Genetic determinants of antibiotic resistance evolution

by

Marta Lukačišinová

December, 2018

A thesis presented to the
Graduate School
of the
Institute of Science and Technology Austria, Klosterneuburg, Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

The dissertation of Marta Lukačišinová, titled *Genetic determinants of antibiotic resistance evolution*, is approved by:

Supervisor: Tobias Bollenbach, University of Cologne, Cologne, Germany

Signature: _____

Committee Member: Nick Barton, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Jonathan Bollback, University of Liverpool, Liverpool, United Kingdom

Signature: _____

Committee Member: Jürg Bähler, University College London, London, United Kingdom

Signature: _____

Exam Chair: Jozsef Csicsvari, IST Austria, Klosterneuburg, Austria

Signature: _____

I hereby declare that this dissertation is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: _____

Marta Lukačišinová

December 28, 2018

Abstract

Antibiotic resistance can emerge spontaneously through genomic mutation and render treatment ineffective. To counteract this process, in addition to the discovery and description of resistance mechanisms, a deeper understanding of resistance evolvability and its determinants is needed. To address this challenge, this thesis uncovers new genetic determinants of resistance evolvability using a customized robotic setup, explores systematic ways in which resistance evolution is perturbed due to dose-response characteristics of drugs and mutation rate differences, and mathematically investigates the evolutionary fate of one specific type of evolvability modifier - a stress-induced mutagenesis allele.

We find several genes which strongly inhibit or potentiate resistance evolution. In order to identify them, we first developed an automated high-throughput feedback-controlled protocol which keeps the population size and selection pressure approximately constant for hundreds of cultures by dynamically re-diluting the cultures and adjusting the antibiotic concentration. We implemented this protocol on a customized liquid handling robot and propagated 100 different gene deletion strains of *Escherichia coli* in triplicate for over 100 generations in tetracycline and in chloramphenicol, and compared their adaptation rates. We find a diminishing returns pattern, where initially sensitive strains adapted more compared to less sensitive ones. Our data uncover that deletions of certain genes which do not affect mutation rate, including efflux pump components, a chaperone and several structural and regulatory genes can strongly and reproducibly alter resistance evolution. Sequencing analysis of evolved populations indicates that epistasis with resistance mutations is the most likely explanation. This work could inspire treatment strategies in which targeted inhibitors of evolvability mechanisms will be given alongside antibiotics to slow down resistance evolution and extend the efficacy of antibiotics.

We implemented a stochastic population genetics model, to verify ways in which general properties, namely, dose-response characteristics of drugs and mutation rates, influence evolutionary dynamics. In particular, under the exposure to antibiotics with shallow dose-response curves, bacteria have narrower distributions of fitness effects of new mutations. We show that *in silico* this also leads to slower resistance evolution. We see and confirm with experiments that increased mutation rates, apart from speeding up evolution, also lead to high reproducibility of phenotypic adaptation in a context of continually strong selection pressure. Knowledge of these patterns can aid in predicting the dynamics of antibiotic resistance evolution and adapting treatment schemes accordingly.

Focusing on a previously described type of evolvability modifier – a stress-induced mutagenesis allele – we find conditions under which it can persist in a population under periodic selection akin to clinical treatment. We set up a deterministic infinite population continuous time model tracking the frequencies of a mutator and resistance allele and evaluate various treatment schemes in how well they maintain a stress-induced mutator allele. In particular, a high diversity of stresses is crucial for the persistence of the mutator allele. This leads to a general trade-off where exactly those diversifying treatment schemes which are likely to decrease levels of resistance could lead to stronger selection of highly evolvable genotypes.

In the long run, this work will lead to a deeper understanding of the genetic and cellular mechanisms involved in antibiotic resistance evolution and could inspire new strategies for slowing down its rate.

Acknowledgments

First of all, I would like to thank Tobias, for his supportiveness, reliability and flexibility. It was always encouraging, when he was relentlessly highlighting what was accomplished, when I could only think of what seemed like never-ending problems. I also thank Tobias for being so extremely reliable in replying to e-mails and setting up meetings and giving fast and thorough feedback. This is especially striking, since the same cannot be said about me. Tobias's flexibility with when and how Martin and I work and arrange family life was instrumental in finishing this PhD.

I would like to thank all members of our group for many things, I will only choose one for each. Guillaume, for being the robot pioneer. Karin, for being an amazing friend, always available for consultation, science related or not. Marjon, for her overflowing honest excitement for science. Veronika, for boldly fighting for the happiness of everyone around her. Marcin, for being an example of how gracefully one can build a star scientific career. Martin, for providing insight into both academic and non-academic troubles, which I encountered during this PhD. Bor, for being an inspiration in how one can bring precision in microbial growth experiments to a whole new level. Andreas and his family, for saving us from loneliness when coming to Cologne. Qi Qin, for all his exceptional linguistic and musical talents, and his unbelievable modesty about them. Booshini, for the great amount of work and dedication she put into experiments that stem from this thesis, despite being overworked already without them. Gabriela, for her unquenchable thirst for everything I have learned during this PhD. Gerrit, Yuval, Cristoph, Janina, Sakshi, Lukas and Sarah, for creating an atmosphere in Cologne, in which it has been a pleasure to work, despite my skepticism at the time of the move.

With great emphasis, I would like to thank Sebastian and Tiago for our collaboration on the paper, which is included here as Chapter 4. Thank you for your patience with the many questions and comments I had, and for involving me in meetings even at times when I was not able to add too much. I learned a great deal from both of you.

A large thank you goes also to IST facilities and the grad school. Traude, for her patience and relentless negotiations with Tecan. Todor and the machine shop for being on call for robot repairs, especially when time for experiments was running out. Hania, for helping us out with several aspects of the move to Cologne.

I would like to thank my thesis committee: Tobias, Jon and Jürg for being there since the beginning and providing support and helpful comments, and special thanks to Nick for joining the committee when the remaining members moved away and thus allowing me to finish my degree at IST.

I would also like to thank many friends at IST, too many to all name here, for countless encouraging discussions at the bridge, for listening again and again to my rants about the broken robot, and for providing an atmosphere at IST that made me really enjoy coming to work.

And, most importantly, I want to thank Martin. For not believing me, when I thought I cannot finish and for being patient through me whining about this. For always making sure I have enough time to write this thesis, which meant taking care of literally everything else that could and needed to be done, for weeks at a time, and still managing to make revolutionary scientific discoveries at night.

This thesis is dedicated to Dorotka.

About the Author

Marta Lukačišinová completed a BSc in Natural Sciences at University College London, UK in 2011. Her specializations were Molecular and Cell Biology and Mathematics and Statistics. During her undergraduate studies, she did a summer internship with Jürg Bähler modelling oxidative stress response in fission yeast. At IST Austria, she continued in combining experimental and theoretical methods. Her research interests focus on mechanisms which potentiate or inhibit the ability of bacteria to become resistant to antibiotics.

List of Publications Appearing in Thesis

1. **Lukačišinová M.**, Bollenbach T. (2017) “Toward a quantitative understanding of antibiotic resistance evolution” Current Opinion in Biotechnology [Review]
2. **Lukačišinová M.***, Novak S.*, Paixão T. (2017) “Stress-induced mutagenesis: Stress diversity facilitates the persistence of mutator genes” PLoS Computational Biology
3. Chevereau G., **Dravecká M.**, Batur T., Guvenek, A, Ayhan D. H., Toprak E., Bollenbach T. (2015) “Quantifying the determinants of evolutionary dynamics leading to drug resistance” PLoS Biology

*equal contribution

Table of Contents

List of Figures

List of Tables

List of Symbols/Abbreviations

DFE Distribution of fitness effects

TET Tetracycline

CHL Chloramphenicol

SNP Single nucleotide polymorphism

IS Insertion sequence

1 Introduction

Some parts of this chapter are heavily inspired by the published review article (Lukačišinová and Bollenbach, 2017).

1.1 Resistance to antibiotics occurs as a result of genomic mutation

Antibiotics are clinically useful molecules that can kill bacteria or inhibit their growth. This is often achieved by targeting structures or functions which are specific to bacterial cells. Such targets include the cell wall (e.g. betalactams), the bacterial ribosome, RNA polymerase, and specific metabolic enzymes such as dihydrofolate reductase in the case of trimethoprim (Kohanski et al., 2010; Walsh, 2003; Yonath, 2005). By being specific to bacterial cells and not toxic for humans and animals, they are ideal for the treatment of bacterial infections. Medical uses of antibiotics are becoming more challenging due to the rise of drug resistance (Blair et al., 2015; Levy and Marshall, 2004). In order to successfully fight bacterial infections despite antibiotic resistance development, the process of resistance acquisition and spread needs to be understood in detail.

Mutations are changes in the genetic makeup of the cell. By modifying the DNA, the changes get inherited by the next generation, potentially causing lasting changes to whole populations. Initially, mutations happen spontaneously due to DNA damage or erroneous replication. Mutations can be single nucleotide polymorphisms (SNPs), deletions or insertions of just one or few nucleotides, but also larger rearrangements when long pieces of DNA get deleted, inverted or inserted/duplicated, either from another place in the genome, from the environment or from another cell (Koonin et al., 2001). In this thesis, I will focus on mutations that do not require the uptake of foreign DNA.

In the classical paradigm, mutations happen spontaneously and randomly and do not depend on the environment (Luria and Delbrück, 1943). Most mutations are neutral with respect to growth rate and survival (Eyre-Walker and Keightley, 2007a). In harsh conditions, rare variants with strong beneficial effects can establish in populations at the expense of others, producing the impression of an orchestrated response to challenges (Luria and Delbrück, 1943).

Rare beneficial mutations which can render bacterial cells resistant to antibiotics can do so through several main mechanisms (Blair et al., 2015). Target modification mutations change the target protein in such a way that the antibiotic cannot bind (Palmer et al., 2015). Antibiotic degrading enzymes can deactivate the drug (Jacoby, 2009). Certain mutations modify structures on the membrane and cause a decrease in influx of antibiotic (e.g. mutations in outer membrane porins), or an increase of efflux (e.g. increased expression of efflux pumps) which decreases the intracellular concentration of the antibiotic (Fernandez and Hancock, 2012).

A particularly common mechanism of resistance against many antibiotics and other harmful molecules is the elevated expression or modification of efflux pumps. Efflux pumps are protein complexes spanning the bacterial membranes. They are powered either by ATP or the proton motive force to pump out harmful molecules out of the cell (Du et al., 2018). The deletion of certain pumps has little effect on growth or survival in conditions where few toxins are present. However, in a range of stresses including ethidium bromide, ribosomal or cell-wall synthesis inhibitors, the deletion of any of the pump components (e.g. the

AcrA/B-TolC pump) causes severe fitness defects (Nichols et al., 2011). In studies investigating resistance acquisition in the lab and in the clinic, mutations related to the AcrA/B-TolC and other efflux pumps repeatedly come up spontaneously (Anes et al., 2015; Fernandez and Hancock, 2012). What is more, the mutations do not confer resistance only to the antibiotic that the population has been exposed to, but to a wider range of drugs (Lázár et al., 2014).

The level of antibiotic resistance conferred by resistance mutations is typically measured by the minimal inhibitory concentration (MIC) for a given antibiotic and bacterial strain. The MIC is the minimal concentration of antibiotic at which the given bacterial strain does not grow. An increase in resistance occurs when the population can grow in higher concentrations of antibiotic. Another common measure of resistance is the IC_{50} (50% inhibitory concentration), which is the concentration at which the exponential growth rate of the strain is half of what it is without any antibiotic present. When the whole dose-response curve is measured, this provides the entire growth rate (and death rate in the case of bactericidal drugs) profile corresponding to a wide antibiotic concentration gradient (Regoes et al., 2004). To extract summary parameters from these curves, most often a hill function

$$g = \frac{g_{\max}}{1 + \left(\frac{c}{c_0}\right)^n}$$

is used, where g_{max} corresponds to the maximal growth rate in the absence of drug, n roughly determines the steepness of the sigmoidal function and c_0 denotes the IC_{50}. One advantage of using the whole growth profile along the antibiotic gradient is that concrete predictions can be made about the population dynamics of mixed populations during particular treatment regimens (Yu et al., 2018).

Although resistance is by definition a genetic trait, the level of resistance is often not entirely determined genetically, but can be heterogeneous within a population, depend on the environment, on the population structure, or on the physiological state of the cell. In addition, usual methods of resistance measurements make it difficult to distinguish between resistance and other forms of decreased sensitivity to antibiotics including tolerance and persistence (Brauner et al., 2016). In this thesis, the focus is on the classic narrow definition of resistance brought about by mutation, causing a lasting heritable growth advantage in the presence of antibiotic.

1.2 Antibiotic resistance acquisition can be studied using experimental evolution

Experimental evolution is a research approach taken to study evolutionary processes in real-time and consists of propagating experimental populations in defined conditions for many generations and measuring changes which occurred as a result (Kawecki et al., 2012). Its roots can be traced all the way to breeding efforts in agriculture or with domesticated animals, where selective conditions (particularly selection of individuals for mating) were designed in a controlled way. The goal in breeding was to channel the changes happening in domesticated organisms to suit particular needs of the breeder. Later, similar approaches were taken to elucidate basic research questions. For example the Illinois long-term selection experiment attempts to test the limits of selection for oil and protein content in

maize, and has been running almost without interruption since 1896 (Dudley and Lambert, 2010). In these early experiments, the focus was mainly on measuring phenotypes, describing their changes and exploring selection limits.

The re-discovery of Mendel's laws and its reconciliation with Darwin's theory of natural selection turned the mainly descriptive study of natural variation into a research field of detailed mechanistic explanations and testable quantitative predictions (Huxley et al., 2009). The transition was aided, for example, by controlled experimentation with fruit flies by Morgan focusing on the mechanisms of heredity (Morgan, 1915) and rigorous mathematical treatment of population genetics pioneered by Wright, Haldane, Fisher and others (Crow, 1987). The modern synthesis paved the way for quantitative genetics becoming an established field. With the advent of advanced genotyping technologies including whole-genome sequencing a whole new class of theoretical predictions were ready to be put to the test by controlled laboratory experiments (Desai, 2013; Kawecki et al., 2012).

Due to their short generation times, small genomes, large population sizes and relative ease of handling and genetic manipulation, microbes have been especially suited for observing evolutionary processes in the laboratory and testing long-standing theoretical predictions (Van den Bergh et al., 2018). Questions such as the plausibility of parallel evolution, rules governing the effects of mutation combinations, the spontaneous occurrence and success of hyper-mutators, the advantages of sexual reproduction or evolution of multicellularity and cooperation, have all been investigated by evolution experiments in microbes (Kawecki et al., 2012).

The same experimental evolution techniques which can be used to systematically investigate general evolutionary phenomena can also be used to study the evolution of antibiotic resistance. The general setup of such an experiment is to expose a bacterial population to an antibiotic for an extended period of time and observe its adaptation by measuring its fitness (in competition with ancestor, or by measuring other related parameters such as survival rate, exponential growth rate and resistance level)(Hegreness et al., 2008; Toprak et al., 2012). By combining phenotypic measurements with inexpensive whole-genome sequencing of evolved strains, a detailed description of antibiotic resistance evolution emerges, and questions of its reproducibility, speed, molecular origins, and constraints can be addressed in a systematic fashion (Lukačišinová and Bollenbach, 2017; MacLean et al., 2010).

The simplest experimental evolution protocol that is useful for studying resistance evolution is the serial transfer protocol (Elena and Lenski, 2003). In this protocol, bacterial cultures grow in flasks or on microtiter plates and are diluted into fresh medium by a fixed factor at regular time intervals (e.g. every 24 hours). These experiments can be continued virtually indefinitely: Richard Lenski's seminal long-term evolution experiment (Good et al., 2017) has exceeded 60 thousand generations in 28 years and is still ongoing. Due to the relative simplicity of a serial transfer protocol, it is feasible to run hundreds of evolution experiments in parallel. Together with increasingly inexpensive whole genome sequencing techniques (Baym et al., 2015), this opens the door for a statistical investigation of the intrinsically stochastic evolutionary dynamics and for identifying general principles governing microbial evolution (Desai, 2013; Levy et al., 2015). A drawback of serial transfer protocols is their inability to keep key parameters that affect the evolutionary process well-

controlled: the population size fluctuates considerably and cultures differ in the time they spend in stationary phase and in their growth rates; this complicates the quantitative investigation of the evolutionary process and its comparison among different cultures. Furthermore, it is not straightforward how the antibiotic concentration should be chosen to gain maximum insight into the process of resistance evolution: if it is too low, there is virtually no selection for resistance; if it is too high, cells cannot grow at all, preventing them from evolving at a significant rate.

Recently developed techniques in which bacteria are exposed to increasing antibiotic concentrations solve this problem. Theoretical work suggested that temporal and spatial selection gradients can facilitate the sequential emergence and fixation of multiple resistance mutations leading to increasingly higher resistance levels (Greulich et al., 2012; Hermsen et al., 2012). Consequently, advanced protocols that gradually increase antibiotic concentrations in time or space have been developed (Baym et al., 2016a; Toprak et al., 2013). A notable example is the "morbidostat": this feedback-controlled device keeps cultures growing in exponential phase and automatically increases the antibiotic concentration during the experiment such that they keep growing at a pre-defined rate despite their increasing resistance. In this way, strong selection pressure for resistance is constantly maintained. For some antibiotics, this protocol enabled the highly reproducible evolution of a ~1000-fold resistance increase in just a few weeks (Toprak et al., 2012).

Spatial antibiotic gradients have also been used to speed up evolution experiments. A recent study followed evolution on a huge, meter-scale agar plate (the "MEGA plate") (Baym et al., 2016a). In contrast to small agar plates where rapid diffusion quickly destroys spatial drug concentration gradients, they remain relatively stable on this larger plate. Further, the size of the plate allows for large bacterial population sizes that should accelerate the occurrence of resistance mutations. Fast resistance evolution reaching extremely high levels within weeks was observed for different antibiotics (Baym et al., 2016a).

Apart from experiments where the antibiotic concentration increases monotonously, resistance evolution has been studied under different temporal sequences of single antibiotics or in various combinations. For example, an evolution experiment where bacteria were exposed to an antibiotic for defined lengths of time interspersed with periods of uninhibited growth enabled the evolution of antibiotic tolerant strains, which developed longer lag times before starting exponential growth which matched the length of antibiotic exposure. This was an extremely successful survival strategy, since the antibiotic used only kills growing cells (Fridman et al., 2014). Several studies also explored the evolution of strains exposed to antibiotic exposure regimes combining various drugs simultaneously (Hegreness et al., 2008) or in sequence (Kim et al., 2014).

A challenge remains to combine the rigor of setups with highly controlled evolutionary parameters, such as the morbidostat, with the statistical power of simple serial transfer protocols, in order to successfully perform truly controlled evolutionary screens.

1.3 Mutation and selection are the main forces determining evolutionary dynamics

In this thesis, we are mainly concerned with mechanisms affecting the evolvability of antibiotic resistance. Evolvability, in this context, is the ability of a particular genotype to develop antibiotic resistance under antibiotic selection. To understand the ways in which a

genotype can affect resistance evolvability, we need to describe the major forces that affect adaptation.

Evolutionary dynamics in general are influenced by the forces of mutation, selection, and random drift. Their interplay in particular conditions can be described using mathematical models, and quantitative measurements. Both are heavily used in researching antibiotic resistance in order to describe, predict and eventually control its spread (Furusawa et al., 2018; Lukačišinová and Bollenbach, 2017; Palmer and Kishony, 2013).

The speed of resistance evolution can be defined at the phenotypic and genotypic levels. The rate of phenotypic adaptation can be measured as the change in fitness over time (Wilke, 2004). In the context of resistance evolution, this would be the change in competitive advantage, growth rate or survival under antibiotic exposure. The speed of evolution at the genotypic level is given by the fixation or substitution rate. Fixation is the event of a new mutation reaching 100% frequency in a population. The fixation (or substitution) rate is the rate at which fixation events occur (Hartl and Clark, 2007). This measure is related to the phenotypic adaptation rate, though their values can be very different. As will be explained below, populations under no selection pressure at all, would still have a substitution rate and possibly high (Kimura, 1962). Briefly, this is because pure randomness in the number of surviving offspring (termed genetic drift) can cause a relative increase in the frequency of a particular genotype which, in the long-run, can lead to a fixation event.

1.3.1 Mutation rate strongly influences evolutionary dynamics

An important factor that sets the speed of evolution is mutational supply. Spontaneously generated genetic variants provide the "raw material" of adaptation. The supply of mutations is simply the product of the per-individual mutation rate μ and the population size N. This product is important in determining the rate of adaptation under most conditions (Hall et al., 2010; Wahl et al., 2002)

In the absence of selection (under the neutral model of evolution), the mutation rate sets the speed of adaptation independently of population size (Hartl and Clark, 2007). Under the neutral model, random fluctuations in progeny numbers and the irreversibility of extinction/fixation of any particular mutation can cause the genetic makeup of the whole population to change over time. In this case, the fixation probability, i.e. the probability the mutation will reach 100% frequency, for any particular mutation is given by its frequency in the population. For new spontaneous mutations as they occur, this is 1/N. Since one individual mutates with rate μ, the population sees new mutations at the rate Nμ and the rate of fixation of new mutations in the population is μ. So under the neutral model, the rate of evolution (more precisely fixation of new mutations) is exactly set by the mutation rate.

High mutation rates, however, can also slow down fixation rates. When mutation supply is too high, especially when μ is comparable to 1/N or higher, clonal interference occurs (Gerrish and Lenski, 1998). Clonal interference means there is competition among lineages with beneficial mutations which are spreading through the population. This competition can prevent a particular fixation event if, during the spreading of a mutation through the population, a fitter mutation appears in a different lineage and overtakes. Such dynamics can lead to a situation where, for stretches of time no mutations fix, followed by a fixation

of a particular lineage with one or more strong beneficial mutations. Due to the high mutation rate, it is common that the successful lineage would carry many other mutations apart from the beneficial one(s). Mutations which fix on the background of a different beneficial mutation are termed hitchhiker mutations (Barton, 2000; Lang et al., 2013; Raynes et al., 2012; Tenaillon et al., 1999).

Since high mutation rates also mean high deleterious mutation rates, in the extreme case, they can cause an extinction of the population by disrupting functions essential for survival in every individual. This phenomenon is called error catastrophe and has been described theoretically as well as observed empirically in certain viruses (Crotty et al., 2001; Eigen and Schuster, 1977; Pariente et al., 2005).

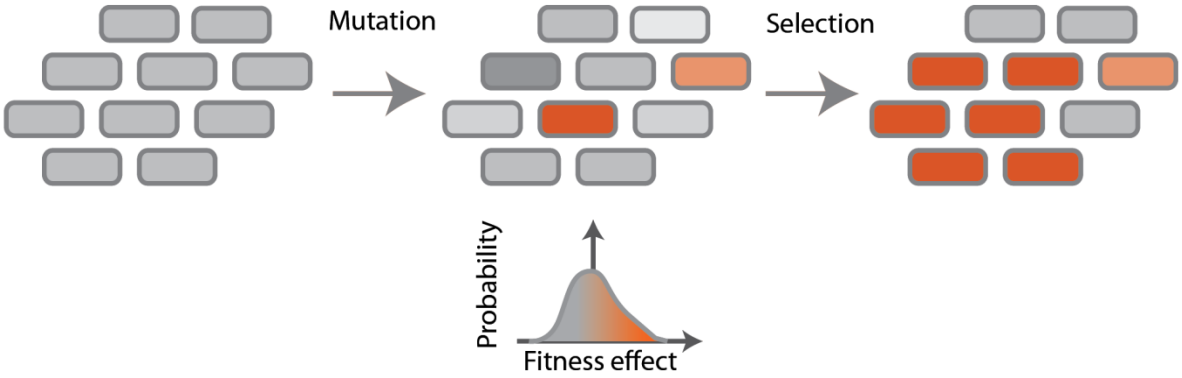


Figure 1 ***Scheme of resistance evolution dynamics***

A clonal bacterial population becomes genotypically diverse due to mutations. Depeding on the environment and genetic background, the fitness effects of these new mutations come from a corresponding distribution of fitness effects (DFE). In the scheme, fitness goes from grey (unfit) to orange (fit). Due to selection, more fit individuals increase in frequency at the expense of less fit individuals.

1.3.2 The distribution of fitness effects strongly influences evolutionary dynamics

The second major force determining evolutionary dynamics apart from mutation is selection. The relation between mutation and selection is given by the distribution of fitness effects (Figure 1). The distribution of fitness effects (DFE) is the distribution determining the probability of a new mutation having a certain effect on survival and/or growth (fitness) (Eyre-Walker and Keightley, 2007b). In particular, it gives the proportion of lethal, deleterious, neutral and beneficial mutations available to a population. The DFE is defined for a particular genotype in a particular environment. Although its form can vary considerably, and can be prohibitively laborious to determine, several simplifications are common to assume (Orr, 2005).

One common framework for determining characteristics of the DFE is Fisher's Geometric Model (FGM). In this framework, mutations represent small steps in a multi-dimensional phenotypic space. In this space a simple smooth concave fitness function is defined (Fisher, 1930). In particular, there is one combination of phenotypic values which represents the optimal phenotype (maximal fitness). The further one is from the optimum, the lower the fitness. In this model, the distribution of fitness effects would be given by the immediate neighborhood of a point in the phenotype space and the corresponding fitnesses. In other

words, mutations generally cause a predictable change in phenotype, which can then be translated to fitness through a smooth function. It can be shown in this framework, that as populations approach an optimum, the proportion of beneficial mutations decreases. Also, given the concave shape of the phenotype-fitness function, as populations approach the optimum, mutational effect sizes tend to decrease. FGM has been successfully used to describe this dynamic in microbial evolution experiments (Harmand et al., 2016; Tenaillon, 2014; Trindade et al., 2012).

The shape of the DFE, in particular its width was shown to affect the rate of evolution. In particular, Fisher's fundamental theorem states that "The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time" (Fisher, 1930; Li, 1967). The genetic variance in a population is intimately linked to the fitnesses of mutations available to it. If the DFE is extremely narrow, the variance in fitness will be severely limited. If the DFE is wide, the more extreme fitnesses will be possible, speeding up adaptation (Chevereau et al., 2015; Li, 1967).

A study argued, that for asexual populations, there is little importance in the particular type of underlying distribution, and the effect on adaptation rate can be condensed into one parameter (Hegreness et al., 2006). In either case, the typical increase in fitness depends on the mutational effects available to the population and is characterized either by the variance or its mean effect (Hegreness et al., 2006; Li, 1967).

In the context of antibiotic resistance, although some DFEs have been experimentally measured (Jacquier et al., 2013; Schenk et al., 2012), a systematic study comparing the DFEs for various drugs has been lacking.

1.4 Both mutation rates and the distribution of fitness effects are partly determined by the genotype

The basic parameters of evolution which we focus on here, namely mutation rate and the distribution of fitness effects, depend on genotype. Some genotypes which affect resistance evolvability are described here, but there may be many more that remain to be discovered.

1.4.1 The genotype can determine the mutation rate

Mutation rates can vary by orders of magnitudes for different bacterial species (Lynch, 2010) but also for different genotypes within the same species. Mutations which increase mutation rate can be commonly found in clinical, environmental and laboratory settings (Sniegowski et al., 1997; Taddei et al., 1997). Most often these are loss of function mutations in mismatch repair genes, for example *mutS* and *mutL*. Mismatch of bases between the two DNA strands happens often due to damage of one of the strands, mismatch repair recognizes such places and exchanges the appropriate base. When mismatch repair has a defect, the wrong base from the damaged strand copies into the other and cements in the mutation, which can now also be inherited to the next generations. In such a way, defects in *mutS* and *mutL* genes can cause mutation rates to increase 100 to 1000-fold (Schofield and Hsieh, 2003).

Despite the costs of high mutation rate, mutators can have a strong advantage in certain situations and achieve high frequencies in a population (Chao and Cox, 1983). Mutators have the greatest advantage in contexts when the population is maladapted to the environment, i.e. when mutations can bring a large fitness benefit. When a strong beneficial

mutation happens in a mutator background, and rate of recombination is very low (or zero, in purely asexual populations), the mutator allele can hitchhike with the beneficial mutation and increase in frequency (Tenaillon et al., 2001).

The dynamics of adaptation as a function of mutation rate has been studied in various contexts. However, little attention has been given to the variability of evolutionary outcome as a function of mutation rate.

1.4.2 Stress-induced mutagenesis is a type of mutation rate variation with unique potential to influence antibiotic resistance evolution

There are multiple known mechanisms that result in elevated rates of general mutagenesis or an increase in the rate of specific genetic changes during stress. These mechanisms are often referred to as stress-induced mutagenesis (SIM) (Bjedov et al., 2003; Rosenberg, 2001). When encountering a range of environmental stresses, several species of bacteria activate SOS responses that – in addition to stimulating various repair mechanisms – activate error-prone DNA polymerases, which have been linked to a faster evolution of antibiotic resistance (Bjedov et al., 2003; Cirz et al., 2005; Do Thi et al., 2011). This activation of error-prone DNA polymerases in response to a wide range of environmental stresses is a thoroughly studied SIM mechanism (Bjedov et al., 2003; Devon M. Fitzgerald et al., 2017). This mechanism has been linked to faster evolution of antibiotic resistance (Cirz et al., 2005).

Since higher mutation rates are beneficial especially when adaptive potential is high, stress-induced mutagenesis has the potential of providing higher mutation rates when they are beneficial, but not incurring them when organisms are well adapted (not stressed) and the cost of high mutation is greater. This intriguing dynamic can be especially relevant for antibiotic resistance evolvability, since many antibiotics upregulate known SIM mechanisms (Cirz et al., 2005). The evolutionary dynamics of such alleles are underexplored, especially in the context of more complex antibiotic treatment schemes.

1.4.3 Mutational effects depend on genetic background

The DFE generally depends on the genetic background; it can thus change as soon as the first mutation has occurred. The general phenomenon where the effect of mutations depends on the presence of other mutations is termed “epistasis” (Poelwijk et al., 2016). Epistasis can be “global” and affect the whole distribution of fitness effects in a systematic way (Kryazhimskiy et al., 2014; Otwinowski et al., 2018). It can also happen among individual pairs or small groups of mutations largely independently of other mutations. Measuring the extent of epistasis is important for evolutionary predictions because pervasive epistasis often leads to multiple fitness peaks and can prevent a population from reaching the global fitness maximum (de Visser and Krug, 2014); in particular, this is the case for reciprocal sign epistasis where the effect of a mutation changes sign depending on the background (Poelwijk et al., 2011). It can also constrain phenotypic adaptation to reproducible outcomes (Kryazhimskiy et al., 2014).

Epistasis can be analyzed using discrete fitness landscapes which are a powerful concept for assessing the constraints and predictability of mutational paths in evolution experiments (Palmer et al., 2015; Weinreich et al., 2006). A discrete fitness landscape is a graph in which the vertices are genotypes, and two genotypes are connected by an edge if they are a single

mutational event apart. The landscape is completed by assigning a fitness value to all genotypes. Paths on the landscape are accessible by natural selection if they represent sequences of genotypes with monotonically increasing fitness, i.e. all mutations along the path are beneficial. If only few of the possible paths are accessible, evolutionary trajectories become more constrained and predictable. The fitness of all mutations and their combinations is experimentally inaccessible due to the astronomically large number of possible genotypes, even for short sequences. Therefore, studies have focused on full landscape reconstructions of just a few mutations relevant for drug resistance (Gabryszewski et al., 2016; Palmer et al., 2015) and proposed biophysical models to predict epistatic interactions from protein structure and function (Figliuzzi et al., 2016; Rodrigues et al., 2016).

Certain specific ways in which epistasis affects resistance evolution have been described (Gifford et al., 2018; Vogwill et al., 2016). However, a more systematic exploration of how various cellular mechanisms can interact with potential beneficial mutations and thus constrain adaptation is needed.

1.4.4 Aims of this thesis

- Develop a method to quantitatively compare dozens of different genotypes in how well they can evolve resistance to antibiotics (Chapter 2)
- Identify genes that strongly modulate resistance evolvability (Chapter 2)
- Use population genetic simulations to explore the effects of general dose-response characteristics and mutation rate on the evolvability of antibiotic resistance (Chapter 3)
- Find conditions under which a specific evolvability modifier (a stress-induced mutagenesis allele) can persist in a bacterial population exposed to periodic stresses. (Chapter 4)

2 Genetically perturbing the evolvability of antibiotic resistance

2.1 Introduction

A promising but under-developed approach to counter antibiotic resistance evolution is the identification of genetic determinants and cellular mechanisms that do not immediately increase a pathogen's resistance but increase its ability to evolve. In general, evolvability, the ability of a genotype to evolve resistance when exposed to antibiotics, can be compared among different genotypes by exposing them to the same well-defined selection pressure and comparing their evolutionary outcomes. These outcomes can depend strongly on the mutation rate of the starting genotype or the distribution of fitness effects of new mutations in that particular background.

A prominent example of a mutation rate altering mechanism which can be inhibited to decrease evolvability is stress-induced mutagenesis. It is a mechanism in which bacteria increase their mutation rate when exposed to certain stresses including antibiotics (Cirz et al., 2005; Devon M. Fitzgerald et al., 2017). Furthermore, highly-mutating strains have often been observed in experimental evolution or even clinical settings, with abundant evidence that higher mutation rates can significantly speed up resistance evolution (Méhi et al., 2013, 2014; Sniegowski et al., 1997). In principle, targeted genetic perturbations could affect the fitness effects of mutations just like they affect the mutation rate. However, the systematic investigation of genetic determinants that alter the evolvability of antibiotic resistance beyond mutation rate has so far been limited (Gifford et al., 2018; Vogwill et al., 2016).

It is clear that mutational effects of new mutations often depend on genetic background – the additive model of mutation effects is, although useful as a null hypothesis, a simplification. It is not clear to what extent these effects vary in systematic, predictable ways, and how this is relevant in particular to antibiotic resistance evolution. One type of global epistasis is diminishing returns epistasis, where effects of beneficial mutations weaken as fitness approaches a maximum. This is a type of antagonistic epistasis, where beneficial mutations have a smaller beneficial effect on backgrounds with higher fitness, which would also explain the common pattern from evolution experiments where fitness seems to level-off with time, albeit not exactly reaching a plateau (Wiser et al., 2013). Such a pattern could be explained by the "running-out" of beneficial mutations where the most potent are more likely to fix first. Alternatively, the same beneficial mutations can have a lower effect when appearing on a fitter background – diminishing returns epistasis. Several recent studies confirm the second explanation in various contexts, including antibiotic resistance evolution (Barrick et al., 2010; Couce and Tenaillon, 2015; Kryazhimskiy et al., 2014). Also, the diminishing returns pattern among beneficial mutations along an adaptive trajectory as well as among randomly generated mutations has been observed (Schenk et al., 2013; Schoustra et al., 2016). Diminishing returns epistasis is a pervasive phenomenon and a type of "global epistasis" which, if well characterized, allows predictions of evolutionary outcomes without prohibitively painstaking evaluation of all possible mutations and their combined effects.

