The maintenance and evolution of antibiotic resistance genes in the absence of antibiotic selection

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Microbiology & Genetics

at Massey University, Albany Campus, New Zealand.

Huei-Yi Lai

2023

Abstract

The rise of new antibiotic resistance in pathogenic bacteria combined with the stagnation of drug development has led to a crisis in treating bacterial infection. An understanding of factors that can influence the development and prevalence of antibiotic resistance in bacteria can help combat resistance. Bacteria can acquire antibiotic resistance via two types of genetic changes—antibiotic resistance mutations (ARMs) and acquisition of antibiotic resistance genes (ARGs). In the presence of antibiotics, these genetic changes are beneficial to bacteria but in the absence of antibiotic any fitness cost of the resistance genotype imposed on the bacteria is uncovered. The fitness cost of a resistance genotype creates a fitness difference between resistant and susceptible bacteria leading to purifying selection against resistant bacteria. As a result, the fitness cost of a resistance genotype can play an important role in the maintenance of resistant bacteria in a population, especially when the antibiotic selection is absent or weak. Previous studies have focused on the degree and mechanistic basis of the fitness cost of ARMs and of ARGs embedded in mobile genetic elements (MGEs), such as plasmids. Little is known about the fitness cost of individual ARGs, let alone its mechanistic basis. Moreover, ARGs are often associated with MGEs, which subject ARGs to frequent gene flow between bacteria. Because of this movement between host strains, any variation in the fitness cost of an ARG between different strains can influence its prevalence at the population level. Despite the potential importance of this effect in determining the success of ARGs, direct measurements of host-specific fitness costs have been made for only a few distinct ARGs. Finally, compensatory evolution can alleviate the fitness cost of resistance genotypes so that both immediate and long-term costs of ARGs must be considered. In this thesis, I aim to investigate the fitness cost of individual ARGs and test its evolutionary significance. In **Chapter 2**, I quantify the fitness

costs of six ARGs prevalent in published *Escherichia coli* genomes and determine the variation in costs across twelve *Escherichia* strains. While on average the fitness cost of the six ARGs is small, consistent with their high prevalence, the costs of most ARGs vary between hosts. I show that this variation can be consequential, resulting in host-dependent evolutionary dynamics of an ARG plasmid. In **Chapter 3**, I use whole genome sequencing and reverse genetics to dissect the genetic basis of the compensatory evolution observed in **Chapter 2**. I identify a mutation on a phage gene that can alleviate the fitness cost of a β-lactamase, and moreover, I demonstrate that the host-dependent cost of the β-lactamase is due to the negative interaction between the β-lactamase and the phage gene. **Chapter 4** extends work on measuring ARG costs and determining their effect on ARG maintenance to investigate the influence of costs on the molecular evolution of an ARG. In **Chapter 4**, I examine if the host-dependent fitness cost of the β-lactamase can influence the accumulation of genetic variation in that gene. Together, these chapters characterize the influence of the fitness cost of ARGs on their maintenance and evolution and demonstrate that, even without antibiotic selection, other selective forces continue to influence the persistence of antibiotic resistance genes in bacterial populations.

Acknowledgement

This thesis cannot be completed without the supervision of Professor Tim Cooper and Professor Xue-Xian Zhang. I am grateful for the opportunity to conduct my study under their supervision. A special thanks to Tim for the freedom he had given me to enjoy doing the research and for his constructive criticism to help me see the bigger picture of my study.

I would like to thank the members in the Cooper Lab–Dr. Bao-Jun Wu, Courtney Armstrong and Razvan Hancu–and people in the molecular biology group for helping me to run experiments smoothly in the lab. I also want to thank my office mates, Arielle Sulit and Emma Holvast, for being my mental support. Your presence was one of the highlights at the time of my study. I vividly remember the night before the lockdown we all rushed back to the office to grab computers and said goodbye to each other without knowing when we would meet again.

Finally, I thank my brothers for taking caring of home while I was on my adventure and my parents for respecting and supporting my life decisions.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Preface

Bacterial infections caused by antibiotic resistant pathogens are becoming a common threat to human health decades after the first clinical introduction of antibiotics in the 1930s, which had seemed to signal a victory in the war against bacterial disease [7]. The continuing rise of bacterial resistance to both commonly used and newly introduced antibiotics reflects partly a stagnation of drug development [8]. In recent years, the narratives on human-centered health has extended to integrate the health of humans, animals, the environment, and their interactions— a concept of "One Health" [9]. The One Health concept emphasizes an integrated approach to understanding and combating the problem of antibiotic resistance. For

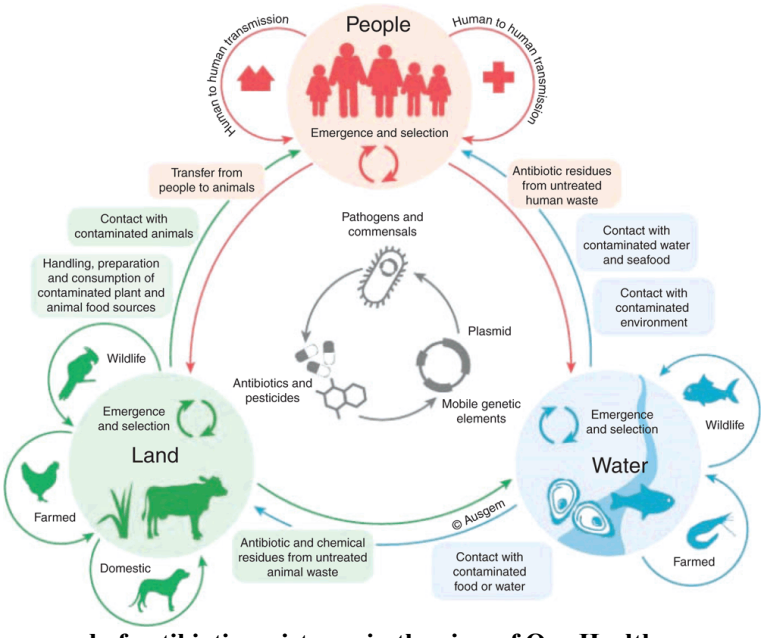

Figure 1.1. The spread of antibiotic resistance in the view of One Health.

Antibiotic resistance can rise and be selected in environmental, commensal, and zoonotic bacteria. To combat antibiotic resistant pathogen under the view of One Health, it is crucial to consider the interactions within and between pathogenic and non-pathogenic bacteria. The image is adapted from Djordjevic et al., 2019 [6].

example, when facing an emergent zoonotic disease in humans, a full understanding requires not only a development of treatment strategies but also an understanding of the interaction

between human and animals, and the influence of environmental changes that lead to the disease emergence [10]. This concept extends to efforts to combat the rise of antibiotic resistant pathogens (Figure 1.1) [6, 11]. In the absence of new antibiotics, there is increasing emphasis on the need to improve our knowledge of the evolutionary and ecological forces driving the emergence and prevalence of antibiotic resistant pathogens. In this introductory chapter, I summarize the genetic basis of antibiotic resistance and discuss key empirical and theoretical findings describing genetic and environmental factors that can influence the selection of the genetic determinants of antibiotic resistance.

1.2 Genetic basis of antibiotic resistance

Bacteria can acquire antibiotic resistance via two main genetic mechanisms: antibiotic resistance mutations (ARMs) and antibiotic resistance genes (ARGs) (Table 1.1). The two mechanisms differ in their genetic origins and evolutionary implications. Below, I discuss each in turn.

1.2.1 Antibiotic resistance mutations

Mechanisms of action of ARMs. ARMs are spontaneous mutations that alter the nucleotide sequence of chromosomal genes to confer antibiotic resistance (AR) [12]. The gene targets of ARMs fall into two categories: antibiotic specific and non-specific. Specific targets are gene products that an antibiotic directly interacts with. AR mutations in these targets often change the target molecule's binding affinity to the relevant antibiotic, thereby reducing the deleterious effect of that antibiotic on cell growth. For example, rifampin inhibits bacterial growth by binding to the β-subunit of DNA-dependent RNA polymerase (encoded by *rpoB*) and

inhibiting transcriptional elongation [13]. Rifampin-resistant mutations in RpoB occur more frequently in a 27 amino acid-long region, called the rifampin resistance determining region (RRDR) [14]. This region contains around 96% of rifampin-resistant mutations in RpoB in *Mycobacterium tuberculosis* [14]. The mutations may interfere with the protein binding to the drug [15]. Other drug targets, such as quinolone-targeted DNA gyrase and topoisomerase IV, have a similar mutational hotspot near the drug-binding sites [16, 17]. Co-crystallization of topoisomerases with quinolone showed that resistance mutations may interfere with their binding [17]. The mutational hotspots in genes encoding drug targets can make the evolution and emergence of resistance predictable [18, 19], which may make it possible to design personalized antibiotic treatments based on a patient's medical history [20].

Antibiotic resistance caused by modification of a drug target usually confers resistance to a limited spectrum of drugs that interact similarly with the same cellular target. By contrast, ARMs that act through non-specific mechanisms, for example, reducing the ability of antibiotics to enter cells by altering the activity or expression of genes that regulate membrane permeability, confer a broader resistance spectrum. For instance, in some Gram-negative bacteria, mutations that reduce the expression level or alter the channel size of substrate size-selective porins can provide resistance to different drugs. For example, mutations in OmpA, OmpC, and OmpF can decrease the import of structurally different drugs, including β-lactams, tetracyclines, quinolones, streptomycin, and fusidic acid [21-24]. Such mutations confer resistance by causing bacteria to be exposed to a lower intracellular concentration of antibiotics. Similarly, mutations that cause an increase in the export of antibiotic molecules can reduce the intracellular accumulation of an antibiotic [24]. Mutations in substrate-promiscuous multidrug resistance (MDR) efflux pumps can increase the expression of the pumps and elevate their

efflux activity [25-27]. Moreover, the gene regulation of MDR efflux pumps is embedded in a complex transcriptionally network. Mutations in several transcriptional regulators can upregulate downstream efflux genes to confer antibiotic resistance [25-27]. The presence of several types of MDR pumps and associated transcriptional regulators expand the mutational targets available to confer antibiotic resistance, and this has rendered efflux-mediated resistance an important contributor to antibiotic resistance [24, 28].

The rise of ARMs. ARMs can exist in a bacterial population as standing genetic variation before an antibiotic treatment or arise *de novo* during an antibiotic treatment. Neutral and slightly deleterious genetic variants can segregate in a population at a low frequency due to a steady mutational supply from spontaneous mutations and the constant gene flow between bacteria [29, 30]. Upon antibiotic treatment, segregating ARMs can provide an immediate selective advantage. However, resistance-conferring variants that segregate in a non-antibiotic environment usually confer low levels of resistance and may be outcompeted by other variants in environments with high antibiotic concentrations [31].

Some antibiotics increase cellular mutation rates, increasing the rate at which mutational resistance occurs, via activation of the SOS and RpoS-regulated stress responses at sublethal concentrations [32, 33]. These responses activate error-prone DNA repair, including the expression of the error-prone DNA polymerase IV, and result in elevated mutation rates, and thus an increase in the origination of ARMs [34]. Such antibiotic-induced mutagenesis can facilitate the emergence of resistance to the inducing antibiotic but can also collaterally increase resistance to other antibiotics [34].

In natural bacterial populations, mutators, bacteria that have a 10- to 1000-fold elevated mutation rate due to defects in DNA mismatch repair, continually arise [35], and the increased mutational supply can provide a selective advantage, especially in the face of a new environmental challenge [36]. In the challenge of an antibiotic treatment, mutators play a role in the emergence of antibiotic resistance. Bacterial mutators colonized around 36% of *Pseudomonas aeruginosa*-infected cystic fibrosis patients, and the mutator isolates were more resistant than non-mutators to various antibiotics [37]. A high frequency of mutators and a close association with multi-drug resistance are common in infections caused by different pathogens [38, 39].

1.2.2 Antibiotic resistance genes

Mechanism of action of ARGs. ARGs, unlike ARMs, are generally non-native to a bacterial chromosome and are usually acquired via horizontal gene transfer (HGT) from other bacteria. Collectively, ARGs encode genes that confer antibiotic resistance through a variety of mechanisms, including target modification, target protection, target replacement, antibiotic efflux, and antibiotic modification (Table 1.1).

Target modification can confer antibiotic resistance by interfering with the binding of antibiotics to their cellular targets [40]. As described in the section above, modification can occur through genetic mutations in genes encoding target molecules, but ARGs can also non-genetically modify the target molecules. Two target modifying enzymes— erythromycin ribosomal methylation (*erm*) and chloramphenicol-florfenicol resistance (*cfr*) genes— encode rRNA methyltransferases that add methyl groups to 23S rRNA [41, 42]. These modifications decrease binding to macrolide and chloramphenicol classes of drugs, respectively. The interaction between a drug and its targets can also be inhibited by protecting targets with surrogate molecules or producing alternative target molecules. Well-known examples of target

Table 1-1. Examples of ARMs and ARGs mentioned in this chapter.

Genetic basis of resistance	Genetic source of resistance	Mode of resistance	Examples
Antibiotic resistance mutation	Chromosomal mutations	Modification of the drug-target interaction	Mutations in the resistance determinizing region of the β subunit of RNA polymerase *rpoB* [15] and DNA topoisomerase/gyrase [16] interfere with drug binding.
		Changes in permeability to drugs	Mutations in porin encoding genes can decrease the import of drug molecules [21].
			Mutations in efflux proteins and their regulators can increase the export of drug molecules [26].
Antibiotic resistance gene	Horizontally acquired genes	Target modification	rRNA methyltransferases, such as *erm* and *cfr*, can methylase 23S rRNA and decrease its binding to macrolide and chloramphenicol, respectively [40, 41].
		Target protection	TetM and TetO displace tetracycline molecules from ribosomes [42]; Qnr protects DNA topoisomerases and gyrases from quinolones [43].
		Target replacement	Dfr proteins substitute the functionality of the cellular DHFR but are not inhibited by trimethoprim [44].
		Drug-modifying enzymes	Aminoglycoside acetyltransferases, nucleotidyltransferases, and phosphotransferases can chemically modify and inactivate aminoglycosides [45]; β-lactamases cleave the β-lactam ring [46].
		Changes in permeability to drugs	Efflux proteins increase the export of drug molecules [47, 48].

protection are those mediated by the tetracycline resistance proteins TetM and TetO, and the quinolone resistance protein Qnr. TetM and TetO are ribosome protection proteins that bind to the small ribosomal subunit and dislodge the tetracycline molecule from the ribosome [43, 44]. Qnr proteins are pentapeptide repeat proteins (PRPs) that have a structure similar to DNA. They can mimic DNA to interact with DNA gyrases and topoisomerases and protect the enzymes from quinolone binding [45, 46]. Alternatively, an antibiotic-proof homolog of a drug target can replace the drug-susceptible target. Dfr proteins with similar enzymatic activities to dihydrofolate reductase (DFHR), involved in folate synthesis, can replace DFHR without the inhibition from trimethoprim [47].

HGT-acquired efflux proteins can confer resistance to metal ions and toxic chemicals in natural environments [48-50]. However, they have become an important mechanism of antibiotic resistance since the introduction of antibiotics [51]. Around 26 efflux pumps, that differ in protein identity and efflux mechanism, can confer resistance to tetracycline in both Gram positive and negative bacteria, and about half of them are plasmid-borne ARGs [52]. Efflux pumps can have promiscuous substrate specificity so that their expression can non-selectively export useful nutrients and metabolites important for cell growth [53]. Consequently, some HGT-mediated efflux proteins are under tight transcriptional regulation and their expression is induced by the cognate antibiotic. For instance, the expression of the tetracycline efflux TetR is de-repressed by the binding of tetracycline to the transcription repressor TetA [54].

