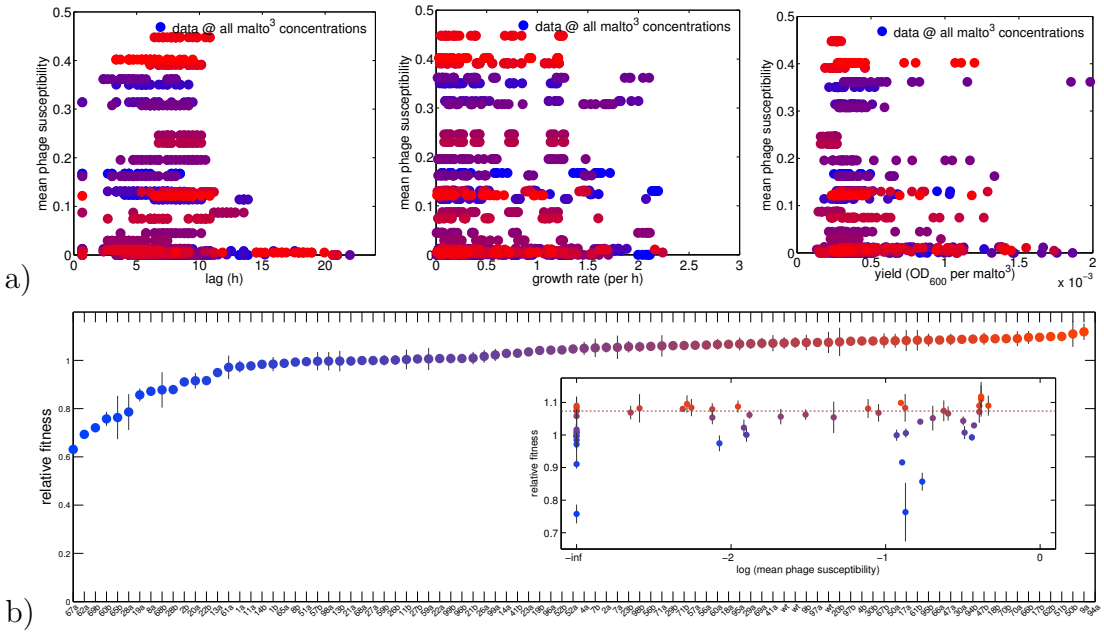


Fig. D.3: **Three vegetative growth traits (rate, yield and lag) measured in laboratory conditions are plotted against their respective library phage susceptibilities**. (a) This data indicates no evidence to support a trade-off hypothesis that rate, yield or lag are traded by the bacteria for phage resistance. (b) This statement is supported by a second dataset whereby *E.coli* relative fitnesses are plotted against phage susceptibility. The fitness are shown in the main plot, (b-inset) then shows the relationship between relative fitness and mean phage susceptibility for each bacterial strain. Relative fitness is defined by growth rate for this plot.

Appendix E

Chapter 6: Appendix

E.1 Materials and Methods

Strains

E.coli strains AG100, AG100A and MG1655 were cultured overnight in M9 minimal media, (per litre, 1.75g K_2HPO_4, 0.5g KH_2PO_4, 0.15g $Na_3C_6H_5O_7$, 0.25g NH_4NO_3, 0.025g $MgSO_4$) supplemented with 0.2% glucose and 0.1% casamino acid. Cultured cells were taken from single colonies isolated on LB agar from stocks stored at -80°C. Soft agar (5g/L agar) was made with the same proportion of M9 salts, and inoculated with 5% overnight culture ($\sim 10^9$ cells). Later studies were conducted in the same manor but using a range of glucose concentrations from 0.05% to 2%. Overnight inocula were grown at the same glucose concentration as was to be used experimentally. Casamino acid concentration was kept constant at 0.1% regardless of glucose level. Strains were kindly provided by Stuart Levy (Tufts) and Tobias Bergmiller (IST) respectively.

Plate construction

150mm square petri dishes were filled with 50ml of soft agar, excluding a circular area in the center of the plate, kept free of agar using a small petri dish. Once the agar was set, the small dish was removed using forceps, and was filled subsequently with the same agar as above, containing no bacteria, but the appropriate concentration of Doxycycline (10μg/ml). This methods was chosen, rather than the traditional paper-disk diffusion method, as it allows an equal diffusion of nutrient and drug across an entirely homogenous media plate. The use of soft agar to aid image capture also necessitated the use of a drug containing region where diffusion would not be biased to the surface of the plate. The plates were grown at 30°C for 48 hours, photographed at zero, 24 and 48 hours in a custom built automated imaging system (BioBox - see supplementary materials).