A specific cellular mechanism that can have a large effect on the fitness effects of many mutations and consequently on evolvability is protein folding. The effects of mutations in protein coding sequences often act through a change of the structure of the resulting protein and this in turn depends to a large extent on the activity of protein-folding

chaperones (Rutherford and Lindquist, 1998). In particular, chaperones provide a buffering mechanism, by enabling the folding of peptide chains into their usual conformations despite variability in sequence. Their main role in evolvability is thus thought to be through masking phenotypic variability (Queitsch et al., 2002; Rutherford, 2003). In certain conditions, the chaperones may not fold efficiently, for example in cases of severe heat shock, when protein misfolding is more abundant. When the chaperone action is insufficient to fold proteins properly, the masked variability is exposed for selection to act on it, possibly increasing the speed of adaptation (Rutherford, 2003). Apart from serving as phenotypic buffers, chaperones also contribute to the folding of proteins into a functional conformation and without them these proteins could become irreparably dysfunctional. In other words, the deletion of a chaperone can expose beneficial phenotypic variation, but can also heavily skew phenotypes of coding mutations to loss of function. Both have been observed in experimental studies (Aguilar-Rodríguez et al., 2016; Kadibelban et al., 2016).

Apart from the more systematic epistasis phenomena, several idiosyncratic accounts of epistatic interactions with beneficial mutations with considerable effect on the course of adaptation have been reported in various microbial evolution experiments. One of the reported phenomena is that a mutation with relatively small beneficial effect (or even a cost) provides a background on which some mutations have a disproportionately strong positive effect, sometimes leading to a higher fitness than can be achievable without it (Blount et al., 2008). Particular examples relating to antibiotic resistance include several strong sign-epistasis interaction between mutations in particular genes important for drug resistance: Beta-lactamase (Jacquier et al., 2013) and DHFR (Palmer et al., 2015). Another notable example is the observation that the deletion of a global transcriptional regulator ampR decelerates adaptation to ceftazidime in *Pseudomonas Aeruginosa* (Gifford et al., 2018). It is conceivable that, similarly to the case in Gifford et al. there exist many more perturbations to transcriptional regulators which, by changing the expression profile of the cell, also modify the phenotypes of many possibly crucial mutations. To extend these accounts and discover new mechanisms which modify resistance evolvability, a protocol that can, in high-throughput, compare evolutionary outcomes of a broad range of starting genotypes is needed.

A quantitative investigation of antibiotic resistance that could be used to fairly compare a broad range of mutants in how well they evolve, requires high replication and protocols that tightly control key evolution parameters. A morbidostat-like protocol (see Section 1.2) is well-suited to this purpose, especially if done in higher throughput. Firstly, the population size and growth rate are tightly controlled as opposed to simpler serial transfer protocols. The disadvantages of a simple serial transfer setup include that it is difficult to control the population size precisely and thus also the mutational availability. In addition, the population is in a complex dynamic environment, often including lag phase, exponential growth and entry to stationary phase, as well as severe bottlenecks. Therefore, selection is complex and thus mutational effects can be hard to analyze. In the morbidostat, cells are growing exponentially throughout the experiment, allowing for a more tightly controlled environment, which is likely to lead to higher reproducibility. Second, the protocol tightly constrains selection to be purely for antibiotic resistance and thus allows fairer comparison between different strains and along longer evolutionary trajectories. A common problem in protocols with a fixed antibiotic concentration is that the first resistance mutation comes up quickly, increases the growth rate to close to no-drug levels, and then the selection pressure

for resistance is comparable to the pressure to grow better in the medium itself. In a morbidostat protocol, since the antibiotic concentration is tuned to increase the pressure as resistance develops, longer mutational paths can be observed and compared. The disadvantage of a classical morbidostat is that it is relatively difficult to perform in high-throughput since each experiment is a tube which needs to be fabricated and maintained separately with its accessories and measuring equipment (Toprak et al., 2013). Biofilms often form on the sides of the flask, introducing undesired spatial structure. A realistic number of experiments that can be done in parallel is in the lower tens. However, with the use of a liquid-handling robot a similar protocol is conceivable in 96-well plates increasing the throughput (and size of screen) drastically.

Here, we aim to identify genes that affect drug resistance evolution, new mutational paths to resistance, and ways of blocking or unlocking these paths. To this end, we developed a high-throughput experimental evolution protocol that tightly controls both population size and selection pressure. We quantified the evolvability of ~100 *Escherichia coli* single gene deletion strains and revealed a general trend of "diminishing returns" where more sensitive genotypes increase in resistance faster than resistant ones. We identified several genes that have a drastic effect on antibiotic resistance evolvability via altering epistatic relationships with common resistance mechanisms. Our approach lays the foundation for improved drug treatments with slower resistance evolution.

2.2 Results

2.2.1 Automated high-throughput evolution protocol reproducibly measures evolvability in highly controlled conditions

To quantify the dynamics of evolution for many different genotypes and replicates, we developed an automatized experimental evolution platform using a dedicated robotic system that keeps hundreds of cultures in tightly controlled and monitored conditions (Figure 2). In order to maintain high selection pressure for antibiotic resistance over longer periods of time, antibiotic concentration is periodically adjusted. Every 3-5 hours, cultures are transferred to new 96-well plates where the volumes of medium, drug and culture are individually tuned to keep each population in exponential phase and at 50% inhibited growth (Figure 2). Keeping the cultures in rich medium and in exponential phase assures that the populations are not under nutrient limitation. This together with continually adjusting the antibiotic concentration assures that the selection pressure is strong and very specifically aimed at faster growth in the antibiotic. The strong selection pressure allows relatively rapid evolution, where up to 100-fold increases in resistance can be observed over a period of 10 days.

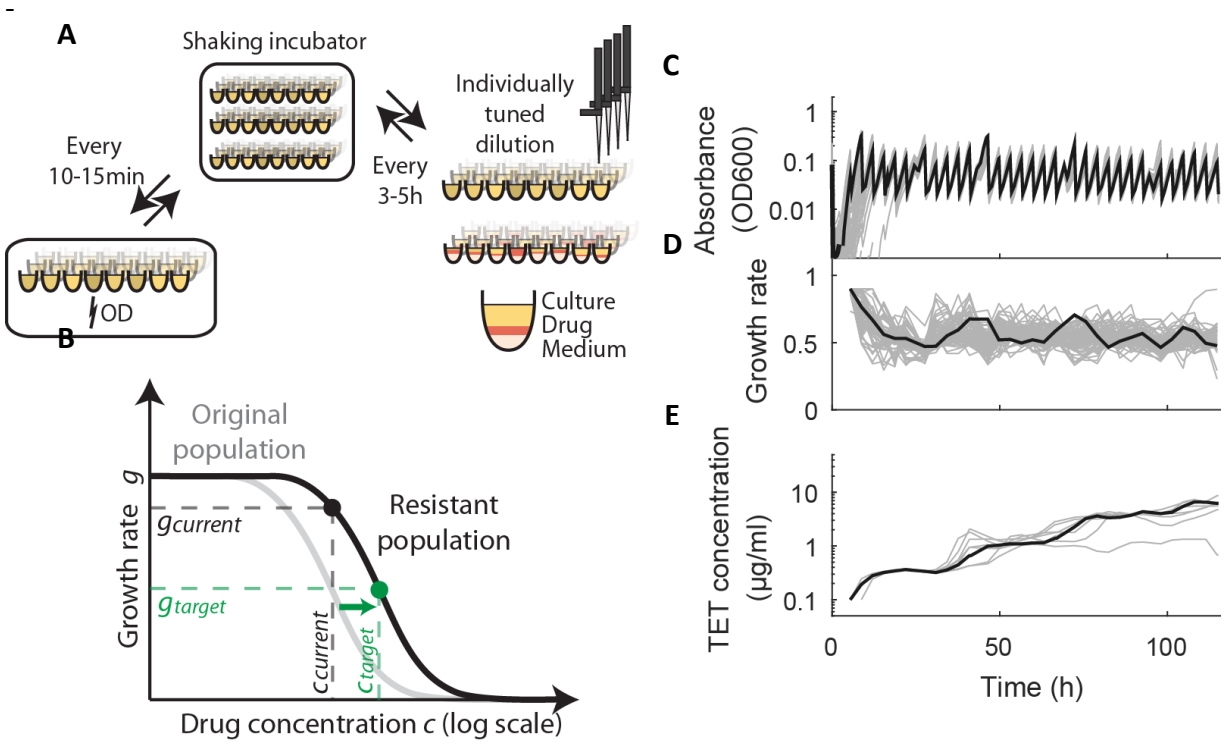


Figure 2 ***High-throughput lab evolution with controlled population size and selection pressure leads to reproducible fast evolution of antibiotic resistance.***

A) Scheme of lab evolution protocol. 96-well plates are shaken in the incubator, the absorbance is measured every 10-15min in a plate reader, and the cultures transferred to new plates every 3-5h. When diluting, the volumes of culture, drug and medium are individually tuned based on the OD values, such that the OD after dilution and the growth rate is always close to the predetermined level (Methods). B) Each dilution, the target antibiotic concentration c_{target} is calculated based on a dose response curve that maintains the same shape as the dose response curve of the ancestral population, but shifts on the log drug concentration axis as the population gets resistant. The growth rate since the last dilution and antibiotic concentration in the given well ($g_{current}$ *and* $c_{current}$*) then define the curve from which* c_{target} *can be calculated. C) Background subtracted OD values a log scale over the course of the experiment for 82 wells from a 96-well plate. The cultures are clearly in exponential phase continuously. D) Growth rates fitted to the OD during the experiment. Values are normalized to growth of the parental strain in no drug. E) The TET concentration in the wells of the parent strain replicates during the experiment. The same replicate of the parent strain of the Keio collection is highlighted in black on all three plots (C-E). All values are from plate 1 of experiment M1 (See Methods)*

Despite simplifying assumptions used in the protocol, the setup is successful at keeping the OD and growth rate within tight bounds throughout the experiment. The antibiotic concentration is adjusted at every dilution step based on the assumption that mutations, initial single gene deletions as well as mutations acquired during the experiment, only change the level of resistance (IC_{50}), but otherwise leave the dose-response curve unchanged. The wild-type values are used for the other two parameters of the assumed hill-shaped dose-response curve, namely, growth rate without drug and the hill-coefficient or "steepness" parameter *n.* The robustness of the steepness of the dose-response curve to genetic perturbation is experimentally well-supported (Chevereau et al., 2015; Woods et al., 2006) and even if a particular strain we are using or particular resistance mutation would greatly influence it, the effect would only be that it would take a few more dilutions to get to the correct antibiotic concentration. The monotonicity of the curve used to calculate the

next antibiotic concentration assures that it will converge to the right value despite small errors in the assumed shape of the curve. The success of this protocol can be seen directly by looking at how much the measured growth rate throughout the experiment departs from the target. The starting strains were chosen to avoid strains with large growth rate defects in no drug. Small departures from the wildtype g_{max} parameter should not affect the accuracy of the concentration update greatly, but do affect the effective strength of selection pressure for resistance and therefore need to be carefully considered when evaluating the speed of adaptation of specific mutants.

The output of the evolution experiment is a real-time estimate of resistance for each culture over the course of the experiment. Since the antibiotic concentration is continually adjusted to reflect the growth rate in the drug and to maintain a 50% inhibition, simply the concentration in the well is a good estimate of the IC_{50}. This "on-the-fly" measurement of the IC_{50} agrees well with a standard IC_{50} measurement performed in a drug concentration gradient after the evolution experiment (Figure 3).

Interestingly, even though evolution is fundamentally a stochastic process, how resistance changes over time for evolutionary replicates starting from the same genotype is often very reproducible (Figure 3). One reason for this could be a high mutational supply of beneficial mutations (Nicoloff and Andersson, 2013; Perfeito et al., 2007), eliminating long waiting times for beneficial mutations. Further, the usual bottleneck size is around 1:10 and that is close to the calculated optimum for such experiments to minimize loss of beneficial mutations (Wahl and Gerrish, 2001).

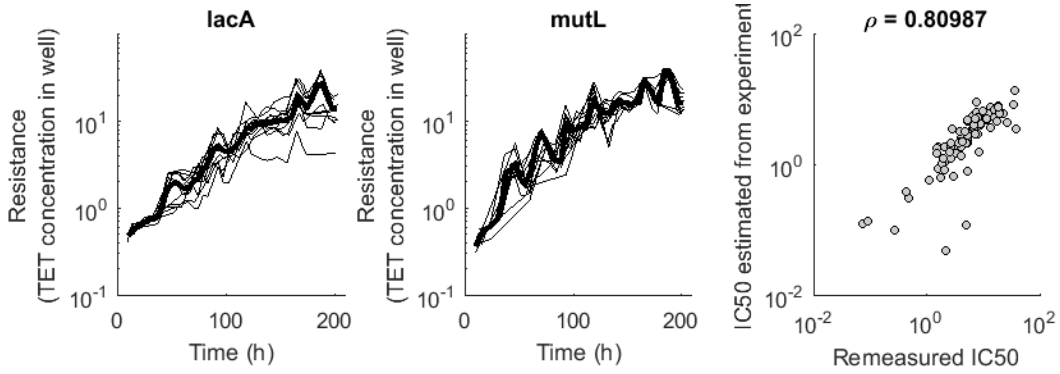


Figure 3 ***Resistance level read out from highly controlled evolution experiment is reproducible and accurate.***

Left: The lacA deletion strain is used as a control in this experiment (M3). The plot shows, on a log scale, the TET concentration in the wells of all replicate evolutions starting from the lacA deletion strain. Since the concentration is constantly adjusted to yield a 50% reduction in growth rate, the concentration itself is an IC_{50} estimate. Middle: Resistance increase (given by TET concentration in the well) is shown for the mutL deletion strain, which is a strain with DNA repair defect, causing an increase in mutation rate. Notably, the resistance trajectories become very reproducible. Right: The concentration of TET in the well at the end of the experiment for many wells is plotted against the fitted IC_{50} values from measuring the growth rate of the same populations in a wide range of TET concentrations (see Methods). The correlation coefficient calculated from the log values of the two measurements is given in the title of the plot.

2.2.2 Mutations acquired during evolution largely affect efflux pump expression and do not depend on genetic background

Generally, mutations gained during the experiment repeatedly hit the same genes. From 149 sequenced population samples from approximately the 100th generation, 84% percent of all (444) mutations identified were either amplifications including a particular genome region, or hit one of the four most often hit genes. The most often hit genes were *lon* and *marR*. In close succession, comes the amplification of a genome region which includes the *acrA/B* operon ("pumps" in Figure 4), then mutations in the coding region of *acrR* and several mutations in the *ybaO* (*decR*) gene. All mutations are consistent with literature apart from the *ybaO* mutations, which we have not found described earlier. In addition to these very common mutations, there were 9 types of mutations which appeared in more than one sample and 39 other mutations which only happened in one sample each.

The mutations related to the Lon protease were almost exclusively IS element insertions in the promoter. These mutations are known to happen at high frequency in response to tetracycline and have been described previously (Nicoloff and Andersson, 2013). The resistance benefit is thought to arise from a lowered expression of the Lon protease, which increases the expression level of its client MarA, leading to a higher expression of efflux pumps.

In the *marR* locus, various types of mutations were seen, SNPs, small indels, as well as IS element insertions (Figure 4). For 50% of the mutations in *marR*, especially the IS insertions and indels a clear loss of function phenotype would be expected. The SNP effects were less clear, though also there, some of the SNPs clearly resulted in a premature stop codon, presumably resulting in a loss of function phenotype. MarR is a transcriptional repressor known to be involved in multiple antibiotic resistance (Alekshun and Levy, 1997). It is the transcriptional repressor of the *marA* gene, which in turn activates over 30 genes involved in oxidative stress response, heavy metal resistance, and also antibiotic resistance (Alekshun and Levy, 1999). Importantly, MarA activates several genes coding for efflux pumps, including the one most relevant for tetracycline resistance, the AcrA/B-TolC efflux pump (Figure 4). A loss of function mutation in this gene therefore causes an upregulation of efflux through MarA.

Next, amplifications in the region from the *lon* gene to the *ybdK* gene were detected in over 40% of the samples. This region includes the *acrA/B* operon coding for efflux pumps and thus is expected to lead to higher levels of transcription due to higher dosage. These and similar amplifications have been reported previously (Roemhild et al., 2015; Sandegren and Andersson, 2009). In contrast to the other detected mutations, in the case of gene amplifications, our sequencing approach (see Methods) cannot distinguish between many-copy amplifications in a subpopulation and lower copy number amplifications in the entire population. Therefore, even though the other mutations described in this section can be considered fixed in the population, the amplifications are at unknown frequency. Due to the possible fast rate of extension and disappearance of amplifications, back-mutations might be much more frequent in this case (Sandegren and Andersson, 2009). Therefore, it might even be the case that when we sequence at the end of the experiment, a previously fixed amplification is already deleted in a part of the population.

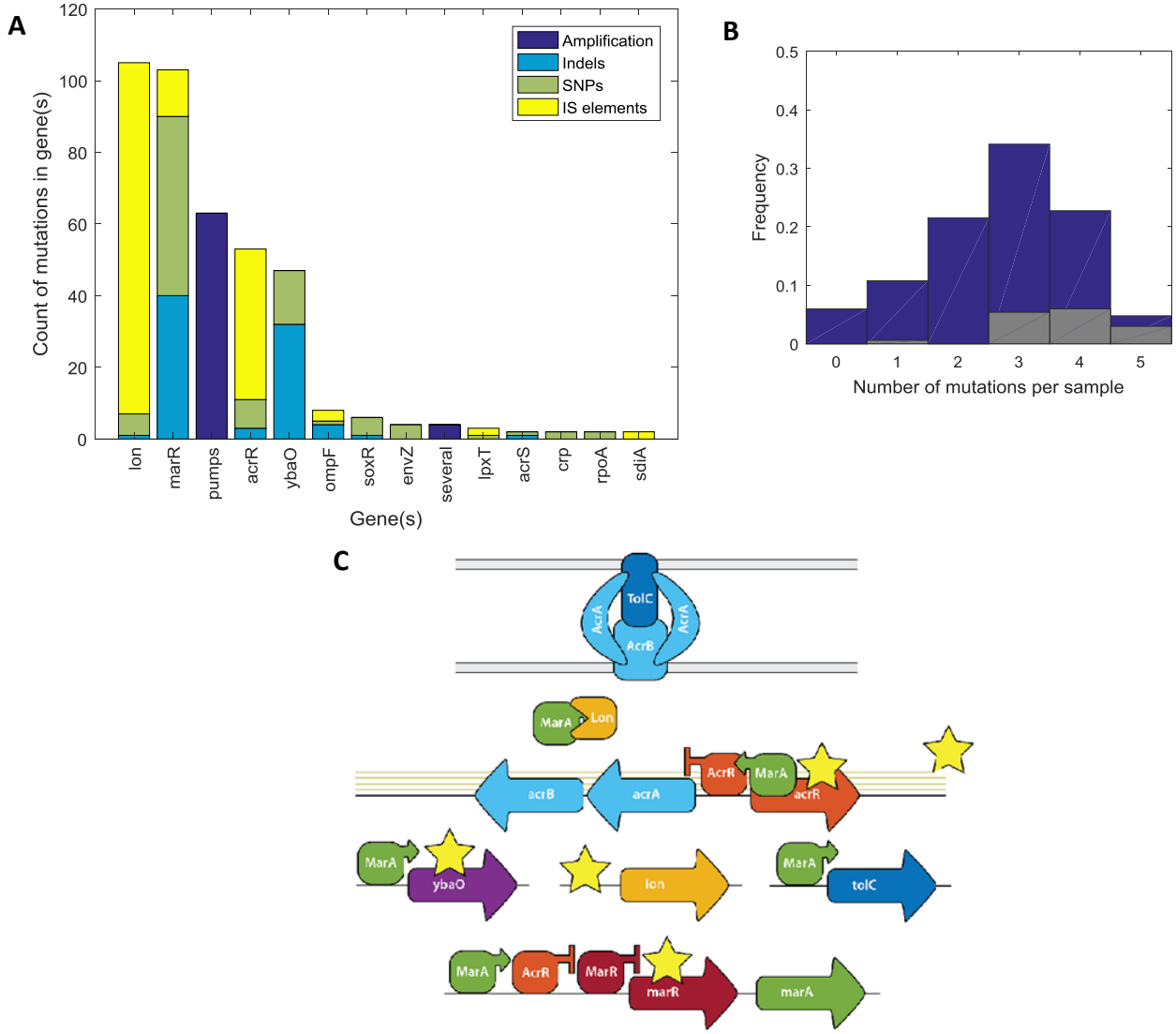


Figure 4 ***Mutations often hit the same genes involved in regulation of efflux pump expression.***

A) Counts and types of fixed mutations found in evolved strains grouped by gene locus where they occurred. Only genes which were hit at least twice in our dataset are shown. B) Histogram of number of mutations detected in each evolved sample. The gray bars represent the number of mutations found in control strains i.e. the keio parent strains or the lacA strain. The typical number of fixed mutations for control strains is 3 or 4. Since the selection sent for sequencing was biased toward slowly evolving strains, the overall counts are biased to fewer mutations. C) The loci of the five most common mutations are shown in this scheme as stars. Amplification of a genome region is shown as several yellow lines across the region with a star attached. Big arrows are genes, squares with pointing arrows are transcriptional activators, squares with blunt arrows are transcriptional repressors and the Pacman shape is a protease cleaving the transcription factor. The top of the diagram shows an AcrA/B-TolC pump in the membrane.

AcrR is another transcriptional repressor of the AcrA/B pumps and of the mar operon and 35% of the sequenced samples had a mutation in the *acrR* coding region. The *acrR* gene is right next to the *acrA/B* operon on the genome, but transcribed in the opposite direction (Figure 4). They have overlapping promoters, but also, the binding region of *marA* overlaps with the *acrR* coding sequence. This means, that mutations close to the start of the coding

sequence could interfere with MarA binding and lower efflux pump expression, whereas loss of function mutations in the later part of the sequence would increase efflux pump expression by interfering with AcrR function, without incurring the cost of disrupting the MarA binding site. Consistent with this prediction, the mutations observed are only in the coding region and avoid the 5' end of the coding sequence. The proximity of *acrR* to *acrA/B* also means that the *acrR* gene is amplified together with the pumps when gene amplification happens. This means that the *acrR* mutation is extremely unlikely to happen after the amplification, since it would have to hit both (or more) copies of the gene to have an effect. Consistent with this idea, almost none of the samples where the "pump" amplification is present have an *acrR* mutation. This sole consideration does not explain though, why samples where *acrR* mutations occurred first, did not subsequently amplify this genomic region.

The gene which is next in the order of number of mutations which hit it is *ybaO* (or *decR*). This transcription factor only has one known binding site, which is within the promoter of the *cyuP* operon coding for cysteine detoxification genes. We are not aware of this mutation having been reported as a resistance mutation previously or seen in comparable lab evolution experiments (Toprak et al., 2012). The exact mutations are interestingly very reproducible. One example is a one base pair insertion in the promoter, which is probably a result of slippage, where a stretch of 7 adenines is extended to 8. This mutation occurs within the promoter (σ70 binding site) of *ybaO*, possibly prohibiting transcription initiation. Another common mutation occurs within the coding region of the gene. It is an in-frame 9 nucleotide insertion within the conserved DNA-binding HTH domain, presumably affecting the activity of the regulator. Several non-synonymous SNPs are also present, some of them (like T18P and T33P) occur repeatedly, and several are clearly loss-of-function mutations (premature stop codons). Since about 15% of the mutations are clearly loss of function (frame-shifting insertion or premature stop codons), we speculate that the ones with unclear effects are also loss of function. The effect of these mutations will therefore likely be comparable to a deletion of the gene.

Loss of function mutations in the outer membrane porin gene *ompF* also happened repeatedly, though much less often compared to the previous five (Figure 4). Mutations in this gene have been reported before and are thought to reduce the uptake of tetracycline into the cell, decreasing the intracellular concentration of the drug (Fernandez and Hancock, 2012).

Generally, the great majority of mutations detected in our lab evolution experiments have been described before (with the notable exception of *ybaO* mutations) and most of them are known to cause increased expression of AcrA/B efflux pumps. Therefore, in the tetracycline experiments, what we describe as "resistance" is, mechanistically, to a large extent the expression level of efflux pumps.

2.2.3 Diminishing returns epistasis in evolution experiments using constant antibiotic concentration

In experiments where populations adapt to fixed antibiotic concentrations, a diminishing returns dynamic would be expected, where less fit genotypes see higher relative increases in fitness compared to more fit genotypes (Couce and Tenaillon, 2015). Our preliminary experiments confirm this.

We conducted simple serial transfer experiments in trimethoprim in M9 supplemented medium and in LB medium (see Methods). In both cases, one could see the diminishing returns dynamic (Figure 5). In these experiments, evolution starts from a selection of single gene knockouts, which is chosen to represent the whole scale of antibiotic sensitivity levels. As the separate populations have time to adapt to the medium – spontaneous mutations arise and fix – the growth rate of the population increases. The faster the growth rate, the smaller potential there is for increasing it further and the fitness levels off.

The data suggests that the negative epistasis relationship holds between random (single gene deletions) vs beneficial mutations as well as among the beneficial mutations that fix during the experiment. This can be seen comparing the initial growth rate of the deletion strains vs their fitness increases as shown in Figure 5, and comparing that to a relationship between early increase and late increase in growth rate. The latter plot corresponds to a well-known pattern from lab evolution experiments, where fitness increases level off when given enough time.

A simple explanation for this phenomenon could come directly from the DFE and dose-response relations described in (Chevereau et al., 2015). It was shown, based on exhaustive measurements of all single-gene deletion strains, that the distribution of resistance effects is more stereotypical than the distribution of fitness effects. It follows, that when predicting the adaptation of the population, one would consider that a new mutation moves a predictable step in resistance space. This would be consistent with a 1-dimensional version of a Fisher's geometric model. If the magnitude of the step taken in resistance space does not change as populations adapt, the changes in associated fitness (growth rate) would change, simply because of the shape of the dose response curve (see Chapter 3). In general, the shallower the dose response curve, the narrower the DFE. As populations adapt to a constant antibiotic concentration, they "move up" the sigmoidal dose-response curve, causing a decrease in DFE width. This, in turn, decreases the expected benefit of a new mutation, explaining the diminishing returns dynamic.

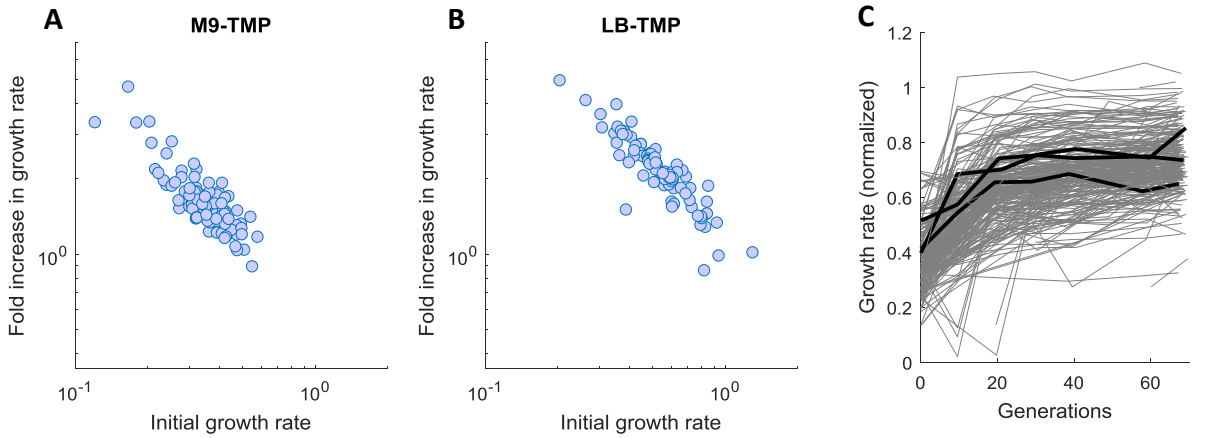


Figure 5 ***Hallmark of diminishing returns during adaptation to a fixed concentration of trimethoprim.***

A-B) Initial growth rate in doublings per hour vs. fold increase in growth rate over approximately 7 days. Each data point is a mean value of three replicate evolution experiments starting from the same initial genotype and plotted on a log-log axis. Clear decreasing trend shows that strains with lower initial growth rate increase their growth rate more than strains starting with a higher growth rate. A) and B) use data from experiments in M9 supplemented medium and LB medium respectively. C) Growth rates over generations from the LB-TMP experiment. Three example curves are shown in black to highlight the saturating shape. Growth rates normalized to the growth rate in no drug for that particular strain.

2.2.4 Diminishing returns epistasis is observed at the level of antibiotic resistance

In experimental evolution protocols akin to a “morbidostat”, there are no grounds for observing a diminishing returns dynamic based on growth rate, such as in the previous section which would be due to the population nearing a growth rate optimum. In “morbidostat” experiments, antibiotic pressure is dynamically adjusted to counteract the resistance that evolves. That is, when growth rate increases due to resistance mutations, the protocol, by changing antibiotic concentration effectively returns the growth rate to its original distance from the optimum. Therefore, a diminishing returns dynamic would not necessarily be expected.

However, one could observe diminishing returns dynamics if the optimized phenotype is not growth rate, but resistance, which is what a morbidostat setup selects for. The observation that resistance increases seem to level off in setups which challenge cells with ever-increasing concentrations has been made for some antibiotics (Toprak et al., 2012). This implies that cells with existing strong beneficial resistance mutations have less potential to adapt. The leveling off effect could simply be due to the “running out” of strong resistance mutations. There could also be a hidden phenotype in which the mutational changes tend to have a stereotypical “step size”. This phenotype is then mapped to resistance in a similar way to how resistance is mapped to growth rate in the previous section. In this way, many simple monotonic functions which would link the hidden phenotype to resistance would produce a diminishing returns pattern at the level of resistance.

To test whether a general trend of diminishing returns epistasis can be observed at the level of resistance, we made a selection of 100 different gene deletion strains and compared their increase in resistance when propagated in the same automatic experimental evolution setup. The selection contains strains with defects in a broad range of cellular mechanisms (Figure 6). About half of the genes were chosen because the mechanism impaired is expected or speculated to have an effect on evolvability. This hypothesis-driven selection includes deletions of genes with altered mutation rates: DNA mismatch repair (Jolivet-Gougeon et al., 2011), SOS response (Cirz et al., 2005), and oxidative stress response (Méhi et al., 2014). Further, it includes deletions of known resistance mechanisms (efflux pumps) and other functions that are thought to interfere with the distribution of fitness effects of available mutations: protein folding, transcription factors, membrane composition (see section 2.1). The rest of the deletion strains were chosen to represent a broad range of pathways expressed in rich medium and deletions of which have negligible fitness costs. The latter is important, because a lower growth rate in no drug (maximal growth rate of the dose response curve) would translate to a weaker selection pressure on resistance, which would slow down the expected adaptation rate for reasons different from the ones we are interested in here.

We used the automated setup described in Section 2.2.1 to evolve tetracycline resistance for the selection of single-gene deletion strains and we observed a hallmark of diminishing returns epistasis at the level of drug resistance: Strains with higher initial resistance undergo lower resistance increases during the experiment compared to strains with lower initial resistance Figure 6. Many gene deletions alter antibiotic resistance (Chevereau et al., 2015; Nichols et al., 2011). Thus, the evolution experiment is started from many different initial resistance levels represented by different deletion strains. However, after 180 hours of

intensive antibiotic exposure in our evolution setup, these differences largely evened out: Strains with an x-fold lower initial resistance (IC_{50}) than the wild type tend to increase their resistance by x-fold more in the evolution experiment Figure 6. This pattern supports the existence of a relatively hard upper bound for the absolute resistance level, which limits the relative resistance increase that is achievable as populations approach this upper bound.

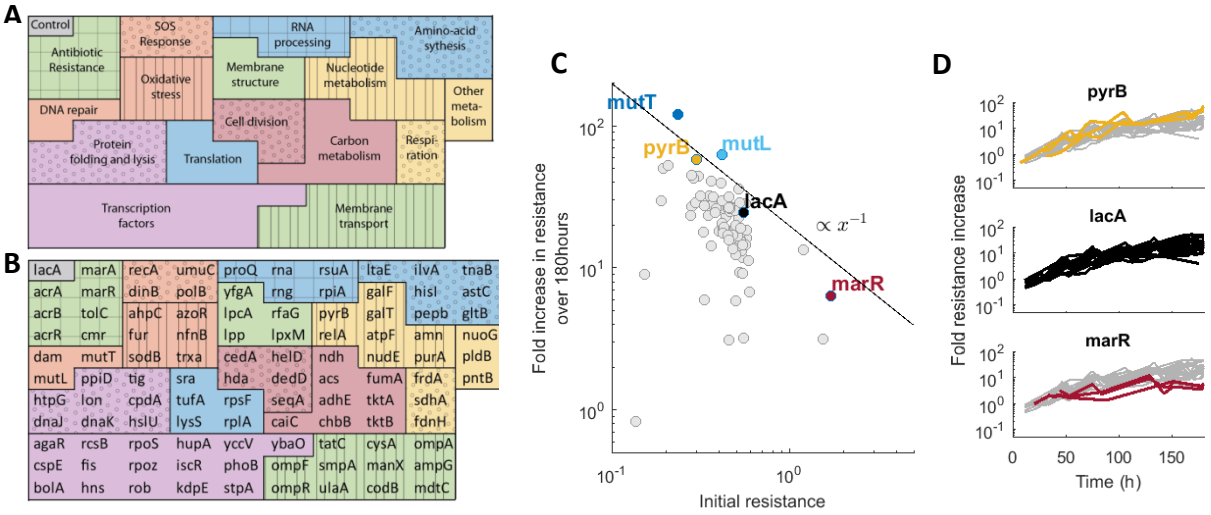


Figure 6 ***Diminishing returns at the level of antibiotic resistance when comparing resistance increases in a broad selection of single gene deletion strains.***

A) Cellular mechanisms represented in the deletion strain selection. B) Gene names which were selected. The arrangement in A) and B) is the same for easy lookup of genes and their functions. C) Mean fold resistance increase after 180 hours of running the experiment plotted against the mean initial resistance for each deletion strain. The strains are named by the gene which is deleted ('lacA' is ΔlacA*). The final and initial resistance measures for each replicate are the mean TET concentrations for appropriate time windows (~12-24h for initial and ~170-180h for final resistance). Only those data points were used where the normalized growth rate was between 0.3 and 0.7. A line proportional to* $y=x^{-1}$ *is shown to indicate the steepness of the slope the points would fall on if all experiments reached the same maximum resistance level irrespective of starting conditions. D) Fold resistance over time for replicates of three strains highlighted in C.) The control (lacA) resistance over time trajectories are plotted in gray for comparison. To compute the relative increase, the resistance values were divided by the mean initial resistance values of all replicate evolutions of that strain. This figure uses data from experiment M2 (see Methods).*

2.2.5 Several outliers to the diminishing returns trend represent particular ways to strongly perturb resistance evolution

Many deletion strains follow the diminishing returns dynamics trend, but several strains are strong outliers: their different increases in resistance could not be explained by their initial resistances. These would be those that perturb resistance evolvability in a particular way and represent cellular pathways, which, when inhibited, could modulate the potential of bacteria to evolve resistance. In particular, mutators, strains with perturbations to efflux pumps and their regulation, strains with chaperone deletions and a few idiosyncratic mutants were observed to strongly affect resistance evolution.