Finally, some antibiotic resistance enzymes can modify the chemical structure of an antibiotic molecule [55]. For instance, three families of transferases—N-acetyltransferases, O-nucleotidyltransferases, and O-phosphotransferases—can modify the same aminoglycoside molecule [56]. These transferases vary in their activity at different positions of an

aminoglycoside and across different aminoglycoside drugs [56]. In addition, some antibiotic modifying enzymes have substrate promiscuity and mutational changes on the enzymes can expand the spectrum of antibiotics they can confer resistance to. For instance, with a few amino acid changes, the cefotaxime-susceptible β-lactamase TEM-1 can increase its resistance to cefotaxime up to 100,000 fold [57]. Over 170 variants of β-lactamase TEM-1 have been identified so far [58]. The substrate spectrum of the TEM-1 variants varies and has compromised the effectiveness of β-lactam chemical derivatives [59]. The gain of substrate specificity can even extend to different drug classes. For instance, two amino acid changes in the aminoglycoside acetyltransferase AAC(6')-Ib can expand its activity to include conferring resistance to ciprofloxacin, in the quinolone class of drugs. The ciprofloxacin-resistant variant (AAC(6')-Ib-cr) has become a major plasmid-mediated determinant of resistance to quinolones since its first discovery in 2003 [60, 61]. Promiscuous enzymatic activities and the continuing rise of mutational variants have made tackling antibiotic-modifying enzymes and new drug development challenging [59].

Species origins of ARGs. ARGs are non-native to most focal bacteria, being acquired from other bacteria. An identification of the species in which an ARG originates can be useful because it can reveal the natural evolution of the ARG and the selective forces and mechanisms underlying its spread into pathogenic bacteria. While tracking the ultimate origin of an ARG is challenging, the proximal source species of some ARGs have been found. For example, homologs of the vancomycin-resistance *vanHAX* cluster have been found in chromosomes of several vancomycin-producing actinomyces and even a non-producer, *Paenibacillus popilliae* [62, 63]. However, the cluster's high GC content compared to nearby chromosomal sequence suggests those proximal species also acquired the gene cluster from other bacteria [62, 63].

Similarly, quinolone-resistance Qnr proteins are found in the chromosomes of several bacterial species [64-66]. Transferring a chromosomal ARG to a mobile DNA element is a singular event making it difficult to capture in nature. Yet, a well-controlled laboratory condition has illustrated the mobilization process in real time. The plasmid-borne CTX-M type β-lactamase, BlaCTX-M, has >95% sequence similarity to a chromosomal β-lactamase, BlaKLUA, found in *Kluyvera ascorbate* [67]. The chromosomal bla_{KLUA} is transcriptionally silent and does not confer β-lactam resistance to *K. ascorbate* [67]. However, transposition of the IS element, *ISEcp1B*, to the upstream of bla_{KLUA} provides an active promoter allowing a β-lactam resistance phenotype to be expressed [68]. Moreover, *ISEcp1B* can combine with bla_{KLUA} and transpose as a unit onto a conjugative plasmid that can transfer to other bacteria via conjugation [68]. This example demonstrates a real-time mobilization and transformation of a cryptic chromosomal ARG to an active and transmissible ARG.

Natural functions of ARGs. Some antibiotics and ARGs are likely to have been selected through a role in mediating microbial competition [69]. Antibiotic-producing bacteria release antibiotics to inhibit their bacterial competitors and express cognate ARGs as a self-protective mechanism [70]. For example, the vancomycin-resistant *vanHAX* operon found in vancomycin-producers and the rRNA methyltransferase encoding gene, *erm*, in macrolide-producing actinomyces can protect the antibiotic producer from the produced antibiotics [71, 72]. However, countering the inhibitory action of antibiotics is only one of the natural functions of ARGs. For example, the aminoglycoside acetyltransferase AAC(2')-Ia can acetylate peptidoglycan in *Providencia stuartii* [73]. This might provide an advantage because acetylation of peptidoglycans can prevent the autolytic action of muramidases and maintain the peptidoglycan structure during cell wall turnover [73]. In addition, environmental bacteria are

known to degrade a variety of chemicals, and some can even consume antibiotics as nutrient sources, though it is unknown how common this is [74-77]. Although the natural functions of many ARGs are unclear, functional metagenomic studies have found that many bacterial genes possess enzymatic activity against antibiotics despite having low or no sequence similarity to clinical ARGs [78, 79]. The discovery of putative ARGs in metagenomic samples has raised concerns about the size and diversity of the available 'resistome', the gene pool of ARGs [80].

Intercellular and intracellular mobility of ARGs. Regardless of the species origin and the gene mobilization process of ARGs, currently circulating genes are often associated with mobile genetic elements (MGEs) (Figure 1.2). MGEs can confer on ARGs an intracellular mobility to move between different DNA elements and an intercellular mobility to transfer between bacterial hosts [4]. Transposable elements (TEs), including insertion sequence (IS) elements and transposons, can transpose to different genetic locations within a bacterial genome. The transposition can capture new cargo genes and relocate from a chromosome to other MGEs, such as plasmids. In addition, some transposons can provide promoters to promoter-less cargo genes and activate their expression [68, 81]. Indeed, after plasmids, TEs are the most common vector of ARGs [82]. Analysis of over 80,000 bacterial genomes found that 8.8% of ARGs are associated with TEs, compared to only 5% with other MGEs [82]. Integrons are a type of MGEs that capture cargo genes to form an array of gene cassettes. They encode an integrase that facilitates recombination between the *attI* site on an integron and the *attC* site on a gene cassette,

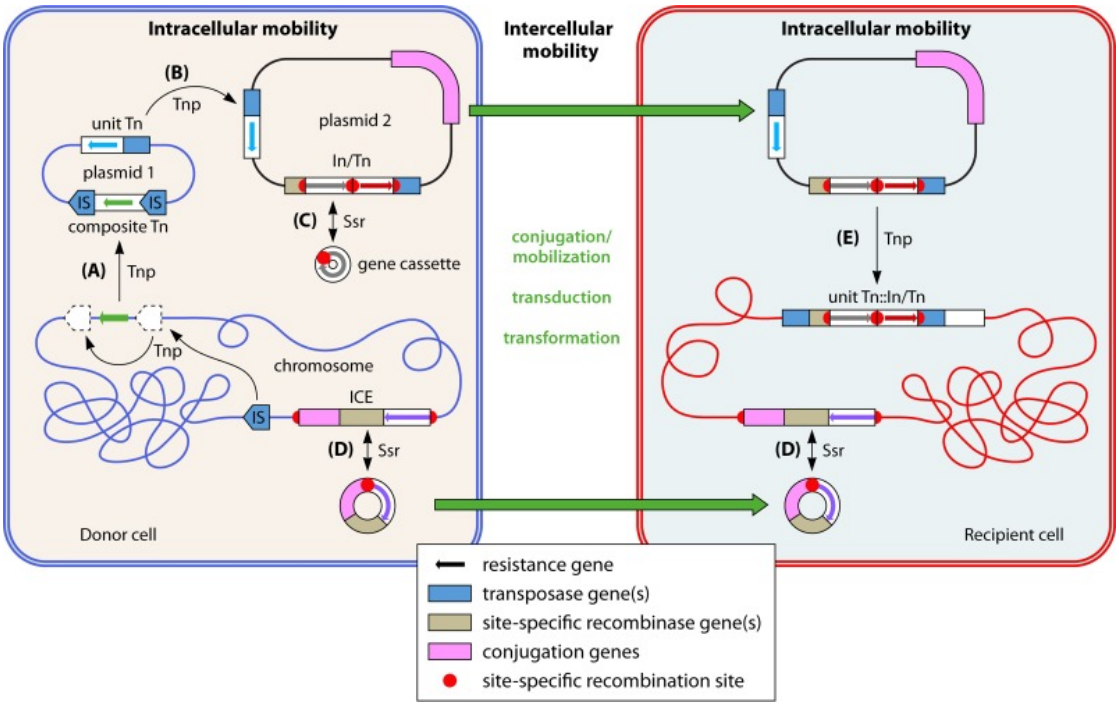

Figure 1.2. Mobile genetic elements (MGEs) provide intercellular and intracellular gene mobility.

Intracellular mobility: transposable elements (Tn and IS) mediate transposition between a chromosome and plasmid (A) and between plasmids (B). Integrons capture gene cassettes (C). Intercellular mobility is mediated by conjugative plasmids and integrative conjugative elements (D). The image is adapted from Partridge et al., 2018 [4].

which can be any gene with an *attC* site, and a Pc promoter that drives the expression of the array of gene cassettes [83]. As a result, integrons are able to capture up to ~180 ARG encoding gene cassettes [84]. Although integrons do not encode transposases and have no intrinsic mobility, around 63% of integrons are nested within other MGEs that confer on them some intracellular and/or intercellular mobility [82].

Plasmids, integrative conjugative elements (ICEs), and bacteriophages are MGEs that provide intercellular mobility [4]. Conjugative plasmids and ICEs encode conjugative machinery that form protein channels between donor and recipient cells and relaxases that mediate the transfer of the conjugative elements. The conjugation efficacy and host-range of a conjugative element

determine its transmission within and between bacterial populations. The significance of conjugative elements for HGT has inspired the evaluation of the possibility that the spread of ARGs can be reduced by inhibition of conjugation [85, 86]. Compared to conjugation, gene transfer via phage-mediated transduction is approximately 1000-fold less efficient [87]. Moreover, ARGs are rarely found in phage genomes as they provide no direct benefit to a phage [88]. However, a recent study on phage-plasmids, temperate phages that possess plasmid features and exist as circular plasmids in their lysogenic phase, showed that a substantial fraction (4%) carry ARGs compared to 20% of plasmids and less than 1% of other phage types [89]. The transition of HGT mode from conjugation to viral infection may confer high transmission ability to phage-plasmids, making them a potentially effective mediator of ARG spread.

1.3 Fitness cost of antibiotic resistance determinants

Antibiotic use is the primary cause of the prevalence of antibiotic resistance in bacteria [90, 91]. Not only does the rise of antibiotic resistance in pathogens coincide with the introduction of antibiotics [92], but also the epidemiology of resistant pathogens correlates with the use of specific drug types [93, 94]. However, factors other than antibiotic selection can also influence the spread and maintenance of AR determinants. One such factor, the fitness cost of an AR determinant in an antibiotic-free environment, has been identified as a possible source of purifying selection against resistance determinants [95]. If an AR determinant is costly, resistant bacteria will be less fit than otherwise isogenic susceptible competitors in the absence of antibiotic selection. This idea has inspired resistance control strategies involving the restriction of the use of antibiotics to limit the spread of resistant bacteria [96, 97]. However, the efficacy of these strategies is equivocal [96, 97]. In this section, I review the evolutionary

concept of the cost of resistance and the mechanistic causes of that cost. Linking the cost's theoretical basis to its physiological mechanism can provide insight into the evolution of resistance and the effective design of antibiotic treatment plans.

1.3.1 The evolutionary concept of the cost of antibiotic resistance

When an organism is well-adapted to one environment, it may have low fitness in another. For instance, a thermophilic bacterium can grow happily at 80°C but cease to proliferate below 40°C [98]. The environment-dependent fitness can be due to a trade-off between fitness optima in different environments arising due to historical selection and intrinsic constraints in the generation of phenotypes representing different fitness optima [99, 100]. The concept extends to predict the cost of resistance to parasites and abiotic molecules, such as insecticides and antibiotics [101-103]. Considering an antibiotic as a selective component of an environment, it is natural to assume the genotype coding for optimum fitness in an antibiotic-free environment will be different from that in an antibiotic-containing environment [104]. Hence, the resistance to an antibiotic may come with a fitness reduction in the antibiotic-free condition and consequently a cost of resistance [104, 105].

Experimental studies on the fitness cost of antibiotic resistance begun in the 1990s [106-108]. These studies measure the fitness cost of resistance genotypes, either ARMs or ARGs. The fitness cost of a resistance genotype is operationally defined as the fitness difference between a resistant genotype and the susceptible wild type in an antibiotic-free environment. Some studies found a fitness cost of a resistance genotype, but others did not [109]. The discrepancy can lie in the pleiotropic effects of different resistance mechanisms. When a new resistant genotype arises in an organism, it may influence many phenotypic traits due to complex genetic

interactions [110-113]. For example, a mutation in the RNA polymerase subunit, RpoB, may prevent the binding and inhibitory effect of an RNA polymerase targeting drug but also compromise transcriptional activity [114]. However, not all the mutations influence the transcription equally [114].

There is a distinction between the fitness cost of antibiotic resistance and the fitness cost of a resistance genotype on how evolutionary changes fine-tune the fitness cost. The fitness cost of a resistance genotype can be modified by subsequent mutations that alter its pleiotropic effects but need not alter the cost of antibiotic resistance, i.e. the fitness optimum in an antibiotic-containing environment [104]. The cost of resistance evolves through changes in the form of its environmental or genetic trade-off. It may require frequent exposure to and selections in both antibiotic-containing and -free environments to achieve fitness optima in both environments [104, 115]. For instance, when unneeded, an overexpression of the tetracycline efflux TetA can cause a significant reduction in cellular growth [108]. To reduce its costly and unneeded expression, the transcription of *tetA* is tightly repressed by a transcriptional repressor, TetR and only de-repressed when tetracycline binds to TetR and dislodge TetR from the *tetA* operators [108]. Thus, the acquisition of tetracycline-mediated regulation of *tetR-tetA* operon achieves the fitness optima in both tetracycline-free and -containing environments.

The distinction between the cost of resistance and a resistance genotype has practical and evolutionary implications. The fitness cost of a resistance genotype does not predict or represent the cost of resistance. Hence, the absence of a fitness cost of a resistance genotype does not necessarily indicate the absence of the cost of resistance or of the fitness cost of other resistance genotypes. Moreover, the pleiotropic effects of a resistance genotype may lead to different evolutionary outcomes, for instance, selection for compensatory mutations. Hence,

success in limiting the spread of antibiotic resistance by restricting antibiotic use is complicated by the pleiotropy of resistance genotypes and the possibility of subsequent compensatory evolution.

1.3.2 Biological basis of the cost of an AR determinant

The fitness cost of an AR determinant can originate from disturbed genetic interactions between a bacterial genome and the newly acquired determinant. Studying the molecular mechanisms of the disturbed interactions can help us understand an AR determinant's spread and evolution.

Fitness cost of ARMs. ARMs often occur in genes encoding antibiotic targets, which are usually genes essential for cell proliferation. Mutations in essential genes are likely to be deleterious [116]. Indeed, empirical studies show that ARMs collectively have a higher fitness cost than ARGs [117]. Mechanisms underlying the cost of ARMs have been studied in some cases. For example, mutations in the rifampin target, the β subunit of RNA polymerase, *rpoB*, often reduce transcriptional efficiency. Differences in the transcriptional efficiencies of nine *rpoB* resistant mutants explain 76% of the variation in their fitness costs [114]. Similarly, resistance mutations occurring in the streptomycin target, the ribosomal protein *rpsL*, are associated with a reduced translation rate [118]. Bacteria can accumulate different ARMs and become multi-drug resistant but at the cost of compounding perturbations in their cellular physiology. For example, bacteria with both a rifampin-resistant *rpoB* mutation and a streptomycin-resistant *rpsL* have a higher risk of double-strand DNA breaks due to a combination of compromised transcription and translation [119, 120]. Another example is the increased expression of efflux pumps conferring a cost due to increases in the energetic expense of driving export (for instance,

ATP-dependent transporters [121]) and increased loss of beneficial metabolites. For instance, in *P. aeruginosa,* overexpression of MexEF-OprN, a proton/substrate antiporter, results in an accumulation of protons in the cytoplasm [122, 123] and increased export of its natural substrate, kynurenine, a quorum sensing precursor [124]. Efflux substrates can be involved in environment-dependent regulation of cellular physiology so that the fitness cost associated with increased efflux is also environmentally dependent. For instance, the reduction of kynurenine does not affect *P. aeruginosa* fitness in normal conditions, yet it can impair the pathogenicity of *P. aeruginosa* to nematodes [124].