Imaging

Images were recorded automatically in a miniature custom built dark-room, using a Canon EOS 1100D digital SLR, (*f=0.4*, ISO=100, exp.=1.2 seconds) controlled by an Arduino controller. Perspective correction was performed using Adobe Lightroom software, and the images were subsequently analysed in MATLAB to obtain pixel intensity as a proxy for cell density.

Bio-Box

The BioBox is a desktop darkroom, custom built to automatically take photos, either under white light or using filters for fluorescent imaging. The box is temperature controlled to provide a consistent environment for multiple day time lapse video capture. The camera and associated servo arm for positioning the filters (two excitation/emission filters pairs can be used, currently for CFP and YFP) are controlled by an Arduino, programmed by Michael Seiber and Rafael Pena-Miller. Programming was later modified by myself and Sam Gardner to optimise settings including exposure time and focus depth. Since being used to generate the images seen in chapter 5, the box has been modified to change the angle of image capture (reducing glare, and removing the need to for perspective correction).

Bibliography

[1] Ram S Abramoff M, Magalhaes P, *Image processing with imagej*, Biophotonics International **11** (2004), no. 7, 36–42.

[2] Angus Buckling Associate Editor: Santiago F. Elena Editor: Judith L. Bronstein Alex R. Hall, Pauline D. Scanlan, *Bacteriaphage coevolution and the emergence of generalist pathogens*, The American Naturalist **177** (2011), no. 1, 44–53.

[3] Ben-Zakour NL Beatson SA Alikhan NF, Petty NK, *Blast ring image generator brig: simple prokaryote genome comparisons*, BMC Genomics **12** (2011), no. 402.

[4] Rybak MJ Allen GP, Kaatz GW, *Activities of mutant prevention concentration-targeted moxifloxacin and levofloxacin against* Streptococcus pneumoniae *in an in vitro pharmacodynamic model.*, Antimicrobial Agents Chemotherapy **47** (2003), 2606–2614.

[5] Diarmaid H Andersson DI, *Antibiotic resistance and its cost: is it possible to reverse resistance?*, Nat Rev Microbiol **8** (2010), no. 4, 260–271 (eng).

[6] Hughes D Andersson DI, *Persistence of antibiotic resistance in bacterial populations*, FEMS Microbiol Rev **35** (2011), no. 5, 901–911.

[7] Levin BR Andersson DI, *The biological cost of antibiotic resistance*, Current Opinion in Microbiology **2** (1999), no. 5, 489–493.

[8] Levin BR Ankomah P, Johnson PJT, *The pharmaco –, population and evolutionary dynamics of multi-drug therapy: Experiments with s. aureus and e. coli and computer simulations*, PLoS Pathog **9** (2013), no. 4.

[9] Best AA DeJongh M Disz T Edwards RA Formsma K Gerdes S Glass EM Kubal M Meyer FO lsen GJ Olson R Osterman AL Overbeek RA McNeil LK Paarmann D Paczian T Parrello B Pusch GD Reich C Stevens R Vassieva O Vonstein V Wilke A Zagnitko O Aziz RK, Bartels D, *The rast server: Rapid annotations using subsystems technology.*, BMC Genomics (2008).

[10] Galán B, García JL, and Prieto MA, *The paax repressor, a link between penicillin g acylase and the phenylacetyl-coenzyme a catabolon of escherichia coli w*, J. Bacteriol. **186** (2004), no. 7, 2215–2200.

[11] Hasegawa M Takai Y Okumura Y Baba M Datsenko KA Tomita M Wanner BL Mori H Baba T, Ara T, *Construction of escherichia coli k-12 in-frame single-gene knockout mutants: the keio collection.*, Mol Syst Biol **2** (2006.), 2006–2008.

[12] Morrison JO Bayer AS, *Disparity between timed-kill and checkerboard methods for determination of in vitro bactericidal interactions of vancomycin plus rifampin versus methicillin-susceptible and -resistant Staphylococcus aureus*, Antimicrob Agents Chemother **26** (1984), no. 2, 220–223 (eng).