To confirm the effects of these mutants, which, in the first experiments were only evolved in triplicate, the experiment was repeated with a selection of the most interesting strains in 8-12 replicates. We also added a few chaperones and efflux pump related deletion strains.

Generally, the outcomes from the two experiments matched very well. In the following sections, the results will include data from both experiments.

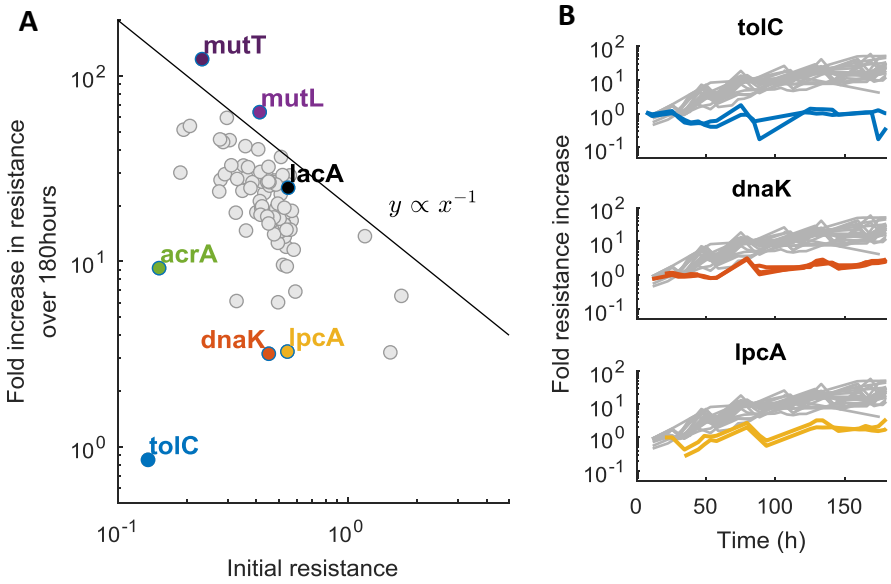


Figure 7 ***Overall diminishing returns pattern sees several strong outliers.***

A) This is the same plot as in Figure 6C. The highlighted examples represent outliers. B) Fold resistance increase over time (on a log scale) for selected highlighted outliers. Control (lacA) resistance over time trajectories are plotted in gray for comparison. The strains are named by the gene which is deleted ('lacA' is ΔlacA*). To compute the relative increase, the resistance values were divided by the mean initial resistance values of all replicate evolutions of that strain. This figure uses data from experiment M2 (see Methods).*

2.2.6 Genetic perturbation of efflux pumps drastically reduces the evolvability of resistance to tetracycline

Considering that the great majority of mutations were related to the regulation of multi-drug efflux pumps, we were curious to find out what happens when the composition or regulation of these is perturbed. One could imagine several scenarios. Generally, when efflux pump expression or function is compromised, the bacteria are much more sensitive to tetracycline. So, based on the diminishing returns dynamic observed for other genes, it would be expected the cells are much further from the imagined optimum resistance, and adaptation would thus proceed faster. On the other hand, deleting a gene crucial for this particular resistance mechanism could effectively block the whole resistance scenario and force the cells to either find a different (likely with more modest resistance gains) path or not evolve at all. This would also mean that the main (or even only) path to resistance of our strains in our setup is fragile enough to be blocked by the perturbation of a single gene.

Indeed, the resistance to tetracycline can effectively be blocked by deleting single genes *acrA* or *tolC* coding for components of the major AcrA/B-TolC efflux pump. The most drastic effect can be observed for Δ*tolC* in tetracycline, where all 7 replicates of the evolution experiment saw no increase in resistance. Only one fixed mutation was detected in one of the replicates. This was a single base-pair substitution in the promoter of the *yhdJ* gene, which also happens to be the terminator sequence of the neighboring *fis* gene. Since the replicate showed no increase in resistance, and no other sequenced sample had a mutation in that locus, we do not expect this mutation to be a resistance mutation. For Δ*acrA*, we

observed one replicate (out of 5) which evolved to a relevant resistance level. This mutant had a IS element insertion close to the *acrE* promoter region and an amplification including the *acrE* and neighboring *acrF* genes, in addition to the typical *lon* and *marR* mutations. This could have led to the upregulation of *acrE/F* genes. The gene *acrE* has 65% sequence similarity with the *acrA* gene and the product also forms an efflux pump with AcrF and TolC (Jellen-Ritter and Kern, 2001; Kobayashi et al., 2001). This is an example of an alternative resistance path when *acrA* is deleted. This alternative path has been observed in experiments previously (Kobayashi et al., 2001). The existence of such alternative paths for other deletions, highlights the uniqueness of Δ*tolC*, for which no resistance evolution was detected.

We also tested the deletions of other efflux pump systems, some of which share the regulation with AcrA/B-TolC and could therefore contribute to resistance. None of these deletions seemed to significantly perturb resistance evolution (Appendix Figure 27).

The deletion of *marR* causes the cells to be initially more resistant to tetracycline and adapt slower. The initial resistance is expected, since MarR is a negative regulator of MarA, which is the main activator of several efflux pumps including the AcrA/B-TolC pump, so its deletion causes transcriptional upregulation of the pumps. Also, many of the evolved strains have loss of function mutations in this gene. Under the experimental evolution protocol, it gains mutations in the same (remaining) genes as the other strains and the final resistance levels are also comparable to the control strains. This suggests that the deletion mimics one of the expected mutations and evolution proceeds the same as in most other strains, just with the *marR* mutation always happening first.

The deletion of *marA* causes a slight increased sensitivity to tetracycline initially, followed by only modest increases in resistance. Since MarA is an important activator of efflux pump expression, albeit not the only one, a slight decrease in resistance as a result of the deletion is expected. The resistance increases are variable, with two replicates almost incapable of adapting and four replicates reaching relatively low but significant resistant levels. The four replicates which evolve more all have a *soxR* mutation whereas the ones that do not evolve don't have it. Since SoxS (repressed by SoxR) is a transcription factor with a large target overlap with MarA, and the mutation in *soxR* is relatively rare for other ancestral strains, it is likely that these mutations are compensatory mutations. Still, even with the compensatory mutation, the final resistance levels reached are significantly lower than those that control strains can reach. The very common and potent loss of function mutation in *marR* is expected to have an effect on resistance through the de-repression of *marA*. By deleting *marA*, it is expected that the cells do not gain any *marR* mutations anymore and thus have limited available fitness gains. It is of interest though, that the observed increase in resistance of the Δ*marA* strain is smaller on average than the increase in resistance of the Δ*marR* strain, so the hindrance in adaptation cannot be purely due to the missing *marR* mutation. Indeed, the evolved populations also did not fix the *ybaO* or *acrR* mutations which could be due to the fact that this as well is only beneficial in presence of MarA activators. It is likely that deleting *marA* also lowers the beneficial effects of other mutations.

The deletion of *acrR* causes an initial sensitivity to tetracycline and compromises resistance evolution. It may seem paradoxical that a loss of function mutation in *acrR* is very common in the evolved strains, but a deletion of the gene causes the cells to be more sensitive. This discrepancy can be explained by the fact, that the N-terminal region of the *acrR* coding

sequence contains the promoter for the neighboring *acrA/B* genes (Figure 4). Thus, the deletion of the whole gene has a combined effect of deleting a repressor (*acrR*), but also an activator binding site. In line with this observation, the loss of function mutations happen closer to the C-terminal region of the gene, compromising AcrR function, but keeping the promoter intact. In terms of adaptation rate, we speculate that the relatively modest increase in resistance is due to the smaller effect of the same resistance mutations. For example, with a missing MarA binding site in the *acrA/B* promoter, the *marR* mutations can only increase efflux pump expression by regulating other pump components, for example, TolC. In addition, AcrR is also a repressor of *marA*, so the deletion strain might have lower *marA* expression, further interfering with mutations which act through MarA.

The Lon protease deletion strain initially has a higher resistance to tetracycline on average, but then reaches very similar resistance levels compared to controls. This points to a similar scenario as in the Δ*marR* strain. The populations behave the same as in controls, except that the *lon* mutation happened first.

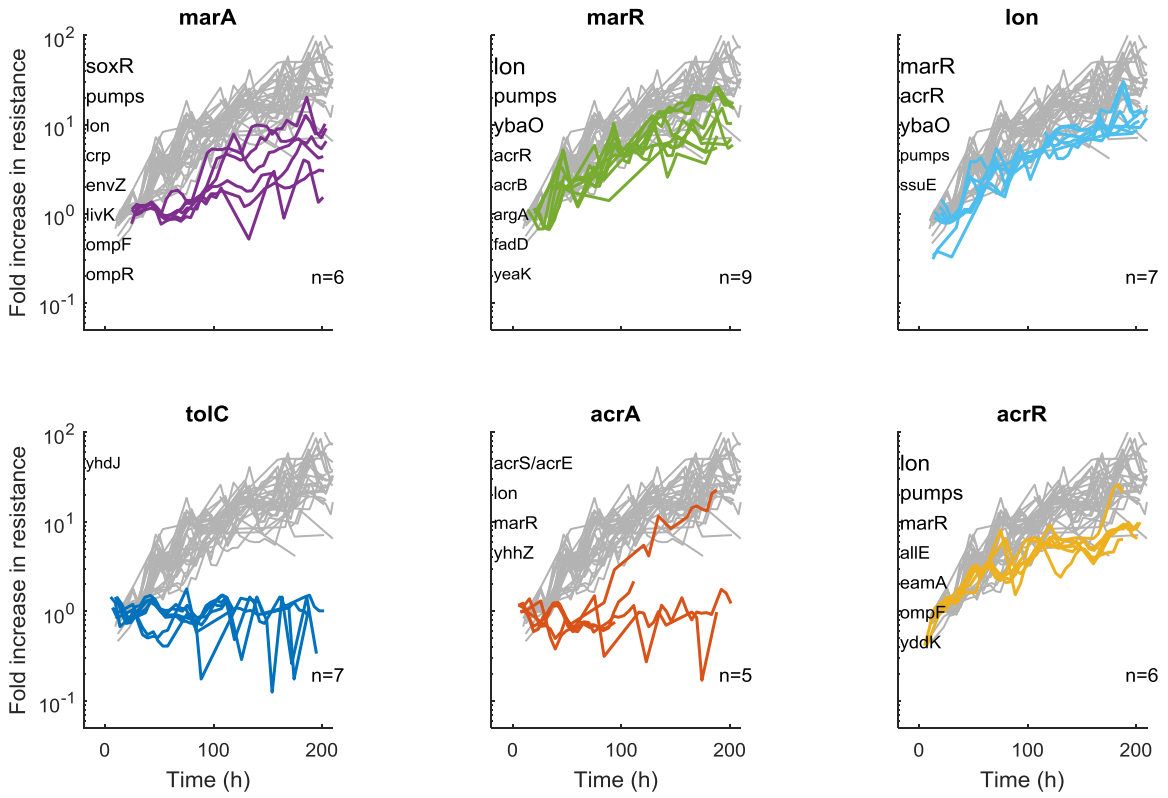


Figure 8 ***Strains with deletions of efflux pump components or regulators show reproducibly impaired adaptation.***

*Fold resistance over time for several strains with deletions of genes related to efflux pumps. The control (lacA) resistance over time trajectories are plotted in gray for comparison. To compute the relative increase, the mean initial resistance values of all replicate evolutions of that particular strain were used. The text in the upper left of the plot indicates loci of mutations found in the evolved strains written in order of frequency. Larger font indicates a more common mutation: font size is equal to 8 + (x*4), where x is the proportion of sequenced strains which had a mutation in the given locus. Gray lines show resistance over time for control (lacA) replicate experiments. The number of sequenced samples is given in the lower right corner of each plot.*

2.2.7 Perturbing chaperones can slow down resistance evolution

The deletion of genes for certain chaperones, especially *dnaK*, coding for the bacterial Hsp70 chaperone, consistently hinders the evolution of resistance. As explained in Section 2.1, chaperones play a role in various different hypotheses for how evolution rate can be changed. In general, the expectation is that since chaperones are responsible for folding proteins, interfering with their function can strongly influence the effects of mutations in coding regions. If they normally mask mutations by folding proteins to the usual conformation despite mutations, perturbing their function may mean that new mutations are uncovered and evolution could proceed faster. We do not see evidence of this in our experiments. We do see, however, a major slowdown in evolution for the deletion strain of *dnaK* and a modest slowdown for the deletion of its co-chaperone DnaJ. This is likely to be due to epistasis with the available resistance mutations. Outer membrane proteins are enriched among clients of *dnaK* and *tolC* is among them. The possible link could therefore be that a portion of expressed efflux pumps do not get folded into a working conformation and thus the same mutations have beneficial, but weaker effects. This is also consistent with a smaller number of fixed mutations on average in these evolved strains.

The Δ*hslU* and Δ*htpG* strains do not show extremely different adaptation trajectories. The *hslU* gene codes for a protein with combined protease and chaperone activity. The deletion strain has slight sensitivity to tetracycline to begin with and keeps this difference in sensitivity until the end of the experiment. The *htpG* gene codes for the Hsp90 chaperone often cited in the context of mutational buffering (Karras et al., 2017; Rutherford, 2003). The deletion of *htpG* seems to have no effect in our experiment. It shows that not all chaperones affect resistance evolution greatly, but the DnaK/J system is a special case.

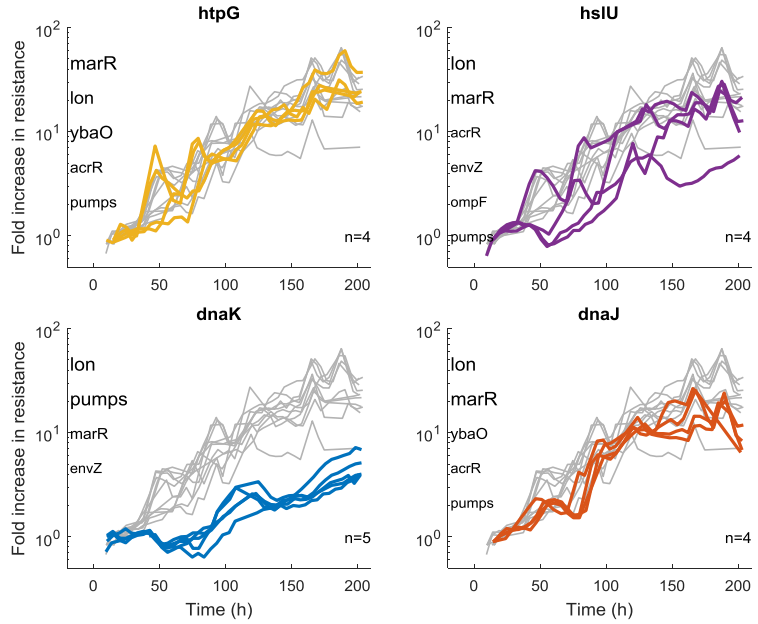


Figure 9 ***Perturbing chaperones can slow down resistance evolution.***

*Resistance (IC_{50} in µg/ml) over time as estimated by the tetracycline concentration in the well during in the automatized evolution protocol is shown for four chaperone deletion strains from experiment M3. The text in the upper left of the plot indicates loci of mutations found in the evolved strains written in order of frequency. Larger font indicates a more common mutation: font size is equal to 8 + (x*4), where x is the proportion of sequenced strains which had a mutation in the given locus. Gray lines show resistance over time for control (lacA) replicate experiments. The number of sequenced samples is given in the lower right corner of each plot.*

2.2.8 Interference with lipopolysaccharide synthesis and protein transport can hinder evolution

Lipopolysaccharide (LPS) is a major component of the outer membrane in gram-negative bacteria and perturbations of it can lead to increased permeability of the membrane (Gunn, 2001). Two strains with deletions in genes coding for different steps in the LPS synthesis pathway (Δ*lpcA* and Δ*lpxM*) start off slightly sensitive and then reach significantly lower levels of resistance than controls. In addition to sensitizing strains to antibiotics, perturbations of the composition of the membrane could also influence efflux activity and thus could interfere with the effects of resistance mutations. A recent study indicates that just perturbing the assembled LPS layer does not change efflux activity (Misra et al., 2015). Another speculative link to efflux could also be that some efflux pumps are used to transport LPS precursors and the pumps could be "clogged" by incorrect products of the LPS synthesis pathway. Either way, a mechanistic explanation would require further experiments.

The Δ*lpcA* strain evolved only approximately two-fold resistance to tetracycline, despite fixing several mutations. In all replicates of the Δ*lpcA* strain, the antibiotic concentration in the well seemed to be imprecise compared to the other strains (Figure 10). The reason for this could be that the strain has a several hours long lag in reacting to the antibiotic. Also, it could be that the real dose response curve of the mutant is steeper than assumed in the protocol. Whichever the reason, the level of inhibition remained 50% on average and despite this, much lower resistance levels were reached than for the other strains (Figure 10).

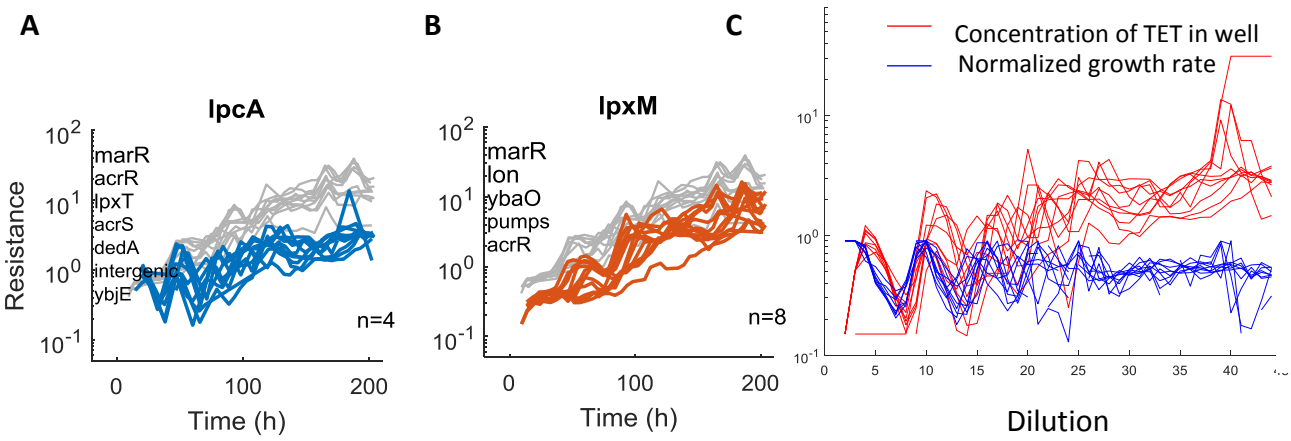


Figure 10 ***Lipopolysaccharide synthesis mutants impede resistance evolution.***

*A) and B) Resistance (IC_{50} in µg/ml) over time in the automatized evolution setup for the lpcA and lpxM deletion strains is shown. The text in the upper left of the plot indicates loci of mutations found in the evolved strains written in order of frequency. Larger font indicates a more common mutation: font size is equal to 8 + (x*4), where x is the proportion of sequenced strains which had a mutation in the given locus. The gray lines show resistance over time for the control (lacA) strains. Data is pooled from experiments M2 and M3. C) Growth rate (blue) normalized to control strain in no drug and tetracycline concentration in well in µg/ml (red) is shown for the lpcA deletion strain.*

2.2.9 Deletions of genes involved in mismatch repair and the deletion of *ybaO* reproducibly leads to faster adaptation

If adaptation rate is limited by the waiting time for the next beneficial mutation, an increase in mutation rate will lead to faster adaptation. In many of the adaptation trajectories, resistance increases tend to be step-like. This suggests that indeed, some time is spent waiting for the next mutation to arrive. Using the Δ*mutL* and Δ*mutT* strains, one can test this and see if increasing the mutation rate is helpful for speeding up adaptation. The mutator strains evolve about as fast as the fastest controls (Figure 11). Therefore, on average, higher mutation rates speed up adaptation in our set up, they do not, however, display adaptation rates that go beyond what is observed for the control strain.

The *ybaO* deletion strain shows signs of positive epistasis with the usual resistance mutations. The *ybaO* gene was a common target of mutation in the sequenced samples (Figure 4) and some of the mutations were clearly loss of function (see Appendix).

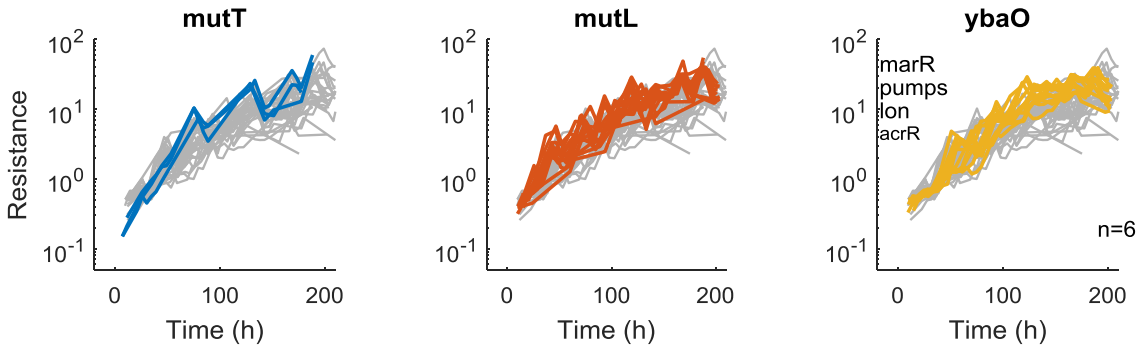


Figure 11 ***Examples of strains which on average adapt faster than controls.***

IC_{50} (TET concentration in well during experiment) over time is shown for mutT, mutL and ybaO deletion strains. For the ybaO strain, also the mutations detected in the evolved strains are shown. These correspond extremely well to the mutations in evolved controls.

2.2.10 Evolution in chloramphenicol shows similar trends to tetracycline experiments

In order to test the generality of our results from the experiment in tetracycline, we have extended our experiments with a lab evolution protocol in chloramphenicol. The two drugs both target the ribosome, although tetracycline targets the 30S and chloramphenicol the 50S subunit. In terms of resistance mutations which occur in lab evolution experiments, both antibiotics prevalently see mutations related to efflux pumps. Apart from the ones that occur for tetracycline, there are also mutations within the coding region of *acrB*, potentially making the efflux pump more permeable to the drug and mutations that potentially increase the expression of the *mdfA* (or *cmr*) pump. It is interesting to note that populations evolved to CHL are also resistant to tetracyclines to a comparable level to those populations evolved to tetracyclines and populations evolved to a tetracycline also have significant resistance to chloramphenicol (Lázár et al., 2014; Toprak et al., 2012), confirming the similarity of the respective resistance mechanisms. There is a large contrast though, in terms of how resistant *E.coli* can get over 2-3 weeks in morbidostat-type experiments. Whereas the usual IC_{50} changes achieved with tetracycline are around 10 to 50-fold, the resistance

gains observed for chloramphenicol can be up to 1000-fold (Toprak et al., 2012). Therefore, the two drugs are comparable in terms of mechanism of action and resistance, but there is reason to believe their evolutionary dynamics to be radically different.

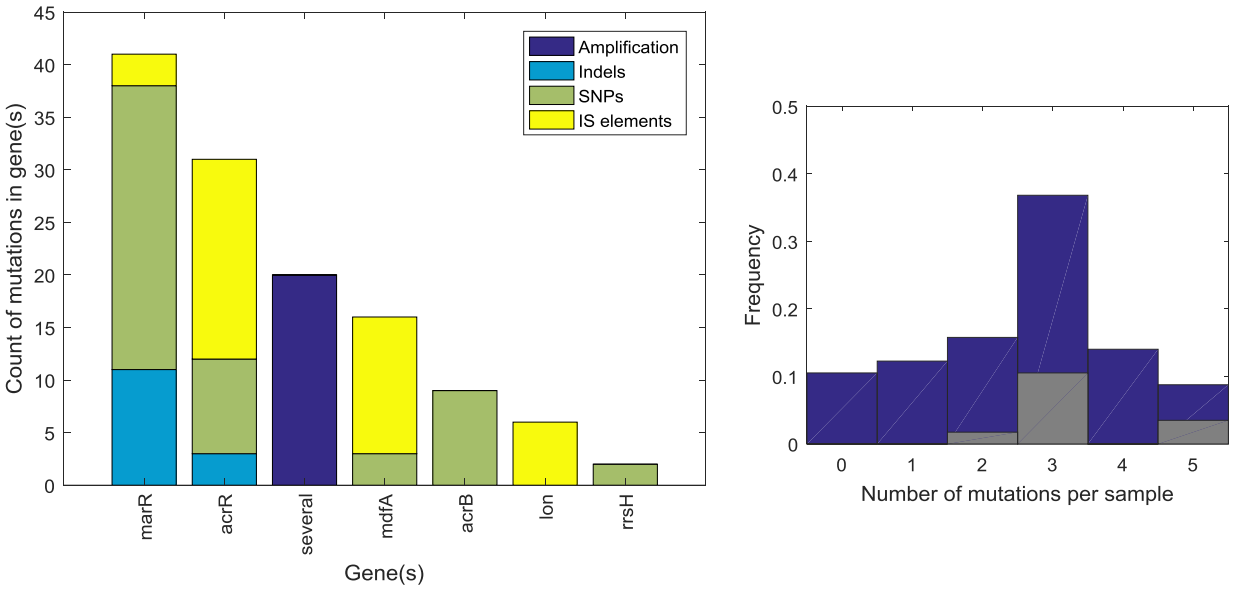


Figure 12 ***Mutations found in samples evolved to chloramphenicol.***

A) Counts and types of fixed mutations found in evolved strains grouped by gene locus where they occurred. Only genes which were hit at least twice in our dataset are shown. "Several" refers to amplifications of regions which include more than one gene. B) Histogram of number of mutations detected in each evolved sample. The gray bars represent the number of mutations found in the control (lacA) strain.

The overall pattern of CHL resistance evolution in our experiment is a fast and relatively steady increase (Figure 13B). This is in stark contrast to the evolution of TET resistance, where a clear slow-down in resistance increase is noticeable for a majority of strains Figure 6. Therefore, it is unclear whether the diminishing returns pattern would also be noticeable for this very different dynamic. Indeed, no clear negative trend can be seen when plotting increase over initial resistance, and the average fold increase over the second half of the experiment is not lower than the average fold increase over the first half of the experiment (data not shown).

These results are still consistent with chloramphenicol mutations exhibiting diminishing returns epistasis, despite not observing the pattern in our data. Considering the non-saturating resistance increases for a great majority of strains, the experiment was probably simply not run long enough to start seeing the diminishing returns effect.

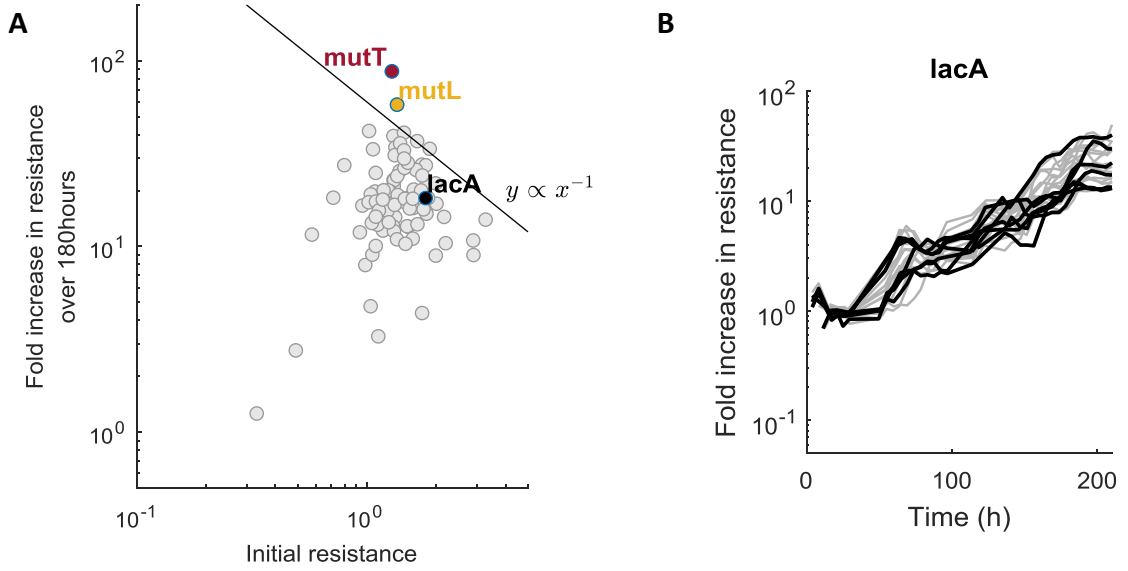


Figure 13 ***Experimental evolution in CHL does not show clear signs of diminishing returns.***

A) Mean fold resistance increase after 180 hours of running the experiment plotted against the mean initial resistance for each deletion strain. Mean values are taken over replicates of same deletion strain. The final and initial resistance measures for each replicate are the mean CHL concentrations for appropriate time windows (~12-24h for initial and ~170-180h for final resistance). Only those data points were used where the normalized growth rate was between 0.3 and 0.7. A line proportional to $y=x^{-1}$ is shown to indicate the steepness of the slope the points would fall on if all experiments reached the same maximum resistance level irrespective of starting conditions. Mutator strains and control lacA strain is highlighted. B) Resistance increase over time for lacA strain in CHL shows steady increase over the timeframe of the experiment. Replicates which were also sent for sequencing are highlighted in black.

The chloramphenicol experiment consisted of comparing the same 98 Keio collection mutants which were used in the TET experiment, in triplicates, in our automated setup for 10 days. The effects in the two drugs were correlated (Correlation coefficient 0.77). The correlation was especially driven by the extremes, which were common to both antibiotics: mutator strains (Δ*mutT* and Δ*mutL*) adapted faster and previously mentioned strains such as Δ*tolC*, Δ*dnaK* and Δ*marR* adapted less in both antibiotics (Figure 14).

Interestingly, the resistance increase seems to be larger for the Δ*mutT* strain than for Δ*mutL* although both of them have been reported as similarly strong mutators. This is likely to be due to the importance of mutational spectra in resistance evolution (Couce et al., 2013).

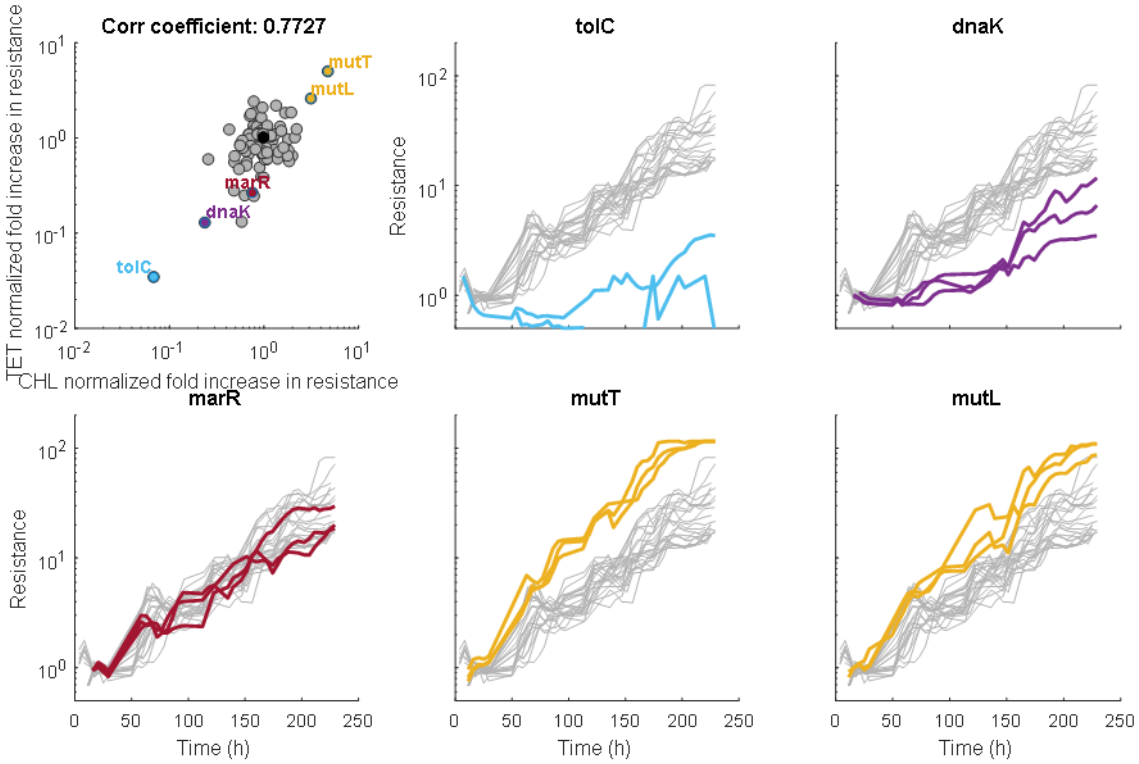


Figure 14 ***Similar strain-specific effects on adaptation can be seen in CHL as were seen in TET.***

Upper left: fold increases in resistance normalized to the mean fold increase in resistance for the control strain (lacA) shown in black. The correlation coefficient is given in the title of the plot. The remaining plots show examples of resistance increase over time for CHL for strains highlighted in the upper left plot. Even though the mutT plot seems to show a saturation of the resistance, this is not what happened, the concentration needed to keep the strains inhibited simply reached the stock concentration used on that day.

2.2.11 Quantitative model suggests that differences in extent of adaptation are mainly due to epistasis between the deleted gene and the resistance mutations

Our speculations that epistasis with resistance mutations is the driver of our strongest evolution inhibiting effects need to be further supported. One way to do this is to use the plentitude of sequencing data combined with the resistance levels measured in our experiment and attempt to predict the resistance benefit of each mutation. Having estimated the usual benefit brought about by the mutations, we can then find the deletion strains on whose background these mutations have a smaller (or larger) than expected resistance benefit. This will point to magnitude epistasis between the deletion and the acquired mutations.

To first obtain the predicted resistance benefits of the most common mutations, we constructed a simple linear regression model to predict resistance effects from sequence. The main simplifying assumption we have made, is that all mutations that happen in the same gene and all amplifications (of a similar region) have the same effect on resistance and, for the purposes of this model, we consider them to be the same mutation. Further we assume that those mutations that happen only once have no effect on resistance, i.e. resistance can be predicted well from the presence or absence of the 5 most common mutations.

The null expectation is that the effects of mutations on resistance are additive on a log scale, i.e. each mutation brings a fixed relative resistance increase irrespective of which other mutations are present. Therefore, the log resistance y can be expressed as a linear model:

$$\vec{y} = b_0 + \vec{b} \cdot \vec{x} + \epsilon,$$

where $\vec{y}$ is the log of the increase in resistance observed for the individual evolving populations, b_0 is a fitted coefficient corresponding to the resistance increase common to all evolved populations not predicted by the 5 most common mutations, $\vec{b}$ is the vector of fitted coefficients which correspond to the effects of the individual mutations, $\vec{x}$ is a vector of 1s or 0s determining the presence or absence of that particular mutation in the given evolved population.