Fitness cost of ARGs. In contrast to the occurrence of ARMs' in native target genes, ARGs are not native to a bacterial genome, and an acquisition of an ARG may be expected to cause less disturbance to the physiology of a bacterial recipient. Phylogenetic analyses show that successful HGT-acquired non-homologous genes are biased toward genes involved in cellular operation, such as metabolism, that are presumably less integrated into the genetic network of a cell than genes involved in cellular information, such as transcription [125, 126]. Consistent with this hypothesis, a study measuring the fitness costs of ~200 ARGs in a naïve bacterial host found that drug-interacting ARGs, such as drug-inactivating enzymes, imposed lower fitness costs and had higher resistance activity than cell-interacting ARGs, such as efflux proteins [127].

In addition to protein functions, sequence features of an ARG that influence expression–for example, GC content, codon usage, and RNA stability–can also affect the fitness cost of an ARG [128-130]. This suggests that the expression of an acquired gene can impose a general cost on a host, perhaps through the direct metabolic cost of expression or production of toxic misfolded proteins [131, 132]. However, there was no correlation between the fitness cost of

~200 ARGs and their sequence features like codon usage, GC content, and RNA stability [127]. There was, however, a positive correlation between the cost of cell-interacting, but not drug-interacting, ARGs, and the phylogenetic distance between donor and recipient bacteria [127]. More specific mechanistic causes underlying the cost of an ARG have been found. For instance, the signal peptide of several β-lactamases can confer a fitness cost, likely due to a disruption of cell membrane and cell wall during the transport of the β-lactamases to the periplasm [133-136].

1.3.3 Fitness cost of an AR determinant varies with its genetic background and environment

As for mutations, the fitness effects of AR determinants can depend on their broader genetic background and the environment in which they are tested [1, 127, 137, 138]. For instance, the consequence of the MexEF-OprN overexpression hinges on the presence of environmental oxygen and the availability of specific bacterial niches [122-124]. Other environmental factors, such as nutrient level, carbon sources, and temperature, can influence the fitness cost of rifampin-resistant mutations of *rpoB* and streptomycin resistant mutations on *rpsL* [139-142]. Environment-dependent effects can explain the discrepancy on the fitness cost of an AR determinant measured *in vivo* and *in vitro* in some cases [143].

Differences in the evolutionary history, ecological niche, and physiology of a bacterial strain can shape the genetic interaction between an AR determinant and the host genome [127, 144]. Whatever the basis of these interactions, host-dependent costs may lead to the association of an AR determinant with a specific bacterial species or strain. For instance, the fitness cost of a β-lactamase OXA48-carrying plasmid, pOXA-48_K8, explains its prevalence among some *K.*

pneumoniae isolates [145]. The naïve isolates in which the plasmid confers a large cost during laboratory growth are genetically close to the clinical isolates that seldom carry the plasmid [145]. Given that both host genetic background of a bacterium and environment can influence the fitness cost of an AR determinant, fitness costs should be examined in various conditions to understand and generalize their influence on the dissemination and evolution of an AR determinant [146].

1.3.4 Compensation evolution to the fitness cost of an AR determinant

The fitness cost of an AR determinant is a result of its pleiotropic effect interfering with essential aspects of cell physiology, and hence, other genes that affect the same physiology have the potential to evolve to modify its fitness cost. In an experimental evolution study on the rifampin-resistant *rpoB* mutants of *M. tuberculosis*, compensatory mutations in the α and β′ subunits of RNA polymerase (*rpoA* and *rpoC*, respectively) could compensate the fitness cost of *rpoB* mutations in the evolved clones [147, 148]. Moreover, those compensatory mutations were highly associated with rifampin-resistant clinical isolates—around 27% of rifampin resistant isolates have mutations in *rpoA* and *rpoC* [147]. The *rpoA* and *rpoC* mutations presumably restore the reduced transcription activity caused by *rpoB* mutations by fine-tuning the interaction among the nascent RNA, incoming nucleotides, and other subunits of the RNA polymerase [148, 149]. The fitness cost of mupirocin-resistant mutations occurring in isoleucyl–tRNA synthetase, IleRS, can also be compensated by mutations occurring in interacting genes [118]. Both secondary mutations on the mutated IleRS that restore its enzymatic activity and mutations on its transcriptional regulators that increase the expression of the mutated IleRS can alleviate its cost [118].

The pleiotropic effect of an AR determinant increases the size of gene targets available for compensation evolution to act on. Hence, in the absence of antibiotic selection, costs of an AR determinant are less likely to be reduced by a mutational reversion of an ARM to the susceptible genotype than origination of some new compensatory mutation [95, 150, 151]. Moreover, environments and genetic backgrounds can influence compensatory evolution [152]. For instance, compensatory mutations to the fusidic acid-resistant mutations of *fusA*, an elongation factor G that mediates ribosome translocation during translation, are mainly via a reversion in mouse but via secondary mutations on *fusA* in vitro [153, 154]. The pleiotropy of an AR determinant suggests that there are multiple genetic solutions to compensate its fitness cost. Exploring the possible compensatory mechanisms may predict the prevalence and evolution of an AR determinant.

1.4 Ecological and evolutionary factors on the prevalence of ARGs

The genetic origin and transmission route differences between ARMs and ARGs can influence their ecological and evolutionary dynamics in bacterial populations. ARMs originate in the genome of a focal bacterium, and their transmission is largely confined to vertical transmission to daughter cells. In contrast, ARGs are acquired from other bacteria and their association with mobile DNA elements equip them to transfer horizontally. The transient genetic exchange between a donor and a recipient makes it challenging to track the spread of an ARG. In this section, I review studies on the possible sources of ARGs and biological and ecological factors that influence the distribution of ARGs.

1.4.1 A large gene pool of ARGs

Antibiotics have selected and enriched the gene pool of ARGs in pathogens, yet the actual pool is challenging to estimate. ARGs existed in soil and human microflora long before the discovery and medical use of antibiotics. DNA fragments with high sequence similarity to modern ARGs were discovered in 30,000-year-old permafrost sediments [155] and in ancient human microbiomes [156, 157]. Functional analysis of some ancient ARGs confirmed their enzymatic activity and resistance to modern antibiotics [155, 158]. Their ancient existence supports the view that ARGs are naturally present in microbial populations. This is further confirmed by the widespread presence of ARGs in microbiomes with little or no exposure to anthropogenic antibiotics [159-163].

Although ARGs are ancient and indigenous, the introduction of antibiotic use has imposed positive selection on antibiotic resistance and changed the prevalence of ARGs in bacterial populations. A study on soil samples collected since the 1940's found that the copy number of ARGs per bacterial genome has increased up to 15-fold [164]. In addition to commonly circulating ARGs, many putative genes can confer antibiotic resistance. Functional metagenomics studies on microbiomes from various sources have uncovered not only genes with a high similarity to ARGs in pathogens but also novel putative ARGs [165, 166]. The prevalence of clinical and putative ARGs in microbial genomes, either due to antibiotic selection or other natural selection forces, points to an accessible toolbox for pathogenic bacteria to source the armoury to resist antibiotic treatment.

1.4.2 Gene flow between microbiomes potentiates but also limits the spread of ARGs

Despite a large gene pool of ARGs, there are barriers that inhibit their transfer. For example, bacteria need to share ecological niches and have compatible DNA protection systems for successful direct genetic exchange to occur. The frequency of HGT between human-associated bacteria is 25-fold higher than that between human-associated and environmental bacteria [167]. Within human bodies, microbiomes from different body parts are under similar selection pressures, which can facilitate gene flow, making commensal bacteria a reservoir of ARGs for pathogenic bacteria to acquire antibiotic resistance [168, 169]. Although HGT between human-associated and environmental bacteria is less frequent than transfer within those groups, antibiotic selection can facilitate the gene flow and selection of ARGs among environmental isolates. For instance, food or farm-associated bacteria carry more clinically relevant ARGs than other environmental bacteria [167], which could be a result of agricultural antibiotic use [170]. With a rare HGT event and timely selection for antibiotic resistance, an ARG from environmental bacteria may emerge and spread into human commensal and pathogenic bacteria [167, 169]. The rare emergence may explain why many clinically relevant ARGs' source species have yet to be identified [171].

1.4.3 Spread and maintenance of ARGs via plasmids

Co-selection of ARGs with plasmid-associated traits. Plasmids play a significant role in the dissemination of ARGs. Apart from antibiotic selection, the selfish nature of plasmids and co-selection of other plasmid-borne adaptive traits can facilitate the spread and maintenance of ARGs [172-174]. Without positive selection for plasmid-borne adaptive traits, a plasmid can sometimes persist in bacterial populations via the ability to transfer between cells by

conjugation [175]. Once entering a new host, active plasmid segregation and the activity of post-segregational killing systems can establish stable maintenance of a plasmid within a bacterial population [176, 177].

Although active segregation and killing systems effectively decrease plasmid missegregation, the formation of plasmid-free cells is probably inevitable. Moreover, many plasmids are not conjugative, reducing the chance that they can re-infect host cells after loss [178]. In that case, any fitness cost of a plasmid puts a plasmid carrier at a fitness disadvantage relative to competing plasmid-free cells and stable plasmid persistence requires occasional positive selection for environmentally adaptive genes encoded by the plasmid, such as genes involved in antibiotic resistance, virulence, metabolism, and detoxification of metal ions [179]. Positive selection for one plasmid-borne adaptive gene can also co-select other adaptive genes co-localized on the same plasmid [179]. For instance, plasmids that carry multiple drug resistance genes or ARGs and virulence genes may become prevalent in pathogenic bacteria due to the co-selection of the plasmid-borne adaptive genes [180, 181].

Compensation of the fitness cost of a plasmid enhances maintenance. The plasmid transfer machinery, plasmid-borne genes that are only adaptive in certain environments and the genetic conflict between a plasmid and a host genome can confer a fitness cost on a host [182]. Countering this, mutations that compensate for plasmid costs can arise and increase plasmid persistence [173, 183]. In *P. fluorescens* SBW25, the fitness cost of the mercury-resistance plasmid pQBR originates from activation of the SOS response and its downstream toxin gene [183]. Mutations on *gacA/S* and a putative gene, *PFLU4242*, can suppress the SOS response and consequently reduce the fitness cost of the plasmid [173, 183]. Alternatively, changes in activation of the SOS response can compensate fitness costs. During plasmid replication, an

imbalanced stoichiometric binding of plasmid replication protein RepA with the plasmid replication origin can result in stalled DNA replication folks, which triggers the SOS response [184, 185]. A resolution of the DNA replication stresses by either an overexpression of the host DNA replication proteins, such as DNA primase and helicase, or an increase of the binding sites for the RepA can relieve a plasmid-imposed fitness cost [185, 186]. This type of cause and resolution of plasmid cost may be general to different plasmids as compensatory mutations to various plasmids have happened in DNA helicases [185, 187-189].

Although compensatory evolution can slow down the loss of a plasmid from a bacterial population, anything but complete compensation of plasmid cost will end in an inevitable loss of the plasmid from a population in the absence of plasmid selection or reinfection [190, 191]. San Millan et al. (2014) found a synergy that complicates this simple expectation. When plasmid-containing cells were exposed to periodic positive selection for a plasmid encoded gene and compensatory evolution to plasmid costs, the benefits reinforced each other to facilitate plasmid persistence [191]. While compensatory evolution slowed the loss of plasmid in a bacterial population, occasional positive selection for the plasmid eliminated plasmid-free cells, creating a selective advantage for plasmid-carrying cells [191]. Consequently, the interaction between positive selection and compensatory evolution increased the half-life of a plasmid in populations and may explain the abundance of plasmids in nature.

Genetic background of a bacterial host impacts the ecological dynamics of a plasmid. A plasmid undergoes frequent HGT between bacteria, and consequently, the genetic background of a host can influence the transmission and the fitness cost of a plasmid. The stability of a mercury-resistant conjugative plasmid pQRB57 differs in two *Pseudomonas* populations. In *P. fluorescens*, the plasmid stably persisted in a population regardless of the presence of

environmental mercury. In contrast, in *P. putida*, the plasmid was either lost quickly in the absence of mercury or transposed the mercury resistance gene to the bacterial chromosome and lost the plasmid backbone in mercury selective conditions [192]. The difference in the fate of the plasmid lies in the conjugation rate, which was high in *P. fluorescens* but not in *P. putida*. The disparity in the conjugation rate can create an unbalanced gene flow between bacteria. When the two species coexist in the same environment, *P. fluorescens* becomes a source species of the plasmid and transfers the plasmid to the sink species, *P. putida* [192].

A source-sink relationship of bacteria facilitates the spread of a plasmid to a bacterium with a poor conjugation ability. However, the differences in conjugation ability may impede plasmid transmission in a complex bacterial community. In a community with multiple bacterial members, the variation in conjugation rate and plasmid maintenance between members may curtail the plasmid transmission because conjugation and maintenance-incompetent members dilute conjugation opportunity and decrease transmission efficiency [193]. In addition, variation in the fitness cost of a plasmid between bacterial hosts can influence its maintenance in a bacterial community. Without antibiotic selection, the fitness effect of a β-lactamase OXA-containing plasmid pOXA-48_K8 ranges from up to a 20% fitness cost to a 20% fitness benefit across different host strains [145]. A simulation model predicts that the variability in the plasmid cost can unexpectedly enhance its maintenance in a bacterial community where members that benefit from plasmid carriage act as a reservoir [145]. Similar variability of plasmid cost is observed in the persistence of a broad-host-range plasmid in a natural microbial community, where 17% of phylotypes suffer from the plasmid cost but 24% can profit from the plasmid [194]. Natural host variation in plasmid transmission and maintenance highlights

a need to compare a plasmid's behaviour between bacterial hosts and study plasmid dynamics in a bacterial community.

Uncoupling of an ARG from its residential plasmid. The evolutionary and ecological dynamics of an ARG are tightly contingent on its plasmid vector, yet their dynamics can diverge in environments where they have distinct selective interests. Under antibiotic selection, if the selective advantage of an ARG outweighs selfishness or the net benefit of the plasmid ARGs, an ARG can transpose to the host chromosome and be lost from the plasmid vector [173, 190, 195]. Correspondingly, in antibiotic-free conditions, if the fitness cost of an ARG is higher than that of the plasmid, deletion of the ARG can be favoured [135, 196-198]. An ARG often embeds in a transposable element and thus gains a transposable activity [199, 200]. Furthermore, plasmid recombination can mobilize an ARG to different plasmids and shape its intercellular mobility to different hosts [201-204]. A phylogenetic analysis testing for an association of carbapenemases with plasmids and bacterial hosts showed different patterns of plasmid-associated transmission [204]. For instance, bla_{OXA-48}-like genes are mainly embedded in one plasmid, which has spread into several lineages of *Klebsiella pneumoniae*. In contrast, bla_{VIM} and bla_{NDM} associate with multiple plasmids and spread into several bacterial lineages [204]. Thus, although an ARG is strongly associated with its plasmid vector, it is an independent segregational unit. Its transmission and maintenance should be considered beyond its physical linkage with a plasmid.