[13] Laehnemann D Jansen G Fuentes-Hernandez A Rosenstiel P Schulenburg H Beardmore RE, Pena-Miller R, *When the most potent combination of antibiotics selects for the greatest bacterial load: the smile-frown transition.*, PLoS Biology **114** (2013), 1–13.

[14] Swerdlow HP Smith GP Milton J Bentley DR, Balasubramanian S, *Accurate whole human genome sequencing using reversible terminator chemistry*, Nature **456** (2008), 53–59.

[15] Hall BG, *Predicting the evolution of antibiotic resistance genes.*, Nature Reviews Microbiology **2** (2004), no. 5, 430–435.

[16] Andersson DI Bjorkman J, Hughes D, *Virulence of antibiotic resistant salmonella typhimuirium.*, Proc Natl Acad Sci U S A **95** (2008), 3949–3953.

[17] Andersson DI Hughes D Bjorkman J, Samuelsson P, *Novel ribosomal mutations affecting translational accuracy: antibiotic resistance and virulence of salmonella typhimurium.*, Mol Microbiol **31** (1999), 53–58.

[18] Ricci V Modi N Cacciotto P-Kleinekathoefer U Ruggerone P Vargiu AV Baylay AJ Smith HE Brandon Y Galloway D Piddock LJV Blair JMA, Bavro VN, *Acrb drug-binding pocket substitution confers clinically relevant resistance and altered substrate specificity*, Proceedings of the National Academy of Sciences **112** (2015), no. 11, 3511–3516.

[19] Disz T Edwards RA Gerdes SO-lsen GJ Olson R Overbeek R Parrello B Pusch GD Shukla M Thomason JA Stevens R Vonstein V Wattam AR Xia F Brettin T, Davis JJ, *Rasttk: A modular and extensible implementation of the rast algorithm for building custom annotation pipelines and annotating batches of genomes.*, Sci Rep (2015).

[20] Snyder M Gerstein M Byzov A, Urban AE, *Cnvnator: An approach to discover, genotype, and characterize typical and atypical cnvs from family and population genomic sequencing.*, Genome Research **21** (2011), no. 6, 987–984.

[21] Sanchez C, *Bacterial evolution: Phage resistance comes at a cost*, Nat Rev Micro **9** (2011), no. 6, 398–399.

[22] Evans ME Campion JJ, McNamara PJ, *Pharmacodynamic modeling of ciprofloxacin resistance in s. aereus.*, Antimicrobial Agents Chemotherapy **49** (2005), 209–219.

[23] McNamara PJ Titlow WB Evans ME Campion JJ, Chung P, *Pharmacodynamic modeling of the evolution of levofloxacin resistance in s. aereus.*, Antimicrobial Agents Chemotherapy **49** (2005), 2189–99.

[24] Lee DG Choi H, *Synergistic effect of antimicrobial peptide arenicin-1 in combination with antibiotics against pathogenic bacteria*, Research in Microbiology **163** (2012), 479 – 486.

[25] CLSI, *Harmonized terminology database.*

[26] Tasan M Mutlu B Weinstein ZB-Suzuki Y Nergiz ME Costanzo M Baryshnikova A Giaever G Nislow C Myers CL Andrews BJ Boone C Roth FP Cokol M, Chua HN, *Systematic exploration of synergistic drug pairs.*, Mol Syst Biol **7** (2011), 544 (eng).

[27] Kenny M Colebrook L, *Treatment of human puerperal infections, and of experimental infections in mice, with prontosil.*, The Lancet **227** (1936), 1279–1281.

[28] Duclos B Cenatiempo Y Gautier C-Prato JL Cozzone AJ Cortay JC, Bleicher F, *Utilization of acetate in escherichia coli: structural organization and differential expression of the ace operon.*, Biochimie **71** (1989), no. 9-10, 1043–1049.

[29] Epstein W Curtis SJ, *Phosphorylation of d-glucose in escherichia coli mutants defective in glucose phosphotransferase mannosephosphotransferase and glucokinase.*, J Bacteriol. **122** (1975), 1189–1199.

[30] Sage D, *Local normalization*, http://bigwww.epfl.ch/sage/soft/localnormalization/ (Last Accessed: Oct 2012).

[31] Armstrong SD Hartley CS Xia D-Wastling JM Makepeace BL Darby AC, Gill AC, *Integrated transcriptomic and proteomic analysis of the global response of wolbachia to doxycycline-induced stress.*, ISME J **84** (2014), no. 925937.