Using the mutations from the complete collection of our sequenced strains evolved in tetracycline and predicting the log increase in resistance, our simple model can account for 55 percent of the variance (see Methods). The unexplained variance can stem from three major sources apart from measurement error. The first is the deviation from our assumption that mutations in the same gene really have the same effects. The second is that even if the different mutations that occur in the same gene are indistinguishable, they can interact either with the initial deletion or among each other, and therefore have different effects in different contexts. And the third is the contribution of mutations which are not included in the model, and, if present, their potential beneficial effects get falsely attributed to the common mutations we model or to the first fitted coefficient, which represents the resistance increase with no mutations.

Related to the third source of error, we indeed see the first coefficient (b_0) is non-zero and the fitted value is around 1.35 which is the same order of magnitude as the fitted mutational effects. It represents the average effect of the idiosyncratic mutations which do not occur repeatedly in the samples, but are still likely to have significant effects on resistance. Possibly, this value can also include increases in resistance due to longer-term gene expression changes which are not the result of mutations.

Despite these limitations, we can look at the approximate predictions of the effects of each mutation and see whether they correspond to what we would expect. The *marR* mutation is predicted to have the largest effect of the mutations included in the regression and it is also one of the two most common mutations. The *lon* and *ybaO* mutations are also predicted to have pretty strong effects and their p-values suggest their effect on resistance is significant (10^{-9} and 10^{-5} respectively). Interestingly, the fits of the *acrR* and pump amplification mutations do not have significant p-values despite occurring very frequently. A version of the regression analysis which included the *ompF* mutation (which occurred in 8 evolved samples) also did not attribute a significant p-value ($p>0.1$ Student's t-test) to the effect of this mutation. It is very unlikely that mutations which happened repeatedly would bring no resistance benefit, it is therefore puzzling that no significant effect was found in this analysis. It shows that indeed, the simple linear model of mutational effects of only the most common mutations is a gross simplification.

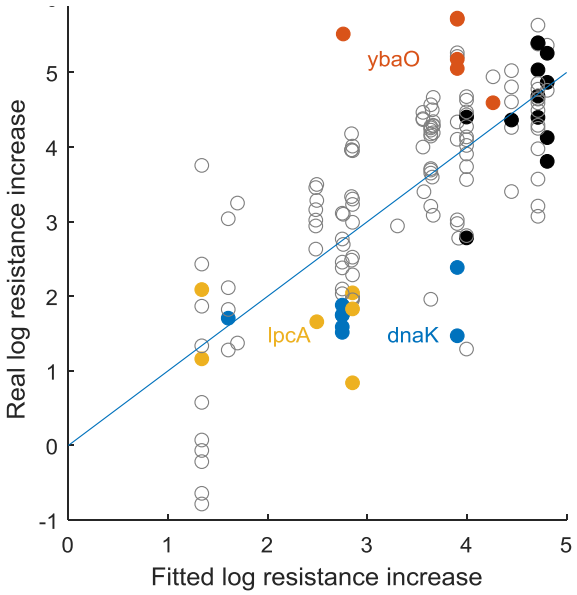


Figure 15 ***Linear regression model fitting resistance increase shows epistatic effects between resistance mutations and initial genotypes.***

The real vs fitted log resistance increase is shown. For the prediction, the 5 most common mutations were considered in this order (coefficients b2-b6): marR*,* lon*, pumps,* acrR*,* ybaO*. The black circles represent control (*lacA*) strains.* ybaO*,* lpcA *and* dnaK *samples are highlighted to show that their predicted effects systematically depart from predictions.*

Using this model, we can now compare the expected and real resistance increases based on the acquired mutations for each sample. Interestingly, those deletions which we identified as slow evolvers (Δ*dnaK*, Δ*lpcA*) consistently show lower resistance increases compared to the model prediction and one notable outlier which evolved faster than expected (Δ*ybaO*) consistently shows higher resistance values than expected based on the mutations it acquired. This points to epistasis between the gene deletions and the mutations which they acquired, which may explain the differences in adaptation we observed.

2.2.12 Adding diminishing returns slightly but significantly improves the prediction power of the model

We extended the model to enable a comparison between the simple additive model and one which includes diminishing returns epistasis. To do this, known outliers from the diminishing returns pattern were excluded from this analysis, in addition to strains with fitness defects stronger than 15% (Δfis, ΔdedD, ΔtolC, ΔacrA, ΔacrB, ΔdnaK, ΔmutL, ΔybaO, ΔlpcA, ΔdnaJ). Also, instead of the relative log resistance increase, the absolute final resistance was predicted. This way, the model included information about the initial resistance of that particular strain, which also goes into the diminishing returns prediction. It is of note that the samples chosen for sequencing were not chosen randomly and were also not chosen to represent a large range of initial resistance, therefore, no large improvement in the prediction is expected.

Diminishing returns was modelled by channeling all the linear terms through a simple monotonically increasing concave function and simultaneously fitting all parameters. Two functions were used. One was an exponential function with exponent smaller than 1. The other was a Michaelis Menten equation (see Methods). All parameters were fitted

simultaneously, the explained variance improved from 0.667 to 0.683 (power law model) and 0.687 (Michaelis-Menten model). The p-values (Student's t-test) of the exponent in the power law and the parameters of the Michaelis-Menten model were all significant (p<<0.001).

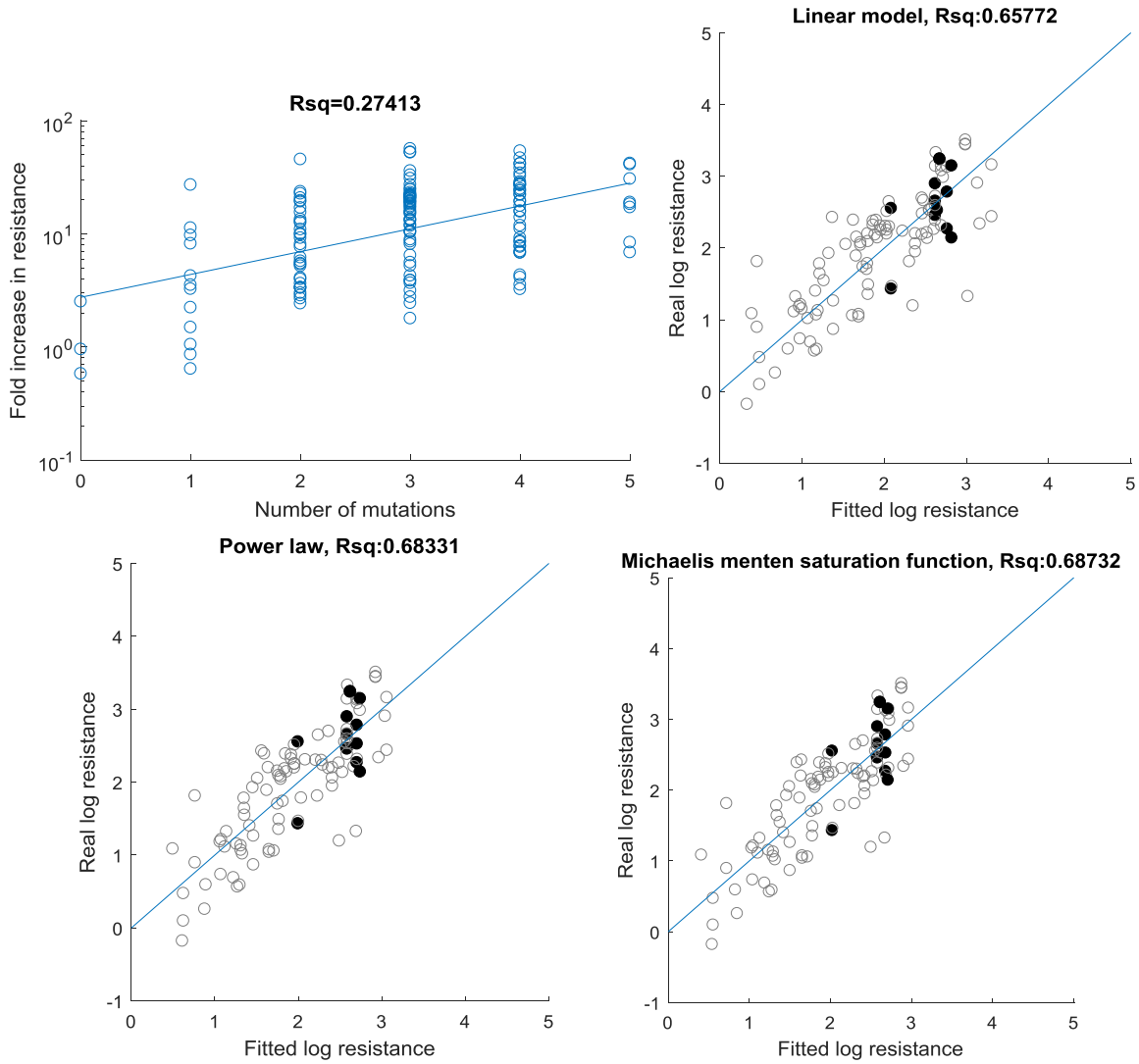


Figure 16 ***Including diminishing returns in regression model improves resistance prediction.***

First model (upper left) assumes each mutation (regardless of locus) has the same effect on resistance. The constant and slope are fitted, representing the resistance increase without mutation and the benefit of a mutation (equal for all mutations). This model results in the worst fit, with an R^2 of 0.274. Upper right: Measured vs fitted values for log resistance using a simple linear model. The model tested is one where only the 5 most common mutations are modelled, each has a different coefficient fitted. The predicted response variable is log resistance (as opposed to fold log resistance increase in the previous model). The fit also uses the initial resistance of each particular modelled strain (see Methods). Lower left: Measured vs fitted values of log resistance for a power law diminishing returns model. The linear model is channeled through a power law and an exponent and constant are fitted in addition to the five mutational effects coefficients. The p-values of the exponent and constant are <<0.001 and the variance explained increased modestly. Lower right: Measured vs fitted values of log resistance for a Michaelis Menten diminishing returns model. In all plots the black circles represent lacA samples.

2.3 Conclusions and discussion

Using automated high-throughput evolution experiments, we identified general patterns guiding resistance evolution and examples of particular genes that perturb antibiotic resistance evolvability most drastically.

We were able to establish a high-throughput experimental evolution protocol that keeps hundreds of cultures in parallel in exponential phase under controlled selection pressure. The protocol enabled us to see a wide dynamic range of rates of adaptation which were also very reproducible.

A limitation of our high-throughput evolution protocol is the inability to correctly adjust selection pressure for genotypes exhibiting strong fitness defects. The reason for this is that the same formula for the dose-response curve (dynamically fitting only the half-inhibitory concentration parameter) is used for the antibiotic concentration adjustment in each culture. The dose response curve has three parameters: c_0 (the half-inhibitory concentration), g_{max} (the growth rate without drug) and n (the hill coefficient determining the steepness of the curve). Measuring the hill coefficient requires many measurements on a concentration gradient and comes with rather large error. In contrast, measuring g_{max} during the experiment for each culture would be possible, thus catering to the differences in the dose-response curve for each individual genotype as well as for how the curve changes during evolution (when cultures experience a fitness cost associated with resistance mutations).

We identified particular genes which had a strong effect on resistance evolvability. Our analysis suggests epistasis with resistance mutations is the main mode of evolvability modulation. However, more empirical data is needed to confirm this. For example, deleting the candidate evolvability modifying genes (Δ*dnaK*, Δ*lpcA*, etc.) in resistant strains and comparing the effect on resistance of the deletions alone, the resistant strains alone and those with all mutations to quantify any departures from additivity.

An extension of our screen to many more genotypes would lead to a comprehensive map of genetic determinants of resistance evolvability and provide a list of possible targets which could be chemically inhibited to increase the long-term efficacy of antibiotics.

2.4 Methods

2.4.1 Whole genome sequencing analysis

Whole genome sequencing was successfully performed for 380 samples altogether as listed in Table 3 in the appendix. For all evolved population samples, also the ancestral clone was sent for sequencing and mutations analyzed, to distinguish clearly between mutations acquired before and during the experiment. Genomic DNA was purified directly from thawed glycerol stocks using the GenElute 96 Well Tissue Genomic DNA Purification Kit. Library preparation, multiplexing, and sequencing were performed by LGC Genomics GmbH. The samples were sequenced on an Illumina NextSeq500 V2 (paired-end sequencing, 150bp read length, ~230-fold coverage on average, but ranging from about 70 to about 800 fold due to the multiplexing protocol). Sequencing data were analyzed using Breseq (Barrick et al., 2014) (Version 0.32.0). Reads were aligned to the deposited Keio parent reference (Accession: CP009273) using Bowtie2. The mutations identified by Breseq were manually inspected for false positives; all validated mutations are listed in Table 3 and Table 4 in the

Appendix. Even though the samples were expected to be heterogeneous (they were not isolated clones), the "clonal" mode of Breseq was used. Therefore, the mutations detected only represent fixed mutations. Amplifications were noted if the coverage of a multi-genic region exceeded twice the average coverage of that sample. Since an IS insertion in the *lon* promoter region was very commonly among the "unassigned new junction evidence" but at very high frequency, this type of mutation was assumed to be fixed if the frequency exceeded 90%. For each evolved sample, it was verified whether the gene which should be deleted is indeed deleted. If any reads in the deletion locus were present, which would suggest a cross-contamination with another strain, however small, the sample was excluded from the analysis.

2.4.2 Automatized experimental evolution

The chosen starting deletion strains from the Keio collection (Baba et al., 2006) as listed in Table 1 were all streaked for single colonies and clonal cultures frozen with 15% glycerol at -80°C. The glycerol stocks were used to assemble the starting 96-well plates. Each plate had at least 12 empty wells which were filled with clear medium, and handled like the other wells throughout the experiment to monitor cross-contamination. Replicates of the same ancestor, if on the same plate, were placed far from each other. Every plate contained at least two replicates (usually more) of the control strain (either the Keio parent or the Δ*lacA* strain).

Table 1 ***Lists of strains used in evolution experiments***

Experiment code	Deletion strains from the Keio collection
M1	*acrA, ΔacrR, ΔahpC, ΔampG, ΔatpF, ΔazoR, ΔcedA, ΔcodB, ΔcspE, ΔcysA, Δdam, ΔdedD, ΔdinB, ΔdnaK, Δfis, ΔfrdA, Δfur, ΔgalT, Δhda, ΔhisI, Δhns, ΔhslU, ΔhtpG, ΔhupA, ΔilvA, ΔkdpE, Δlon, ΔlpcA, Δlpp, ΔlpxM, ΔmarR, ΔmdtC, ΔmutL, ΔmutT, Δndh, parent, ΔpolB, ΔppiD, ΔproQ, ΔrecA, ΔrelA, ΔrfaG, Δrng, ΔrpoS, ΔrpsF, ΔseqA, ΔsodB, ΔtatC, ΔtolC, ΔumuC, ΔyccV*
M2	*acrA, ΔacrR, Δacs, ΔadhE, ΔagaR, ΔahpC, Δamn, ΔampG, ΔastC, ΔatpF, ΔazoR, ΔbolA, ΔcaiC, ΔcedA, ΔchbB, ΔcodB, ΔcpdA, ΔcspE, ΔcysA, Δdam, ΔdedD, ΔdinB, ΔdnaK, ΔfdnH, Δfis, ΔfrdA, ΔfumA, Δfur, ΔgalF, ΔgalT, ΔgltB, Δhda, ΔhelD, ΔhisI, Δhns, ΔhslU, ΔhtpG, ΔhupA, ΔilvA, ΔiscR, ΔkdpE, ΔlacA, Δlon, ΔlpcA, Δlpp, ΔlpxM, ΔltaE, ΔlysS, ΔmanX, ΔmarR, ΔmdtC, ΔmutL, ΔmutT, Δndh, ΔnfnB, ΔnudE, ΔnuoG, ΔompA, ΔompR, Δpepb, ΔphoB, ΔpldB, ΔpntB, ΔpolB, ΔppiD, ΔproQ, ΔpurA, ΔpyrB, ΔrcsB, ΔrecA, ΔrelA, ΔrfaG, Δrna, Δrng, Δrob, ΔrpiA, ΔrplA, ΔrpoS, Δrpoz, ΔrpsF, ΔrsuA, ΔsdhA, ΔseqA, ΔsmpA, ΔsodB, Δsra, ΔstpA, ΔtatC, Δtig, ΔtktA, ΔtktB, ΔtnaB, ΔtolC, Δtrxa, ΔtufA, ΔulaA, ΔumuC, ΔyccV, ΔyfgA*
M3	*acrA, ΔacrE, ΔacrR, Δcmr, ΔdedD, ΔdnaJ, ΔdnaK, ΔemrB, Δfis, ΔgalT, Δhda, ΔhslU, ΔhtpG, ΔlacA, Δlon, ΔlpcA, Δlpp, ΔlpxM, ΔmarA, ΔmarR, ΔmdtC, ΔmdtE, ΔmutL, ΔompF, ΔrecA, ΔseqA, ΔtatC, Δtig, ΔtktA, ΔtolC, ΔybaO*
M4	*acrA, ΔacrB, ΔacrR, Δacs, ΔadhE, ΔagaR, ΔahpC, Δamn, ΔampG, ΔastC, ΔatpF, ΔazoR, ΔbolA, ΔcaiC, ΔcedA, ΔchbB, Δcmr, ΔcodB, ΔcpdA, ΔcspE, ΔcysA, Δdam,*

	ΔdedD, ΔdinB, ΔdnaJ, ΔdnaK, ΔfdnH, Δfis, ΔfrdA, ΔfumA, Δfur, ΔgalF, ΔgalT, ΔgltB, Δhda, ΔhelD, ΔhisI, Δhns, ΔhslU, ΔhtpG, ΔhupA, ΔilvA, ΔiscR, ΔkdpE, ΔlacA, Δlon, ΔlpcA, Δlpp, ΔlpxM, ΔltaE, ΔlysS, ΔmanX, ΔmarA, ΔmarR, Δmdh, ΔmdtC, ΔmutL, ΔmutT, Δndh, ΔnfnB, ΔnudE, ΔnuoG, ΔompA, ΔompF, ΔompR, Δpepb, ΔphoB, ΔpldB, ΔpntB, ΔpolB, ΔppiD, ΔproQ, ΔpurA, ΔpyrB, ΔrcsB, ΔrecA, ΔrelA, ΔrfaG, Δrna, Δrng, Δrob, ΔrpiA, ΔrplA, ΔrpoS, Δrpoz, ΔrpsF, ΔrsuA, ΔsdhA, ΔseqA, ΔsmpA, ΔsodB, Δsra, ΔstpA, ΔtatC, Δtig, ΔtktB, ΔtnaB, ΔtolC, Δtrxa, ΔtufA, ΔulaA, ΔumuC, ΔyccV, ΔyfgA

The protocol was performed using the Tecan Freedom Evo 150 liquid handling platform. The 200µl cultures were kept in 96-well plates (Nunc, transparent flat-bottom) in a shaking incubator (Liconic Storex, 30°C, >95% humidity, 720rpm). Every 10-15min, each plate was transferred to a plate reader (Tecan Infinite F500) using a robotic manipulator arm (RoMa) and the absorbance (600nm) was measured. Every 3-5h, the cultures were transferred to new plates. They were not diluted in the same plates due to large errors in volumes left in the wells after pipetting out most of the culture. The new plate was filled in three steps. First, pure LB medium (v_{med}) was pipetted, than medium with antibiotic (v_{ab}) and last the culture ($v_{culture}$) from the previous plate. Each culture had a dedicated 200µl disposable tip for a day, which was washed in ethanol after every dilution. LB medium and antibiotic stock were multi-pipetted into the new plates using 1000µl tips. All tips were exchanged once a day. The reservoirs with media had lids that were only taken off using the RoMa arm just before usage.

Every day of the experiment, the penultimate plates of that day were left in the incubator to grow out: the next day 70µl of 50% glycerol was added to each well and frozen at -80°C. Fresh antibiotic stocks and medium reservoirs were filled. There were always two different concentrations of antibiotic stocks available, the protocol always only chose one of the available stocks to pipette from. The concentrations of the antibiotic stocks were chosen each day depending on how resistant the populations became.

Every 3-5h the cultures from each plate were transferred to new plates using the Air LiHa robotic arm. The appropriate volumes of culture, medium and antibiotic to use were calculated at each dilution step and for each culture using the OD values obtained since the last dilution. The growth rate was obtained from 18 consecutive OD measurements by obtaining the slope of the least squares linear fit (numpy.polyfit function) to the log2 of those background subtracted OD values which were between 0.01-0.1. All growth rates were normalized to growth rate without inhibition (1.7 doublings per hour). The volumes in µl were calculated separately for each well using the last background subtracted OD measurement (d), normalized fitted growth rate ($g_{current}$), concentration of antibiotic stock (c_{stock}), current antibiotic concentration in the well ($c_{current}$) and hill coefficient of the dose response curve ($n_{TET} = 1.8, n_{CHL} = 2.4$) in order to reach the target OD ($d_{target} = 0.01$), growth rate ($g_{target} = 0.5$) and total volume ($v_{total} = 200$) according the equations:

$$v_{culture} = v_{total} * d_{target}/d$$

$$c_{target} = c_{current}(\frac{g_{current}}{1 - g_{current}})^{\frac{1}{n}}$$

$$v_{ab} = (v_{total} * c_{target} - b * c_{current})/c_{stock}$$

$$v_{med} = v_{total} - v_{culture} - v_{ab}$$

Several measures were implemented to deal with atypical input values. If the concentration $c_{current}$ is zero, c_{target} is set to a default concentration of 0.1µg/ml for tetracycline and 0.5µg/ml for chloramphenicol. If the sum of squared residuals from the fit to obtained the growth rate is greater than 0.8 then c_{target} is set to $c_{current}$. If the measured normalized growth rate is larger than 0.9, it is set to 0.9 to avoid very large or undefined values for c_{target}. If the calculated volume v_{ab} is smaller than 5µl, v_{ab} is set to zero (only medium is used to dilute the culture) and concentrations are updated accordingly. $v_{culture}$ is capped at 140µl, to assure accurate pipetting from the small 200µl culture. There were two available reservoirs of antibiotic stocks, the higher concentration was only used if, for the lower stock concentration $v_{ab} > v_{total} - v_{culture}$.

Table 2 ***Differences between instances of the automatized evolution experiments***

	M1	M2	M3	M4
Antibiotic	TET	TET	TET	CHL
Number of different strains	51	99	31	104
Replicates of strains	3	3	8-12	3
Control strain	Keio parent	Δ*lacA*	Δ*lacA*	Δ*lacA*
Replicates of control	10	22	11	23
Medium	LB	LB+50µg/ml kan	LB+50µg/ml kan	LB+50µg/ml kan

2.4.3 Serial transfer experiments with constant antibiotic concentration

Overnight cultures were inoculated from a selection of 84 deletion strains from the Keio collection. The cultures were assembled in a 96-well "selection" plate and frozen with glycerol (final concentration 15%). The first day, ~0.2µl from each well was transferred (using the VP 408 pinner) to several fresh plates with 150µl medium with or without antibiotic. The conditions depend on the experiment (Table 2). Every 24hours, the OD was measured in a plate reader and all cultures transferred with the VP 408 pinner to fresh plates with the same conditions as the previous day. 65µl of 50% glycerol was added to all the wells and the plates were frozen at -80°C. The growth rate of all the cultures in the same conditions as they were evolved in was then measured on the same day using the liquid handling robot as described in (Chevereau and Bollenbach, 2015).

2.4.4 Antibiotic stocks

Tetracycline stock solutions of 7mg/ml were prepared by diluting tetracycline hydrochloride in 90% ethanol at room temperature. Chloramphenicol stocks of 10mg/ml were prepared by diluting powder in 99% ethanol. Trimethoprim was dissolved in ethanol to prepare 10mg/ml solution. All stocks were stored at -20°C.

2.4.5 Regression model of mutational effects

Mutations from all sequenced samples evolved in Tetracycline which passed contamination and quality control (145) were included in the analysis from Figure 15. The selection was biased toward mutants that evolved slower than expected. The function fitnlm (Matlab R2016b) was used to fit different models of mutational effects. The predicting variables were the presence and absence of mutations in the 5 most commonly hit genes (*marR*, *lon*, pump amplification, *acrR*, *ybaO*), the fitted parameters were the multiplicative effects of mutations in those loci and the response variable was the log resistance increase over the course of the experiment.

	Estimate	SE	tStat	pValue
b0	1.3476	0.19756	6.821	2.55E-10
b1 (marR)	1.1539	0.17973	6.4201	2.00E-09
b2 (*lon*)	1.1439	0.18006	6.3528	2.81E-09
b4 (pumps)	0.26567	0.19734	1.3462	0.18042
b5 (*acrR*)	0.35865	0.19679	1.8225	0.070526
b6 (*ybaO*)	0.80957	0.18072	4.4798	1.55E-05

In a variation of the analysis, another predictor variable was added (and replaced one of the fitted parameters) which corresponded to the measured initial resistance (IC_{50} from the concentration of antibiotic in the well in the first hours of the experimental evolution protocol) of the ancestral clone and the final absolute log resistance level instead of the relative increase was predicted. To model diminishing returns, the whole linear formula was channeled through a nonlinear function: either $f(x)=a*x^b$ or $f(x)=(a*x)/(b+x)$.

3 Quantitatively predicting the evolutionary dynamics of antibiotic resistance based on mutation rate and dose-response characteristics

This chapter is partly based on (Chevereau et al., 2015). The simulations, the whole-genome sequencing analysis and the experimental evolution of mutator strains were done by ML. The growth rate experiments and analysis of dose response curves, distributions of fitness effects and their relationship, which represent the bulk of the published paper were done by Guillaume Chevereau. The analysis of mutational diversity was done by Tobias Bollenbach. The "morbidostat" evolution experiments of *E.coli* in nitrofurantoin and chloramphenicol were performed by Tugce Batur, Aysegul Guvenek, Dilay Hazal Ayhan, and Erdal Toprak (and published in (Chevereau et al., 2015)).

3.1 Introduction

The dynamics of resistance evolution differs greatly between antibiotics and explanations remain elusive. For example, spontaneous resistance to chloramphenicol evolves steadily and rapidly under laboratory conditions, but resistance to trimethoprim seems much to develop in large steps with long waiting times in between (Toprak et al., 2012). It is intriguing to investigate whether differences in evolution can be explained by systematic differences in the inherent properties of the antibiotics.

Highly mutating strains pose a threat to successful treatment and understanding their behavior is crucial for attempts to avoid them (Jolivet-Gougeon et al., 2011). The mutation rate of bacteria profoundly affects the dynamics of resistance evolution, in terms of speed, genetic and phenotypic reproducibility and eventual fate of resistance alleles (Sniegowski et al., 1997).

Quantitative population genetic models, coupled with highly controlled experimental evolution, provide a way to investigate general relationships between measureable antibiotic or bacterial parameters and resistance evolution outcome. On the one hand, classic theoretical results shed light on general behaviors; on the other hand, targeted models closely following the conditions under which populations evolve in the laboratory can provide more specific almost mechanistic insight into antibiotic and bacteria specific dynamics. In this chapter, with the help of quantitative models and experimental evolution, we investigate the ways in which antibiotic dose-response characteristics and bacterial mutation rates affect evolutionary dynamics and reproducibility.

A basic functional measure of antibiotic action is the concentration needed to inhibit bacteria. The most common measure is the MIC, which is the minimum concentration needed to inhibit bacterial growth (over ~24hours). Another common measure is the IC_{50}, which is the concentration at which bacteria grow at half their uninhibited rate. This measure differs greatly among antibiotics as well as bacterial strains, and when a certain strain exceeds an agreed on clinical threshold, it is considered resistant. The MIC and IC_{50} do not give any information on how sensitive cell growth would be to small change in concentration.

The detailed relationship of how cell growth depends on changes in concentration is given by the dose-response curve. The dose-response curve is simply a measurement of growth rates for a range of concentrations of antibiotic, capturing the whole range from uninhibited

growth to no growth, for bactericidal drugs even extending to negative (or death) rates. To summarize the curve in just a few parameters, a hill function fit is often used:

$$g = \frac{g_{max}}{1 + \left(\frac{c}{c_0}\right)^n}$$

In this form, $g_{max}$is the uninhibited growth rate, c_0 is the IC_{50} concentration and n is the hill coefficient, roughly representing the slope of the curve, or the sensitivity to changes in dose.

Antibiotics can be very different in their dose sensitivity (Figure 17). Several beta-lactam antibiotics have extremely steep dose response curves, where very slight changes to concentration mean bacteria go from full to no growth; in contrast, trimethoprim has an extremely shallow dose response curve, and therefore small value for *n*.

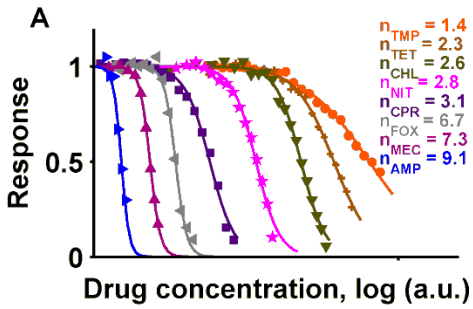


Figure 17 ***Antibiotics differ in the hill coefficient of their dose response curves.***

Figure adapted from Figure 3 in (Chevereau et al., 2015). Dose-response curves are shown for 8 different antibiotics: trimethoprim (TMP), tetracycline (TET), chloramphenicol (CHL), nitrofurantoin (NIT), ciprofloxacin (CPR), cefoxitin (FOX), mecillinam (MEC) and ampicillin (AMP). Response is given by the growth rate normalized to growth rate in now drug. The drug concentration is shown on a log-scale in arbitrary units. The curves are shifted on the concentration axis to better show the differences in their shape. The fitted hill-coefficient (n) is given for each curve in the upper right corner of the plot.

It is useful to note that the steepness of the dose-response curve is a robust property of the antibiotic rather than the particular bacterial strain. This is supported by an experiment, where the dose-response curve was measured for a selection of 78 *E.coli* single gene deletion strains (Chevereau et al., 2015). When the growth rates where rescaled to the growth rates without drug and the concentrations to the respective IC_{50}, the curves generally collapsed onto one curve, showing that the shape is conserved despite the genetic perturbations (Figure 18). Another way to formulate the result is that the dose sensitivity remains approximately constant for a range of bacterial mutants.

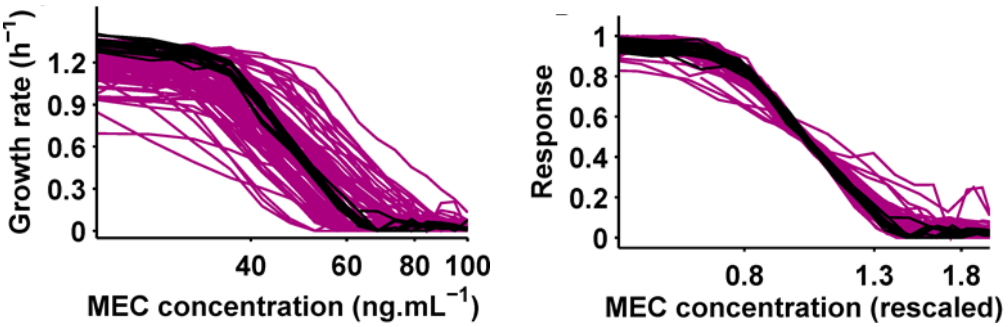


Figure 18 ***Dose response curves of a range of deletion strains collapse on one curve after linear rescaling***

Figure adapted from Figure 3 in (Chevereau et al., 2015). Left: Mecillinam dose-response curves for 78 arbitrary deletion mutants (purple, see Methods) and 17 WT replicates (black). Right: Same data with concentration rescaled to IC_{50} and growth rate response rescaled to g_0 (Materials and Methods).

Both mutation rate and dose-response characteristics could have a systematic impact on evolutionary dynamics. In this chapter, the relationship between these determinants and speed and/or reproducibility of adaptation will be explored with experimental evolution and population genetic simulations closely matching experimental conditions.

3.2 Results

3.2.1 Population genetic simulations model evolution in a morbidostat

In order to predict and better interpret results from evolution experiments with dynamically increasing antibiotic dose similar to those described in Section 2.2.1, we developed stochastic population genetic simulations which model this scenario. They take the mutation rate and the distribution of resistance effects of mutations as input and, as output, give the corresponding resistance increase over time.

The simulation scheme was adapted from (Fogle et al., 2008) and it is a discrete generation, fixed population size stochastic model which models drift explicitly. The population is modelled as a set of subpopulations, representing different genotypes, each with its own resistance level. The antibiotic dose is adjusted every generation based on the mean growth rate of the population, so that the new mean growth rate of the population is expected to be a fixed value (usually 0.5 of the wild type growth rate). For each subpopulation, the number of individuals, growth rates and resistances are tracked. Each generation, the population is approximately halved, in a way where for each individual the probability of survival into the next generation is weighed by their current growth rate, but on average 0.5. This is done to model the effects of drift and selection. To keep the population size approximately constant, the whole population is then doubled irrespective of fitness. In the next step, mutations occur and found new genotypes. The way the new fitness of each mutation is selected is based on the results in (Chevereau et al., 2015). Each mutation first has an effect on resistance (or IC_{50}) according to a given lognormal distribution, which is then translated into fitness using the dose-response curve typical for the appropriate drug (see Figure 19 or Figure 2B) and the current antibiotic concentration. The antibiotic concentration (common for the whole population) is then adjusted so that the mean growth rate of the population is equal to the fixed growth rate selected (usually 0.5 of the wild type).

3.2.2 Analysis of high-throughput experiments shows clear relationship between dose-sensitivity and DFE width

Based on the experiments in (Chevereau et al., 2015) a clear correlation can be seen between the dose sensitivity (hill coefficient n of the dose-response curve) and the width of the DFE. The DFE in this case was estimated by measuring the growth rate of the whole single-gene deletion library (Baba et al., 2006) in a given drug concentration. The widths of these distributions varied considerably among antibiotics.

In contrast to the DFE, the distribution of resistance effects of the gene deletion library exhibited a stereotypical width. The distribution of resistance effects was estimated by computing the "effective concentration" of antibiotic that the population senses when growing at a particular rate in a particular concentration (for details see (Chevereau et al., 2015)). This knowledge is interesting in its own right, suggesting that the steps that mutations take are more stereotypical in terms of resistance changes rather than fitness changes. This is in line with the idea of Fisher's Geometric Model (FGM), which assumes mutations represent steps in a multi-dimensional phenotype space.

Given that the changes in effective drug concentration caused by mutations are similarly strong in many antibiotics, the widths of the DFEs can be explained by a simple mapping through their respective dose-response curve (Figure 19). In cases of high dose sensitivity (and steep dose-response curve), a narrow distribution on the concentration axis translates to wide distribution on the growth rate (fitness) axis and conversely, a shallow dose response curve leads to a narrow DFE given the same distribution on the concentration axis.