1.5 Research objectives

The One Health plan advocates that human wellness should include consideration of the health of animals and the environment [9]. In the face of infections caused by antibiotic resistance

pathogens, a sole reliance on new drug development has proven to be an ineffective tactic to tackle antibiotic resistance [8, 11]. Pathogenic bacteria can acquire antibiotic resistance through evolution and genetic exchange with non-pathogens. Therefore, knowledge of the evolution and selection of antibiotic resistance in non-pathogenic bacteria is a necessary part of a full understanding of the emergence, persistence and dissemination of antibiotic resistance. For example, antibiotic resistance can impose a fitness cost on bacteria [95, 109]. This cost can lead to purifying selection against resistant bacteria in the absence of antibiotics but also drives compensatory evolution to reduce the cost. Together with antibiotic selection, purifying selection and compensatory evolution can influence the evolution and persistence of antibiotic resistance in bacteria.

Mutations in chromosomal genes and horizontally transferred antibiotic resistance genes can confer antibiotic resistance to bacteria. ARGs are associated with mobile DNA elements that mediate the gene transfer between bacteria. The mobility of an ARG makes it difficult to predict changes in frequency in susceptible bacteria. Moreover, genetic linkage of ARGs with mobile DNA elements couples the ecological and evolutionary dynamics of ARGs to selfish DNA elements. Consequently, our knowledge of the selective coefficients of an individual ARG is scarce. In this thesis, I aim to study the fitness cost of individual ARGs and the influence of these costs on the persistence and evolution of ARGs. To do this, I investigate the variation of the fitness cost of a range of ARGs between bacterial hosts and the consequences of variation in costs on an ARG-carrying plasmid's ecological and evolutionary dynamics (**Chapter 2**). In **Chapter 3**, I dissect the genetic basis of compensatory evolution to the cost of an ARG. Finally, I conduct an evolution experiment to evaluate the influence of the genetic background of bacterial hosts on the evolution of an ARG **(Chapter 4)**.

Aim 1. Quantify the variability of an ARG's fitness cost and explore its biological significance. In **Chapter 2**, I estimate the fitness costs of six ARGs across a set of 12 *Escherichia* bacteria. Next, I investigate the variation in cost of one ARG on its maintenance and evolution in three different host strains.

Aim 2. Determine the genetic basis of mutations that compensate for the fitness cost of an ARG. As a follow-up on **Chapter 2**, I conduct whole genome sequencing to dissect the genetic basis of the compensation evolution in evolved ARG-carrying cells (**Chapter 3**).

Aim 3. Investigate the influence on host genetic background on the evolvability of an ARG. In **Chapter 2**, I discover that the fitness cost of an ARG is host dependent. In **Chapter 4**, I investigate if the host-dependency cost can influence the genetic variation of a substrate-promiscuous ARG in different hosts. The size and diversity of the genetic variation may influence the gain of new antibiotic resistance in a promiscuous antibiotic deactivating enzyme.

The fitness cost of an ARG is an often-neglected selection parameter, but it can play an important role in an ARG's prevalence in environments with no or weak antibiotic selection. Completion of my three aims has the potential to broaden our knowledge on the persistence, prevalence and evolution of antibiotic resistance in non-clinical environments—an understudied field integral to the One Health plan [11].

Chapter 2. Costs of antibiotic resistance genes depend on the host strain

2.1. Abstract

The prevalence of antibiotic resistance genes (ARGs) is a major contributor to bacterial antibiotic resistance. ARGs are selected when host cells grow in environments containing corresponding antibiotics, but it is less well understood how they are maintained in environments where antibiotic selection is weak or sporadic. In particular, relatively little is known about the potential for ARGs to impose costs in the absence of direct antibiotic selection and whether such costs are fixed or depend on the host bacterial strain. In this chapter, I quantify the fitness effects of six ARGs in each of 12 diverse *Escherichia* strains. Four genes, *aadA*, *cat*, *dfrA5*, and $bla_{TEM\text{-}116}$*, confer a small cost on average, but $bla_{SHV\text{-}12}$ and $bla_{CTX\text{-}M\text{-}15}$ do not. Moreover, there is a variation of the cost between strains that can be partially explained by the interaction between an ARG and its host. To investigate the consequences of the host-dependent cost of an ARG, I follow the evolutionary dynamics of a $bla_{TEM\text{-}116}$*-encoding plasmid in three hosts with various degree of $bla_{TEM\text{-}116}$* cost. While the host in which the gene initially conferred a large cost evolved to compensate the cost and slow plasmid loss, plasmid dynamics remain largely unchanged in two other hosts in which the gene had a zero or small cost. Overall, these results suggest that although costs may select against carriage of ARGs in some strains, other strains might be available to act as reservoirs, allowing ARGs to be maintained in complex communities. Furthermore, compensatory evolution can quickly evolve when the fitness cost of an ARG is large to change even immediate barriers to ARG carriage in a host-dependent manner.

2.2. Introduction

The increasing prevalence of antibiotic-resistant bacteria is recognized by the World Health Organization as a major global health concern [205]. Indeed, a recent comprehensive analysis estimated 1.27 million deaths were directly attributable to antimicrobial resistance in 2019 [205]. Antibiotic resistance genes (ARGs) associated with mobile genetic elements play an important role in the spread of resistance by enabling gene transfer between strains, and even species, including from environmental to pathogenic species [171]. While ARGs likely evolved to provide resistance to antibiotics and associated chemicals produced in natural environments, their rapid spread over the last several decades is strongly linked with anthropogenic use of antibiotics [171, 206-208]. For example, the usage of antibiotics and the rise of the resistance are positively correlated within and between countries, and ARGs are often first detected shortly after introduction of cognate antibiotics [94, 209, 210]. Even low levels of antibiotics, below minimal inhibitory concentrations, can confer significant advantages to resistant cells, potentially allowing selection in a wide range of environments [211].

The fitness benefit conferred by ARGs to host cells during antibiotic exposure is, however, only one factor determining their overall success, and the fitness cost of ARGs, revealed in the absence of antibiotic selection, may also be important [95, 212-215]. Of 22 ARGs for which a recent origin could be inferred, most were from species with some human or domestic animal association [171]. While this link underlines the importance of direct selection driving the success of ARGs, many of these species transit through environments where antibiotic concentrations are low [11]. For other ARGs, no originating source could be inferred, consistent with them originating in poorly characterized environmental species [171]. These findings underlie the view that selection of ARGs in environments with low or no antibiotic

exposure will be influential and, therefore, that costs associated with the carriage and expression of ARGs can play a major role in determining their relative success [212]. Consistent with this possibility, antibiotic resistance can plateau despite continuing use of an antibiotic, suggesting some counteracting selection [216, 217].

Several studies have measured the fitness cost of spontaneously occurring antibiotic resistance alleles [218-223]. In one experiment, costs associated with rifampicin resistance were sufficient to drive selection for frequent reversion to sensitivity in a mouse infection model [219]. In other cases, secondary mutations were selected that compensated for costs of resistance (reviewed in [218]). Spontaneous resistance to an antibiotic often occurs in essential genes encoding proteins targeted by that antibiotic [224]. Resistance often arises through changes in protein structure that reduce affinity with the antibiotic, but such changes are likely to also have some deleterious effect on the original protein function, introducing a cost to resistant strains [117]. By contrast, less work has focused on costs of introduced ARGs. A meta-analysis found that costs associated with plasmids generally increase with the number of encoded ARGs, suggesting that ARGs do, at least on average, confer a fitness cost to host cells [117]. Direct measurement of the effect of *tetAR*, encoding tetracycline resistance, found a cost of around 1% in rich medium [108, 197]. Similarly, β-lactamases, a large class of resistance genes that confer resistance to one or more of the β-lactam antibiotics, can decrease host fitness, though different variants have different effects in different hosts [24, 135, 225]. Notably, in at least some situations, ARGs, including *tetAR*, can evolve to provide a benefit to host cells [143, 226].

A factor making it especially important to understand the costs of ARGs is that their association with horizontally mobile elements means they can be readily transmitted among different

bacterial populations (reviewed in [11]). Studies on the fitness effects of mutations, including those leading to antibiotic resistance, have shown that host background can play an important role in determining their fitness effects [220, 227-231]. Although not isolating the effect of the ARG itself, a *bla*$_{\text{OXA-48}}$-carrying plasmid was found to have a mean negative effect across a panel of host strains but with a range between a ~20% cost and a ~20% benefit [145]. Simulations predict that this high variation in fitness effect can promote plasmid persistence in a bacterial community, even in the absence of positive selection [145]. During intermittent selection for resistance, hosts in which the plasmid confers a low fitness cost can act as a refuge from which the plasmid can subsequently transfer to hosts in which it confers higher costs when the environment changes to include the antibiotic, so that the plasmid confers a net advantage [145].

Here, I isolate ARGs from their original contexts and examine their fitness effects when expressed from a common reference plasmid in the absence of cognate antibiotics. I measure the variation in fitness cost of six ARGs conferring resistance to four different classes of antibiotics in 12 genetically diverse *Escherichia* strains. I find that the costs vary across ARGs and strains, and that they interact so that different strains have the potential to act as reservoirs for different ARGs. Finally, I investigate the influence of an ARG's cost on the persistence and evolution of an ARG-carrying plasmid in different bacterial hosts. My results reveal the importance of variation in fitness effects of ARGs in bacterial populations on the prevalence of ARGs in the absence of antibiotics.

2.3. Materials and Methods

2.3.1. Construction of plasmids and bacterial strains.

I chose six ARGs, *aadA*, *cat*, *dfrA5*, $bla_{TEM-116}$*, $bla_{CTX-M-15}$ and bla_{SHV-12} to examine based on their prevalence in sequenced *E. coli* strains and their availability in our laboratory. The prevalence of the ARG class (and specific type) within sequenced *E. coli* genomes is: *aad* 15.3% (*aadA* 1.73%), *cat* 5.3% (*cat* 4.1%), *dfrA* 19.4% (*dfrA5* 1.8%), bla_{TEM} 22.7% ($bla_{TEM-116}$ 0.07%) bla_{CTX-M} 17.1% ($bla_{CTX-M-15}$ 0.42%), and bla_{SHV} 5.8% (bla_{SHV-12} 0.59%) [232]. I initially cloned the six ARGs into a pUA66 derivative containing a promoterless *gfp* [233]. However, I found a leaky expression of GFP in these constructs. To avoid the influence of any cost of GFP expression on later experiments, I removed *gfp* from the pUA66 derivative by PCR amplifying the vector (forward primer pUA_minusFP_F: atgtccagacctgcaggcatg; reverse primer pUA_minusFP_R: ggatccatcgaggtgaagacg) and closing the ends with a linker oligo (cgtcttcacctcgatggatccatgtccagacctgcaggcatg) using the NEBuilder HiFi DNA Assembly (NEB) kit. The resulting *gfp*-free plasmid, pmFP, was used to construct ARG-carrying plasmids used

Table 2-1. Strains used in this study.

Species	Strain*
Escherichia albertii	B156
Escherichia coli	M863
Escherichia coli	H442
Escherichia coli	M114
Escherichia coli	M056
Escherichia coli	B354
Escherichia coli	M605
Escherichia coli	H305
Escherichia coli	H504
Escherichia coli	H588
Escherichia coli	M718
Escherichia coli	REL606

* Source: Michigan State University Shiga Toxin-Producing Escherichia coli (STEC) Center

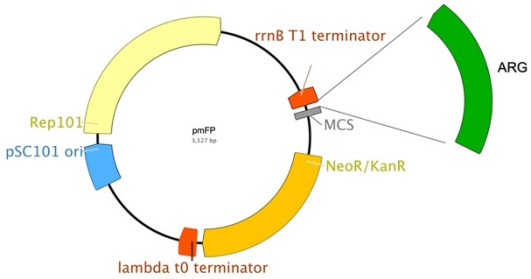

Figure 2.1. Plasmid map of the pUA66-derived plasmids.

The pmFP vector was derived from the pUA66 plasmid with the promoter-less GFP removed [233]. DNA fragments of the selected ARGs were subcloned onto the pmFP plasmid as described in Materials and Methods. The pUA66 plasmid is a low copy number plasmid with a pSC101 replication origin. The plasmid does not encode genes required for conjugation and an origin of transfer, and thus, it is expected to be non-conjugative and non-mobilizable.

in this study (Figure 2.1). The six ARGs were PCR amplified (Table 2.1) and cloned into the amplified pmFP backbone (primers: pUA66_F: aataggcgtatcacgagg and pUA66_EcoRI_R: gaattcatggtttcttagacgtcgg) using NEBuilder® HiFi DNA Assembly (NEB). All PCR reactions were performed using Q5® High-Fidelity 2X Master Mix (NEB) unless otherwise specified.

All inserts were checked by sequencing. I found that the $bla_{TEM\text{-}116}$ cloned from portMAGE2 has a non-synonymous mutation (Q274R) different from $bla_{TEM\text{-}116}$. To indicate this distinction, I denote the variant as $bla_{TEM\text{-}116}$*. The pmFP-ARG plasmids described above were used to transform 12 divergent *Escherichia* strains (Table 2.2 and Figure 2.2).

Table 2-2. Source of test ARGs and primers used to amplify them.

ARG	Primer pairs	Source of PCR templates	upstream of the CDS	downstream of the CDS
aadA	ccgacgtctaagaaaccatgaattcctactcgagttcatgtgcagc cctcgtgatacgcctatttctgacgctcagtggaacg	pTargetF plasmid (Addgene, #62226)	100**	33**
catI	ccgacgtctaagaaaccatgaattcttctggtcggcgcatagctg cctcgtgatacgcctattgcaagcagcagattacgcg	portMAGE4 plasmid (Addgene, #72679)	178	25
dfrA5	ccgacgtctaagaaaccatgaattccatggcttgttatgactg cctcgtgatacgcctatttgtgcttagtgcatctaacg	*E. coli* strain B706	158	93
bla $_{TEM-116}$*	ccgacgtctaagaaaccatgaattcttctggtcggcgcatagctg cctcgtgatacgcctattgcaagcagcagattacgcg	portMAGE2 plasmid (Addgene, #72677)	72	147
bla $_{CTX-M-15}$	ccgacgtctaagaaaccatgaattc cctcgtgatacgcctatt	Sequence was obtained from gene synthesis and based on GenBank: MK125035.1.	250	150
bla $_{SHV-12}$	ccgacgtctaagaaaccatgaattc cctcgtgatacgcctatt	Sequence was obtained from gene synthesis and based on GenBank: CP048293.1.	237	150

The portMAGE4 encoded BlaTEM (BlaTEM-116) has more similarity to BlaTEM-116 than BlaTEM-1, but has one amino acid difference from BlaTEM-116 at amino acid 274 (Q in BlaTEM-116 but R in BlaTEM-116*). BlaTEM-116 differs from BlaTEM-1 at amino acid 82 and 184 (I and A respectively in BlaTEM-1, but V and V in BlaTEM-116).

** The upstream and downstream regions of aadA on pTargetF were contributed by aminoglycoside N-acetyltransferase AAC(3)-Iva (AAC(3)-IV5a).