[32] Andersson DI, *Persistence of antibiotic resistant bacteria*, Current Opinion in Microbiology **6** (2003), 452–456.

[33] Deziel M Gumbo T Drusano GL, Louie A, *The crisis of resistance: Identifying drug exposures to surpress amplification of resistant mutant subpopulations.*, Antimicrobial Resistance **42** (2006), no. 525-532.

[34] Bartlett JG Kaplan JE Pau AK Dybul M, Fauci AS, *Guidelines for using antiretroviral agents among HIV-infected adults and adolescents*, Ann Intern Med **137** (2002), no. 5 Pt 2, 381–433 (eng).

[35] Fischer E, *The lock and key model*, Ber, 1984.

[36] Soupene E, *Physiological studies of escherichia coli strain mg1655: growth defects and apparent cross-regulation of gene expression.*, J. Bacteriol. **185** (2003), 5611–6526.

[37] Tabor E, *Bacteria by the book*, Science **322** (2008), no. 5903, 853–854.

[38] Levy M Eagle H, Fleischman R, *Development of increased bacterial resistance to antibiotics i. continuous spectrum of resistance to penicillin, chloramphenicol, and streptomycin.*, J. Bacteriol. **63** (1952), no. 5.

[39] Musselman AD Eagle H, *The rate of bactericidal action of penicillin in vitro as a function of its concentration, and its paradoxically reduced activity at high concentrations against certain organisms.*, J Exp Med **88** (1948), no. 1, 99–131 (eng).

[40] EBI, *Ag100 fasta*, http://www.ebi.ac.uk/ena/data/view/PRJEB7832 (Last Accessed: 15 March 2015).

[41] ENSEMBLE, *Mg1655 fasta*, ftp://ftp.ensemblgenomes.org/pub/bacteria/release-28/gff3/bacteria-0-collection/escherichia-coli-str-k-12-substr-mg1655 (Last Accessed: 22 March 2015).

[42] EUCAST, *Antimicrobial susceptibility testing*, European Committee on Antimicrobial Susceptibility Testing, 2010.

[43] Daschner F, *Tetracyclines: bacteriostatic or bactericidaldrugs? invitro studies with tetracycline, minocycline and doxycycline. (authors translation).*, Zentralbl Bakteriol Orig A. **239** (1977), no. 4, 527–534.

[44] Walsh F, *Antibiotics resistance 'as big a risk as terrorism' - medical chief.*, Available: http://www.bbc.co.uk/news/health-21737844. **Last accessed 18/03/13.** (2013).

[45] Odds FC, *Synergy, antagonism, and what the chequerboard puts between them.*, J Antimicrob Chemother **52** (2003), no. 1, 1 (eng).

[46] Nielsen UB Sorger PK Fitzgerald JB, Schoeberl B, *Systems biology and combination therapy in the quest for clinical efficacy*, Nat Chem Biol **2** (2006), no. 9, 458–466.

[47] Taubes G, *The bacteria fight back*, Science **321** (2008), no. 5887, 356–361.

[48] Wright GD, *Mechanisms of resistance to antibiotics*, Curr Opin Chem Biol **7** (2003), no. 563, 569.

[49] Parsons JC Greco WR, Bravo G, *The search for synergy: a critical review from a response surface perspective.*, Pharmacol Rev **47** (1995), no. 2, 331–385 (eng).

[50] Berg OG Ilbäck C Sandegren L-Hughes D Andersson DI Gullberg E, Cao S, *Selection of resistant bacteria at very low antibiotic concentrations*, PLoS Pathog **7** (2011), no. 7, e1002158.

[51] Morgan AD Buckling A Hall AR, Scanlan PD, *Host-parasite coevolutionary arms races give way to fluctuating selection*, Ecology Letters **14** (2011), 635–642.

[52] Westmacott D Hall MJ, Middleton RF, *The fractional inhibitory concentration (fic) index as a measure of synergy*, J. Antimicrob. Chemother **11** (1983), 427–433.

[53] William D. Hamilton, *Sex versus non-sex versus parasite*, Oikos **35** (1980), no. 2, 282–290.

[54] Damian D Hart lD Kishony R Hegreness M, Shoresh N, *Accelerated evolution of resistance in multidrug environments.*, PNAS **105** (2008), no. 37, 13977–13981 (eng).