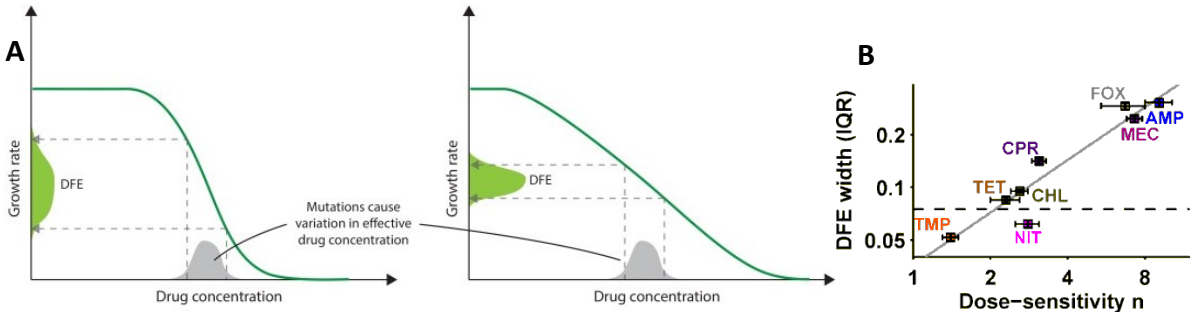


Figure 19 ***Steeper dose response curves lead to wider DFEs***

A) Schematics of two different dose response curves: the left curve is steep, i.e. the growth rate is sensitive to small changes in drug concentration; the right curve is shallow. Mutations cause shifts in the effective drug concentration a bacterium experiences; the typical magnitude of these shifts is surprisingly similar for diverse antibiotics (Chevereau et al., 2015). The distribution of effective drug concentrations resulting from many different mutations is shown in gray. These mutations produce distributions of growth rates (fitness) that are wide for the steep dose-response curve (left) and narrow for the shallow dose-response curve (right).B) Figure adapted from Figure 3 (Chevereau et al., 2015). Fitted hill coefficients of dose response curves (dose-sensitivity n) for different antibiotics plotted against the interquartile range of their DFEs.

3.2.3 Simulations predict faster evolution to antibiotics with steeper dose-response curves

Together with classical evolutionary biology results, specifically, Fisher's Fundamental Theorem, it is expected that a wider DFE would lead to faster evolution. We used the population genetic simulations which model an experimental setup with dynamically increasing antibiotic doses to test if steeper dose-response curves lead to faster evolution also in this testable context. We find that, indeed, higher values of dose-sensitivity (n) lead to faster resistance evolution in a morbidostat setting (Figure 20).

Even though the distribution of resistance changes of new mutations was observed to be stereotypical for various drugs (Figure 19), there were also differences observed. In particular, the distribution for the drug nitrofurantoin was considerably narrower than for the other drugs tested (Chevereau et al., 2015). Assuming, the same steepness of the dose-response curve, a narrow distribution of resistance effects is expected to lead to a narrower distribution of fitness effects and therefore slower adaptation.

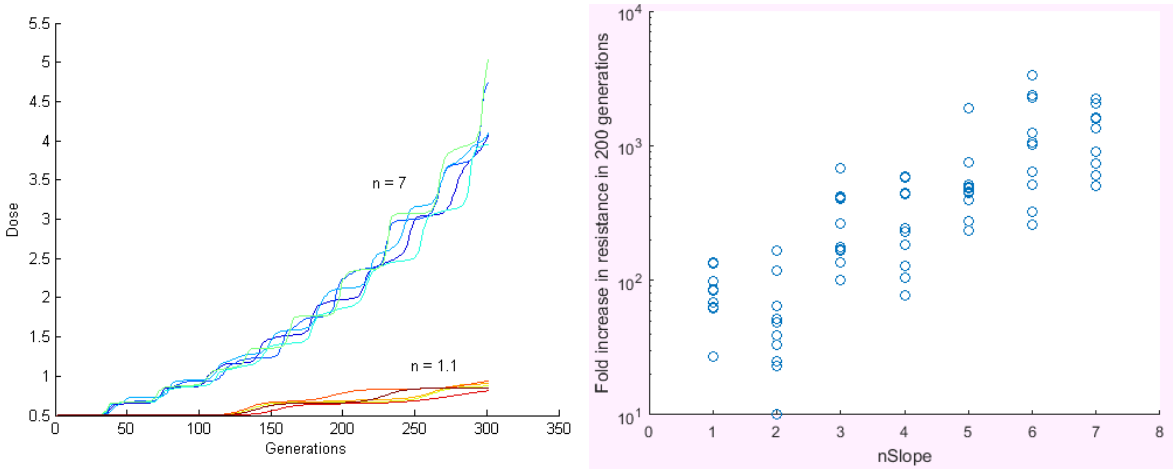


Figure 20 ***Higher dose-sensitivity leads to faster adaptation.***

Left: The increase in dose (mimicking resistance increase over time, since the dose increases in a way to keep the mean growth rate inhibited by 50%) for two dose-response curve shapes. High dose sensitivity (n=7) is shown in blue-green and low dose sensitivity (n=1.1) is shown in orange and red. Right: Fold increase in resistance after 200 generations for a range of dose sensitivities.

3.2.4 Experimental evolution in morbidostats is consistent with population genetic simulations in two antibiotics with different evolutionary determinants

Nitrofurantoin and chloramphenicol represent a drug pair for which the estimates for the dose sensitivity (n) are essentially the same, but the distribution of resistance effects vary significantly (Chevereau et al., 2015). Therefore, they are good candidates for testing the relevance of our pre-measured evolutionary determinants for experimental evolution.

The dynamics of resistance adaptation to the two drugs were compared using the morbidostat protocol (Toprak et al., 2012). In this protocol (similarly to the one described in Chapter 2), the antibiotic concentration is continually adjusted to keep the culture 50% inhibited. In this way, the dynamics can be observed over a longer time frame and differences can be seen more clearly than in a protocol with fixed concentration.

We observed that evolution of chloramphenicol resistance followed a steadily increasing trajectory, whereas the evolution of nitrofurantoin resistance consisted of an initial fast increase followed by a more modest increase. A possible reason for this difference in dynamics can be inferred from whole genome sequencing data. The initial increase in resistance to nitrofurantoin can be ascribed to a handful of very reproducible mutations which arise within the first 10 days. Among these mutations, there are loss of function mutations of *nfsA* and *nfsB*. These two enzymes are responsible for modifying nitrofurantoin into its active form. Loss of function mutations in them are known to provide large resistance benefits (Breeze and Obaseiki-Ebor, 1983). In contrast, adaptation to chloramphenicol sees a higher diversity of mutations throughout the experiment.

We adjusted our simulations to incorporate these few very beneficial mutations which can strongly influence adaptation. To cater for the strong mutations, the DFE could be modelled as a smooth distribution with a very long tail, but in that case, also the effects of the fixed mutations would vary enormously between replicates. Instead, we decided to update the

model with explicitly adding a few extremely strong mutations (see Methods). After this modification, the simulation predictions match experimental data very well Figure 21.

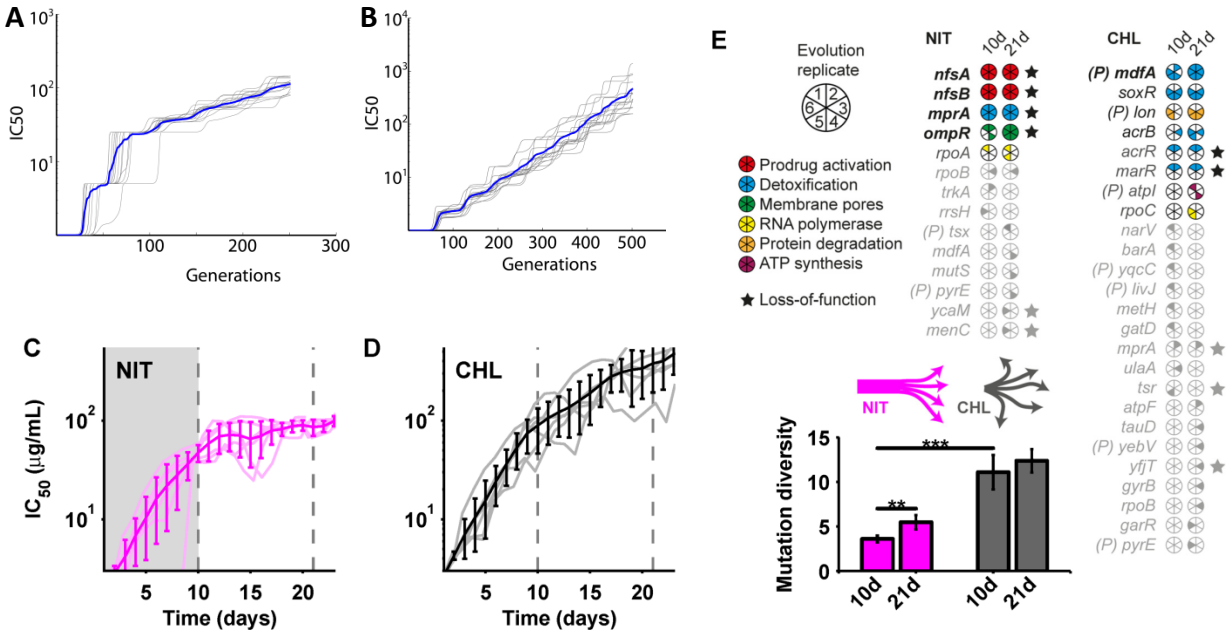


Figure 21 ***Resistance variability affects the dynamics of adaptation***

A) Simulation results from a theoretical model of resistance evolution in a morbidostat: IC50 increase over time for a drug with narrow distribution of resistance effects and two available large-effect mutations; light lines are sample runs; dark lines are mean of 15 runs (Materials and Methods). (B) Same as in panel A for wider distribution of resistance effects (Methods). (C, D) Results from morbidostat laboratory evolution experiments: IC50 increase over time for nitrofurantoin (C) and chloramphenicol (D); light lines are individual runs; dark lines are mean, error bars standard deviation; shaded region in C indicates early phase during which large-effect mutations fix (Methods). (E) Mutated loci in nitrofurantoin (left) and chloramphenicol (right)-resistant clones after 10 and 21 d, respectively. Filled pie segments show evolution replicates in which genes were mutated; (P) indicates promoter mutations. Bar chart shows diversity (entropy) of mutations under nitrofurantoin (magenta) and chloramphenicol (gray); p < 0.002 (**) and p < 0.0003 (***) from two-sample t test; error bars show jackknife standard error (Methods). (C, D, E) are taken from (Chevereau et al., 2015)

3.2.5 Simulations predict higher phenotypic reproducibility for highly mutating genotypes.

By definition, higher mutation rates are associated with higher genetic variability, but also with more thorough sampling of genotypes, which can, counter-intuitively, lead to higher reproducibility of evolutionary outcomes. When following how a population becomes resistant over time, at the level of phenotype, one can observe several sources of randomness. One comes from random drift, where the size and eventual fate of mutations is influenced by random sampling effects. This could be regular dilutions in various experimental evolution setups, for example. Another source depends on mutational availability and the resulting variability in time until the next mutation appears. It is this second source of variability that is diminished with higher mutation rate. In the extreme case of "perfect" mutational availability, all possible mutations would be present at all times

and, at each point in time, one of the strongest available would win. In a context of no epistasis, this would lead to reproducible phenotypic outcomes.

To verify this intuition, we used the simulations described in the previous sections to model evolution in the morbidostat for a range of mutation rates. The results indeed show a decrease in the variability of adaptation rates as mutation rates increase (Figure 22).

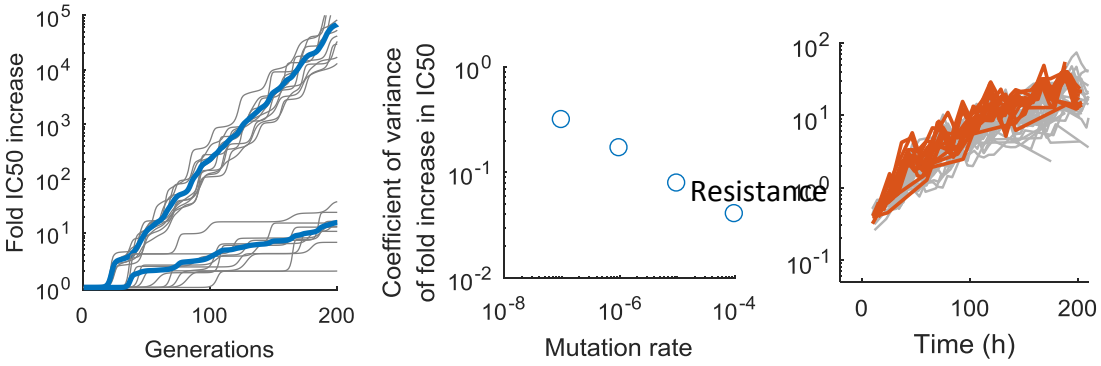


Figure 22 ***Increased mutation rates bring about more reproducible phenotypic adaptation.***

Population genetic simulations were performed with varying mutation rates (Methods). Left: The IC_{50} increase over time for two different mutation rates: 10^{-7} and 10^{-5}. Grey lines represent individual simulation runs; the blue line is the mean over 20 replicate runs. Middle: The coefficient of variance of the log values of IC_{50} from generation 200 is given for different mutation rates. Higher mutation rates lead to lower variability of outcome. Right: Example from evolution experiment M3 in increasing antibiotic concentration (see Methods of Chapter 2). Fold increase in IC_{50} (given as antibiotic concentration in well during dynamically adjusted evolution experiment) for a highly mutating strain (mutL deletion in orange) and control (lacA deletion strain).

3.2.6 Predictions of reproducibility are largely confirmed by experimental evolution

As part of the experiments in Chapter 2, we tested deletion strains of DNA mismatch repair which are known to increase mutation rates around 100-fold (*mutT* and *mutL*). In line with simulation predictions, the adaptation trajectories of these strains show an increase in adaptation rate and decrease in the variability of adaptation (Figure 22).

3.3 Conclusions

In this chapter, we have presented a population genetic model which is specially catered to predict or analyze outcomes of morbidostat-type experiments. Into it, we incorporated a recently established framework, which predicts evolutionary determinants from dose-response characteristics of drugs. Using this framework, we illustrate that steepness of the dose-response curve can increase adaptation speed. In addition we show that the distribution of resistance effects, and in particular its variance, can determine the speed of adaptation. We illustrate this with simulations as well as experimental evolution results and find that in this case, a lower width of the distribution of resistance effects can lead to more reproducible genotypic adaptation. We further model the effects of mutation rate on the

reproducibility of phenotypic adaptation and see that higher mutation rates can lead to higher phenotypic reproducibility.

3.4 Methods

3.4.1 Population genetic simulations

Simulations were setup in a similar manner to those in (Fogle et al., 2008). For each genotype that exists, its resistance, their growth rate and the number of individuals which belong to it are being saved. A fixed population size and synchronized generations are assumed. Every generation, the population is essentially halved: the probability of being chosen into the next generation for each individual is weighted by their growth rate, but 0.5 on average. To make the calculation more efficient, for larger subpopulations (>300 individuals), the number of individuals surviving to the next generation is taken from a Poisson distribution with the mean $0.5*n_i*g_i/\langle g \rangle$. Then all subpopulations are exactly doubled. This scheme ensures an approximately fixed population size, a realization of fitness effects and random drift. The next step which also happens every generation, are mutations. First, the number of mutations is drawn from a Poisson distribution with mean corresponding to the mutation rate. Then the cumulative probability that each subpopulation experiences a mutation is calculated ($n_i*\mu$). For each mutation, a random number is drawn, which is used to sample the cumulative distribution function and determine which subpopulation the mutation happened in. Then, for each mutation, one individual is subtracted from the parental subpopulation and one new subpopulation consisting of one individual is founded. A new resistance (c0) is drawn from a log-normal distribution centered on the c0 of the parental genotype, for the newly created genotype. Based on the dose response curve (with given parameters n and the new c0), and the current antibiotic dose (c), the growth rate or fitness (g) is calculated for each genotype (g=1/(1 + (c./c0).^n)).Then the antibiotic dose is adjusted according to the same dose response curve such that the mean growth rate of the entire population is 0.5.

3.4.2 Whole genome sequencing analysis

Five clones were isolated from each of the six replicates and their dose response curves compared with those of the evolved populations. Whole genome sequencing was performed for one clone from each population for which the resistance matched the population and also for the MG1655 ancestor. Genomic DNA was purified from overnight cultures using the Promega Wizard Genomic DNA Purification Kit (catalogue number A1120). Library preparation, multiplexing, and sequencing were performed at the EMBL GeneCore facility. The samples were sequenced on an Illumina HiSeq2000 (paired-end sequencing, 100bp read length, ~140-fold coverage). Sequencing data were analyzed using Breseq (Barrick et al., 2014) (Version 0.25) and Geneious (Kearse et al., 2012) (Version 7, http://www.geneious.com). Reads were aligned to the deposited MG1655 reference (NC_000913) using Bowtie2. The mutations identified by Breseq were manually inspected for false positives; regions with ambiguous evidence were further examined in Geneious; all validated mutations are listed in Table S2. We identified several mutations in the ancestor (Table S3); these were included in the reference sequence and reads from the ancestor realigned to this new reference until no additional mutations were identified by Breseq. A nitrofurantoin resistant clone from day 21 was sequenced in duplicate to verify

reproducibility of sequencing results; the mutations identified in both sequencing replicates agreed perfectly.

3.4.3 Experimental evolution of mutator strains

Experiments were done as described in Section 2.4.2.

4 Stress-induced mutagenesis: Stress diversity facilitates the persistence of mutator genes

This chapter has been published in PLOS Computational Biology (Lukačišinová et al., 2017), where Sebastian Novak gave equal contribution as ML. The analytical treatment of the dynamical system was done by Sebastian Novak. The simulation code, analysis, writing and conceptualization were done in collaboration with Sebastian Novak and Tiago Paixão.

4.1 Introduction

New mutations are the ultimate source of the variation that fuels adaptation. Accordingly, any mechanism that affects the rate at which new mutations are produced will impact a population's ability to adapt. In a constant environment, well adapted populations are expected to evolve to lower mutation rates (Altenberg and Feldman, 1987), while maladapted populations (e.g. under stress) frequently evolve mutator phenotypes (Taddei et al., 1997). These mutator phenotypes rise due to the short-term benefit of hitchhiking with beneficial mutations they induce, but suffer a penalty as the population approaches a fitness peak and the availability and/or effect of beneficial mutations decreases (Tenaillon et al., 2001). Consequently, elevated mutation rates persist most easily when selection pressure can be sustained despite adaptation. This is the case when populations experience fluctuating selection, which includes cells experiencing bursts of stresses during antibiotic treatment or cancer chemotherapy. Intuitively, if there exist mechanisms that increase mutation rates specifically during periods of stress, they could be selected since they provide the benefits of elevated mutations rates under stress while not incurring an additional mutation load in times when the population is well adapted.

There are multiple known mechanisms that result in elevated rates of general mutagenesis or an increase in the rate of specific genetic changes during stress. These mechanisms are often referred to as stress-induced mutagenesis (SIM). When encountering a range of environmental stresses, several species of bacteria activate SOS responses that—in addition to stimulating various repair mechanisms—activate error-prone DNA polymerases, which have been linked to a faster evolution of antibiotic resistance (Bjedov et al., 2003; Cirz et al., 2005; Do Thi et al., 2011). This activation of error-prone DNA polymerases in response to a wide range of environmental stresses is a thoroughly studied SIM mechanism (Bjedov et al., 2003; Devon M. Fitzgerald et al., 2017). The mutations are incorporated in proximity to DNA double-strand breaks under the condition that both the DNA damage activated SOS response and general stress response are active (Ponder et al., 2005). This mechanism has been linked to faster evolution of antibiotic resistance (Cirz et al., 2005).

Several other such mechanisms have been identified. It has been shown that *Streptococcus pneumonia*e activates the expression of competence genes when treated with various antibiotics (Prudhomme et al., 2006). These genes allow the bacteria to take up DNA from the environment and incorporate it into its genome. Another example is the beneficial excision of a genomic region in the plant pathogen *Pseudomonas syringae* in response to the host's immunity (Pitman et al., 2005). Similar mechanisms that link certain stresses to an increase of mutation rates have been found in *Drosophila melanogaster* (Agrawal and Wang, 2008) and yeast (Heidenreich, 2007).

Several hypotheses may explain the prevalence of stress-induced mutagenesis. The first is a pleiotropic argument, presuming that SIM mechanisms are primarily due to first order

selection for faster repair or nutritional gain (uptake of foreign DNA); then, the elevation of mutation rates is a side effect (Torres-Barceló et al., 2015). MacLean et al. (MacLean et al., 2013) suggest an alternative hypothesis to explain the stress-linked induction of error-prone DNA polymerases: DNA polymerases that are linked to specific stress situations and that are used less often may be subject to weaker selection, and become error-prone by accumulation of slightly deleterious mutations. Another hypothesis, the second-order selection hypothesis, states that stress-induced mutagenesis has evolved due to its advantage of combining elevated mutation rates with those situations when they give most benefit (Ram and Hadany, 2012; Rosenberg, 2001). An allele that causes elevated mutation rates hitchhikes with the beneficial mutations it produces. By keeping mutation rates down at times of no stress, it reduces mutational load from excessive deleterious mutations compared to unconditionally increased mutation rates. There is no reason to think that only one of these hypotheses is correct; it is plausible that the interplay of these factors is responsible for the prevalence of SIM mechanisms in many organisms.

We explore the basic principle behind the second-order selection hypothesis of stress-induced mutagenesis: under which conditions and at what levels can a mechanism that increases mutation rates under stress be sustained in a population? What stress patterns and regimes promote it most? The relevance of these questions is imminent: stress-induced mutagenesis facilitates the adaptation of a population subjected to changing conditions. This is critical for cancer therapy or antibiotic treatment. Much effort goes into identifying strategies that keep the treatment effective for as long as possible, i.e., that impede the evolution of resistance (Bollenbach, 2015; Bonhoeffer et al., 1997). If second-order selection is a key factor in the emergence and maintenance of SIM genes, however, different treatment regimes also affect the evolution of mutagenesis, and thus the evolvability towards resistance in the long term. It is therefore essential to understand to what extent different patterns of changing conditions cause second-order selection on stress-induced mutagenesis.

Previous studies have analyzed the evolution of mutator alleles by second-order selection in constant and variable environments. Some have focused on the evolution of mutation rates and the fate of constitutive mutator alleles (Ishii et al., 1989; Sniegowski et al., 1997), which are predicted to be lost in constant and persist in fluctuating environments (Tanaka et al., 2003). Closer to the system we study here, several models have investigated the evolution of fitness-dependent mutagenesis, where a decrease in fitness due to any deleterious mutation causes mutation rate to increase (Agrawal, 2002; Ram and Hadany, 2012; Tenaillon et al., 2004). Interestingly, fitness-dependent mutator alleles are predicted to persist under a wide range of parameters in variable as well as constant environments (Ram and Hadany, 2012). To complement existing studies, we explore the persistence of a SIM allele which is strictly conditional on a stressful environment and cannot be triggered by a genetic change, since we focus on those SIM mechanisms which are dependent on environmental stress responses (Devon M. Fitzgerald et al., 2017; Mamun et al., 2012; Prudhomme et al., 2006).

We apply a mathematical model to investigate the plausibility of the second-order hypothesis for the evolution and maintenance of stress-induced mutagenesis (SIM). We show that populations subjected to diverse stresses can maintain SIM alleles as long as the period between exposure to these stresses is below a critical threshold. We provide analytical expressions for this critical threshold and show that there is an upper limit to the

prevalence of such an allele, irrespective of the number of stressful environments a population is exposed to. Finally, in the context of the evolution of antibiotic resistance, we evaluate different scheduling alternatives of antibiotic therapies for their ability to prevent the maintenance of SIM alleles.

4.2 Results

4.2.1 A population genetic model for stress induced mutagenesis (SIM) alleles

We set up a model of a hypothetical stress-induced mutator (SIM) allele; its properties are based on the features of existing SIM mechanisms, yet focusing on the essence of a SIM mechanism independent of the molecular implementation. We are interested in exploring specifically the effectiveness of second-order selection in the evolution of a SIM allele. To do so, we need to isolate second-order selection from any direct benefit of the SIM system. Direct effects, for example faster DNA repair, are likely co-determinants of the persistence of SIM mechanisms in the wild, but such dynamics have also been extensively studied with existing evolutionary models, e.g. (Hegreness et al., 2006). We therefore assume that the SIM allele does not confer any direct fitness cost or benefit, and consider a population of haploid individuals with two non-recombining loci. At the first locus, the SIM allele can be present or absent (alleles M or m, respectively), and the second locus carries alleles that may or may not grant resistance to a given stress (alleles R or r). The resulting four possible genotypes are displayed in Figure 23.

In the absence of stress, we assume that transitions between the genotypes are only due to mutations, as indicated by the arrows in Figure 23; thus in particular, we assume that there is no cost to being resistant. Individuals may lose or gain resistance at rates μ_R and ν_R, respectively. The SIM allele M may lose its function at rate μ_M; since we are interested in conditions for the ultimate loss of the SIM allele, we neglect back-mutation from m to M.

In the stress environment, genotypes containing the resistance allele R have increased fitness $w = 1 + s$ relative to susceptible genotypes. Furthermore, the Mr genotype increases all outgoing mutation rates by a factor $\sigma > 1$ due to stress-induced mutagenesis, see Figure 23B. Key assumptions behind this modelling approach are: First, stress does not activate the SIM allele in resistant individuals. This is reasonable if, for example, the stressor is effective inside the cell but the resistant mutation makes the cell membrane impermeable to it. Second, the only cost of an active SIM allele is that it increases the rate of its own loss. This at best partially represents the detrimental effects of elevated mutation rates not considered in this model. However, artificially creating an idealized situation for the SIM allele allows us to keep the model tractable.

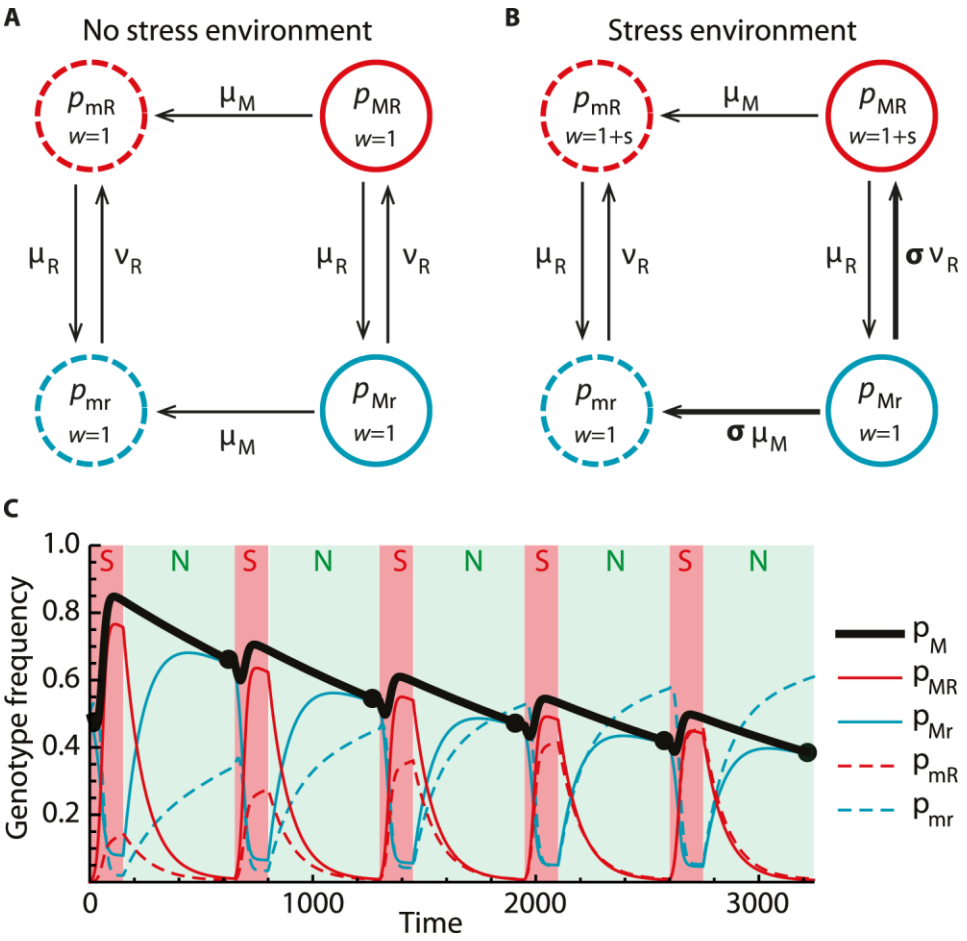


Figure 23 ***Schematic illustration of the SIM dynamics.***

A) Under no stress, all genotypes have the same fitness w = 1 and transitions between the states are solely due to mutations. Resistance is lost and gained at rates μR and vR, respectively. Furthermore, the SIM allele degrades at a rate μM. B) In the stress environment, individuals that are resistant to the stress gain a selective advantage s (fitness w = 1 + s). In addition, the genotype that is susceptible to the stress and carries the SIM allele (pMr) increases its outgoing mutation rates by a factor σ > 1. C) Periods of stress (S, red shading) and no stress (N, green shading) are alternated and the dynamics Eq (1) of genotype frequencies is simulated according to the schematic A) and B). During stress, resistant genotypes increase in frequency (red lines), and the SIM allele frequency pM hitch-hikes (black line). If there is no stress, both resistance and SIM allele frequency levels decay. Over time, the SIM allele frequency thus fluctuates, possibly converging to stable oscillations. We sample the SIM allele frequency pM at the end of each no-stress phase (black points), thus obtaining a discrete system in which the time between two successive measurements is given by the iteration of one cycle of stress and no stress.

We cast the schematic dynamics of Figure 23 into two sets of differential equations for the variables p = {p_{mr}, p_{Mr}, p_{mR}, p_{MR}}. Using the classical mutation-selection dynamics of population genetics, they take the form (Bürger, 2000)

$$\dot{p} = p(w - \overline{w}) + M.p \qquad (1)$$

where w is the vector containing the marginal fitnesses of the genotypes, $\overline{w}$ is the mean fitness of the population, and M is a matrix encoding the mutation scheme (see Methods for the explicit set of equations). In order to make analytical progress, we make a number of simplifying assumptions. First, we assume that selection under stress is strong compared to

mutation. This is justified in treatment-like scenarios we consider here. Second, we assume that the mutation rate leading to a loss of a resistance mechanism is larger than the mutation rate leading to its gain. This is plausible, since by random genetic modifications it is more likely to disable a functional mechanism than to create one. Hence, it seems reasonable to assume the following hierarchy among the parameters:

$$s \gg (\mu_M, \mu_R) \gg \nu_{r=R} \quad (2)$$

Given this hierarchy, it can be shown that the SIM allele is not maintained in any of the two environments separately. Switching between the stress and no-stress environments, however, gives rise to non-trivial dynamics. During a stress phase, the SIM allele may increase in frequency along with the resistance mutations it produces. As resistance levels in the population rise, this effect weakens and the SIM allele frequency falls off because of mutations degrading the SIM mechanism. Periods of no stress allow resistance levels to decline, such that the SIM allele becomes effective again in the next stress phase.

4.2.2 SIM alleles are maintained at higher frequencies under diverse stresses

In order to obtain analytical insight into the dynamics of SIM alleles we first consider two extreme scenarios: the recurrent (R) and non-recurrent (NR) stress scenarios. In both scenarios, an infinitely large population is subjected to an environment that alternates between periods of stress (for τS time units) and of no stress (for τNS time units). In the recurrent scenario, the stress periods are assumed to be all the same (i.e. resistance acquired in the previous stress period carries over to the next stress period). In the non-recurrent stress scenario we assume that each new stress period is different such that resistance acquired in previous stress periods is not beneficial in any subsequent stress period.

In both regimes, the genotype frequencies evolve as described by the dynamical system in Equation 1 and according to the schematics in Figure 23. Iterating this procedure leads to oscillations in the SIM allele frequency $p_M = p_{Mr} + p_{MR}$ as depicted in Figure 23C. We measure genotype frequencies at discrete time points directly before the onset of each stress period (bold points in Figure 23C). The long-term equilibria of this time series thus describe the long-term prevalence of the SIM allele, which we denote by $\hat{p}_M$. Since our model assumes an effectively infinite population, the SIM allele cannot be lost within one cycle. Nevertheless, it is possible that the SIM allele frequency asymptotically declines to zero as the cycles are iterated (i.e., that $\hat{p}_M = 0$).

We assume that during stress selection is strong relative to mutation, and that the effect of the SIM allele is large. As a consequence, the stress dynamics has two phases; during the rapid first phase, genotype frequencies are almost exclusively due to selection (s) and those mutation rates that are amplified by the SIM allele ($\sigma\mu_M$ and $\sigma\nu_R$). At the end of the first phase, almost all individuals have acquired resistance to the stress. In the second, slower, phase, resistance levels remain high and the SIM allele slowly degrades due to mutation (μ_M). We further assume that stresses are of short duration, so that we may ignore this

second phase. Mathematically, we replace $s \mapsto \alpha s$ and $\sigma \mapsto \alpha\sigma$, rescale time by $dt \mapsto dt/\alpha$, divide by α, and let $\alpha \to \infty$ (see the Methods section for details).

We aim to calculate the SIM allele frequencies $p_M = p_{Mr} + p_{MR}$ before the onset of each stress period, i.e., at the end of each cycle of stress followed by no stress. Under stress, the relative proportions of the mR and MR genotypes are maintained except for an excess of MR genotypes being generated by amplified mutation from the Mr genotypes. This excess is $\nu_R/(s/\sigma + \mu_M+\nu_R)$. In the absence of stress, resistance levels relax to $p_R(\tau_{NS})$, which approaches mutation balance ($\nu_R/(\mu_R + \nu_R)$) for long periods without stress ($\tau_{NS} \to \infty$). At the same time, the frequency of the SIM allele decays exponentially due to mutations from its initial value $p_M(0)$ to $p_M(0)\exp(-\mu_M \tau_S)$. Heuristically, the SIM allele frequency before the next stress is thus obtained from the SIM allele frequency p_M before the current stress as

$$p'_M = p_M\, e^{-\mu_M \tau_{NS}} \frac{1+\lambda}{1+p_M\lambda}\,,$$

(3)

where $\lambda = (1 - p_R(\tau_{NS}))/p_R(\tau_{NS})\ \nu_R/((s/\sigma + \mu_M + \nu_R)$. This intuitive derivation of the dynamics is made precise in Methods, where we also calculate $p_R(\tau_{NS})$ for the recurrent stress (R) scenario. In the non-recurrent (NR), we have $p_R(\tau_{NS}) = \nu_R/(\mu_R + \nu_R)$, since resistance levels to yet unknown stresses can be assumed to be at mutation balance. Solving Eq(3) for equilibria yields the long-term prevalences of the SIM allele in the (R) and (NR) scenarios as

$$\hat{p}_M^{(R)} = max\left\{0, e^{-\mu_M t} - \Gamma(1 - e^{-\mu_M t}).\left(1 + \frac{\mu_R+\nu_R}{\nu_R}\left(e^{(\mu_R+\nu_R)t} - 1\right)^{-1}\right)\right\},$$

(4)

$$\hat{p}_M^{(NR)} = max\{0, e^{-\mu_M t} - \Gamma(1 - e^{-\mu_M t})\}$$

(5)

(see Appendix), where $\tau = \tau S + \tau NS$ is the length of one cycle of stress and no-stress, and

$$\Gamma = \frac{\frac{s}{\sigma} + \mu_R + \nu_R}{\nu_R}$$

(6)

In particular, we thus see that the stress intensity s and the strength of the SIM allele σ enter the long-term SIM allele prevalences only via their ratio s/σ.