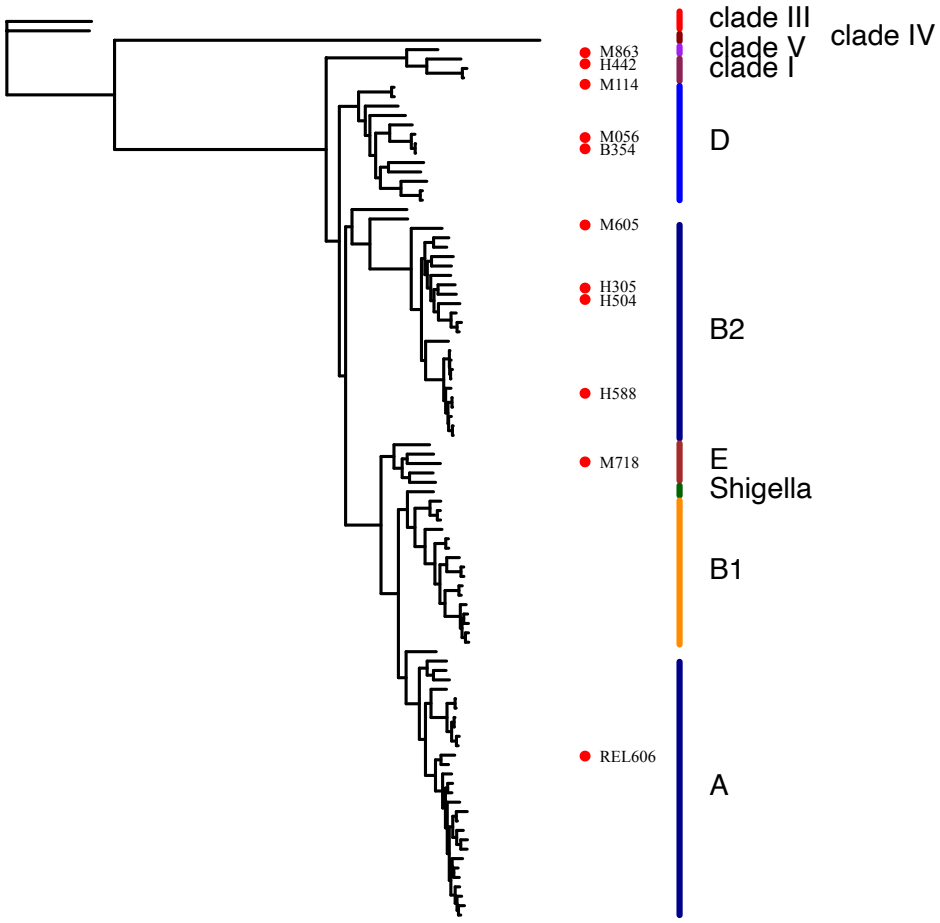

Figure 2.2. Phylogeny of the 12 *Escherichia* strains.

The phylogeny was constructed based on the core genomes of 96 *E. coli* isolates, and this figure was adapted from Wang et al., 2016 [1].

2.3.2. Fitness competitions.

The fitness effect of each ARG was estimated indirectly as the fitness difference between cells carrying an ARG plasmid (W_{P_ARG}) and cells carrying the control plasmid, pmFP (W_{P_Vec}) separately grown in competition with cells carrying a reference GFP expressing plasmid (pUA66-P_{rpsL}GFP [233]). Cells carrying the ARG plasmid, pmFP, or pUA66-P_{rpsL}GFP were inoculated from frozen stocks into 200 μL of Davis minimal broth supplemented with 250 μg/ml glucose (DM250) and 50 μg/ml kanamycin and cultured at 37ºC overnight. Cultures were then pre-conditioned through two 24-hour growth cycles consisting of 1:100 dilution into 200 μL of fresh DM250. Fitness competition assays were initiated by first mixing 5 μL samples of cells containing either the ARG plasmid or the control plasmid with cells containing the reference pUA66-P_{rpsL}GFP plasmid in 40 μL of fresh DM250. From each mix, 40 μL was mixed with 160 μL of 1× phosphate buffered saline and assayed by flow cytometry to determine the frequency of each competitor at the start of each competition. The remaining 10 μL was added to 190 μL of DM250 and incubated to allow competition between reference plasmid-containing cells and control or ARG plasmid-containing cells. At the end of competitions a sample was again analyzed by flow cytometry to determine the frequency of each competitor.

Cell mixtures were diluted in 1X phosphate buffered saline (PBS) and sampled by BD FACSCanto™ II with a high throughput sampler platform. The sampling rate was controlled around 1,000~3,000 events per second and 10,000~30,000 events were collected for each sample. The obtained data was analysed by the R package, flowCore. To reduce the noises of the flow events, the events were filtered on the FSH-H and SSC-H channels and only events

fit to a bivariate normal distribution with one standard deviation were passed to examine the GFP-H channel (Figure 2.3A). An arbitrary value that distinguished GFP positive and GFP negative cells was set to estimate the frequencies of the GFP-positive and GFP-negative events in the samples (Figure 2.3B). The application of the bivariant gating reduced the GFP-negative events of GFP-expressing cells by 60% (Figure 2.3C).

The fitness effect of an ARG (W_{ARG}) was estimated as: $W_{ARG} = W_{P\text{-}ARG} / W_{P\text{-}Vec}$ [234]. Where $W_{P\text{-}ARG} = \ln(ARG_1 * 100 / ARG_0) / \ln(GFP_1 * 100 / GFP_0)$. $W_{P\text{-}Vec} = \ln(Vec_1 / Vec_0) / \ln(GFP_1 / GFP_0)$. ARG, Vec, and GFP indicate the frequency of cells containing ARG-encoding, control, and reference plasmids, respectively, subscripts indicate the number of days of competition at which those frequencies were estimated, and the 100 accounts for the dilution factor between the beginning and end of competitions.

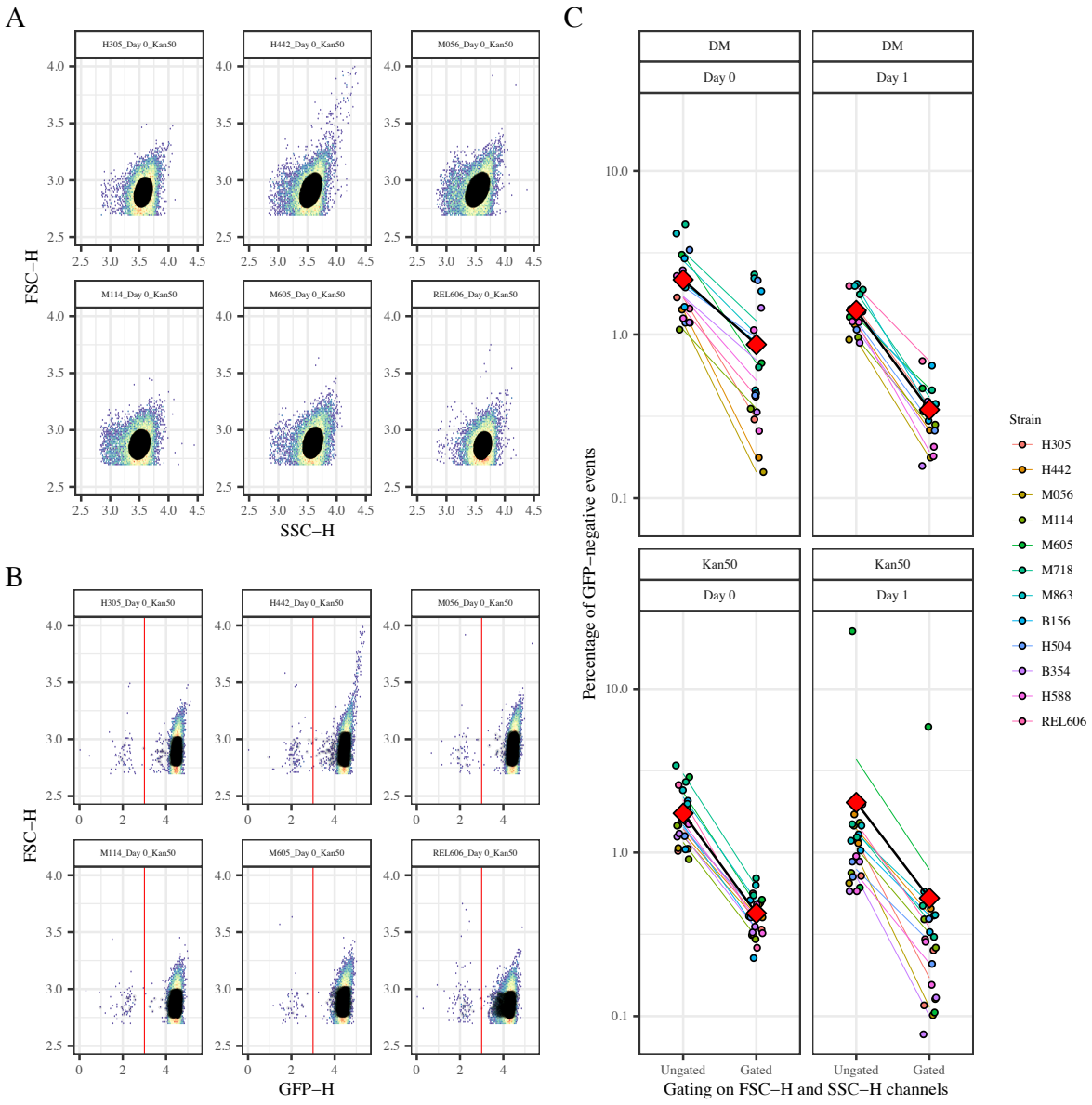

Figure 2.3. Application of a post-sampling filter on the flow cytometry events reduced the sampling noise.

To reduce the sampling noises from flow cytometry, a bivariant filter was applied in the post-sampling analysis. (**A**) The flow events were filtered by a bivariant gating (mean ± one standard deviation) on the FSC-H and SSC-H channels. The events passed the gating (black dots) were then used to analyzes the GFP signals. The shown samples were cells carrying the pUA66-PrpsLGFP plasmid. (**B**) The gated and ungated events from (**A**) were plotted against the FSC-H and GFP-H channels. The gated events (black dots) were enriched in the population with strong GFP signal. An arbitrary threshold (red line) on the GFP channel was used to distinguish GFP-positive and GFP-negative events. (**C**) Application of the bivariant filter reduced the GFP-negative events in the gated samples. The cells carrying the pUA66-PrpsLGFP plasmid were pre-conditioned in DM or DM mdium with 50mg/ml kanamycin for two 24hours growth cycles and subjected to the same conditions as in the competition fitness assays. Day 0 and Day 1 were the sampling times of the competition assay. Only one experimental block was shown here. Dots were colored as the strains (n=2) and lines represents the mean percentage of each strain. Red diamonds were the mean percentage of all the strains.

In the antibiotic-free assay environment it is possible that plasmid missegregation could influence fitness estimates through production of relatively higher fitness plasmid-free sub-populations. To estimate the extent of plasmid loss occurring in the competition environment I started monocultures of each strain used in this study containing the reference pUA66-$_{PrpsL}$GFP plasmid and estimated changes in the fraction of GFP- events, indicating cells that lost the plasmid, during growth. Following the same pre-conditioning protocol used in the competition assays I found that GFP- events increased by only a small amount over the competition period (Figure 2.4, mean increase in GFP- event frequency was 0.25% (95% CI: 0.05-0.44) (comparison of nested ANOVA with and without day term: $\chi^2 = 8.12$, P = 0.015). An increase of 0.25% of GFP-negative events would cause me to underestimate the fitness cost of vector and ARG plasmids by 0.2% relative to the reference plasmid. If the ARG effect is neutral, this underestimate cancels out. Even if the ARG imposes a true cost of 10%, comparable to the largest cost I see in this study, the difference between actual and estimated

costs is only ~0.1%. This consideration does not account for the higher fitness of plasmid-free missegregants, but this effect is expected to be negligible [235].

To further confirm the low frequency of plasmid-losing cells over the competition assays, I plated out the monocultures of cells carrying the *bla*$_{TEM\text{-}116}$* plasmid and the pUA66-*rpsLGFP* plasmid in the absence of positive selection for the plasmid (red dots and lines in Figure 2.5). Cell cultures were diluted and plated on LB and LB with 50mg/ml kanamycin to estimate the plasmid frequency in a population. After three daily transfer in DM medium, which corresponds to the time point Day 0 in the competition assays, the plasmid free cells remained low (*bla*$_{TEM\text{-}116}$* plasmid loss in H588: 102±6.9%; *bla*$_{TEM\text{-}116}$* plasmid loss in M114: 96.8±5.9%; *bla*$_{TEM\text{-}116}$* plasmid loss in REL606: 88.7±10.3%; pUA66-*rpsLGFP* plasmid loss in REL606: 99.5±0.03% (mean±standard deviation, n=4)). In summary, the influence of

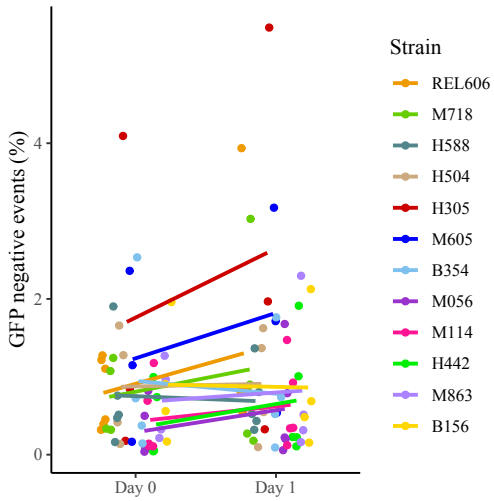

Figure 2.4. Plasmid loss in each host population.

The loss of the GFP plasmid (pUA66-rpslGFP) in antibiotic-free medium after 24 hrs of growth as determined by flow cytometry. The GFP-negative events slightly increased 0.25% (95% CI: 0.05-0.44) after one day of growth but not to an extent that significantly effects fitness estimates (M&M for details). Dots indicate independent replicate estimates and lines connect means of those at the two time points.

plasmid-losing cells over the one-day competition assay has little influence on our estimates of ARG fitness effects.

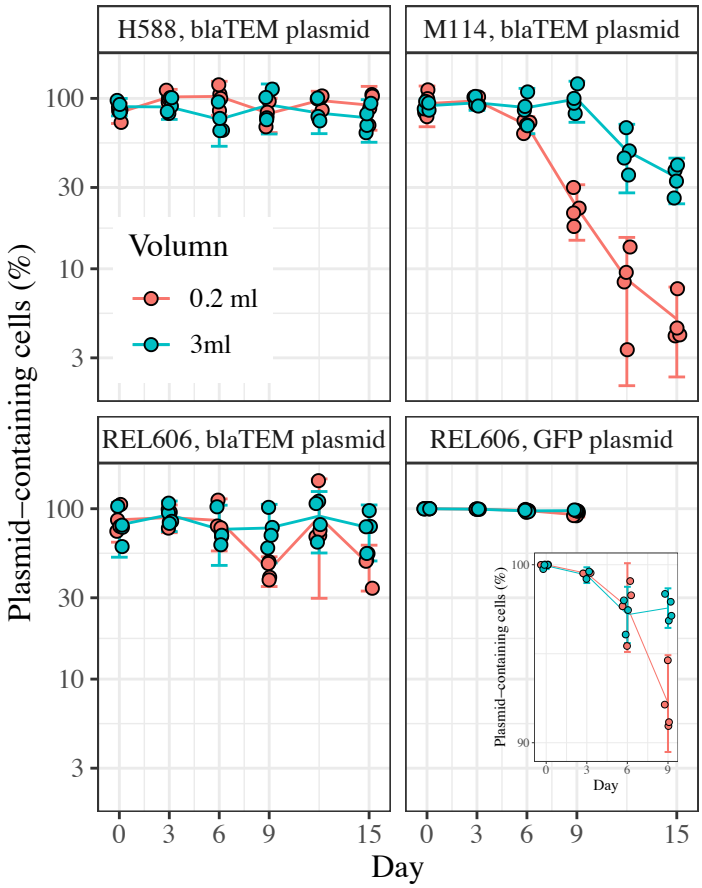

Figure 2.5. Colony formation assays showed that the frequencies of plasmid-containing cells decreased over time in the absence of positive selection for the plasmid and were influenced by the volumes of the cell cultures.