[55] Poot MC deJonge MJ Degener JE-Dzoljic-Danilovic G Horrevorts AM, deRidder CM, *Checkerboards titrations: the influence of serial dilutions of antibiotics on the fractional inhibitory concentration index and fractional bactericidal concentration index*, Antimicrob Chemother **19** (1987), 119–125.

[56] http://ecoliwiki.net/colipedia/index.php/Bacteriophage lambda, *Phage lambda*, Last accessed: 23rd July 2015.

[57] http://ecoliwiki.net/colipedia/index.php/MG1655cite-note PMID:12949114-5, *Strain: Mg1655*, Last accessed: 15th August 2015.

[58] https://commons.wikimedia.org/wiki/File:Glyoxylate cycle.jpg, *Glyoxylate cycle schematic*, Last accessed: 12th Jan 2015.

[59] Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q andPaley S, Schröder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, and Karp PD, *Ecocyc: fusing model organism databases with systems biology*, Nucl. Acids Res. **41** (2013), no. D1, 605–612.

[60] Andrews J, *Determination of minimum inhibitory concentrations*, J. Antimicrob. Chemother. **48** (2001), 5–16.

[61] Silva J, *Mechanisms of antibiotic resistance*, Current Therapeutic Research **57** (1996), no. Supp. Ac.

[62] Acar JF, *Antibiotic synergy and antagonism.*, Medical Clinics of North America **84** (2000), no. 6, 1391–1406.

[63] Lewis K, *Platforms for antibiotic discovery*, Nat Rev Drug Discov **12** (2013), 371–387.

[64] Shimizu K Kabir MM, *Metabolic regulation analysis of icd-gene knockout escherichia coli based on 2d electrophoresis with maldi-tof mass spectrometry and*

enzyme activity measurements., Appl Microbiol Biotechnol **65** (2004), no. 1, 84–96.

[65] Blake WJ Collins JJ Kaern M, Elston TC, *Stochasticity in gene expression: from theories to phenotypes.*, Nat Rev Genet **66** (2005), 451–464.

[66] Liao JC Kao KC, Tran LM, *A global regulatory role of gluconeogenic genes in escherichia coli revealed by transcriptome network analysis.*, J Biol Chem **280** (2005), 36079–36087.

[67] Michael A Kohanski, Daniel J Dwyer, and James J Collins, *How antibiotics kill bacteria: from targets to networks*, Nat Rev Micro **8** (2010), no. 6, 423–435.

[68] Michael Kopp and Sergey Gavrilets, *Multilocus genetics and the coevolution of quantitative traits*, Evolution **60** (2006), no. 7, 1321–1336.

[69] Hardy KR., *Plasmids*, vol. Bacterial Plasmids, American Society of Microbiology, Washington DC, 1986.

[70] Rosenstiel P Beardmore R Jansen G-Schulenburg H Laehnemann D, Pena-Miller R, *Genomics of rapid adaptation to antibiotics: convergent evolution and scalable sequence amplification.*, Genome Biol Evol **6** (2014), no. 6, 1287–1301 (eng).

[71] Rice LB, *Bacteria by the book: Response*, Science **322** (2008), 853–854.

[72] Bryan LE, *General mechanisms of resistance to antibiotics*, Journal of Antimicrobial Chemotherapy **22** (1988), 1–15.

[73] Wysoker A Fennell T Ruan J Homer N Marth G Abecasis G Durbin R Li H, Handsaker B, *The sequence alignment/map format and samtools.*, Bioinformatics **25** (2009), 2078–2079.

[74] Nikaido H Li XZ, *Efflux-mediated drug resistance in bacteria: an update.*, Drugs **69** (2009), no. 12, 1555–1623 (eng).

[75] Baquero F Martinez JL Linares JF, Gustafsson I, *Antibiotics as inter-microbial signalling agents instead of weapons.*, PNAS **103** (2006), no. 51, 19484–19489.

[76] Levin BR Lipsitch M, *The population dynamics of antimicrobial chemotherapy.*, Antimicrob Agents Chemother **41** (1997), no. 2, 363–373 (eng).

[77] Piddock LJV, *Multidrug-resistance efflux pumps - not just for resistance.*, Nat Rev Microbiol **4** (2006), no. 8, 629–636 (eng).