To test our analytical predictions, we explicitly simulate the dynamics (Eq (1)) of a population experiencing stress and no-stress phases according to the schematics in Figure 23 without the simplifications that lead to the above formulae. Figure 24A shows the long-term SIM prevalences as functions of the cycle length τ for a representative choice of the remaining parameters. For both the (R) and (NR) regimes, the simulated values (points) align well with the above formulae (solid lines). In the non-recurrent regime, the SIM allele is maintained in the population as long as stresses occur frequently enough; more precisely,

there is a critical cycle length τ_c such that the SIM allele is not maintained for cycle lengths exceeding τ_c,

$$\hat{p}_M^{(NR)} = 0 \quad if \quad \tau > \tau_C = \frac{1}{\mu_M} \log\left(1 + \frac{1}{\Gamma}\right) \tag{7}$$

Furthermore, in this regime there is a strictly monotone dependence between the steady state SIM allele frequency and the frequency of stress occurrence; in particular, the SIM allele becomes fixed in the population in the limit of infinitely rapid stress occurrence (i.e., $\hat{p}_M^{(NR)} \to 1$ for $\tau \to 1$).

In the recurrent regime, the dependence of the equilibrium SIM levels on the cycle length τ is less simple. If the rate of gaining resistance without the SIM allele is sufficiently low (i.e., $\nu_R \ll 1$, in particular $\nu_R \ll \mu_R$), the SIM allele is not maintained in the population for any choice of τ (Figure 24A). Note that in general there are conditions that do lead to the maintenance of SIM alleles in the recurrent regime. Such cases, however, are not in concordance with our basic ranking of parameters, inequality Eq(2) . Furthermore, we show in Appendix that the non-recurrent regime generally maintains a higher SIM prevalence than the recurrent regime, i.e. $\hat{p}_M^{(NR)} \geq \hat{p}_M^{(R)}$, for any choice of parameters.

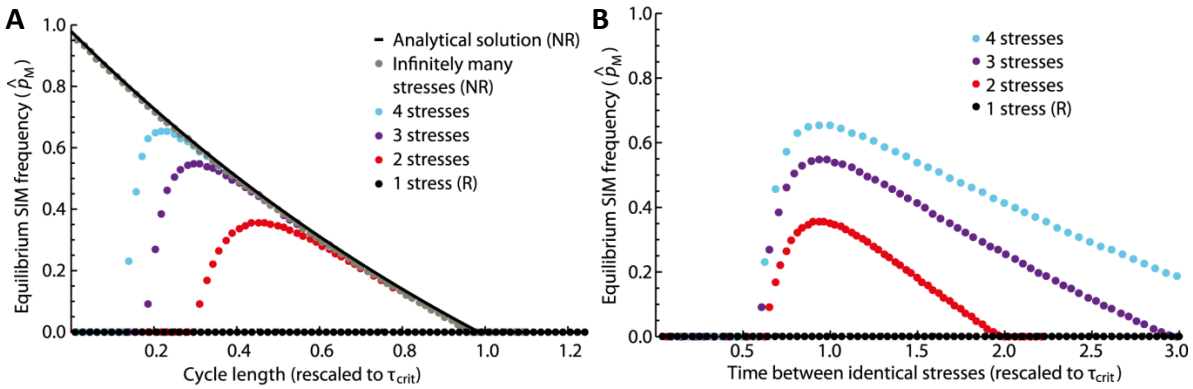


Figure 24 ***SIM prevalences increase with stress diversity.***

A representative parameter set (σ = 100, s = 1, $\mu_M = 10^{-3}$, $\mu_R = 10^{-2}$, and $\nu_R = 10^{-4}$) was simulated for a range of values of $\tau/\tau_c = (\tau_S + \tau_{NS})/\tau_c$. The solid black lines represent the analytical predictions from Eq (4) for the (R) and (NR) regimes. For the numerical simulations, we chose $\tau_S = 10$ and varied τ_{NS} accordingly. The simulation results of the (R) and (NR) regimes (black and grey points) fit their corresponding predictions well. The red, purple, and blue points represent simulation results for two, three, and four different stresses occurring cyclically. Increasing the number of stresses increases the SIM prevalences up to a maximum given by the prediction for the (NR) regime. (A) The critical cycle length τ_c determines the maximal stress re-occurrence time that allows for the maintenance of the SIM allele. (B) The minimal stress re-occurrence time that permits positive long-term SIM prevalences is determined by the time between identical stresses. This time is 2τ (3τ, 4τ) for two (three, four) stresses.

We can extend our basic model to include additional biologically relevant factors, such as cost of resistance or the presence of lethal mutations (see Figure 25). These factors change the long-term SIM prevalences in intuitive ways, yet leave our qualitative statements unchanged. For example, maintaining a resistance mechanism in the absence of stress may

incur a fitness cost. Consequently, resistance levels decrease faster in the no-stress environment if resistance is costly, which increases the benefit of increased mutation rates to acquire resistance under stress. Accordingly, including a cost of resistance to our model increases the long-term SIM prevalences (see Figure 25).

We observe the opposite effect if we consider a mutational load due to lethal mutations. The greater risk for mutator phenotypes to acquire deleterious mutations can be expected to cause indirect selection pressure against the SIM allele. Describing a gradual accumulation of deleterious mutations requires the consideration of multiple fitness classes, which is infeasible in our approach. Instead, we show in Figure 25 that lethal mutations translate into selection against the SIM allele and thus reduce long-term SIM prevalences.

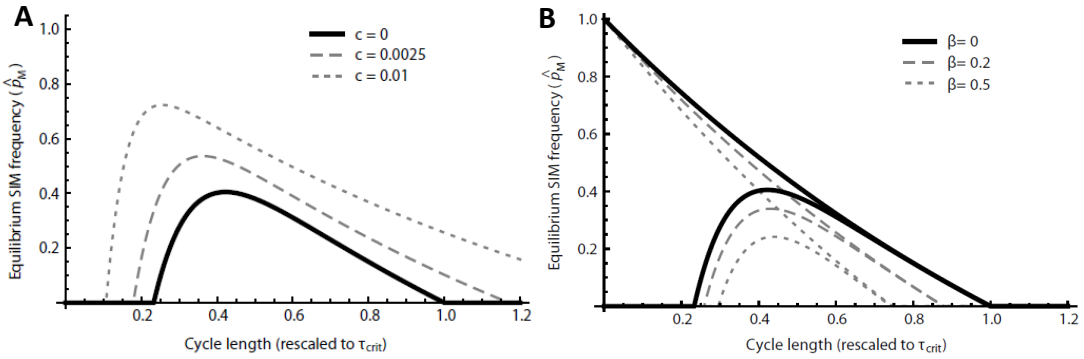


Figure 25 ***Long-term prevalence of the SIM allele when cost of resistance and lethal mutations are included in the model.***

A) Long-term prevalence of the SIM allele with a cost of resistance. The solid black line shows the long-term SIM prevalence in the (R) regime as a function of cycle length τ without a cost of resistance (c = 0). A cost of resistance causes resistance levels in the no-stress phase to decay more rapidly, which increases the benefit provided by the SIM allele. Consequently, we observe elevated equilibrium SIM frequencies (dashed and dotted lines for c = 0:0025 and c = 0:01, respectively). B) Long-term prevalence of the SIM allele with lethal mutations. The solid black lines show the long-term SIM prevalences in the (R) and (NR) regimes as functions of cycle length τ in the absence of lethal mutations (β = σδ = 0). Increased mutation rates during stress also increase the deleterious mutation load for the Mr genotype, hence reducing the equilibrium frequency of the SIM allele (dashed and dotted lines for β = 0.2 and β = 0.5, respectively).

4.2.3 SIM prevalence increases with number of sequentially applied stresses

We explore the prevalence of the SIM allele when subjected to a finite number of stresses. To this end, we simulate the full system as explained earlier for the (R) and (NR) regimes, but for a finite number χ of challenges. This is done by taking into account a separate resistance locus for every challenge. Each of these extra resistance loci is neutral during non-cognate environmental challenges. During this time period, their frequency changes only by mutational degradation or if they are associated to the resistance allele that is under selection. As in the (NR) regime, we assume no cross-resistance and there is complete linkage between all loci. Hence, there are $2^{\chi+1}$ different genotypes to consider. The stresses are applied in a deterministically cycling manner. Each stress period is of constant length τ_S, and is followed by a no-stress period of length τ_{NS}.

The results interpolate between the (R) and (NR) regimes, where every increase in the number of stresses, χ, also increases the SIM allele equilibrium frequency and the parameter regime where it is maintained (Figure 24). In particular, the simulations suggest a simple classification of the possible dynamical regimes, based on the length of the stress periods ($\tau = \tau_S + \tau_{NS}$).

First, for small values of τ, we observe the loss of the SIM allele. The upper bound of this region is inversely proportional to the time it takes for the same stress to recur. Keeping the cycle length τ constant and increasing the number of challenges χ also increases this time and therefore allows for the maintenance of the SIM allele for smaller values of τ. The scaling with the time between two stresses of the same type can be seen clearly in Figure 24B. We may deduce that a too frequent occurrence of the same stress is not beneficial for the SIM allele. This is not surprising; the SIM allele has no fitness advantage on its own and therefore can only rise in frequency if the relevant resistance levels in the population are low. When stresses re-occur frequently, resistance levels are kept high, preventing the SIM allele from hitchhiking.

Second, if the number of different stresses is high enough, a SIM allele can be kept for intermediate frequencies of stress occurrence. The size of this region expands with increasing stress diversity up to the level of the (NR) regime of infinite stress diversity. The maximum allele frequency that can be kept also increases with increasing stress diversity, geometrically approaching the analytically determined value of the (NR) case.

Third, if stresses occur too infrequently, the SIM allele is lost. The critical time between two consecutive stresses, above which the SIM allele is lost for any number of stresses χ, was calculated analytically as τ_c, see Eq (7)

With χ different stresses, each particular stress occurs every $\chi\tau$ time units. Assuming that resistance alleles to different stresses do not interact, we thus may replace $p_R(\tau_{NS})$ by $p_R(\chi\tau_{NS})$ in the heuristic derivation of the recursion Eq (3) to obtain an approximation to the dynamics of SIM allele frequencies with χ different stresses. In our actual model, however, resistance mutations to different stresses do not evolve independently since they are linked to the genetic backgrounds they appear on and cross resistance against multiple stresses is possible. The approximation thus captures the qualitative behaviour of the long-term SIM allele frequencies for multiple stresses, yet overestimates the numerical results for the parameters used in our simulations, (see Appendix)

To relax our assumption of stresses occurring in a strict cycle, we randomize our model by choosing one out of the χ stresses at each iteration of the simulation. Qualitatively, this leaves the picture unchanged, see Figure 26: The SIM prevalence levels and the interval of stress occurrence times τ that maintain the SIM allele both increase with increasing stress diversity, though not as readily as in the deterministic case. However, a shift can be seen in which values of τ make maintenance of SIM possible, leading to a small interval of cycle lengths τ when randomization enables SIM allele maintanence. This happens because the effective time interval between two identical stresses is now a random variable, and there is some probability that the same stress is seen sooner than in the deterministic regime. This

means that a cycle length, that in the deterministic regime is not conducive to the maintenance of the SIM allele, can now sustain it because there is some probability that the same stress is seen at an interval that does support it. One important point is that the minimum time interval between two identical stresses is the time of cycle. This means that the distribution of time intervals is right-skewed, which explains why the "shift" seen on the simulation curves is to the left (the simulations "sample" times to the right).

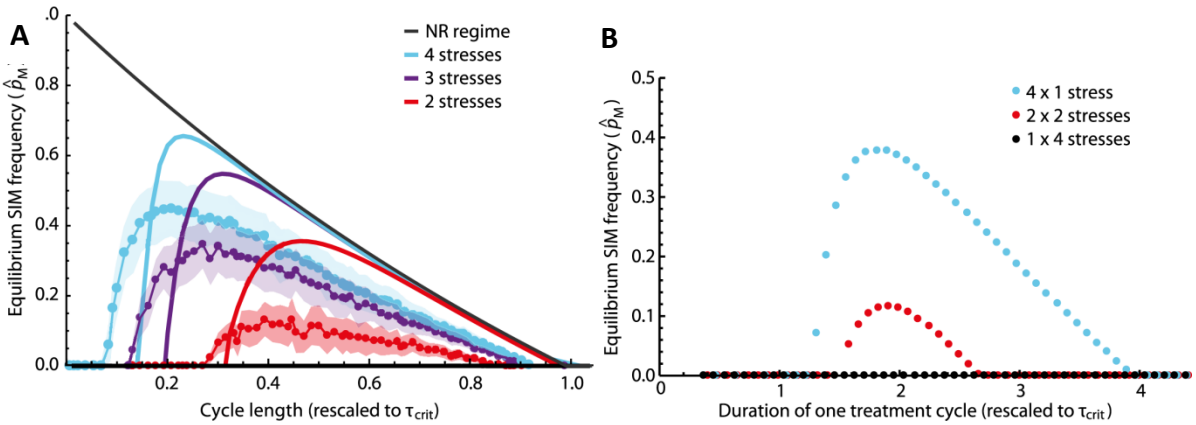


Figure 26 ***The effect of stochasticity and grouping of stresses on SIM prevalences.***

(A) The solid lines represent the same data as Figure 24. In addition, we randomized the simulation by choosing the next stress randomly from the set of available stresses (instead of a deterministic periodic stress cycle). 10,000 iterations of randomly chosen stress and no stress were performed, and the SIM prevalences over the last 1,000 were calculated (red, purple, and blue points for two, three, and four different stresses). The shaded areas indicate the standard deviations in the samples and there is a clear increase in SIM prevalence levels for increasing numbers of different stresses. However, the mean SIM prevalences are significantly lower than corresponding long-term SIM prevalences from deterministic simulations. (B) To mimic treatment regimes, we simulated the simultaneous versus sequential occurrence of four different stresses. We assumed that resistance to each stress confers a selective advantage s. If multiple stresses occur simultaneously, their effects add up such that, for instance, being resistant against two simultaneously occurring stresses provides an advantage of 2s. If all four stresses occur simultaneously (combined treatment) every τ time units, the SIM allele is not maintained for our chosen set of parameters (black points). In contrast, if the four stresses are applied in sequence (sequential treatment) with τ/4 time units between consecutive stresses—such that one cycle through all stresses takes τ time units—the SIM allele is maintained at considerable frequency for a wide range of values of τ (blue points). Grouping the four stresses in two pairs and alternating those at half the previous rate (τ/2; each pair of stresses re-occurs every τ time units) leads to intermediate SIM allele maintenance levels (red points). (Parameters: σ = 100, s = 0.5, μM = 0.001, μR = 0.005, vR = 0.0001.)

4.2.4 Combination treatments prevent the rise of SIM alleles

For practical questions in antibiotic therapy, it is of interest to investigate treatment scenarios in which a set of pharmaceuticals is administered simultaneously or separately over a given period of time (combined versus sequential treatment (Bonhoeffer et al., 1997; D'Agata et al., 2008; Perron et al., 2013)). To this end, we simulate and compare four stresses either occurring simultaneously, being grouped in two pairs, or being applied separately. We assume that the stresses do not allow for cross-resistance mutations (i.e., single mutations that provide resistance to multiple stresses), that their effects on fitness in genotypes with multiple resistance mutations are additive, and that one cycle through all stresses or stress combinations takes τ time units in each case. The results of our simulations are depicted in Figure 26B: while applying all stresses at once does not maintain

the SIM allele for our choice of parameters (Figure 26B, black), the SIM allele prevalence increases if stresses occur more frequently, yet in a less clustered fashion (Figure 26B, red and blue). We have also measured the levels of multi-resistance in these scenarios (Appendix). Interestingly, we find that for short treatment cycles (in which the SIM allele is not expected to be maintained) a mixed strategy in which different sets of multiple stresses are applied sequentially seems to perform best at avoiding multi-resistance, even if at the cost of a higher prevalence of single resistant strains (see Appendix).

4.3 Discussion

Our study investigates the fate of a hypothetical stress-induced mutagenesis (SIM) mechanism under various schemes of environmental fluctuation. We assume that stress-induced mutagenesis is brought about by an active mechanism that increases mutation rates as a response to stress, modeled by a modifier allele for stress-induced mutagenesis that is much easier lost than gained. As a consequence, it decays over time unless maintained by recurrent second-order selection due to changes in the environment. This is what would be expected under the adaptive hypothesis which we are testing. Our results indicate that there are plausible regimes under which the SIM allele could be kept purely by second-order selection under the adaptive hypothesis. What is needed is that the basic hierarchy of parameters outlined in Eq (2) is met. This is reasonable if one considers relatively strong stress episodes and a resistance mechanism such as an antibiotic degrading beta-lactamase enzyme, which is difficult to acquire, but easier to degrade by mutation (Schenk et al., 2012). Furthermore, a regime of environmental fluctuations is needed such that resistance levels are not kept very high between repeated strikes of the same stress, which can be aided in natural populations by fitness costs associated with resistance mutations (Andersson and Levin, 1999) (see also Appendix). Also, stresses in which SIM helps bring about a beneficial mutation need to occur frequently enough to prevent the degradation of SIM. Finally, stress diversity greatly facilitates the maintenance of SIM by requiring resistance mutations that are new or less prevalent in the population. Considering that bacteria in a human body can often experience starvation, acid stress, inflammation, or treatment induced antibiotic stress, these conditions are also plausible (Medzhitov and Janeway Jr., 1997; Weinstein et al., 2013). Although the maintanance of SIM due to second order selection is plausible, our model tends to underestimate SIM frequencies observed in natural populations which are close to 100% (Bjedov et al., 2003). Direct benefits of SIM mechanisms (Torres-Barceló et al., 2015) or a high cost of resistance mutations are common phenomena which are expected to increase the frequency of SIM alleles and could explain the higher frequency found in nature.

Under our assumptions, environmental fluctuations are essential for the SIM allele to be maintained in the population: in the absence of environmental challenges (stresses), the SIM allele is lost due to the neutral accumulation of loss-of-function mutations. Repeatedly occurring stresses, however, give rise to second-order selection on the SIM allele. Under reasonable assumptions on the model parameters, c.f. Eq (2), we show that simple fluctuations caused by a repetitive stress generally fail to maintain the SIM allele. As the stress diversity—i.e., the number of different stresses available—increases, the SIM allele

may be maintained at increasingly high levels (see Figure 24). In the limit of infinite stress diversity, the SIM allele is maintained for any frequency of stress occurrence above a given threshold, which we characterized analytically by τ_c.

Interestingly, when a fixed number of stresses are applied in a random order, the prevalence levels of the SIM allele generally decrease, and the parameter region conducive to maintenance shifts: maintanence can happen at shorter time intervals, and τ_c is apparently reduced (Figure 26A). This is because in this scenario, the time between two stresses of the same kind is now stochastic: there is a probability distribution for the time a particular stress is re-applied. This effectively "smoothes" the deterministic expectation for the steady state frequency. This leads to the "shift" of the simulation curves seen in Figure 25.

It should be noted that we model the dynamics of an infinite population which prevents the examination of the stochastic effects introduced by genetic drift. In our model the SIM allele can never truly fix or be lost from the population. The first point is not very consequential, since it is natural to assume that deleterious mutation will always act to degrade the SIM mechanism and lower its frequency from fixation. However, the second point may be more important since mutations that reintroduce the SIM mechanism after it has been lost may be rare. However, our results can still provide some insights: if the frequency of the SIM allele drops below 1/N, where N is the population size, one can say that it is effectively lost. Furthermore, it is not clear if the rate of back-mutations in nature is effectively zero. If indeed there is some probability of reintroducing the SIM mechanism then our deterministic results provide an expectation for its long-term frequency.

Our results focus on how the maintenance of a SIM allele depends on the frequency and diversity of stresses. We find that in the case of cycling a finite number of stresses, the SIM allele can only be maintained at intermediate stress frequencies. Irrespective of the number of available stresses, a lower bound for the stress frequency can be determined analytically as $1/\tau_c$. For the upper bound, we find that the time between two stresses of the same kind is crucial (Figure 24B). This could inform the choice of therapeutic strategies by identifying treatment schedules that exert extensive selection pressure to keep a SIM allele and possibly strengthen its effect.

To date, various temporal treatment strategies have been investigated to counter the current antibiotic resistance crisis (Kim et al., 2014; Roemhild et al., 2015). To prevent the emergence of resistant strains, one approach is to inhibit known resistance mechanisms directly (Reading and Cole, 1977). Another is to use combinations of existing drugs in treatment regimes that are rationally designed to suppress resistance levels (Baym et al., 2016b; Bergstrom et al., 2004). However, to keep drugs effective in the long term, it is desirable to develop strategies that not only decrease resistance levels, but also restrict evolvability. To this end, there have been efforts to directly inhibit SIM mechanisms (Alam et al., 2016; Cirz et al., 2005). Our study complements this approach by assessing how temporal treatment schemes prevent second-order selection on a SIM mechanism. We find that an increasing diversity of stresses encountered increases long-term SIM frequencies

(see Figure 24 and Figure 26B). This suggests a trade-off between controlling resistance and controlling evolvability when designing multi-drug therapies: in most proposed schemes, one tries to prevent the evolution of resistance by diversifying the stresses (antibiotics) (Kim et al., 2014; Roemhild et al., 2015). However, our findings suggest that this is precisely the scenario in which SIM alleles are more likely to persist and hence promote the evolvability of the population. Experimental work is needed to further characterize this trade-off and assess its relevance in a clinical setting. Currently, it is known that SIM mechanisms are common in bacteria, vary greatly in their potency (Bjedov et al., 2003) and can be lost due to a variety of mutations (Mamun et al., 2012). Selection pressures we describe here could therefore favor those strains that have a significantly higher mutation rate in stress also in the clinic. To confirm this relevance, studies measuring temporal dynamics of SIM alleles in a clinical setting are needed. Also, long-term microbial evolution experiments in a more controlled setting that would follow the prevalence of a synthetic or natural SIM allele over time under different treatment schemes are plausible. Our results may inform such experiments to confirm the suggested trade-off between the evolution and evolvability of resistance.

It has been proposed that the simultaneous application of drugs that exhibit no cross-resistance may be more effective against resistant strains than their sequential application (Bonhoeffer et al., 1997; D'Agata et al., 2008; Perron et al., 2013), but also the opposite (Obolski and Hadany, 2012). In our model, the same applies to reducing positive second-order selection on SIM alleles. Exploring this finding further may provide a resolution of the trade-off between fighting resistance and evolvability, at least for those drug combinations that allow for simultaneous application despite common toxicity or dosage problems.

4.4 Methods

4.4.1 Differential equation model

Casting the schematic dynamics of Figure 23 into differential equations of the form Eq (1) yields

$$\begin{aligned}
\dot{p}_{mr} &= \mu_M p_{Mr} + \mu_R p_{mR} - \nu_R p_{mr} \\
\dot{p}_{Mr} &= \mu_R p_{MR} - (\mu_M + \nu_R) p_{Mr} \\
\dot{p}_{mR} &= \nu_R p_{mr} + \mu_M p_{MR} - \mu_R p_{mR} \\
\dot{p}_{MR} &= \nu_R p_{Mr} - (\mu_M + \mu_R) p_{MR}
\end{aligned} \tag{8}$$

For the stress environment, we have

$$\begin{aligned}
\dot{p}_{mr} &= -s\, p_{mr}(p_{mR} + p_{MR}) + \sigma\mu_M p_{Mr} - \nu_R p_{mr} + \mu_R p_{mR} \\
\dot{p}_{Mr} &= -s\, p_{Mr}(p_{mR} + p_{MR}) + \sigma(\mu_M + \nu_R) p_{Mr} + \mu_R p_{MR} \\
\dot{p}_{mR} &= s\, p_{mR}(1 - p_{mR} - p_{MR}) + \nu_R p_{mr} - \mu_R p_{mR} + \mu_M p_{MR} \\
\dot{p}_{MR} &= s\, p_{MR}(1 - p_{mR} - p_{MR}) + \sigma\nu_R p_{Mr} - (\mu_R + \mu_M) p_{MR}
\end{aligned}$$

(9)

Assuming that stress is strong and of short duration, and that the SIM allele has a large effect, we may replace $s \mapsto \alpha s$, $\sigma \mapsto \alpha\sigma$, and rescale time $dt \mapsto dt/\alpha$. Dividing by α and letting $\alpha \to \infty$, Eq (8) simplifies and permits an approximation for the SIM allele frequency after a short period of stress. We write $p_M^* = \mathcal{F}(p_M)$ for the SIM allele frequency after stress; the mapping depends on whether stress is recurrent or non-recurrent (the (R) and (NR) regimes). Measuring genotype frequencies directly before each stress, we thus obtain a recursion for the SIM allele frequency p_M as

$$p_M' = (\mathcal{G} \circ \mathcal{F})(p_M)$$

which can be written as Eq (3). Solving this recursion for $p_M' = p_M$ leads to the long-term prevalences in Eq (4).

4.4.2 Simulations

Our numerical simulations were implemented using the software *Mathematica*. For a single recurrent stress (the (R) regime), we alternate periods of stress (dynamics Eq (9)) for τ_S time units with periods of no stress (dynamics Eq (8)) for τ_{NS} time units. Genotype frequencies are recorded before each stress period, and the procedure is stopped after 10^4 iterations or once the genotype values reach an equilibrium. To simulate the (NR) regime, we proceed likewise but replace the genotype frequencies $\{p_{mr}, p_{Mr}, p_{mR}, p_{MR}\}$ before every stress by

$$\{(1-\varepsilon)(p_{mr}+p_{mR}), (1-\varepsilon)(p_{Mr}+p_{MR})\}, \varepsilon(p_{mr}+p_{mR}), \varepsilon(p_{Mr}+p_{MR})\}$$

before every new stress, where $\varepsilon = \nu_R/(\mu_R + \nu_R)$. Since the particular kind of stress has never occurred before, the probability of being resistant to it is given by the balance ε between the rates of gaining and losing resistance due to mutation.

With $\chi > 1$ different stresses, there are $2^{\chi+1}$ different genotypes. We consider only single point mutations; the SIM allele is lost at rate μ_M, and each resistance allele is gained (lost) at rate ν_R (μ_R) independently. The fitness of genotypes is $w = 1$ under no stress. In the presence of a stress, the corresponding resistance mutation provides a selective advantage $s > 0$. If multiple stresses occur simultaneously (as is the case in Figure 25B), the fitness advantages due to resistance to the individual stresses are assumed to be additive. There are no cross-resistances, i.e., each resistance allele confers resistance against exactly one stress.

5 Conclusions and discussion

In this thesis, we have set out to find new determinants of antibiotic resistance evolution and quantify the evolutionary dynamics of those already known. We have succeeded in finding new genetic determinants of resistance evolution to tetracycline and chloramphenicol using a highly-controlled automatized evolution protocol (Chapter 2). Further, we quantitatively described the role of general determinants such as the dose-response relationship of a drug and the mutation rate on the extent and reproducibility of resistance evolution and find that steeper dose-response curves lead to faster evolution, and higher mutation rates to more reproducible phenotypic trajectories (Chapter 3). We further focused on a particular previously described mechanism that increases evolvability: stress-induced mutagenesis and describe antibiotic treatment schemes that allow the evolvability allele to persist in a population despite providing no direct benefit to the population (Chapter 4). These results point to new ways in which the evolvability of resistance can be understood, directly inhibited or controlled.

We have established an automatized high-throughput method well-suited to screen for genotypes with modified evolvability (Figure 2). We chose to implement a dynamically adjusting selection pressure in order to observe and compare multi-sweep evolution trajectories. Carefully controlling the population size and growth rate makes the evolutionary dynamics reproducible and directly comparable between genotypes (Figure 3). Due to idiosyncratic technical limitations (space on the deck of the robotic setup), we were only able to follow the evolution of cultures in four 96-well plates at a time. With slight modification, this could be extended to dozens, allowing the simultaneous comparison of thousands rather than hundreds of genotypes, opening the door to exhaustive library screens of genetic determinants of resistance.

Using this method we observed a general pattern of diminishing returns, where more sensitive genotypes were able to undergo larger resistance increases in the course of our experiment compared to more resistant genotypes (Figure 6). It is particularly interesting that certain gene deletion strains were outliers to this pattern and did not evolve significant resistance despite being sensitive to the drug and continually exposed to strong selection pressure for resistance (Figure 7). These genes include efflux pump components and regulators, but also other mechanisms that were not expected to strongly influence evolution such as LPS synthesis, protein transport and one particular chaperone mechanism – DnaK/J (Figure 9). On the opposite side, the *ybaO* gene deletion strain evolved resistance faster than average, ata a similar rate as known mutator strains (Figure 11).

Unlike for previously reported perturbations of evolvability, the most likely explanation for the differences in adaptation rate that we observe here are not mutation rate differences. Rather, epistatic interaction between the particular gene deletion and the resistance mutations can account for the change in adaptation rate. Although it is predicted that strong genetic interactions constrain evolution (Weinreich et al., 2006), the understanding of which mechanisms interact with antibiotic resistance mutations in particular and exploiting this knowledge to constrain resistance evolution has so far been limited (Gifford et al., 2018). The genes we successfully identified in this thesis can serve as motivation to expand this search and create a comprehensive mapping of cellular mechanisms which genetically interact with resistance mechanisms at a genome-wide scale. By expanding also to more antibiotics and more organisms, one can envision the creation of a road-map which

can, for a given drug and organism, provide a list of possible cellular functions which need to be inhibited to constrain resistance evolution. In the long-run, such a comprehensive map could inform how to combine antibiotics and other inhibitory molecules to most successfully cripple resistance evolution.

Using population genetic simulations, we predict that steeper dose-response curves lead to faster resistance evolution (Figure 20). Experimental evidence supports the effect of dose-response curve steepness on the distribution of fitness effects, which could lead to a change in evolutionary dynamics (Chevereau et al., 2015). It would be informative to test the relevance of this relationship by measuring the speed of adaptation to antibiotics with large differences in dose-sensitivity in natural settings. In addition, since shallower dose-response curves are predicted to result in slower adaptation, quantitative models of antibiotic action could be used to devise ways in which dose-response characteristics can be directly perturbed, presenting another strategy to inhibit resistance evolution (Greulich et al., 2015).

Higher mutation rates lead to more reproducible phenotypic adaptation. We see this in population genetic simulations as well as in evolution experiments (Figure 22). Highly mutating bacterial strains are a threat to antibiotic therapies and tackling such infections presents unique challenges (Jolivet-Gougeon et al., 2011; Maciá et al., 2006). Predicting the long-term population dynamics of hyper-mutating strains is important for understanding their success in clinical contexts. Extensions of our model could include deleterious mutations (Orr, 2000) and differences in mutational spectra (Couce et al., 2013) which would provide a better basis to address questions about the reproducibility of mutator strain emergence in chronic infection (Oliver and Mena, 2010).

We show that the exposure to diverse stresses enables the long-term persistence of stress-induced mutagenesis alleles in a bacterial population (Figure 24). This theoretical result together with the indication that diversification of selective pressures leads to better resistance prevention (Kim et al., 2014; Roemhild et al., 2015) indicates the existence of an underappreciated trade-off, where precisely those treatment schemes which select against resistance, can select for highly evolvable genotypes which may represent a great obstacle to treatment in the long-run. The experimental verification of the relevance of such predictions by tracking the frequency of evolvable genotypes in clinical populations is needed to design treatment schemes that would not only inhibit bacterial growth but also prevent the rise of highly evolvable strains.

Altogether, the environmental and genetic determinants of antibiotic resistance evolution are far from understood. This thesis exposes several new ways in which the dynamics of resistance evolution are modulated by genetic determinants and points to how these could be exploited to prolong the efficacy of antibiotic treatments. A comprehensive map of genetic determinants of resistance evolution together with dynamical *in silico* models will provide a powerful platform to predict resistance dynamics in the clinic. With the advent of increasingly affordable high-throughput sequencing and phenotyping technology, extremely precise tracking of natural populations can also test the clinical relevance of these predictions (Lyu et al., 2018; Quainoo et al., 2017). The combination of systematic lab-based and quantitative clinic-based approaches can then lead to highly effective personalized strategies to combat resistance evolution.

References

Agrawal, A.F. (2002). Genetic loads under fitness-dependent mutation rates. Journal of Evolutionary Biology *15*, 1004–1010.

Agrawal, A.F., and Wang, A.D. (2008). Increased transmission of mutations by low-condition females: evidence for condition-dependent DNA repair. PLoS Biology *6*, e30.

Aguilar-Rodríguez, J., Sabater-Muñoz, B., Montagud-Martínez, R., Berlanga, V., Alvarez-Ponce, D., Wagner, A., and Fares, M.A. (2016). The Molecular Chaperone DnaK Is a Source of Mutational Robustness. Genome Biol Evol *8*, 2979–2991.

Alam, M.K., Alhhazmi, A., DeCoteau, J.F., Luo, Y., and Geyer, C.R. (2016). RecA Inhibitors Potentiate Antibiotic Activity and Block Evolution of Antibiotic Resistance. Cell Chemical Biology *23*, 381–391.

Alekshun, M.N., and Levy, S.B. (1997). Regulation of chromosomally mediated multiple antibiotic resistance: the mar regulon. Antimicrob. Agents Chemother. *41*, 2067–2075.

Alekshun, M.N., and Levy, S.B. (1999). The mar regulon: multiple resistance to antibiotics and other toxic chemicals. Trends Microbiol. *7*, 410–413.

Altenberg, L., and Feldman, M.W. (1987). Selection, Generalized Transmission and the Evolution of Modifier Genes. I. The Reduction Principle. Genetics *117*, 559–572.

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

Anes, J., McCusker, M.P., Fanning, S., and Martins, M. (2015). The ins and outs of RND efflux pumps in Escherichia coli. Front Microbiol *6*.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

Barrick, J.E., Kauth, M.R., Strelioff, C.C., and Lenski, R.E. (2010). Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. Mol. Biol. Evol. *27*, 1338–1347.

Barrick, J.E., Colburn, G., Deatherage, D.E., Traverse, C.C., Strand, M.D., Borges, J.J., Knoester, D.B., Reba, A., and Meyer, A.G. (2014). Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. BMC Genomics *15*, 1039.

Barton, N.H. (2000). Genetic hitchhiking. Phil. Trans. R. Soc. Lond. B *355*, 1553–1562.

Baym, M., Kryazhimskiy, S., Lieberman, T.D., Chung, H., Desai, M.M., and Kishony, R. (2015). Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes. PLOS ONE *10*, e0128036.

Baym, M., Lieberman, T.D., Kelsic, E.D., Chait, R., Gross, R., Yelin, I., and Kishony, R. (2016a). Spatiotemporal microbial evolution on antibiotic landscapes. Science *353*, 1147–1151.

Baym, M., Stone, L.K., and Kishony, R. (2016b). Multidrug evolutionary strategies to reverse antibiotic resistance. Science *351*, aad3292.