Monocultures of cells carrying the *bla*$_{TEM\text{-}116}$* plasmid or the pUA66-*rpsLGFP* plasmid were cultured in DM medium. Cells were periodically plated out on LB (lysogeny broth) agar plates and LB plates with 50 μg/ml kanamycin to estimate the plasmid frequency in the population (average colony counts on LB plates: 189, minimal counts: 57; maximal count: 555). The monocultures were set up two different volumes: 200μl (red dots and red lines) and 3ml (green dots and green lines). Four independent replicates were carried out.

2.3.3. Determination of the relative copy number of the bla$_{TEM-116}$ plasmid.*

The copy number of plasmids encoding the *bla$_{TEM-116}$** ARG was measured by the comparative Ct (ΔΔCt) method of quantitative polymerase chain reaction (qPCR**)**. Plasmid carrying cells were grown overnight in DM250 supplemented with 50 μg/ml kanamycin. DNA of the overnight cultures was isolated using Wizard® Genomic DNA Purification Kit (Promega) and amplified using a SYBR Green based qPCR mixture consisting of Q5 High-Fidelity 2X Master Mix (NEB), SYBR Green (final concentration: 0.5×), and oligo mix (final concentration of 0.25 μM). One microliter of DNA (1-2 ng/ul) was used per 10 μl reaction. The Q-PCR was performed using a Thermo Scientific PikoReal Real-Time PCR System with default settings. Oligos were designed using PrimerQuest (Integrated DNA Technologies, Inc.) to amplify the chromosomal *dxs* gene, in order to determine the absolute Ct value of the bacterial chromosome (Ct_{ch}) (dxs_qpcr_F1:cgagaaactggcgatcctta; dxs_qpcr_R1: cttcatcaagcggtttcaca [236]) and the pUA66 plasmid encoded *aph(3')-IIa* gene to determine the Ct value of the plasmid (Ct_p) (pUA_Kan_qpcr_F1: ctcgtcaagaaggcgatagaag; pUA_Kan_qpcr_R1:cgttggctacccgtgatatt). The relative copy number of the *bla$_{TEM-116}$** plasmid to the chromosome (ΔCt) was calculated as $2^{-(Ct_p - Ct_{ch})}$. The ΔCt of different bacterial strains was then normalized to the lab strain, REL606, to obtain the ΔΔCt value.

2.3.4. Evolutionary dynamics of the blaTEM-116 plasmid*

The plasmid-carrying cells were cultured in DM with 50μg/ml kanamycin for overnight. The overnight cultured were then 1:100 diluted and cultured in 200μl DM medium. The cultured were then 1:100 transferred into fresh DM medium every ~24 hours. After 61 transfers, the cultures were grown in DM with 50μg/ml kanamycin for two growth cycles to select plasmid-

carrying cells. The kanamycin-treated evolved populations were then cultured in 3ml DM medium to track the plasmid loss. To count the plasmid loss in a population, cell cultures were plated on LB agar plates and LB with 50μg/ml kanamycin to count the plasmid-carrying and plasmid-free colonies. I noted that the volume of cell culture can influence the plasmid dynamics as seen in Figure 2.8B (200μl) and Figure 2.8C (3ml) where the plasmid loses in the ancestral cells in two volumes (200μl vs. 3ml) differed. To further confirm the difference was not an artefact, I sat up another experiment to verify the plasmid loss in two culture volumes (Figure 2.5). The loss of the $bla_{TEM\text{-}116}$* plasmid in M114 was faster in 200μl than in 3ml culture. An estimation of plasmid loss by simultaneously plating cells on selective and non-selective plates may not be able to distinguish small differences in plasmid loss. Yet, counting the loss of the pUA-$_{PrpsL}$GFP plasmid in REL606 by counting the GFP-positive and GFP-negative colonies grown on LB plates showed that cells grown in 200μl cultures had a higher plasmid loss rate than 3ml cultures (Figure 2.5).

2.3.5. Phylogeny construction and phylogenetic signal testing.

Core and accessory genomes of a collection of 96 environmental *E. coli* isolates were used to build phylogenies as described previously [1]. These phylogenies were used to test for a phylogenetic signal in the effect of ARG effects using Pagel's lambda and Blomberg's K metrics as implemented in the phylosig function of the phytools package v1.0-3 [237].

2.3.6. Statistical analysis.

Statistical analyses were performed using R 4.1.0 [238]. Flow cytometry data was analyzed using the flowCore package and mixed-model ANOVA were performed using the packages lme4 and lmerTest [239-241].

2.4. Results

2.4.1. ARGs can impose costs and benefits in the absence of antibiotic selection.

I cloned six ARGs that belonged to four distinct mechanistic classes separately into a low-copy number plasmid (Table 2.3). The effect of each ARG on the fitness of each of 12 divergent

Table 2-3. Characteristics of ARGs used in this study.

Gene	Enzyme type	Conferring resistance to:	Mechanism
aadA	Aminoglycoside adenyltransferase	Aminoglycosides such as spectinomycin and streptomycin	Drug-modifying enzymes
bla $_{TEM-116*}$	Beta lactamase	β-lactam such as ampicillin	Drug-modifying enzymes
bla $_{SHV12}$	Beta lactamase	β-lactam such as ampicillin and cefotaxime	Drug-modifying enzymes
bla $_{CTX-M-15}$	Beta lactamase	β-lactam such as ampicillin and cefotaxime	Drug-modifying enzymes
cat	Chloramphenicol acetyl trasferase	Chloramphenicol	Drug-modifying enzymes
dfrA5	Dihydrofolate reductase	Trimethoprim	Target replacement

Escherichia strains was measured in an antibiotic free environment. I found significant differences in the mean effect of the different ARGs (χ^2=28.7, P<0.0001; Figure 2.6A). Of the six tested ARGs, four conferred an overall fitness cost measured across all host strains: *aadA* = 1.2% (95% CI: 1.6-2,1%, Dunnett's test P = 0.014), *dfrA5* = 1.2% (95% CI: 0.2-2.2%, Dunnett's test P = 0.008), *cat* = 1.7% (95% CI: 0.7-2.7%, Dunnett's test P <0.001) and *bla*$_{TEM1-116}$* = 2.5% (95% CI: 1.6-3.3%, Dunnett's test P <0.001) (Table 2.4). Together these results

Table 2-4. Average fitness cost of the ARGs in Escherichia isolates.

ARG	Fitness effect	95% CI	pvalue*	Environment
aadA	**-0.012**	**-0.021; -0.002**	**0.014**	**DM**
bla CTX-M-15	-0.004	-0.014; 0.006	0.8	DM
bla SHV12	-0.006	-0.016; 0.003	0.308	DM
bla TEM-116*	**-0.025**	**-0.033; -0.016**	**<0.001**	**DM**
cat	**-0.017**	**-0.027; -0.007**	**<0.001**	**DM**
dfrA5	**-0.012**	**-0.022; -0.002**	**0.008**	**DM**
aadA	**-0.021**	**-0.038; -0.005**	**0.004**	**Kan50**
bla CTX-M-15	-0.006	-0.023; 0.011	0.874	Kan50
bla SHV12	-0.006	-0.023; 0.01	0.844	Kan50
bla TEM-116*	**-0.102**	**-0.118; -0.085**	**<0.001**	**Kan50**
cat	**-0.026**	**-0.043; -0.01**	**<0.001**	**Kan50**
dfrA5	**-0.023**	**-0.04; -0.006**	**0.003**	**Kan50**

* Dunnett's test; significance at 0.05 is in bold

clearly indicate that ARGs can confer significant effects on host fitness in the absence of direct selection for their resistance phenotypes. Moreover, the three β-lactamases had significantly different effects on host fitness, suggesting that effects might depend strongly on subtle differences between ARGs ($\chi 2 = 26.9$, $P<0.001$).

2.4.2. Fitness effects of ARGs depend on the host strain.

The results presented above consider the mean effect of each ARG across 12 host strains. To determine if effects differed between host strains, I extended the analysis of variance to include an ARG effect-by-strain interaction term. This extended analysis provided a significantly improved fit (effect of strain-ARG interaction term: $\chi^2= 140.6$, $p<0.0001$). All ARGs imposed a significant fitness effect in at least one strain (*aadA* - 2 strains; $bla_{TEM\text{-}1116*}$ - 4 strains; *cat* - 2 strains; $bla_{ctx}15$ - 2 strains, *dfr*A5 - 2 strains; $bla_{shv}1\text{-}12$ - 2 strains) (Table 2.5). In all but two cases these effects were negative. Of the 12 tested strains, all except H305 and M605 had a significant change in fitness caused by at least one ARG. Mean fitness difference between different ARG-Strain combinations was 2.1% (95% CI: 1.4%-2.6%), which is substantially

larger than mean differences in ARG effects (0.74%; 95% CI: 0%-1.9%) and thus indicates the potential for some strains to act as reservoirs for ARGs in antibiotic-free environments. Indeed, although all ARGs were costly in at least one strain, all were also neutral or beneficial in at least one other host strain.

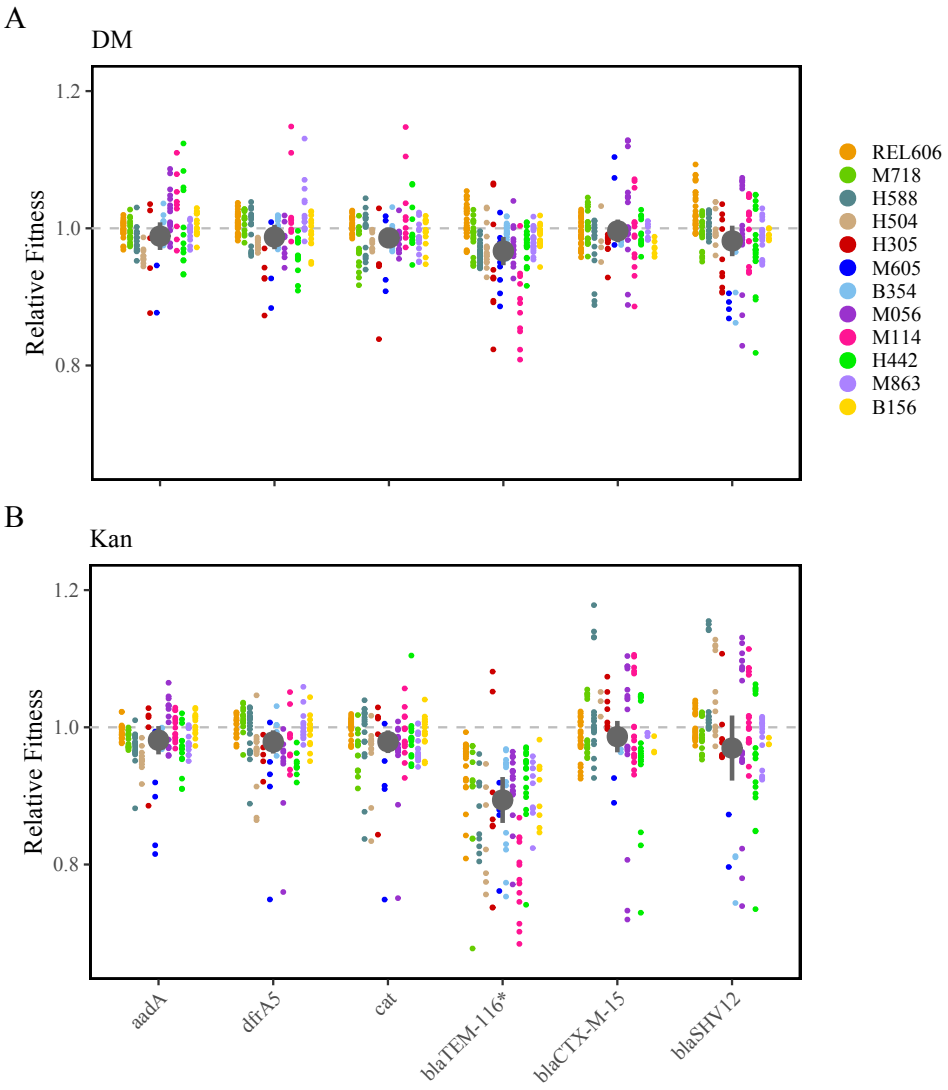

Figure 2.6. Fitness costs of the ARGs.

The fitness costs of the ARGs in the 12 strains was measured in DM **(A)** and DM+kanamycin **(B)** media. Competitive fitness of the ARG plasmid-carrying cells (or the control plasmid) was measured by their fitness in competition to pUA66-rpslGFP-carrying cells. The fitness cost of an ARG is reported as the fitness of the competitive fitness of ARG plasmid-carrying cells relative to that of the control plasmid. The mean cost of each ARG was averaged from the median cost measured in each *Escherichia* strain (gray large dot and gray solid line (95% CI)). The 12 strains are indicated using colored symbols with each symbol representing an independent measurement (n=2-32).

Table 2-5. Fitness cost of the ARGs in Escherichia strains.