[78] Bahniuk N Van Scoy B Brown DL Kulawy R Drusano GL Louie A, Grasso C, *The combination of meropenem and levofloxacin is synergistic with respect to both pseudomonas aeruginosa kill rate and resistance suppression*, Antimicrobial agents and chemotherapy **54** (2010), no. 6, 2646–2654.

[79] Collins J Lu TK, *Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy*, PNAS (2009).

[80] Nunn WD Maloy SR, *Genetic regulation of the glyoxylate shunt in escherichia coli k-12.*, J Bacteriol. **149** (1982), no. 1, 173–180.

[81] Roilides E Walsh TJ Meletiadis J, Pournaras S, *Defining fractional inhibitory concentration index cutoffs for additive interactions based on self-drug additive combinations, monte carlo simulation analysis, and in vitro-in vivo correlation data for antifungal drug combinations against aspergillus fumigatus[down-pointing small open triangle]*, Antimicrob Agents Chemother. **54** (2010), no. 2, 602–609.

[82] Kassen R Melnyk A, Wong A, *The fitness costs of antibiotic resistance mutations.*, Evolutionary Applications **8** (2014), no. 3, 273–283.

[83] Beardmore R Meyer JR, Gudelj I, *Biophysical mechanisms that maintain biodiversity through trade-offs.*, Nat Commun **6** (2015), DOI:10.1038/ncomms7278.

[84] Cars O Derendorf H Drusano GL Mouton JW, Dudley MN, *Standardization of phamacokinetic/pharacodynamic (pk/pd) terminology for anti-infective drugs: an update.*, Journal of Antibiotic Chemotherapy **55** (2005), 601–607.

[85] Hervé Nicoloff, Vincent Perreten, Laura M McMurry, and Stuart B Levy, *Role for tandem duplication and lon protease in acrab-tolc- dependent multiple antibiotic resistance (mar) in an escherichia coli mutant without mutations in marrab or acrrab*, Journal of Bacteriology **188** (2006), no. 12, 4413–4423.

[86] Spande TF Shiloach J Noronha SB, Yeh HJ, *Investigation of the tca cycle and the glyoxylate shunt in escherichia coli bl21 and jm109 using 13c-nmr.*, Biotechnol Bioenge **68** (2000), 316–327.

[87] Scott L. Nuismer, Benjamin J. Ridenhour, and Benjamin P. Oswald, *Antagonistic coevolution mediated by phenotypic differences between quantitative traits*, Evolution **61** (2007), no. 8, 1823–1834.

[88] Sorkine O, *Least-squares rigid motion using svd*, Technical notes, ETH Zurich, February 2009.

[89] Pusch GD Olsen GJ Davis JJ Disz T Edwards RA Gerdes S Parrello B Shukla M Vonstein V Wattam AR Xia F Stevens R Overbeek R, Olson R, *The seed and the rapid annotation of microbial genomes using subsystems technology rast.*, Nucleic Acids Research (2014).

[90] Ehrlich P, *Address in pathology, on chemiotherapy: Delivered before the seventeenth international congress of medicine.*, BMJ **2** (1913), no. 2746, 353–359.

[91] Harpe SE Johnson C Edmond M Polk RE Pakyz A, Powell JP, *Diversity of antimicrobial use and resistance in 42 hospitals in the United States.*, Pharmacotherapy **28** (2008), no. 7, 906–912 (eng).

[92] Kishony R Palmer AC, Angelino E, *Chemical decay of an antibiotic inverts selection for resistance*, Nat Chem Biol **6** (2010), no. 3, 244–244.

[93] Andersson DI Paulander W, Maisnier-Patin S, *The fitness cost of streptomycin resistance depends on rpsl mutation, carbon source and rpos.*, Genetics **183** (2009), no. 2, 539–546.

[94] Holmes DJ Pompliano L Payne DJ, Gwynn NM, *Drugs for bad bugs: confronting the challenges of antibacterial discovery*, Nature Reviews Drug Discovery **6** (2007), 29–40.

[95] Jansen G Fuentes-Hernandez A Rosenstiel P Schulenburg H Beardmore R Pena-Miller R, Laehnemann D, *When the most potent combination of antibiotics selects for the greatest bacterial load: the smile-frown transition*, PLoS Biol **11** (2013), no. 4, in press.

[96] van Moorsel CJ Kroep JR Bergman AM Ackland SP Peters GJ, van der Wilt CL, *Basis for effective combination cancer chemotherapy with antimetabolites.*, Pharmacol Ther **87** (2000), no. 2-3, 227–253 (eng).