Bergstrom, C.T., Lo, M., and Lipsitch, M. (2004). Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. PNAS *101*, 13285–13290.

Bjedov, I., Tenaillon, O., Gérard, B., Souza, V., Denamur, E., Radman, M., Taddei, F., and Matic, I. (2003). Stress-induced mutagenesis in bacteria. Science *300*, 1404–1409.

Blair, J.M.A., Webber, M.A., Baylay, A.J., Ogbolu, D.O., and Piddock, L.J.V. (2015). Molecular mechanisms of antibiotic resistance. Nat Rev Micro *13*, 42–51.

Blount, Z.D., Borland, C.Z., and Lenski, R.E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. PNAS *105*, 7899–7906.

Bollenbach, T. (2015). Antimicrobial interactions: mechanisms and implications for drug discovery and resistance evolution. Current Opinion in Microbiology *27*, 1–9.

Bonhoeffer, S., Lipsitch, M., and Levin, B.R. (1997). Evaluating treatment protocols to prevent antibiotic resistance. PNAS *94*, 12106–12111.

Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nat Rev Micro *14*, 320–330.

Breeze, A.S., and Obaseiki-Ebor, E.E. (1983). Mutations to nitrofurantoin and nitrofurazone resistance in Escherichia coli K12. J. Gen. Microbiol. *129*, 99–103.

Bürger, R. (2000). The mathematical theory of selection, recombination, and mutation (Wiley Chichester).

Chao, L., and Cox, E.C. (1983). Competition Between High and Low Mutating Strains of Escherichia coli. Evolution *37*, 125–134.

Chevereau, G., and Bollenbach, T. (2015). Systematic discovery of drug interaction mechanisms. Mol Syst Biol *11*.

Chevereau, G., Dravecká, M., Batur, T., Guvenek, A., Ayhan, D.H., Toprak, E., and Bollenbach, T. (2015). Quantifying the Determinants of Evolutionary Dynamics Leading to Drug Resistance. PLOS Biol *13*, e1002299.

Cirz, R.T., Chin, J.K., Andes, D.R., de Crécy-Lagard, V., Craig, W.A., and Romesberg, F.E. (2005). Inhibition of mutation and combating the evolution of antibiotic resistance. PLoS Biology *3*, e176.

Couce, A., and Tenaillon, O.A. (2015). The rule of declining adaptability in microbial evolution experiments. Front. Genet. *6*.

Couce, A., Guelfo, J.R., and Blázquez, J. (2013). Mutational Spectrum Drives the Rise of Mutator Bacteria. PLoS Genetics *9*.

Crotty, S., Cameron, C.E., and Andino, R. (2001). RNA virus error catastrophe: Direct molecular test by using ribavirin. PNAS *98*, 6895–6900.

Crow, J.F. (1987). Population Genetics History: A Personal View. Annual Review of Genetics *21*, 1–22.

D'Agata, E.M.C., Dupont-Rouzeyrol, M., Magal, P., Olivier, D., and Ruan, S. (2008). The impact of different antibiotic regimens on the emergence of antimicrobial-resistant bacteria. PLoS One *3*, e4036.

Desai, M.M. (2013). Statistical questions in experimental evolution. J. Stat. Mech. *2013*, P01003.

Devon M. Fitzgerald, P.J. Hastings, and Susan M. Rosenberg (2017). Stress-Induced Mutagenesis: Implications in Cancer and Drug Resistance. Annual Review of Cancer Biology *1*, 119–140.

Do Thi, T., López, E., Rodríguez-Rojas, A., Rodríguez-Beltrán, J., Couce, A., Guelfo, J.R., Castañeda-García, A., and Blázquez, J. (2011). Effect of recA inactivation on mutagenesis of \em Escherichia coli exposed to sublethal concentrations of antimicrobials. Journal of Antimicrobial Chemotherapy *66*, 531–538.

Du, D., Wang-Kan, X., Neuberger, A., van Veen, H.W., Pos, K.M., Piddock, L.J.V., and Luisi, B.F. (2018). Multidrug efflux pumps: structure, function and regulation. Nat. Rev. Microbiol. *16*, 523–539.

Dudley, J.W., and Lambert, R.J. (2010). 100 Generations of Selection for Oil and Protein in Corn. In Plant Breeding Reviews, (Wiley-Blackwell), pp. 79–110.

Eigen, M., and Schuster, P. (1977). The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Naturwissenschaften *64*, 541–565.

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

Eyre-Walker, A., and Keightley, P.D. (2007a). The distribution of fitness effects of new mutations. Nat. Rev. Genet. *8*, 610–618.

Eyre-Walker, A., and Keightley, P.D. (2007b). The distribution of fitness effects of new mutations. Nat. Rev. Genet. *8*, 610–618.

Fernandez, L., and Hancock, R.E.W. (2012). Adaptive and Mutational Resistance: Role of Porins and Efflux Pumps in Drug Resistance. Clin. Microbiol. Rev. *25*, 661-+.

Figliuzzi, M., Jacquier, H., Schug, A., Tenaillon, O., and Weigt, M. (2016). Coevolutionary Landscape Inference and the Context-Dependence of Mutations in Beta-Lactamase TEM-1. Mol Biol Evol *33*, 268–280.

Fisher, R.A. (1930). The Genetical Theory Of Natural Selection (At The Clarendon Press).

Fogle, C.A., Nagle, J.L., and Desai, M.M. (2008). Clonal Interference, Multiple Mutations and Adaptation in Large Asexual Populations. Genetics *180*, 2163–2173.

Fridman, O., Goldberg, A., Ronin, I., Shoresh, N., and Balaban, N.Q. (2014). Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. Nature *advance online publication*.

Furusawa, C., Horinouchi, T., and Maeda, T. (2018). Toward prediction and control of antibiotic-resistance evolution. Current Opinion in Biotechnology *54*, 45–49.

Gabryszewski, S.J., Modchang, C., Musset, L., Chookajorn, T., and Fidock, D.A. (2016). Combinatorial Genetic Modeling of pfcrt-Mediated Drug Resistance Evolution in Plasmodium falciparum. Mol Biol Evol *33*, 1554–1570.

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

Gifford, D.R., Furió, V., Papkou, A., Vogwill, T., Oliver, A., and MacLean, R.C. (2018). Identifying and exploiting genes that potentiate the evolution of antibiotic resistance. Nature Ecology & Evolution.

Good, B.H., McDonald, M.J., Barrick, J.E., Lenski, R.E., and Desai, M.M. (2017). The dynamics of molecular evolution over 60,000 generations. Nature *551*, 45–50.

Greulich, P., Waclaw, B., and Allen, R.J. (2012). Mutational Pathway Determines Whether Drug Gradients Accelerate Evolution of Drug-Resistant Cells. Phys. Rev. Lett. *109*, 088101.

Greulich, P., Scott, M., Evans, M.R., and Allen, R.J. (2015). Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics. Mol. Syst. Biol. *11*, 796.

Gunn, J.S. (2001). Bacterial modification of LPS and resistance to antimicrobial peptides. Journal of Endotoxin Research *7*, 57–62.

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

Harmand, N., Gallet, R., Jabbour-Zahab, R., Martin, G., and Lenormand, T. (2016). Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients. Evolution n/a-n/a.

Hartl, D.L., and Clark, A.G. (2007). Principles of Population Genetics (Sunderland, Mass: Sinauer).

Hegreness, M., Shoresh, N., Hartl, D., and Kishony, R. (2006). An equivalence principle for the incorporation of favorable mutations in asexual populations. Science *311*, 1615–1617.

Hegreness, M., Shoresh, N., Damian, D., Hartl, D., and Kishony, R. (2008). Accelerated evolution of resistance in multidrug environments. PNAS *105*, 13977–13981.

Heidenreich, E. (2007). Adaptive mutation in Saccharomyces cerevisiae. Critical Reviews in Biochemistry and Molecular Biology *42*, 285–311.

Hermsen, R., Deris, J.B., and Hwa, T. (2012). On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient. PNAS *109*, 10775–10780.

Huxley, J., Pigliucci, M., and Müller, G.B. (2009). Evolution: The Modern Synthesis (Cambridge, Mass: The MIT Press).

Ishii, K., Matsuda, H., Iwasa, Y., and Sasaki, A. (1989). Evolutionarily stable mutation rate in a periodically changing environment. Genetics *121*, 163–174.

Jacoby, G.A. (2009). AmpC beta-Lactamases. Clin. Microbiol. Rev. *22*, 161-+.

Jacquier, H., Birgy, A., Nagard, H.L., Mechulam, Y., Schmitt, E., Glodt, J., Bercot, B., Petit, E., Poulain, J., Barnaud, G., et al. (2013). Capturing the mutational landscape of the beta-lactamase TEM-1. PNAS *110*, 13067–13072.

Jellen-Ritter, A.S., and Kern, W.V. (2001). Enhanced Expression of the Multidrug Efflux Pumps AcrAB and AcrEF Associated with Insertion Element Transposition in Escherichia coli Mutants Selected with a Fluoroquinolone. Antimicrobial Agents and Chemotherapy *45*, 1467–1472.

Jolivet-Gougeon, A., Kovacs, B., Le Gall-David, S., Le Bars, H., Bousarghin, L., Bonnaure-Mallet, M., Lobel, B., Guillé, F., Soussy, C.-J., and Tenke, P. (2011). Bacterial hypermutation: clinical implications. Journal of Medical Microbiology *60*, 563–573.

Kadibelban, A.S., Bogumil, D., Landan, G., and Dagan, T. (2016). DnaK-dependent accelerated evolutionary rate in prokaryotes. Genome Biol Evol evw102.

Karras, G.I., Yi, S., Sahni, N., Fischer, M., Xie, J., Vidal, M., D'Andrea, A.D., Whitesell, L., and Lindquist, S. (2017). HSP90 Shapes the Consequences of Human Genetic Variation. Cell.

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

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England) *28*, 1647–1649.

Kim, S., Lieberman, T.D., and Kishony, R. (2014). Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance. PNAS *111*, 14494–14499.

Kimura, M. (1962). On the probability of fixation of mutant genes in a population. Genetics *47*, 713.

Kobayashi, K., Tsukagoshi, N., and Aono, R. (2001). Suppression of Hypersensitivity ofEscherichia coli acrB Mutant to Organic Solvents by Integrational Activation of the acrEF Operon with the IS1 or IS2 Element. Journal of Bacteriology *183*, 2646–2653.

Kohanski, M.A., Dwyer, D.J., and Collins, J.J. (2010). How antibiotics kill bacteria: from targets to networks. Nat. Rev. Microbiol. *8*, 423–435.

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001). Horizontal gene transfer in prokaryotes: Quantification and classification. Annu. Rev. Microbiol. *55*, 709–742.

Kryazhimskiy, S., Rice, D.P., Jerison, E.R., and Desai, M.M. (2014). Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science *344*, 1519–1522.

Lang, G.I., Rice, D.P., Hickman, M.J., Sodergren, E., Weinstock, G.M., Botstein, D., and Desai, M.M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature *500*, 571–574.

Lázár, V., Nagy, I., Spohn, R., Csörgő, B., Györkei, Á., Nyerges, Á., Horváth, B., Vörös, A., Busa-Fekete, R., Hrtyan, M., et al. (2014). Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. Nat Commun *5*, 4352.

Levy, S.B., and Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. Nat. Med. *10*, S122-129.

Levy, S.F., Blundell, J.R., Venkataram, S., Petrov, D.A., Fisher, D.S., and Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature *519*, 181–186.

Li, C.C. (1967). Fundamental Theorem of Natural Selection. Nature *214*, 505–506.

Lukačišinová, M., and Bollenbach, T. (2017). Toward a quantitative understanding of antibiotic resistance evolution. Curr. Opin. Biotechnol. *46*, 90–97.

Lukačišinová, M., Novak, S., and Paixão, T. (2017). Stress-induced mutagenesis: Stress diversity facilitates the persistence of mutator genes. PLOS Computational Biology *13*, e1005609.

Luria, S.E., and Delbrück, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics *28*, 491–511.

Lynch, M. (2010). Evolution of the mutation rate. Trends in Genetics *26*, 345–352.

Lyu, F., Pan, M., Patil, S., Wang, J.-H., Matin, A.C., Andrews, J.R., and Tang, S.K.Y. (2018). Phenotyping antibiotic resistance with single-cell resolution for the detection of heteroresistance. Sensors and Actuators, B: Chemical *270*, 396–404.

Maciá, M.D., Borrell, N., Segura, M., Gómez, C., Pérez, J.L., and Oliver, A. (2006). Efficacy and potential for resistance selection of antipseudomonal treatments in a mouse model of lung infection by hypermutable Pseudomonas aeruginosa. Antimicrob. Agents Chemother. *50*, 975–983.

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

MacLean, R.C., Torres-Barceló, C., and Moxon, R. (2013). Evaluating evolutionary models of stress-induced mutagenesis in bacteria. Nat Rev Genet *14*, 221–227.

Mamun, A.A.M.A., Lombardo, M.-J., Shee, C., Lisewski, A.M., Gonzalez, C., Lin, D., Nehring, R.B., Saint-Ruf, C., Gibson, J.L., Frisch, R.L., et al. (2012). Identity and Function of a Large Gene Network Underlying Mutagenic Repair of DNA Breaks. Science *338*, 1344–1348.

Medzhitov, R., and Janeway Jr., C.A. (1997). Innate Immunity: The Virtues of a Nonclonal System of Recognition. Cell *91*, 295–298.

Méhi, O., Bogos, B., Csörgő, B., and Pál, C. (2013). Genomewide Screen for Modulators of Evolvability under Toxic Antibiotic Exposure. Antimicrob Agents Chemother *57*, 3453–3456.

Méhi, O., Bogos, B., Csörgő, B., Pál, F., Nyerges, Á., Papp, B., and Pál, C. (2014). Perturbation of Iron Homeostasis Promotes the Evolution of Antibiotic Resistance. Mol Biol Evol msu223.

Misra, R., Morrison, K.D., Cho, H.J., and Khuu, T. (2015). Importance of Real-Time Assays To Distinguish Multidrug Efflux Pump-Inhibiting and Outer Membrane-Destabilizing Activities in Escherichia coli. Journal of Bacteriology *197*, 2479–2488.

Morgan, T.H. (1915). The mechanism of Mendelian heredity, (New York: Holt).

Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K.M., Lee, K.J., Wong, A., et al. (2011). Phenotypic landscape of a bacterial cell. Cell *144*, 143–156.

Nicoloff, H., and Andersson, D.I. (2013). Lon protease inactivation, or translocation of the lon gene, potentiate bacterial evolution to antibiotic resistance. Mol. Microbiol. *90*, 1233–1248.

Obolski, U., and Hadany, L. (2012). Implications of stress-induced genetic variation for minimizing multidrug resistance in bacteria.

Oliver, A., and Mena, A. (2010). Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. Clinical Microbiology and Infection *16*, 798–808.

Orr, H.A. (2000). The Rate of Adaptation in Asexuals. Genetics *155*, 961–968.

Orr, H.A. (2005). The genetic theory of adaptation: a brief history. Nature Reviews Genetics *6*, 119–127.

Otwinowski, J., McCandlish, D.M., and Plotkin, J.B. (2018). Inferring the shape of global epistasis. PNAS *115*, E7550–E7558.

Palmer, A.C., and Kishony, R. (2013). Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. Nat Rev Genet *14*, 243–248.

Palmer, A.C., Toprak, E., Baym, M., Kim, S., Veres, A., Bershtein, S., and Kishony, R. (2015). Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. Nat Commun *6*, 7385.

Pariente, N., Sierra, S., and Airaksinen, A. (2005). Action of mutagenic agents and antiviral inhibitors on foot-and-mouth disease virus. Virus Res. *107*, 183–193.

Perfeito, L., Fernandes, L., Mota, C., and Gordo, I. (2007). Adaptive Mutations in Bacteria: High Rate and Small Effects. Science *317*, 813–815.

Perron, G.G., Kryazhimskiy, S., Rice, D.P., and Buckling, A. (2013). Multidrug Therapy and Evolution of Antibiotic Resistance: When Order Matters. Appl Environ Microbiol *79*, 6521.

Pitman, A.R., Jackson, R.W., Mansfield, J.W., Kaitell, V., Thwaites, R., and Arnold, D.L. (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Current Biology *15*, 2230–2235.

Poelwijk, F.J., Tănase-Nicola, S., Kiviet, D.J., and Tans, S.J. (2011). Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. J. Theor. Biol. *272*, 141–144.

Poelwijk, F.J., Krishna, V., and Ranganathan, R. (2016). The Context-Dependence of Mutations: A Linkage of Formalisms. PLOS Comput Biol *12*, e1004771.

Ponder, R.G., Fonville, N.C., and Rosenberg, S.M. (2005). A Switch from High-Fidelity to Error-Prone DNA Double-Strand Break Repair Underlies Stress-Induced Mutation. Molecular Cell *19*, 791–804.

Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.-P. (2006). Antibiotic Stress Induces Genetic Transformability in the Human Pathogen Streptococcus pneumoniae. Science *313*, 89–92.

Quainoo, S., Coolen, J.P.M., Hijum, S.A.F.T. van, Huynen, M.A., Melchers, W.J.G., Schaik, W. van, and Wertheim, H.F.L. (2017). Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clinical Microbiology Reviews *30*, 1015–1063.

Queitsch, C., Sangster, T.A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. Nature *417*, 618–624.

Ram, Y., and Hadany, L. (2012). The evolution of stress-induced hypermutation in asexual populations. Evolution *66*, 2315–2328.

Raynes, Y., Gazzara, M.R., and Sniegowski, P.D. (2012). Contrasting dynamics of a mutator allele in asexual populations of differing size. Evolution *66*, 2329–2334.

Reading, C., and Cole, M. (1977). Clavulanic acid: a beta-lactamase-inhiting beta-lactam from Streptomyces clavuligerus. Antimicrob. Agents Chemother. *11*, 852–857.

Regoes, R.R., Wiuff, C., Zappala, R.M., Garner, K.N., Baquero, F., and Levin, B.R. (2004). Pharmacodynamic Functions: a Multiparameter Approach to the Design of Antibiotic Treatment Regimens. Antimicrobial Agents and Chemotherapy *48*, 3670–3676.

Rodrigues, J.V., Bershtein, S., Li, A., Lozovsky, E.R., Hartl, D.L., and Shakhnovich, E.I. (2016). Biophysical principles predict fitness landscapes of drug resistance. PNAS *113*, E1470–E1478.

Roemhild, R., Barbosa, C., Beardmore, R.E., Jansen, G., and Schulenburg, H. (2015). Temporal variation in antibiotic environments slows down resistance evolution in pathogenic Pseudomonas aeruginosa. Evol Appl n/a-n/a.

Rosenberg, S.M. (2001). Evolving responsively: adaptive mutation. Nature Reviews Genetics *2*, 504–515.

Rutherford, S.L. (2003). Between genotype and phenotype: Protein chaperones and evolvability. Nature Reviews Genetics *4*, 263–274.

Rutherford, S.L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. Nature *396*, 336–342.

Sandegren, L., and Andersson, D.I. (2009). Bacterial gene amplification: implications for the evolution of antibiotic resistance. Nat Rev Micro *7*, 578–588.

Schenk, M.F., Szendro, I.G., Krug, J., and Visser, J.A.G.M. de (2012). Quantifying the Adaptive Potential of an Antibiotic Resistance Enzyme. PLOS Genet *8*, e1002783.

Schenk, M.F., Szendro, I.G., Salverda, M.L.M., Krug, J., and Visser, J.A.G.M. de (2013). Patterns of Epistasis between Beneficial Mutations in an Antibiotic Resistance Gene. Mol Biol Evol *30*, 1779–1787.

Schofield, M.J., and Hsieh, P. (2003). DNA Mismatch Repair: Molecular Mechanisms and Biological Function. Annual Review of Microbiology *57*, 579–608.

Schoustra, S., Hwang, S., Krug, J., and Visser, J.A.G.M. de (2016). Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. Proc. R. Soc. B *283*, 20161376.

Sniegowski, P.D., Gerrish, P.J., and Lenski, R.E. (1997). Evolution of high mutation rates in experimental populations of E. coli. Nature *387*, 703–705.

Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., and Godelle, B. (1997). Role of mutator alleles in adaptive evolution. Nature *387*, 700–702.

Tanaka, M.M., Bergstrom, C.T., and Levin, B.R. (2003). The Evolution of Mutator Genes in Bacterial Populations: The Roles of Environmental Change and Timing. Genetics *164*, 843–854.

Tenaillon, O. (2014). The Utility of Fisher's Geometric Model in Evolutionary Genetics. Annu. Rev. Ecol. Evol. Syst. *45*, 179–201.

Tenaillon, O., Toupance, B., Le Nagard, H., Taddei, F., and Godelle, B. (1999). Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. Genetics *152*, 485–493.

Tenaillon, O., Taddei, F., Radman, M., and Matic, I. (2001). Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. Research in Microbiology *152*, 11–16.

Tenaillon, O., Denamur, E., and Matic, I. (2004). Evolutionary significance of stress-induced mutagenesis in bacteria. Trends in Microbiology *12*, 264–270.

Toprak, E., Veres, A., Michel, J.-B., Chait, R., Hartl, D.L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. *44*, 101–105.

Toprak, E., Veres, A., Yildiz, S., Pedraza, J.M., Chait, R., Paulsson, J., and Kishony, R. (2013). Building a morbidostat: An automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition. Nature Protocols *8*, 555–567.

Torres-Barceló, C., Kojadinovic, M., Moxon, R., and MacLean, R.C. (2015). The SOS response increases bacterial fitness, but not evolvability, under a sublethal dose of antibiotic. Proc. R. Soc. B *282*, 20150885.

Trindade, S., Sousa, A., and Gordo, I. (2012). Antibiotic Resistance and Stress in the Light of Fisher's Model. Evolution *66*, 3815–3824.

Van den Bergh, B., Swings, T., Fauvart, M., and Michiels, J. (2018). Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. Microbiol. Mol. Biol. Rev. *82*.

de Visser, J.A.G.M., and Krug, J. (2014). Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet *15*, 480–490.

Vogwill, T., Kojadinovic, M., and MacLean, R.C. (2016). Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of Pseudomonas. Proc. R. Soc. B *283*, 20160151.

Wahl, L.M., and Gerrish, P.J. (2001). The Probability That Beneficial Mutations Are Lost in Populations with Periodic Bottlenecks. Evolution *55*, 2606–2610.

Wahl, L.M., Gerrish, P.J., and Saika-Voivod, I. (2002). Evaluating the impact of population bottlenecks in experimental evolution. Genetics *162*, 961–971.

Walsh, C. (2003). Antibiotics (American Society of Microbiology).

Weinreich, D.M., Delaney, N.F., DePristo, M.A., and Hartl, D.L. (2006). Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. Science *312*, 111–114.

Weinstein, D.H., deRijke, S., Chow, C.C., Foruraghi, L., Zhao, X., Wright, E.C., Whatley, M., Maass-Moreno, R., Chen, C.C., and Wank, S.A. (2013). A new method for determining gastric acid output using a wireless pH-sensing capsule. Aliment Pharmacol Ther *37*, 1198–1209.

Wilke, C.O. (2004). The Speed of Adaptation in Large Asexual Populations. Genetics *167*, 2045–2053.

Wiser, M.J., Ribeck, N., and Lenski, R.E. (2013). Long-Term Dynamics of Adaptation in Asexual Populations. Science 1243357.

Woods, R., Schneider, D., Winkworth, C.L., Riley, M.A., and Lenski, R.E. (2006). Tests of parallel molecular evolution in a long-term experiment with Escherichia coli. PNAS *103*, 9107–9112.

Yonath, A. (2005). Antibiotics targeting ribosomes: Resistance, selectivity, synergism, and cellular regulation. In Annual Review of Biochemistry, (Palo Alto: Annual Reviews), pp. 649–679.

Yu, G., Baeder, D.Y., Regoes, R.R., and Rolff, J. (2018). Predicting drug resistance evolution: insights from antimicrobial peptides and antibiotics. Proc. R. Soc. B *285*, 20172687.

A. Appendix 1

Table 3 ***Identified resistance mutations in Tetracycline***

Gene deleted in ancestral strain	Mutation locus	Mutation type	Amplification	Promoter	Loss of function	Experiment ID	ID	Comment	
lacA	lon	IS	0	1	1	2	1		
lacA	ybaO	insertion	0	0	1	2	1		
lacA	acrR	IS	0	0	1	2	1		
lacA	renD	snp	0	0	0	2	1		
lacA	marR	IS	0	0	1	2	1		
lacA	lon	IS	0	1	1	2	2		
lacA	ybaO	insertion	0	0	0	2	2		
lacA	acrR	IS	0	0	1	2	2		
lacA	ycjX	snp	0	0	0	2	2		
lacA	marR	snp	0	0	1	2	2		
lacA	pumps	amplification	5	0	0	2	3		
lacA	lon	IS	0	1	1	2	3		
lacA	ybaO	snp	0	0	0	2	3		
lacA	marR	snp	0	0	1	2	3		
lacA	pumps	amplification	3	0	0	2	4		
lacA	lon	IS	0	1	1	2	4		
lacA	ybaO	insertion	0	0	0	2	4		
lacA	marR	deletion	0	1	0	2	4		
lacA	pumps	amplification	3	0	0	2	5		
lacA	lon	IS	0	1	1	2	5		
lacA	ybaO	insertion	0	1	0	2	5		
lacA	marR	deletion	0	0	1	2	5		
parent	lon	IS	0	1	1	1	6		
parent	ybaO	insertion	0	0	1	1	6		
parent	marR	deletion	0	0	1	1	6		
parent	pumps	amplification	2	0	0	1	7		
parent	lon	IS	0	1	1	1	7		
parent	ydeO	snp	0	0	0	1	7		
parent	marR	snp	0	0	1	1	7		

parent	pumps	amplification	3	0	0	1	8		
parent	lon	IS	0	1	1	1	8		
parent	marR	unclear	0	1	0	1	8		
parent	lon	IS	0	1	1	1	9		
parent	acrR	IS	0	0	1	1	9		
parent	marR	deletion	0	0	1	1	9		
parent	marR	snp	0	0	1	1	10		
tktA	pumps	amplification	2	0	0	2	11		
tktA	marR	deletion	0	0	1	2	11		
tktA	crp	snp	0	0	0	2	11		
tktA	acrR	IS	0	0	1	2	12		
tktA	marR	insertion	0	0	1	2	12		
marR	pumps	amplification	6	0	0	2	13		
marR	lon	IS	0	1	1	2	13		
marR	fadD	snp	0	1	0	2	13		
marR	argA	snp	0	1	0	2	13		
marR	lon	IS	0	1	1	2	14		
marR	acrB	snp	0	0	0	2	14		
marR	acrR	IS	0	0	1	2	14		
marR	pumps	amplification	3	0	0	2	15		
marR	lon	IS	0	1	1	2	15		
acrR	pumps	amplification	3	0	0	1	16		
acrR	envZ	SNP	0	0	0	1	16		
acrR	pumps	amplification	3	0	0	1	17		
acrR	cyoA	SNP	0	1	0	1	17		
acrR	pumps	amplification	4	0	0	1	18		
acrR	acrS	deletion	0	0	1	1	18		
acrR	gadW	IS	0	1	0	1	18		
tolC	yhdJ	snp	0	1	0	2	19		
dnaK	lon	IS	0	1	1	1	20		
dnaK	marR	deletion	0	0	1	1	20		
dnaK	lon	IS	0	1	1	1	21		
ompR	lon	IS	0	1	1	2	22		
ompR	acrR	IS	0	0	1	2	22		
ompR	marR	snp	0	0	0	2	22		
ompR	acrR	IS	0	0	1	2	23		

ompR	marR	deletion	0	0	1	2	23		
dedD	acrR	IS	0	0	1	2	24		
dedD	marR	deletion	0	0	1	2	24		
dedD	acrR	IS	0	0	1	2	25		
dedD	marR	snp	0	0	0	2	25		
kdpE	pumps	amplification	2	0	0	2	26		
kdpE	lon	IS	0	1	1	2	26		
kdpE	acrR	IS	0	0	1	2	26		
kdpE	marR	IS	0	0	1	2	26		
kdpE	pumps	amplification	2	0	0	2	27		
kdpE	lon	IS	0	1	1	2	27		
kdpE	marR	snp	0	0	1	2	27		
kdpE	ybaO	deletion	0	0	0	2	28		
kdpE	acrR	IS	0	0	1	2	28		
kdpE	marR	snp	0	0	1	2	28		
tufA	acrR	IS	0	0	1	2	29		
tufA	ompF	insertion	0	0	1	2	29		
tufA	marR	insertion	0	0	1	2	29		
tufA	pumps	amplification	4	0	0	2	30		
tufA	lon	IS	0	1	1	2	30		
tufA	marR	IS	0	0	1	2	30		
cspE	lon	IS	0	1	1	2	31		
cspE	marR	snp	0	0	0	2	31		
cspE	pumps	amplification	4	0	0	2	32		
cspE	lon	IS	0	1	1	2	32		
cspE	ybaO	snp	0	0	0	2	32		
cspE	marR	IS	0	0	1	2	32		
astC	lon	IS	0	1	1	2	33		
astC	acrR	IS	0	0	1	2	33		
astC	marR	deletion	0	0	1	2	33		
astC	lon	IS	0	1	1	2	34		
astC	ybaO	insertion	0	1	0	2	34		
astC	marR	IS	0	0	1	2	34		
tatC	bamA	snp	0	0	0	2	35		
tatC	acrR	IS	0	0	1	2	35		
tatC	allC	snp	0	0	0	2	35		

tatC	rpoA	snp	0	0	0	2	35		
tatC	acrR	IS	0	0	1	2	36		
tatC	ompF	IS	0	0	1	2	36		
tatC	marR	snp	0	0	0	2	36		
recA						1	37	No mutations	
recA						1	38	No mutations	
recA						1	39	No mutations	
tolC						2	40	No mutations	
acrA						2	41	No mutations	
acrA	marR	deletion	0	0	1	2	42	(A)6→5 coding (377/435 nt)	
acrA	lon	IS	0	0	1	2	42		
acrA	acrS/acrE	IS	0	0	0	2	42		
acrA	yhhZ	snp	0	0	0	2	42		
acrA						2	43	No mutations	
acrA						3	44	No mutations	Earlier time point
acrA						3	45	No mutations	
acrA						3	46	No mutations	Earlier time point
acrR	lon	IS	0	1	1	2	47		
acrR	pumps	amplification	3	0	0	2	47		
acrR	marR	snp	0	0	0	2	48	A→T	D6V (G<u>A</u>T→G<u>T</u>T)
acrR	lon	IS	0	1	1	2	48		
acrR	yddK	snp	0	0	0	3	49	T→G	pseudogene (390/951 nt)
acrR	marR	deletion	0	1	0	3	49	Δ20 bp	intergenic (-192/-1)
acrR	lon	IS	0	1	1	3	49	unclear	
acrR	pumps	amplification	4	0	0	3	49	nested amplification	
acrR	marR	deletion	0	0	1	3	50	Δ1 bp	coding (87/435 nt)
acrR	lon	IS	0	1	1	3	50		
acrR	pumps	amplification	6	0	0	3	50		
acrR	allE	snp	0	0	0	3	51	G→A	A47V (G<u>C</u>G→G<u>T</u>G)
acrR	ompF	deletion	0	0	1	3	51	$(T)_{6\rightarrow5}$	coding (691/1089 nt)
acrR	eamA	snp	0	0	0	3	51	G→A	S167L (T<u>C</u>G→T<u>T</u>G
acrR	lon	IS	0	1	1	3	51	unclear (70%)	
cspE	marR	snp	0	0	0	2	52	T→A	V84E (G<u>T</u>G→G<u>A</u>G)
cspE	lon	IS	0	1	1	2	52		
cspE	marR	snp	0	0	0	2	53	T→G	L78R (C<u>T</u>G→C<u>G</u>G)

cspE	lon	IS	0	1	1	2	53		
dinB	acrR	deletion	0	0	1	2	54	Δ5 bp	coding (422-426/648 nt)
dinB	dtpD	IS	0	0	1	2	54	IS*4* (+) +12 bp	coding (790-801/1482 nt)
dinB	marR	snp	0	0	1	2	54	A→T	K140* (AAG→TAG)
dnaJ	ybaO	insertion	0	0	0	3	55	$(GCGCCTGAA)_{1\rightarrow2}$	coding (119/459 nt)
dnaJ	marR	deletion	0	0	0	3	55	Δ34 bp	coding (132-165/435 nt)
dnaJ	lon	IS	0	1	1	3	55	unassigned (98%)	
dnaJ	acrR	snp	0	0	0	3	56	A→G	Y49C (TAC→TGC)
dnaJ	marR	deletion	0	0	1	3	56	Δ1 bp	coding (126/435 nt)
dnaJ	lon	IS	0	1	1	3	56	unassigned (95%)	
dnaJ	marR	snp	0	0	0	3	57	T→A	L33Q (CTG→CAG)
dnaJ	lon	IS	0	1	1	3	57	unassigned (95%)	
dnaJ	lon	IS	0	1	1	3	58	IS*186* (+) +6 bp :: Δ1	intergenic (+90/-93)
dnaJ	ybaO	insertion	0	1	0	3	58	$(A)_{7\rightarrow8}$	intergenic (+135/-18)
dnaJ	marR	snp	0	0	0	3	58	C→A	R77S (CGC→AGC)
dnaJ	pumps	amplification	2	0	0	3	58		
dnaK	pumps	amplification	3	0	0	2	59	nested amplification	
dnaK	lon	IS	0	1	1	2	59	unclear (60%)	
dnaK	pumps	amplification	2	0	0	2	60		
dnaK	lon	IS	0	1	1	2	60	unassigned (92%)	
dnaK	pumps	amplification	3	0	0	2	61	nested amplification	
dnaK	marR	insertion	0	1	0	3	62	$(TATTATCCCCTGCAACTAATTACTTGCCAGGGCAA)_{1\rightarrow2}$	intergenic (-206/-6)
dnaK	lon	IS	0	1	1	3	62	unclear	
dnaK	pumps	amplification	3	0	0	3	62		
dnaK	lon	IS	0	1	1	3	63	unclear (combination of IS)	
dnaK	pumps	amplification	3	0	0	3	63		
dnaK	marR	IS	0	0	1	3	64	IS*2* (+) +5 bp	coding (81-85/435 nt)
dnaK	lon	IS	0	1	1	3	64		
dnaK	pumps	amplification	3	0	0	3	64		
dnaK	lon	IS	0	1	1	3	65	unassigned(92%)	
dnaK	pumps	amplification	4	0	0	3	65		
dnaK	envZ	snp	0	0	0	3	66	C→G	A102P (GCG→CCG)
dnaK	lon	IS	0	1	1	3	66		