Strain	ARG	Fitness effect	95%CI	pvalue*	Environment
B156	*aadA*	0.004	-0.013; 0.022	0.974	DM
B354	*aadA*	-0.01	-0.028; 0.008	0.561	DM
H305	*aadA*	-0.016	-0.101; 0.069	0.993	DM
H442	*aadA*	-0.013	-0.049; 0.022	0.838	DM
H504	***aadA***	**-0.034**	**-0.055; -0.012**	**<0.001**	**DM**
H588	*aadA*	-0.022	-0.043; -0.001	0.031	DM
M056	*aadA*	0	-0.036; 0.036	1	DM
M114	*aadA*	-0.001	-0.045; 0.042	1	DM
M605	*aadA*	-0.099	-0.219; 0.021	0.144	DM
M718	*aadA*	-0.004	-0.025; 0.016	0.985	DM
M863	*aadA*	-0.008	-0.029; 0.014	0.881	DM
REL606	*aadA*	-0.003	-0.016; 0.011	0.993	DM
B156	***bla CTX-M-15***	**-0.025**	**-0.049; -0.001**	**0.041**	**DM**
B354	*bla CTX-M-15*	-0.003	-0.028; 0.022	0.999	DM
H305	*bla CTX-M-15*	-0.021	-0.092; 0.049	0.927	DM
H442	*bla CTX-M-15*	-0.006	-0.041; 0.029	0.997	DM
H504	*bla CTX-M-15*	-0.026	-0.056; 0.004	0.116	DM
H588	***bla CTX-M-15***	**-0.024**	**-0.044; -0.003**	**0.015**	**DM**
M056	*bla CTX-M-15*	0.012	-0.02; 0.044	0.84	DM
M114	*bla CTX-M-15*	-0.012	-0.049; 0.025	0.907	DM
M605	*bla CTX-M-15*	0.061	-0.043; 0.164	0.455	DM
M718	*bla CTX-M-15*	0.002	-0.016; 0.02	1	DM
M863	*bla CTX-M-15*	0.005	-0.025; 0.034	0.998	DM
REL606	*bla CTX-M-15*	-0.008	-0.02; 0.005	0.386	DM
B156	*bla SHV12*	-0.006	-0.03; 0.018	0.982	DM
B354	***bla SHV12***	**-0.044**	**-0.069; -0.019**	**<0.001**	**DM**
H305	*bla SHV12*	-0.037	-0.102; 0.029	0.501	DM
H442	*bla SHV12*	-0.03	-0.062; 0.001	0.067	DM
H504	*bla SHV12*	-0.003	-0.024; 0.017	0.996	DM
H588	*bla SHV12*	0.004	-0.018; 0.025	0.996	DM
M056	*bla SHV12*	-0.006	-0.039; 0.027	0.996	DM
M114	*bla SHV12*	-0.003	-0.039; 0.034	1	DM
M605	*bla SHV12*	-0.082	-0.172; 0.008	0.085	DM
M718	*bla SHV12*	-0.007	-0.025; 0.011	0.847	DM
M863	*bla SHV12*	-0.014	-0.033; 0.004	0.204	DM
REL606	***bla SHV12***	**0.031**	**0.018; 0.044**	**<0.001**	**DM**
B156	*bla TEM-116**	-0.013	-0.028; 0.003	0.154	DM
B354	*bla TEM-116**	-0.005	-0.021; 0.011	0.946	DM
H305	*bla TEM-116**	-0.051	-0.117; 0.014	0.18	DM
H442	*bla TEM-116**	-0.027	-0.059; 0.005	0.133	DM
H504	***bla TEM-116****	**-0.029**	**-0.046; -0.011**	**<0.001**	**DM**
H588	***bla TEM-116****	**-0.038**	**-0.057; -0.019**	**<0.001**	**DM**
M056	*bla TEM-116**	-0.031	-0.065; 0.002	0.081	DM
M114	***bla TEM-116****	**-0.112**	**-0.151; -0.073**	**<0.001**	**DM**
M605	*bla TEM-116**	-0.05	-0.123; 0.022	0.283	DM
M718	*bla TEM-116**	-0.013	-0.031; 0.005	0.257	DM
M863	***bla TEM-116****	**-0.019**	**-0.037; 0**	**0.048**	**DM**
REL606	*bla TEM-116**	0.009	-0.003; 0.022	0.243	DM
B156	*cat*	-0.008	-0.025; 0.009	0.721	DM
B354	*cat*	-0.015	-0.034; 0.003	0.14	DM
H305	*cat*	-0.049	-0.14; 0.042	0.543	DM
H442	*cat*	-0.012	-0.048; 0.023	0.872	DM
H504	*cat*	-0.013	-0.034; 0.008	0.432	DM
H588	*cat*	-0.012	-0.033; 0.009	0.467	DM
M056	***cat***	**-0.04**	**-0.077; -0.002**	**0.032**	**DM**
M114	*cat*	-0.005	-0.048; 0.038	1	DM
M605	*cat*	-0.045	-0.14; 0.051	0.678	DM
M718	***cat***	**-0.034**	**-0.054; -0.014**	**<0.001**	**DM**
M863	*cat*	-0.018	-0.039; 0.003	0.127	DM
REL606	*cat*	0.006	-0.007; 0.02	0.688	DM
B156	*dfrA5*	-0.008	-0.025; 0.009	0.711	DM
B354	*dfrA5*	-0.008	-0.026; 0.011	0.786	DM
H305	*dfrA5*	-0.061	-0.146; 0.024	0.255	DM
H442	***dfrA5***	**-0.055**	**-0.092; -0.019**	**0.001**	**DM**
H504	*dfrA5*	-0.017	-0.038; 0.005	0.197	DM
H588	*dfrA5*	-0.005	-0.026; 0.016	0.972	DM
M056	*dfrA5*	-0.042	-0.084; 0	0.053	DM
M114	*dfrA5*	-0.003	-0.046; 0.04	1	DM
M605	*dfrA5*	-0.059	-0.155; 0.037	0.394	DM
M718	*dfrA5*	0.009	-0.011; 0.03	0.708	DM
M863	***dfrA5***	**0.025**	**0.003; 0.046**	**0.015**	**DM**
REL606	*dfrA5*	0.007	-0.006; 0.021	0.528	DM

Table 2.5. Fitness cost of the ARGs in Escherichia strains. (continued)

Strain	*ARG*	Fitness effect	95%CI	pvalue	Environment
B156	*aadA*	0.006	-0.02; 0.031	0.987	Kan50
B354	*aadA*	-0.011	-0.046; 0.025	0.945	Kan50
H305	*aadA*	0.013	-0.1; 0.125	1	Kan50
H442	*aadA*	-0.046	-0.101; 0.01	0.145	Kan50
H504	*aadA*	-0.041	-0.088; 0.006	0.112	Kan50
H588	*aadA*	-0.035	-0.076; 0.006	0.131	Kan50
M056	*aadA*	-0.015	-0.086; 0.057	0.988	Kan50
M114	*aadA*	-0.003	-0.047; 0.04	1	Kan50
M605	*aadA*	-0.086	-0.336; 0.163	0.864	Kan50
M718	*aadA*	-0.017	-0.053; 0.019	0.668	Kan50
M863	*aadA*	-0.021	-0.047; 0.004	0.14	Kan50
REL606	*aadA*	-0.002	-0.022; 0.018	1	Kan50
B156	*bla* CTX-M-15	-0.033	-0.083; 0.016	0.336	Kan50
B354	*bla* CTX-M-15	-0.048	-0.115; 0.018	0.258	Kan50
H305	*bla* CTX-M-15	0.053	-0.043; 0.149	0.51	Kan50
H442	*bla* CTX-M-15	-0.044	-0.096; 0.009	0.138	Kan50
H504	*bla* CTX-M-15	0.018	-0.063; 0.099	0.984	Kan50
H588	*bla* CTX-M-15	0.021	-0.02; 0.061	0.6	Kan50
M056	*bla* CTX-M-15	-0.025	-0.092; 0.042	0.834	Kan50
M114	*bla* CTX-M-15	0.004	-0.037; 0.045	1	Kan50
M605	*bla* CTX-M-15	-0.137	-0.494; 0.22	0.807	Kan50
M718	*bla* CTX-M-15	-0.007	-0.04; 0.026	0.984	Kan50
M863	*bla* CTX-M-15	-0.01	-0.058; 0.037	0.987	Kan50
REL606	*bla* CTX-M-15	-0.018	-0.037; 0.001	0.079	Kan50
B156	*bla* SHV12	-0.027	-0.076; 0.023	0.578	Kan50
B354	***bla* SHV12**	**-0.23**	**-0.297; -0.163**	**<0.001**	**Kan50**
H305	*bla* SHV12	-0.012	-0.108; 0.084	0.999	Kan50
H442	*bla* SHV12	-0.05	-0.103; 0.002	0.066	Kan50
H504	***bla* SHV12**	**0.047**	**0.001; 0.093**	**0.045**	**Kan50**
H588	*bla* SHV12	0.037	-0.006; 0.081	0.121	Kan50
M056	*bla* SHV12	-0.006	-0.073; 0.061	1	Kan50
M114	*bla* SHV12	0.015	-0.027; 0.056	0.875	Kan50
M605	*bla* SHV12	-0.21	-0.568; 0.147	0.428	Kan50
M718	*bla* SHV12	-0.012	-0.045; 0.021	0.854	Kan50
M863	*bla* SHV12	-0.02	-0.044; 0.004	0.154	Kan50
REL606	*bla* SHV12	0.004	-0.015; 0.023	0.994	Kan50
B156	***bla* TEM-116***	**-0.096**	**-0.127; -0.065**	**<0.001**	**Kan50**
B354	***bla* TEM-116***	**-0.105**	**-0.14; -0.07**	**<0.001**	**Kan50**
H305	***bla* TEM-116***	**-0.114**	**-0.21; -0.018**	**0.015**	**Kan50**
H442	***bla* TEM-116***	**-0.079**	**-0.131; -0.026**	**0.001**	**Kan50**
H504	***bla* TEM-116***	**-0.164**	**-0.219; -0.109**	**<0.001**	**Kan50**
H588	***bla* TEM-116***	**-0.115**	**-0.163; -0.068**	**<0.001**	**Kan50**
M056	***bla* TEM-116***	**-0.094**	**-0.161; -0.027**	**0.002**	**Kan50**
M114	***bla* TEM-116***	**-0.22**	**-0.264; -0.176**	**<0.001**	**Kan50**
M605	*bla* TEM-116*	-0.037	-0.281; 0.208	0.997	Kan50
M718	***bla* TEM-116***	**-0.1**	**-0.137; -0.064**	**<0.001**	**Kan50**
M863	***bla* TEM-116***	**-0.09**	**-0.118; -0.062**	**<0.001**	**Kan50**
REL606	***bla* TEM-116***	**-0.06**	**-0.079; -0.041**	**<0.001**	**Kan50**
B156	*cat*	-0.005	-0.03; 0.021	0.995	Kan50
B354	*cat*	-0.015	-0.051; 0.02	0.775	Kan50
H305	*cat*	0.001	-0.111; 0.114	1	Kan50
H442	*cat*	-0.027	-0.082; 0.029	0.648	Kan50
H504	*cat*	-0.033	-0.08; 0.013	0.263	Kan50
H588	*cat*	-0.019	-0.06; 0.022	0.682	Kan50
M056	*cat*	-0.069	-0.14; 0.003	0.064	Kan50
M114	*cat*	-0.014	-0.058; 0.031	0.931	Kan50
M605	*cat*	-0.072	-0.321; 0.178	0.935	Kan50
M718	***cat***	**-0.04**	**-0.075; -0.004**	**0.021**	**Kan50**
M863	***cat***	**-0.026**	**-0.052; -0.001**	**0.039**	**Kan50**
REL606	*cat*	0.003	-0.017; 0.023	0.998	Kan50
B156	*dfrA5*	-0.004	-0.03; 0.021	0.997	Kan50
B354	*dfrA5*	-0.008	-0.043; 0.028	0.988	Kan50
H305	*dfrA5*	-0.018	-0.131; 0.094	0.996	Kan50
H442	***dfrA5***	**-0.068**	**-0.131; -0.006**	**0.027**	**Kan50**
H504	*dfrA5*	-0.035	-0.081; 0.01	0.198	Kan50
H588	*dfrA5*	-0.014	-0.055; 0.027	0.905	Kan50
M056	***dfrA5***	**-0.095**	**-0.176; -0.013**	**0.015**	**Kan50**
M114	*dfrA5*	-0.021	-0.071; 0.029	0.774	Kan50
M605	*dfrA5*	-0.067	-0.317; 0.182	0.95	Kan50
M718	*dfrA5*	0.012	-0.023; 0.047	0.881	Kan50
M863	*dfrA5*	0.002	-0.023; 0.028	1	Kan50
REL606	*dfrA5*	0.004	-0.018; 0.025	0.996	Kan50

* Dunnett's test; significance at 0.05 is in bold
Positive fitness effect is highlighted in red.

2.4.3. Fitness effects of ARGs depend on the environment.

Fitness effects reported above were estimated in an environment not containing any antibiotic. However, the focal ARGs were cloned into a vector that included a backbone ARG (*aph3'-II*, encoding resistance to kanamycin). This in-common non-focal ARG allowed us to test if the presence of an antibiotic a cell is resistant to might nevertheless affect physiology in a way that alters the effect of additional non-selected ARGs. I note that kanamycin is an aminoglycoside antibiotic with similarity to the *aadA* substrates spectinomycin and streptomycin, but there is no detectable cross-resistance [242]. When I measured the fitness effects of each ARG-strain combination in the same environment as used previously but supplemented with kanamycin, I again found a significant interaction between ARG and host strain in determining fitness (χ^2 = 74.4, P < 0.001, Figure 2.6B and Table 2.5). Again, most ARGs affected host fitness in at least one strain (all except *aadA* and $bla_{CTX\text{-}M\text{-}15}$) and most host strains suffered a fitness decrease in the presence of at least one ARG (all except M605). Nevertheless, there was a significant effect of environment on fitness (χ^2 = 88.8, P < 0.001). Evidently, a general pattern of ARG fitness effects depending on the strain background is common across environments, although the exact nature of that dependence is different. I note that plasmid loss was rare in the antibiotic free environment and cannot explain observed fitness differences (Materials and Methods).

2.4.4. Fitness effects of ARGs are not explained by plasmid copy number.

A possible explanation for different effects of ARGs across host strains is that the host pmFP plasmid has different copy number in different strains. Different copy number can affect expression levels of plasmid-borne genes and, thus, the phenotypic effect of gene products. One well known example is the effect of bla_{TEM1} copy number on degree of ampicillin

resistance [243]. To test for a link between plasmid copy number and ARG fitness effect I used quantitative PCR to determine the copy number of the *bla*$_{TEM\text{-}116*}$-encoding plasmid across our set of 12 host strains. This ARG showed the highest variation in fitness cost, so represents a good test of the contribution of copy number differences to this underlying this variation. I found that the relationship between *bla*$_{TEM\text{-}116*}$ copy number and fitness effect was positive, but not significant, in both the minimal medium and kanamycin supplemented environments (Pearson correlations: DM medium, r = 0.13, P = 0.68; kanamycin supplemented medium, r = 0.03, P = 0.9; Figure 2.7). This result indicates that differences in plasmid copy number does not explain differences in the fitness effect of *bla*$_{TEM\text{-}116*}$ across host strains.

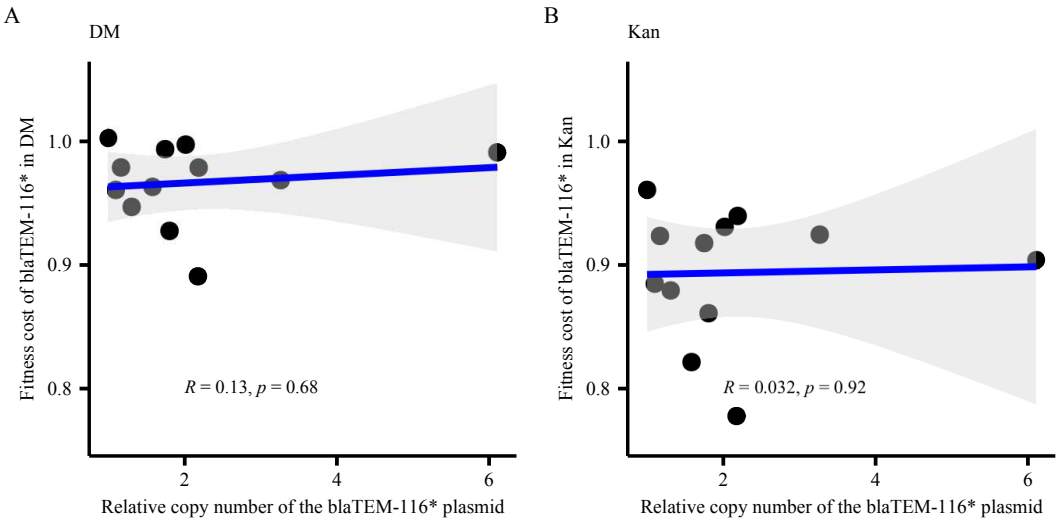

Figure 2.7.Copy number variation of the *bla*$_{TEM\text{-}116}$* plasmid does not explain variation in cost

The copy number of the *bla*$_{TEM\text{-}116}$* plasmid in each strain was measured by quantitative PCR and normalized to that in REL606. The copy numbers were plotted against the fitness cost of *bla*$_{TEM\text{-}116}$* measured in DM **(A)** and kanamycin **(B)** in each strain. Pearson correlation coefficients and the P values were shown in each plot. The linear regression line was plotted as a blue line with gray shading indicating the 95% CI.

Table 2-6. Phylogenetic signal of ARG fitness effects.