[97] Randall WA Price CW, *Studies of the combined action of antibiotics and sulfonamides.*, Am J Public Health Nations Health **39** (1949), no. 3, 340–344 (eng).

[98] Hall L Quinlan A, *Bedtools: A flexible framework for comparing genomic features.*, Bioinformatics **26** (2010).

[99] Maharjan R, *The form of a trade-off determines the response to competition*, Ecol. Lett **16** (2013), 1267–1276.

[100] Wise R, *The relentless rise of resistance?*, J Antimicrob Chemother **54** (2004), no. 2, 306–310.

[101] Jackson RC, *Amphibolic drug combinations: the design of selective antimetabolite protocols based upon the kinetic properties of multienzyme systems.*, Cancer Res **53** (1993), no. 17, 3998–4003 (eng).

[102] Calvard S Ridler T, *Picture thresholding using an iterative selection method*, Man and Cybernetics **8** (1978), no. 8, 630–632.

[103] Harvey RJ, *Interaction of two inhibitors which act on different enzymes of a metabolic pathway*, Journal of Theoretical Biology **74** (1978), no. 3, 411 – 437.

[104] Deresinski S, *Vancomycin in combination with other antibiotics for the treatment of serious methicillin-resistant Staphylococcus aureus infections.*, Clin Infect Dis **49** (2009), no. 7, 1072–1079 (eng).

[105] Loewe S, *The problem of synergism and antagonism of combined drugs*, Arzneimittelforschung **3** (1953), no. 6, 285–290.

[106] Bennett KM Dotson TL Anderson DJ Vaslef SN Sarraf-Yazdi S, Sharpe M, *A 9-year retrospective review of antibiotic cycling in a surgical intensive care unit.*, J Surg Res **176** (2012), no. 2, e73–8 (eng).

[107] Perrot V Schrag S, *Reducing antibiotic resistance.*, Nature **381** (1996), 120–121.

[108] Forde SE, Beardmore RE, Gudelj I, Arkin SS, Thompson JN, and Hurst LD., *Understanding the limits to generalizability of experimental evolutionary models*, Nature **455** (2008), no. 7210, 220–223.

[109] Holm SE, *Interaction between β-lactam and other antibiotics*, Reviews of Infectious Diseases **8** (1986), S305–S314.

[110] Taylor-Blake B Spellberg B, *On the exoneration of dr. william h. stewart: debunking an urban legend.*, Infectious Diseases of Poverty. **2** (2013), 3.

[111] Seemann T, *Prokka: rapid prokaryotic genome annotation*, Bioinformatics **30** (2014), no. 14, 2068–2069.

[112] Lewis RE Melnick DA Boucher AN Tam VH, Schilling AN, *Novel approach to characterization of combined pharmacodynamic effects of antimicrobial agents*, Antimicrob. Agents Chemother. **48** (2004), no. 11, 4315–4321.

[113] Maragakisb LL Tammaa PD, Cosgroveb SE, *Combination therapy for treatment of infections with gram-negative bacteria*, Clin Microbiol Rev **25** (2012), no. 3, 450–470.

[114] Michel JB Chait R Hartl DL Kishony R Toprak E, Veres A, *Evolutionary paths to antibiotic resistance under dynamically sustained drug selection*, Nat Genet **44** (2012), no. 1, 101–105.

[115] van Valen L, *The red queen lives.*, Nature **260** (1976), 575.

[116] William AC Vogelman B, *Kinetics of antimicrobial activity*, The Journal of Pediatrics **5** (1986), no. 2, 835–840.

[117] Manduru M Bosso JA White RL, Burgess DS, *Comparison of three different in vitro methods of detecting synergy: time-kill, checkerboard, and E test*, Antimicrob Agents Chemother **40** (1996), no. 8, 1914–1918 (eng).

[118] Bonhoeffer S Wiesch PS, Engelstädter J, *Compensation of fitness costs and reversibility of antibiotic resistance mutations*, Antimicrob Agents Chemother. **54** (2010), no. 5, 2085–2095.

[119] Aiden AP Kishony R Yeh PJ, Hegreness MJ, *Drug interactions and the evolution of antibiotic resistance.*, Nat Rev Microbiol **7** (2009), no. 6, 460–466 (eng).