dnaK	pumps	amplification	3	0	0	3	66	nested amplification	
fis	marR	snp	0	0	0	2	67	T→G	V84G (GTG→GGG)
fis	ompF	insertion	0	1	0	2	67	$(A)_{5\rightarrow6}$	intergenic (-185/+418)
fis	lon	IS	0	1	1	2	67	unassigned(70%)	
fis	ecpE	snp	0	0	0	2	68	C→T	V143M (GTG→ATG)
fis	ompF	insertion	0	0	1	2	68	+C	coding (365/1089 nt)
fis	marR	deletion	0	0	1	2	68	Δ1 bp	coding (71/435 nt)
fis	lon	IS	0	1	1	2	68	unassigned(93%)	
fis	pumps	amplification	2	0	0	2	68	nested?	
fis	lon	snp	0	0	0	3	69	A→T	D445V (GAT→GTT)
fis	ybaO	insertion	0	1	0	3	69	$(A)_{7\rightarrow8}$	intergenic (+135/-18)
fis	marR	snp	0	0	0	3	69	G→T	G69V (GGA→GTA)
fis	ars	snp	0	0	0	3	69	G→A	G58R (GGG→AGG)
fis	pumps	amplification	2	0	0	3	69		
fis	ybaO	insertion	0	0	0	3	70	$(GCGCCTGAA)_{1\rightarrow2}$	coding (119/459 nt)
fis	marR	snp	0	0	1	3	70	C→T	Q117* (CAG→TAG)
fis	pumps	amplification	2	0	0	3	70	nested amplification	
fis	lon	IS	0	1	1	3	70	95%	
fis	lon	snp	0	0	0	3	71	T→C	S368P (TCC→CCC)
fis	marR	snp	0	0	0	3	71	T→C	F8L (TTC→CTC)
fis	several	amplification	0	0	0	3	71	elsewhere	
fis	pumps	amplification	3	0	0	3	72		
fis	several	amplification	2	0	0	3	72	elsewhere	
fis	rssB	IS	0	0	1	3	72	IS*1* (–) +9 bp	coding (640-648/1014 nt)
fis	wbbI	IS	0	0	1	3	72	unassigned(97%)	
fis	marR	snp	0	0	0	3	73	T→A	V84E (GTG→GAG)
fis	sdiA	IS	0	1	0	3	73		
fis	pumps	amplification	3	0	0	3	73		
fis	sdiA	IS	0	1	0	3	74	IS*5* (+) +4 bp	intergenic (-60/+167)
fis	several	amplification	2	0	0	3	74	elsewhere	
galT	acrR	snp	0	0	0	3	75	A→C	Q26P (CAG→CCG)
galT	marR	snp	0	0	0	3	75	G→A	R77H (CGC→CAC)
galT	lon	IS	0	1	1	3	75	unassigned (90%)	
galT	ybaO	insertion	0	0	0	3	76	$(GCGCCTGAA)_{1\rightarrow2}$	coding (119/459 nt)
galT	safA	snp	0	1	0	3	76	A→G	intergenic (-109/+139)

galT	marR	snp	0	0	0	3	76	A→G	N89D (AAC→GAC)
galT	lon	IS	0	0	0	3	76	unassigned(95%)	
galT	acrR	IS	0	0	0	3	76	unassigned(73%)	
lon	acrR	snp	0	0	0	2	77	C→A	T5N (ACC→AAC)
lon	marR	snp	0	0	1	2	77	G→A	M1V (GTG→GTA) †
lon	acrR	IS	0	0	1	2	78	IS*30* (–) +2 bp :: Δ2	coding (463-464/648 nt)
lon	ybaO	insertion	0	1	0	3	79	$(A)_{7\rightarrow 8}$	intergenic (+135/-18)
lon	acrR	snp	0	0	0	3	79	G→T	W178C (TGG→TGT)
lon	ssuE	IS	0	1	0	3	79	IS*1* (–) +8 bp	intergenic (-75/-274)
lon	marR	IS	0	0	1	3	79	unassigned(93%)	
lon	ybaO	snp	0	0	0	3	80	A→C	T18P (ACC→CCC)
lon	marR	deletion	0	1	0	3	80	Δ20 bp	intergenic (-180/-13)
lon	ybaO	insertion	0	0	0	3	81	$(GCGCCTGAA)_{1\rightarrow 2}$	coding (119/459 nt)
lon	acrR	insertion	0	0	1	3	81	+T	coding (71/648 nt)
lon	marR	snp	0	0	1	3	81	C→A	C47* (TGC→TGA)
lon	ybaO	insertion	0	0	0	3	82	$(GCGCCTGAA)_{1\rightarrow 2}$	coding (119/459 nt
lon	marR	IS	0	0	1	3	82	unassigned(95%)	
lon	pumps	amplification	2	0	0	3	82		
lon	ybaO	snp	0	0	0	3	83	A→C	T33P (ACC→CCC)
lon	acrR	IS	0	0	1	3	83	Δ2 :: IS*30* (+) +2 bp	coding (463-464/648 nt)
lon	marR	snp	0	0	0	3	83	G→T	R16L (CGC→CTC)
lpcA	several	amplification	2	0	0	2	84	insH1-arcB	has tolC
lpcA	rpoB	snp	0	0	0	2	85	C→A	A543E (GCA→GAA)
lpcA	acrR	IS	0	0	1	3	86	IS*5* (–) +4 bp	coding (265-268/648 nt)
lpcA	ybjE	snp	0	1	0	3	86	A→C	intergenic (-342/+153)
lpcA	marR	deletion	0	0	1	3	86	Δ5 bp	coding (155-159/435 nt)
lpcA	lpxT	IS	0	0	0	3	86	unassigned(90%)	
lpcA	marR	deletion	0	0	1	3	87	Δ5 bp	coding (140-144/435 nt)
lpcA	dedA	snp	0	0	1	3	87	A→T	L61* (TTG→TAG)
lpcA	acrR	IS	0	0	1	3	87	unassigned(50%)	
lpcA	marR	snp	0	0	0	3	88	G→T	R94L (CGC→CTC)
lpcA	acrS	snp	0	1	0	3	88	C→T	intergenic (-105/-294)

lpcA	lpxT	IS	0	0	1	3	88	unassigned(90%)	
lpcA	acrR	IS	0	0	1	3	89	IS*5* (–) +4 bp	coding (217-220/648 nt)
lpcA	intergenic	snp	0	0	0	3	89	G→T	intergenic (+47/+11)
lpcA	marR	deletion	0	0	1	3	89	Δ15 bp	coding (207-221/435 nt)
lpcA	lpxT	snp	0	0	1	3	89	C→T	R218* (CGA→TGA)
tolC	yhdJ	snp	0	1	0	2	90		
tolC			0	0	0	2	91	No mutations	
tolC			0	0	0	3	92	No mutations	Earlier time point
tolC			0	0	0	3	93	No mutations	Earlier time point
tolC			0	0	0	3	94	No mutations	
tolC			0	0	0	3	95	No mutations	
tolC			0	0	0	3	96	No mutations	Earlier time point
ybaO	marR	deletion	0	0	1	3	97	Δ1 bp	coding (285/435 nt)
ybaO	lon	IS	0	1	1	3	97	unassigned(92%)	
ybaO	pumps	amplification	3	0	0	3	97	nested	
ybaO	marR	snp	0	0	0	3	98	G→A	G116S (GGC→AGC)
ybaO	lon	IS	0	1	1	3	98		
ybaO	pumps	amplification	6	0	0	3	98		
ybaO	acrR	IS	0	0	1	3	99	IS*1* (+) +9 bp	coding (265-273/648 nt)
ybaO	marR	insertion	0	1	0	3	99	$(TACTTGCCAGGGCAACTAAT)_{1\rightarrow2}$	intergenic (-211/-1)
ybaO	lon	IS	0	1	1	3	99	unassigned(90%)	
ybaO	pumps	amplification	2	0	0	3	99		
ybaO	marR	snp	0	0	0	3	100	T→G	L33R (CTG→CGG)
ybaO	pumps	amplification	3	0	0	3	100		
ybaO	marR	snp	0	0	0	3	101	G→A	G69E (GGA→GAA)
ybaO	lon	IS	0	1	1	3	101	unassigned(96%)	
ybaO	pumps	amplification	3	0	0	3	101		
ybaO	marR	insertion	0	0	0	3	102	57 bp x 2	duplication
ybaO	lon	IS	0	1	1	3	102		
ybaO	pumps	amplification	5	0	0	3	102		
lacA	lon	snp	0	0	0	3	103	C→A	A381E (GCG→GAG)
lacA	ybaO	snp	0	0	0	3	103	A→C	T18P (ACC→CCC)
lacA	marR	deletion	0	0	1	3	103	Δ1 bp	coding (121/435 nt)
lacA	pumps	amplification	3	0	0	3	103		

lacA	acrR	IS	0	0	1	3	104	IS*5* (+) +4 bp	coding (217-220/648 nt)
lacA	ompF	IS	0	0	1	3	104	IS*1* (+) +8 bp	coding (12-19/1089 nt)
lacA	marR	insertion	0	1	0	3	104	(TACTTGCCAGGGCA ACTAAT)$_{1\rightarrow2}$	intergenic (-211/-1)
lacA	lon	IS	0	1	1	3	104	unassigned(95%)	
lacA	dsrB	IS	0	0	1	3	104	unassigned(90%)	
lacA	ybaO	insertion	0	1	0	3	105	(A)$_{7\rightarrow8}$	intergenic (+135/-18)
lacA	acrR	IS	0	0	1	3	105	IS*5* (–) +4 bp	coding (217-220/648 nt)
lacA	marR	snp	0	0	0	3	105	G→C	V66L (GTC→CTC)
lacA	lon	IS	0	1	1	3	105	unassigned(90%)	
lacA	acrR	IS	0	0	1	3	106	IS*1* (+) +9 bp	coding (265-273/648 nt)
lacA	marR	deletion	0	0	1	3	106	Δ1 bp	coding (59/435 nt)
lacA	lon	IS	0	1	1	3	106	unassigned(90%)	
lacA	ybaO	snp	0	0	0	3	107	A→C	T33P (ACC→CCC)
lacA	marR	insertion	0	0	0	3	107	(TATTATCCCCTGCAA CTAATTACTTGCCAG GGCAA)$_{1\rightarrow2}$	intergenic (-206/-6)
lacA	lon	IS	0	1	1	3	107	90%	
lacA	ybaO	insertion	0	0	1	3	108	(AGCGCCTGAAACGG CTGGA)$_{1\rightarrow2}$	coding (128/459 nt)
lacA	acrR	snp	0	0	0	3	108	C→A	A9E (GCG→GAG)
lacA	marR	snp	0	0	1	3	108	G→T	E10* (GAA→TAA)
lacA	lon	IS	0	1	1	3	108	92%	
lacA	ybaO	insertion	0	0	0	3	109	(GCGCCTGAA)$_{1\rightarrow2}$	coding (119/459 nt)
lacA	marR	snp	0	0	0	3	109	A→G	K44E (AAG→GAG)
lacA	yjcF	snp	0	0	1	3	109	C→T	W353* (TGG→TAG)
lacA	lon	IS	0	1	1	3	109	90%	
lacA	pumps	amplification	3	0	0	3	109		
recA	acrR	IS	0	0	1	2	110	IS*5* (–) +4 bp	coding (217-220/648 nt)
recA	lon	IS	0	1	1	2	110	unassigned	
recA	soxR	snp	0	0	0	2	111	G→T	D137Y (GAC→TAC)
recA	lon	IS	0	1	1	2	111	unassigned	
lpxM	lon	snp	0	0	0	3	112	T→G	I391S (ATC→AGC)
lpxM	ybaO	insertion	0	0	1	3	112	(A)$_{6\rightarrow7}$	coding (178/459 nt)

lpxM	acrR	snp	0	0	0	3	112	A→C	I62L (ATC→CTC)
lpxM	marR	insertion	0	0	1	3	112	$(ATGACAA)_{1\rightarrow2}$	coding (278/435 nt)
lpxM	ybaO	snp	0	0	1	3	113	G→T	E102* (GAA→TAA)
lpxM	marR	deletion	0	0	1	3	113	Δ2 bp	coding (90-91/435 nt)
lpxM	lon	IS	0	1	1	3	113	unassigned	
lpxM	pumps	amplification	2	0	0	3	113		
marA	envZ	snp	0	0	0	3	114		
marA	soxR	snp	0	0	1	3	114		
marA	lon	IS	0	1	1	3	114		
marA	pumps	amplification	2	0	0	3	114		
marA	pumps	amplification	3	0	0	3	115		
marA	crp	snp	0	0	0	3	116		
marA	pumps	amplification	3	0	0	3	116		
marA	ompR	snp	0	0	0	3	117		
marA	soxR	snp	0	0	0	3	117		
marA	lon	IS	0	1	1	3	117		
marA	ompF	IS	0	0	1	3	118		
marA	soxR	snp	0	0	0	3	118		
marA	livK	snp	0	1	0	3	119		
marA	soxR	snp	0	0	0	3	119		
marR	ybaO	insertion	0	0	1	3	120		
marR	lon	IS	0	1	1	3	120		
marR	pumps	amplification	6	0	0	3	120		
marR	yeaK	snp	0	0	0	3	121		
marR	lon	IS	0	1	1	3	121		
marR	ybaO	snp	0	0	0	3	122		
marR	acrR	snp	0	0	0	3	122		
marR	lon	IS	0	1	1	3	122		
marR	pumps	amplification	2	0	0	3	122		
marR	ybaO	snp	0	0	0	3	123		
marR	lon	IS	0	1	1	3	123		
marR	pumps	amplification	5	0	0	3	123		
marR	ybaO	insertion	0	1	0	3	124		
marR	acrR	IS	0	0	1	3	124		
marR	lon	IS	0	1	1	3	124		
marR	ybaO	snp	0	0	0	3	125		

marR	lon	IS	0	1	1	3	125		
marR	pumps	amplification	6	0	0	3	125		
galT	ybaO	insertion	0	1	0	3	126	$(A)_{7\rightarrow 8}$	intergenic (+135/-18)
galT	marR	deletion	0	0	0	3	126	Δ3 bp	coding (206-208/435 nt)
galT	lon	IS	0	1	1	3	126		
galT	pumps	amplification	2	0	0	3	126		
galT	acrR	IS	0	0	1	3	127		
galT	marR	IS	0	1	0	3	127		
galT	lon	IS	0	1	1	3	127		
galT	ybaO	snp	0	0	0	3	128		
galT	marR	deletion	0	0	1	3	128		
galT	lon	IS	0	1	1	3	128		
galT	pumps	amplification	2	0	0	3	128		
hslU	acrR	IS	0	0	1	2	129		
hslU	marR	snp	0	0	0	2	129		
hslU	lon	IS	0	1	1	2	129		
hslU	acrR	IS	0	0	1	2	130		
hslU	lon	IS	0	1	1	2	130		
hslU	acrR	IS	0	0	1	2	131		
hslU	marR	snp	0	0	0	2	131		
hslU	dgoK	snp	0	0	0	2	131		
hslU	lon	IS	0	1	1	2	131		
hslU	ompF	snp	0	0	1	3	132		
hslU	lon	IS	0	1	1	3	132		
hslU	pumps	amplification	3	0	0	3	132		
hslU	marR	snp	0	0	0	3	133		
hslU	lon	IS	0	1	1	3	133		
hslU	acrR	IS	0	0	1	3	133	unassigned	
hslU	marR	deletion	0	1	0	3	134		
hslU	envZ	snp	0	0	0	3	134		
hslU	lon	IS	0	1	1	3	134		
hslU	marR	deletion	0	0	1	3	135		
hslU	lon	IS	0	0	0	3	135		
htpG	ybaO	deletion	0	0	1	2	136		
htpG	marR	snp	0	0	0	2	136		
htpG	lon	IS	0	1	1	2	136		

htpG	pumps	amplification	2	0	0	2	136		
htpG	ybaO	insertion	0	0	0	2	137		
htpG	acrR	IS	0	0	1	2	137		
htpG	marR	snp	0	0	1	2	137		
htpG	ybaO	insertion	0	0	0	2	138	$(GCGCCTGAA)_{1\rightarrow2}$	coding (119/459 nt)
htpG	marR	snp	0	0	0	2	138	IS*5* (–) +4 bp	coding (217-220/648 nt)
htpG	lon	IS	0	1	1	2	138	G→A	W83* (TGG→TAG)
htpG	ybaO	snp	0	0	0	3	139		
htpG	marR	snp	0	0	0	3	139		
htpG	acrR	IS	0	0	1	3	139		
htpG	ybaO	insertion	0	0	0	3	140		
htpG	marR	insertion	0	1	0	3	140		
htpG	lon	IS	0	1	1	3	140		
htpG	pumps	amplification	5	0	0	3	140	$(GCGCCTGAA)_{1\rightarrow2}$	coding (119/459 nt)
htpG	lon	snp	0	0	0	3	141	$(TATTATCCCCTGCAACTAATTACTTGCCAGGGCAA)_{1\rightarrow2}$	intergenic (-206/-6)
htpG	acrR	IS	0	0	1	3	141	unassigned	
htpG	marR	snp	0	0	0	3	141		
htpG	ybaO	insertion	0	0	0	3	142		
htpG	marR	snp	0	0	0	3	142		
htpG	lon	IS	0	0	0	3	142		
htpG	pumps	amplification	5	0	0	3	142		
lpxM	ybaO	insertion	0	1	0	3	143		
lpxM	marR	snp	0	0	1	3	143		
lpxM	lon	IS	0	1	1	3	143		
lpxM	pumps	amplification	2	0	0	3	143		
lpxM	ybaO	insertion	0	1	0	3	144		
lpxM	marR	snp	0	0	1	3	144		
lpxM	lon	IS	0	1	1	3	144		
lpxM	marR	IS	0	0	1	3	145		
lpxM	acrR	IS	0	0	1	3	145		
lpxM	lon	IS	0	1	1	3	145		
lpxM	marR	snp	0	0	1	3	146		
lpxM	ybaO	insertion	0	0	0	3	147		
lpxM	marR	snp	0	0	0	3	147		

lpxM	lon	IS	0	1	1	3	147		
lpxM	pumps	amplification	2	0	0	3	147		
mdtC	marR	snp	0	0	1	2	148		
mdtC	lon	IS	0	0	0	2	148		
mdtC	pumps	amplification	2	0	0	2	148		
mdtC	ybaO	snp	0	0	0	3	149		
mdtC	marR	insertion	0	0	1	3	149		
mdtC	hflK	snp	0	1	0	3	149		
mdtC	lon	IS	0	1	1	3	149		
mdtC	ybaO	snp	0	0	0	3	150		
mdtC	acrR	IS	0	0	0	3	150		
mdtC	lon	IS	0	0	0	3	150	unassigned (60%)	
mdtC	acrR	IS	0	0	1	3	151		
mdtC	rpoA	snp	0	0	0	3	151		
mdtC	lon	IS	0	1	1	3	151		
mdtC	lon	IS	0	1	1	3	152		
mdtC	pumps	amplification	3	0	0	3	152		
mdtC	lon	IS	0	1	1	3	153		
mdtC	marR	IS	0	1	0	3	153		
mdtC	lit	insertion	0	1	0	3	154		
mdtC	marR	deletion	0	1	0	3	154		
mdtC	lon	IS	0	1	1	3	154		
mdtC	acrR	IS	0	0	1	3	154		
ompF	rrsH	snp	0	1	0	3	155		
ompF	ybaO	insertion	0	1	0	3	155		
ompF	marR	snp	0	0	0	3	155		
ompF	lon	IS	0	1	1	3	155		
ompF	pumps	amplification	4	0	0	3	155		
ompF	ybaO	insertion	0	1	0	3	156		
ompF	marR	snp	0	0	0	3	156		
ompF	lon	IS	0	1	1	3	156		
ompF	pumps	amplification	2	0	0	3	156		
ompF	soxR	deletion	0	1	0	3	157		
ompF	lon	IS	0	1	1	3	157		
ompF	ybaO	snp	0	0	0	3	158		
ompF	treA	snp	0	0	0	3	158		

ompF	lon	IS	0	1	1	3	158		
ompF	pumps	amplification	3	0	0	3	158		
tatC	acrR	IS	0	0	1	3	159		
tatC	marR	insertion	0	0	1	3	159		
tatC	acrR	IS	0	0	1	3	160		
tatC	marR	deletion	0	0	1	3	160		
tatC	yihU	snp	0	0	0	3	160		
tatC	acrR	IS	0	0	1	3	161		
tatC	nmpC	deletion	0	0	1	3	161		
tatC	marR	snp	0	0	0	3	161		
tatC	lon	IS	0	1	1	3	161		
recA	lon	deletion	0	0	1	3	162		
recA	acrR	deletion	0	0	1	3	162		
recA	marR	IS	0	1	0	3	162		
recA	lon	snp	0	0	1	3	163		
recA	acrR	snp	0	0	0	3	163		
recA	marR	snp	0	0	0	3	164		
recA						3	165	No mutations	
acrR	pumps	amplification	5	0	0	3	166		
lpxM	lon	IS	0	1	1	3	167		
lpxM	marR	IS	0	1	0	3	167		

Table 4 ***Identified resistance mutations in Chloramphenicol***

Gene deleted in ancestral strain	Mutation locus	Mutation type	Amplification	Promoter	Loss of function	ID	Comment	
acrA						1	No mutations	Earlier time point
acrB						2	No mutations	Earlier time point
acrB						3	No mutations	Earlier time point
acrB	several	amplification	2	0	0	4	mdfA	
acrB						5	No mutations	Earlier time point
acrR	several	amplification	3	0	0	6	acrAB	
acrR	several	amplification	4	0	0	7	acrAB	

acrR	gyrB	snp	0	0	0	7	C410G (TGC→GGC)	
acrR	rpoC	snp	0	0	0	7	R1140S (CGC→AGC)	
astC	marR	deletion	0	0	1	8	Δ1 bp	coding (243/435 nt)
astC	cyaA	snp	0	0	0	8		
astC	lpxK	snp	0	0	0	9		
astC	marR	snp	0	0	1	9	Q121* (CAA→TAA)	
cmr	acrR	IS	0	0	1	10	IS1 (+) +8 bp	coding (526-533/648 nt)
cmr	marR	insertion	0	1	0	10	(TACTTGCCAGGGCAACTAAT)1→2	intergenic (-211/-1)
cmr	acrB	snp	0	0	0	11	G570S (GGC→AGC)	
cmr	acrR	IS	0	0	1	11	IS5 (+) +4 bp	coding (217-220/648 nt)
cmr	nagA	snp	0	0	0	11	E161A (GAA→GCA)	
cmr	marR	deletion	0	0	1	11	Δ1 bp	coding (44/435 nt)
cmr	marR	snp	0	0	0	12	R73H (CGT→CAT)	Earlier time point
cspE	acrB	snp	0	0	0	13	S630T (TCC→ACC)	
cspE	acrR	snp	0	0	0	13	L174P (CTG→CCG)	
cspE	marR	snp	0	0	0	13	G69E (GGA→GAA)	
cspE	acrR	snp	0	0	0	14	R13S (CGC→AGC)	
cspE	marR	snp	0	0	0	14	W83G (TGG→GGG)	
cspE	several	amplification	3	0	0	14	mdfA	
cspE	acrB	snp	0	0	0	15	Q369L (CAG->CTG)	
cspE	mdfA	snp	0	1	0	15	intergenic (-273/-12)	
cspE	marR	snp	0	0	1	15	Q117* (CAG→TAG)	
dinB	marR	snp	0	0	0	16	T39A (ACC→GCC)	
dnaK	marR	deletion	0	0	1	17	Δ106 bp	coding (255-360/435 nt)
dnaK	several	amplification	2	0	0	17	insN-mhpC?	
dnaK	several	amplification	12	0	0	18	mdfA	
dnaK	rrsA	snp	0	0	0	18	T→G	noncoding (1537/1542 nt)
dnaK	yjbI	snp	0	0	0	18	T→C	pseudogene (61/75 nt)
galT	acrB	snp	0	0	0	19	F136V (TTC→GTC)	
galT	acrR	IS	0	0	0	19	unclear	
galT	marR	snp	0	0	0	19	V45E (GTG→GAG)	
galT	ycbK	snp	0	0	0	20	R74H (CGC→CAC)	
galT	marR	snp	0	0	1	20	W83* (TGG→TAG)	
galT	acrR	IS	0	0	1	21	IS5 (–) +4 bp	coding (217-220/648 nt)
galT	marR	snp	0	0	1	21	S65* (TCG→TAG)	

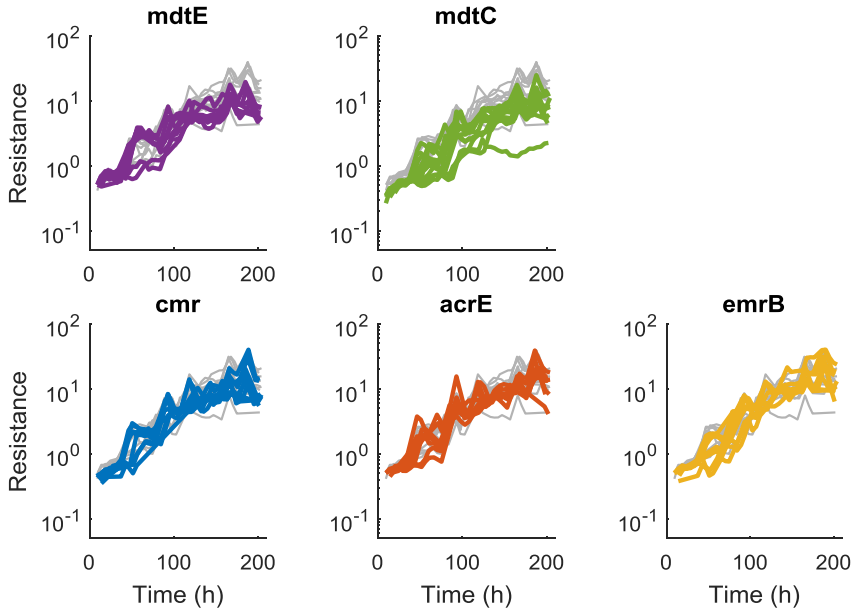
galT	several	amplification	6	0	0	21	mdfA-nfsA	
lacA	acrR	IS	0	0	1	22	IS*5* (–) +4 bp	coding (217-220/648 nt)
lacA	chiP	snp	0	0	0	22	A92A (GCT→GCG)	
lacA	marR	snp	0	0	1	22	M1V (GTG→GTT) †	
lacA	several	amplification	17	0	0	22	mdfA	
lacA	several	amplification	2	0	0	22	insN-mhpC?	
lacA	acrR	IS	0	0	1	23	IS*1* (–) +9 bp	coding (337-345/648 nt)
lacA	marR	snp	0	0	0	23	T101I (ACC→ATC)	
lacA	lon	IS	0	1	0	23	unclear	
lacA	marR	snp	0	0	0	24	V84E (GTG→GAG)	
lacA	acrR	IS	0	0	0	24	unclear	
lacA	several	amplification	2	0	0	24	insN-mhpC?	
lacA	acrR	IS	0	0	1	25	IS*5* (–) +4 bp	coding (217-220/648 nt)
lacA	marR	snp	0	0	1	25	Q113* (CAA→TAA)	
lacA	ydjL	deletion	0	0	1	25	Δ1 bp	coding (108/1077 nt)
lacA	yhiM	snp	0	0	0	25	S141N (AGC→AAC)	
lacA	mdfA	IS	0	1	0	25	unclear	
lacA	several	amplification	2	0	0	25	insN-mhpC	
lacA	acrR	IS	0	0	1	26	IS*1* (–) +8 bp	coding (362-369/648 nt
lacA	marR	snp	0	0	0	26	D92E (GAC→GAG)	
lacA	lon	IS	0	1	0	26	unclear	
lacA	marR	snp	0	0	0	27	T101I (ACC→ATC)	
lacA	lon	IS	0	1	0	27	unclear	
lacA	several	amplification	2	0	0	27	acrAB	
lacA	acrR	snp	0	0	0	28	L86P (CTC→CCC)	
lacA	marR	snp	0	0	0	28	G69E (GGA→GAA)	
lacA	mdfA	IS	0	1	0	28	unclear	
lacA	yeaR	IS	0	0	0	28	unclear	
lacA	several	amplification	3	0	0	28	insN-mhpC?	
lon	rrsH	snp	0	0	0	29	A→G	noncoding (864/1542 nt)
lon	acrR	snp	0	0	0	29	L174Q (CTG→CAG)	
lon	mdfA	IS	0	1	0	29	unclear	
lon	marR	IS	0	0	0	29	unclear	
lon	acrR	deletion	0	0	1	30	$(G)_{5\rightarrow4}$	coding (85/648 nt)
lon	marR	snp	0	0	0	30	D6V (GAT→GTT)	
lon	mdfA	IS	0	1	0	30	unclear	

lon	several	amplification	5	0	0	30	mdfA	
lon	rrsH	snp	0	0	0	31	A→G	noncoding (864/1542 nt)
lon	mrdA	snp	0	0	0	31	Q150P (CAG→CCG)	
lon	soxR	snp	0	0	0	31	G143D (GGT→GAT)	
lon	mdfA	IS	0	1	0	31	unclear	
lon	several	amplification	2	0	0	31	mdfA	
lpcA	marR	insertion	0	0	1	32	(ATTACCGCGGCACAGTTTAA)$_{1\rightarrow2}$	coding (131/435 nt)
lpcA	lpxT	deletion	0	0	0	32	(T)$_{5\rightarrow4}$	coding (540/714 nt)
lpcA	acrR	IS	0	0	1	32	unclear	
lpcA						33	No mutations	Earlier time point
lpcA	marR	snp	0	0	0	34		
lpxM	acrR	IS	0	0	1	35	IS*5* (+) +4 bp	coding (217-220/648 nt)
lpxM	marR	snp	0	0	1	35	Q117* (CAG→TAG)	
lpxM	acrR	IS	0	0	1	36	IS*5* (–) +4 bp	coding (217-220/648 nt)
lpxM	marR	snp	0	0	1	36	Q23P (CAG→CCG)	
lpxM	mdfA	IS	0	1	0	36	unclear	
lpxM						37	No mutations	Earlier time point
marA	several	amplification	2	0	0	37	acrAB	
marA	acrR	IS	0	0	1	37	IS5 (–) +4 bp coding (265-268/648 nt)	
marA	fis	deletion	0	0	1	37	Δ18 bp	coding (262-279/297 nt)
marA	acrR	deletion	0	0	0	38	Δ1 bp	coding (545/648 nt)
marA	several	amplification	2	0	0	38	acrAB	
ppiD	acrB	snp	0	0	0	39	T329S (ACC→AGC)	
ppiD	acrR	snp	0	0	0	39	S33L (TCG→TTG)	
ppiD	marR	IS	0	0	0	39	IS*2* (+) +5 bp	coding (48-52/435 nt)
ppiD	lon	IS	0	1	0	39	unclear	
ppiD	several	amplification	4	0	0	39	hipA	
ppiD	acrR	deletion	0	0	1	40	Δ1 bp	coding (469/648 nt)
ppiD	marR	deletion	0	0	1	40	Δ150 bp	coding (85-234/435 nt)
ppiD	mdfA	IS	0	1	0	40	unclear	
proQ	acrB	snp	0	0	0	41	G288A (GGT→GCT)	
proQ	mdfA	snp	0	1	0	41	A→G	intergenic (-247/-38)
proQ	marR	IS	0	0	0	41	IS*2* (+) +5 bp	coding (81-85/435 nt)
proQ	lptG	snp	0	0	0	41	L146P (CTC→CCC)	
proQ	acrR	IS	0	0	1	42	IS*1* (+) +9 bp	coding (52-60/648 nt)

proQ	opgG	deletion	0	0	0	42	Δ1 bp	coding (1152/1536 nt)
proQ	marR	insertion	0	0	0	42	(ACTGGTAAAACTTACCACC GGCGGCGCGGCAATATG)$_{1 \rightarrow 2}$	coding (323/435 nt)
proQ	frlA	IS	0	0	0	42	unclear	
recA	ahpF	snp	0	0	0	43	T343A (ACC→GCC)	
recA	marR	snp	0	0	0	43	D67G (GAC→GGC)	
recA	lon	IS	0	1	0	43	unclear	
recA	marR	snp	0	0	0	44	D76V (GAT→GTT)	
recA	mdfA	IS	0	1	0	44	unclear	
recA	lon	IS	0	0	0	44	unclear	
rfaG	acrB	snp	0	0	0	45	W634C (TGG→TGT)	
rfaG	acrR	IS	0	0	1	45	IS*1* (+) +8 bp	coding (452-459/648 nt)
rfaG	marR	snp	0	0	0	45	A→C	T72P (ACC→CCC)
rfaG	mtfA	snp	0	0	0	45	S260S (TCG→TCC)	
rfaG	acrB	snp	0	0	0	46	V139F (GTT→TTT)	
rfaG	acrR	snp	0	0	0	46	I62L (ATC→CTC)	
rfaG	marR	snp	0	0	0	46	L46H (CTC→CAC)	
rfaG	acrR	snp	0	0	0	47	A9P (GCG→CCG)	
rfaG	marR	deletion	0	0	0	47	Δ18 bp	coding (113-130/435 nt)
rfaG	intergenic	insertion	0	0	0	47	$(T)_{5 \rightarrow 6}$	intergenic (+34/+22)
sodB	marR	snp	0	0	0	48	E131V (GAA→GTA)	
sodB	acrR	IS	0	0	0	48	unclear	
sodB	mdfA	IS	0	1	0	48	unclear	
sodB	acrB	snp	0	0	0	49	A371T (GCC→ACC)	
sodB	acrR	IS	0	0	1	49	IS*5* (+) +3 bp :: +C	coding (266-268/648 nt)
sodB	marR	snp	0	0	1	49	E122* (GAA→TAA)	
sodB	mdfA	IS	0	1	0	49	unclear	
sodB	several	amplification	2	0	0	49	mdf?	
sodB	acrR	snp	0	0	0	50	L34Q (CTG→CAG)	
sodB	mdfA	snp	0	1	0	50	G→A	intergenic (-259/-26)
sodB	marR	insertion	0	0	1	50	(AGAAAGATCGCCTGCTTA AC)$_{1 \rightarrow 2}$	coding (90/435 nt)
sodB	mdfA	IS	0	1	0	50	unclear	
rplA	acrR	IS	0	0	1	51	IS*5* (–) +4 bp	coding (217-220/648 nt)
rplA	marR	deletion	0	0	1	51	Δ10 bp	coding (105-114/435 nt)
rplA	several	amplification	2	0	0	51	unclear	
rplA	acrR	snp	0	1	0	52	A→G	intergenic (-95/-47)

rplA	marR	deletion	0	1	0	52	Δ20 bp	intergenic (-192/-1)
rplA	menC	snp	0	0	1	52	C→T	W66* (TGG→TAG)
rplA	rpoA	snp	0	0	0	52	G→C	L290V (CTT→GTT)
tatC	mdfA	IS	0	1	0	53	unclear	
tatC	acrR	IS	0	0	1	54	IS*5* (+) +4 bp	coding (217-220/648 nt)
tatC	marR	snp	0	0	1	54	Q117* (CAG→TAG)	
tolC	mdfA	IS	0	1	0	55	unclear	
tolC	mdfA	deletion	0	1	0	55	Δ12 bp	intergenic (-245/-29)
tolC						56	No mutations	Earlier time point
tolC	mdfA	IS	0	1	0	57	unclear	

Figure 27 ***Deletions of efflux pumps other than acrA/B have limited effect on resistance evolution.***



Resistance (IC_{50} in µg/ml) over time as estimated by the tetracycline concentration in the well during in the automatized evolution protocol is shown for five efflux pump deletion strains from experiment M3. Gray lines show resistance over time for control (lacA) replicate experiments.