Tree	ARG	Pagel's lambda*	Blomberg's K**	Environment
Core gene phylogeny	aadA	0.169 (0.723)	0.159 (0.615)	DM
Core gene phylogeny	dfrA5	0 (1)	0.038 (0.625)	DM
Core gene phylogeny	cat	0 (1)	0.02 (0.87)	DM
Core gene phylogeny	blaTEM116	0.9 (1)	0.14 (0.421)	DM
Core gene phylogeny	blaCTXM15	0 (1)	0.349 (0.192)	DM
Core gene phylogeny	blaSHV12	0 (1)	0.056 (0.611)	DM
Accessory gene phylogeny	aadA	0.503 (0.313)	0.737 (0.194)	DM
Accessory gene phylogeny	dfrA5	0 (1)	0.226 (0.615)	DM
Accessory gene phylogeny	cat	0 (1)	0.155 (0.846)	DM
Accessory gene phylogeny	blaTEM116	0 (1)	0.49 (0.408)	DM
Accessory gene phylogeny	blaCTXM15	0 (1)	0.738 (0.087)	DM
Accessory gene phylogeny	blaSHV12	0 (1)	0.323 (0.538)	DM
Core gene phylogeny	aadA	0 (1)	0.02 (0.856)	Kan50
Core gene phylogeny	dfrA5	0 (1)	0.204 (0.279)	Kan50
Core gene phylogeny	cat	0 (1)	0.04 (0.66)	Kan50
Core gene phylogeny	blaTEM116	0 (1)	0.073 (0.54)	Kan50
Core gene phylogeny	blaCTXM15	0 (1)	0.123 (0.445)	Kan50
Core gene phylogeny	blaSHV12	0 (1)	0.14 (0.576)	Kan50
Accessory gene phylogeny	aadA	0 (1)	0.162 (0.818)	Kan50
Accessory gene phylogeny	dfrA5	0 (1)	0.51 (0.325)	Kan50
Accessory gene phylogeny	cat	0 (1)	0.266 (0.604)	Kan50
Accessory gene phylogeny	blaTEM116	0 (1)	0.398 (0.531)	Kan50
Accessory gene phylogeny	blaCTXM15	0 (1)	0.529 (0.347)	Kan50
Accessory gene phylogeny	blaSHV12	0 (1)	0.46 (0.556)	Kan50

*Pagel's lambda metric is 0 for a star phylogeny (with no phylogenetic signal) and 1 for the actual phylogeny.
**Blomberg's K metric

2.4.5. ARG effects are not predicted by evolutionary relationships between host strains.

To test if the dependence of ARG effect on host strain can be predicted by the genetic relationship of different hosts I tested for a phylogenetic signal in the effect of each ARG. I estimated Pagel's λ, a measure of the phylogenetic signal in a response variable, considering the fitness effect of each ARG across phylogenies derived from both the core and accessory genomes of our host strains [244]. In no case was any significant signal found. An alternative measure of phylogenetic signal, Blomberg's K, which has a distinct null expectation, gave a qualitatively consistent result (Table 2.6). With only a small number of host strains, the power of individual tests is not high. However, the lack of signal given that the strains cover a wide

range of the genetic diversity contained in a larger collection of *E. coli* strains and that tests were repeated for each of our six ARGs, indicate that that fitness effects of ARGs appear to be at best weakly predicted by the overall genetic similarity of host strains.

2.4.6. Variation in fitness costs leads to different ecological and evolutionary dynamics of the $bla_{TEM\text{-}116}$ plasmid*

Variation in the fitness cost of ARGs among bacteria suggests the ecological, and perhaps evolutionary, fate of an ARG may depend on its host. For instance, a large cost can impose strong purifying selection against an ARG's carriage but can also favour the rise of compensatory mutations to reduce the cost [245]. In contrast, an ARG with a small fitness cost may be maintained in populations growing in the absence of antibiotic selection. First, I examined the stability of an ARG-carrying plasmid in different hosts. The $bla_{TEM\text{-}116}$* ARG imposes different degrees of cost in three isolates—0.9% in REL606 (95% CI: ±1.2%), 3.8% in H588 (95% CI: ±1.9%) and 11.1% in M114 (95% CI: ±4.0%) (Figure 2.8A)—and consequently serves as an ideal candidate to investigate the effect of host-dependent cost on ARG maintenance and evolution.

The $bla_{TEM\text{-}116}$* plasmid-carrying cells and control plasmid-carrying cells were cultured and transferred daily to fresh DM medium for 61 days without positive selection for the plasmid. The frequency of the plasmid in all evolving populations was monitored periodically. The non-conjugative pmFP plasmid contains no active segregation systems and is subject to random segregation loss during cell division. Nevertheless, the control plasmid stably persisted in all

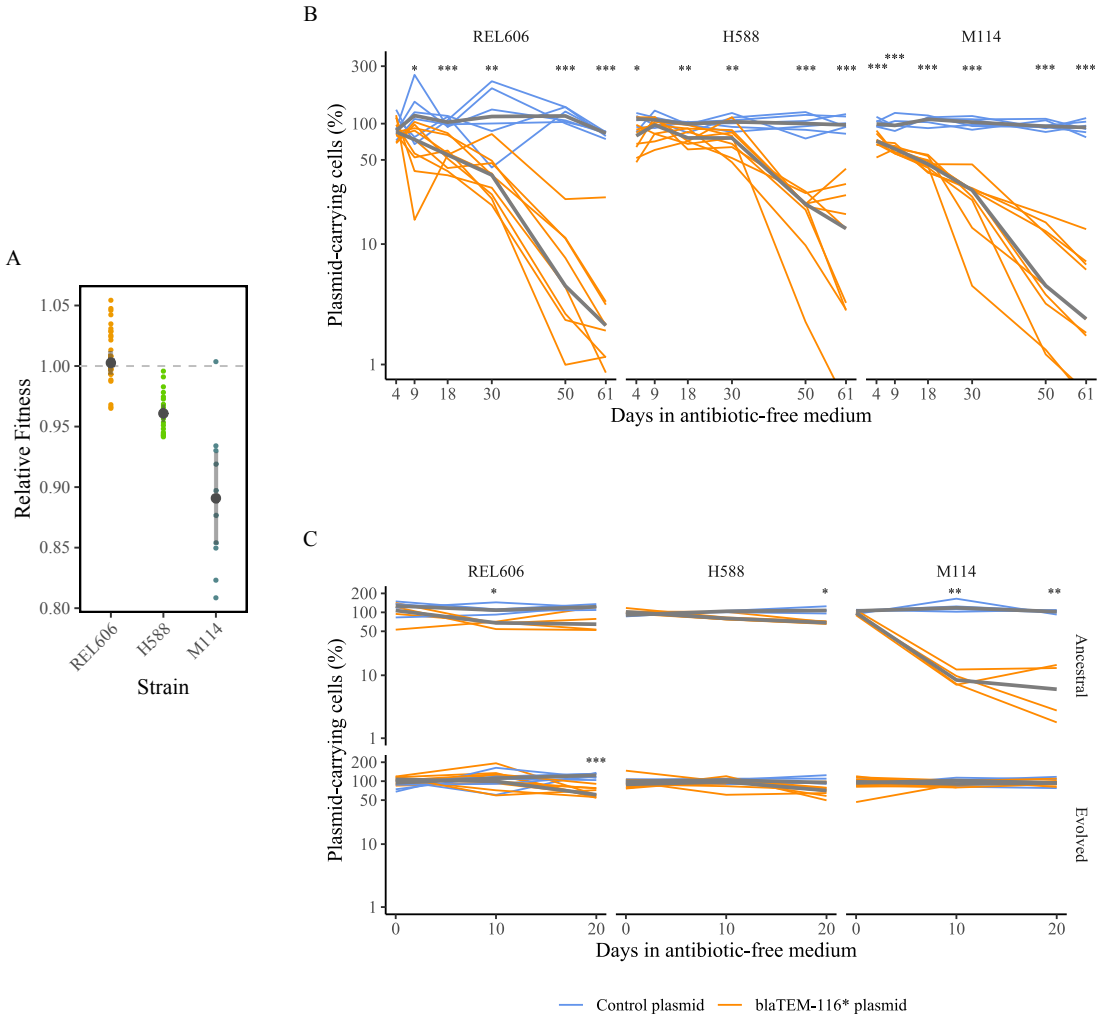

Figure 2.8. The evolutionary dynamics of the *bla*$_{TEM\text{-}116}$* plasmid in different *E. coli* hosts.

(A) Fitness effect of *bla*$_{TEM\text{-}116}$* differ in three strains. REL606: 3.1% (95% CI: 1.8–4.4%); H588: -3.9% (95% CI: -1.9– -5.9%); M114: -11.2% (95% CI: -7.3%– -15.1%). **(B)** Frequency of plasmid-carrying cells in antibiotic-free DM medium over time. Six lines of the control plasmid-carrying cells (blue lines) and nine lines of *bla*$_{TEM\text{-}116}$* plasmid-carrying cells (orange lines) were monitored in each host. The mean loss in each group is indicated by gray lines. The differences in the plasmid loss rates between the control and *bla*$_{TEM\text{-}116}$* populations were tested at each sample time point (t-test: ns $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$). **(C)** Comparison of plasmid dynamics in ancestral and evolved populations. Populations were isolated on Day 61, treated with kanamycin to select plasmid-carrying cells, and then cultured again in antibiotic-free medium for 20 days. Changes in plasmid frequency occurring in these evolved populations (bottom panel) was compared to changes seen in a new set of ancestral populations (top panel) evolved in the same conditions. Experiments in **(B)** were carried out in 200µl medium, and the experiments in **(C)** were done in 3ml medium. The volume of growth cultures influenced the plasmid loss dynamics of the ancestral plasmids in (**B**) and (**C**) and was confirmed in Figure 2.5 and discussed in Materials and Methods.

six independent lines started from each of the three hosts, suggesting the segregational loss rates are low (Figure 2.8B). In contrast, the *bla*$_{TEM\text{-}116}$* plasmid decreased in frequency in lines

started from all three hosts. M114 has the shortest half-life for plasmid maintenance with its frequency in the population declining to 50% by 17.1 days compared to 24.8 days among REL606 lines and 35.7 days among H588 lines (Figure 2.8B). The short half-life of the *bla*$_{TEM\text{-}116}$* plasmid in the M114 population is consistent with its large cost.

Different ecological dynamics across hosts can impose different selection for mutations that compensate for plasmid costs, leading to changes in subsequent dynamics. To test this possibility, I treated the evolved populations with kanamycin to select plasmid-carrying cells and then followed the plasmid loss over a second period of propagation in the absence of antibiotic selection. Compared to a paired set of propagations started with the ancestral plasmid-host combination, the loss of the evolved plasmid was greatly reduced in the evolved M114 cells suggesting the plasmid-carrying population had evolved to compensate for costs of the *bla*$_{TEM\text{-}116}$* plasmid (Figure 2.8C top vs. bottom panel). By contrast, plasmid dynamics in paired H588 and REL606 ancestral and evolved lines did not change significantly, suggesting no mutations occurred to change the dynamics (Figure 2.8C top vs. bottom panel). Together, these results indicate that the host-dependent cost of *bla*$_{TEM\text{-}116}$* can lead to different evolutionary fates of a *bla*$_{TEM\text{-}116}$*-carrying plasmid.

To examine if the slowed *bla*$_{TEM\text{-}116}$* plasmid loss in M114 populations is due to compensation to the cost of the ARG, I compared the fitness between plasmid-carrying and plasmid-cured cells (Figure 2.9A). The result showed no significance fitness difference between plasmid-carrying and plasmid-cured clones isolated from lines evolved for 61 days, indicating that the evolved *bla*$_{TEM\text{-}116}$* plasmid no longer imposes a cost on the evolved clone. Compensatory mutations could occur in the bacterial genome or on the *bla*$_{TEM\text{-}116}$* plasmid. To examine the location of compensatory mutations, I introduced the ancestral plasmids into the cured clones (Figure 2.9B). All nine evolved clones bear no fitness cost of *bla*$_{TEM\text{-}116}$*, indicating that compensatory mutations reside in the bacterial chromosome. In summary, the host-dependent cost of *bla*$_{TEM\text{-}116}$* can lead to different plasmid stability in a bacterial population and incur different evolutionary outcomes of the *bla*$_{TEM\text{-}116}$* plasmid.

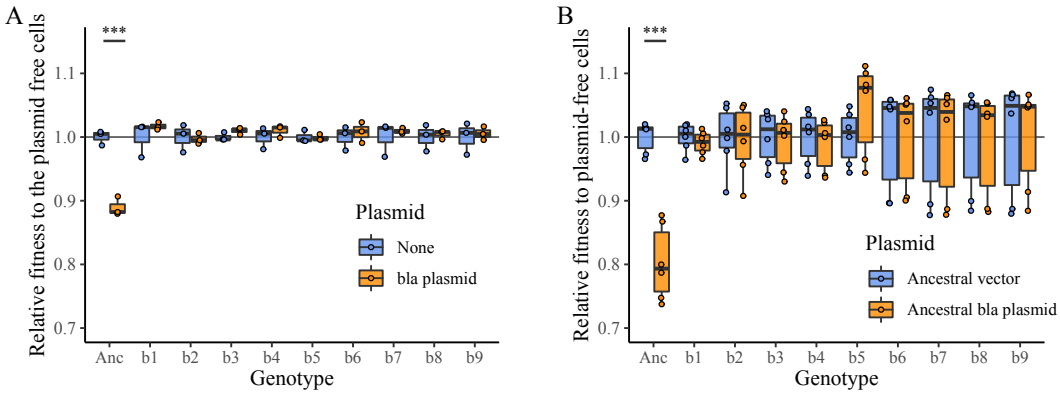

Figure 2.9. Compensatory evolution changes the fitness cost of *bla*$_{TEM\text{-}116}$* in the strain M114.

(A) The fitness cost of the *bla*$_{TEM\text{-}116}$* plasmid was alleviated in the evolved cells. To estimate the fitness cost of the *bla*$_{TEM\text{-}116}$* plasmid in the evolved cells, we passively selected cells that naturally lost the plasmid. One clone from each line of the evolved *bla*$_{TEM\text{-}116}$* plasmid populations (b1-b9) was cultured in antibiotic-free medium for 3-4 rounds of 24 hour culture and plated on non-selective agar plates. The overnight cultured plates were then replicated to kanamycin selection plates to screen plasmid-free colonies. The fitness cost of *bla*$_{TEM\text{-}116}$* plasmids in the evolved clones was estimated as the difference in fitness of the original isolated evolved clones (yellow boxes) and paired derivative clones that had lost the plasmid (blue boxes). Anc: ancestral cells. Symbols represent independent fitness estimates (n = 3). **(B)** Reintroducing the ancestral control and *bla*$_{TEM\text{-}116}$* plasmid into plasmid-free evolved clones derived from Figure 2.6A demonstrated that reduction in the cost of the *bla*$_{TEM\text{-}116}$* plasmid was due to changes in the host genotype. Symbols represent independent fitness estimates (n = 6). Boxes give median and interquartile range of estimates. t-test: ***P<0.001.

2.5. Discussion

The fitness cost of a resistance genotype affects its persistence and evolution. On the one hand, costs promote purifying selection that decreases the expected frequency of ARGs, but the same costs also favour compensatory evolution to alleviate them. In this study, I found four out of six prevalent ARGs have a fitness cost on average, but there is a variation in the cost due to ARG-host interactions. To examine if the variability is consequential, I tracked the persistence of the $bla_{TEM\text{-}116}$* plasmid in three bacterial hosts. The plasmid showed different evolutionary dynamics between hosts, indicating that the host-dependent cost of $bla_{TEM\text{-}116}$* can lead to contrasting evolutionary fates of the plasmid. These results highlight that even without antibiotic selection, selection on the fitness cost can shape the persistence and evolution of an ARG in a host-dependent manner.

Host-dependent costs may correlate with the distribution of an AR determinant in bacterial populations. Epidemiological surveillance data found that there is a non-random association between an AR determinant and its bacterial hosts. For instance, the occurrence of the metallo-β-lactamases VIM-2 and SPM-1 is host-specific. While VIM-2 and SPM-1 are commonly found in *P. aeruginosa*, they are rare in *Enterobacteriaceae* [133]. This ARG-host association is correlated with a host-specific cost of the metallo-β-lactamases. The signal peptides of VIM-2 and SPM-1 induce envelope stresses and compromise cellular growth in enterobacteria, but have no effect in *P. aeruginosa* [133]. A similar association is seen in the carbapenem-resistant pOXA-48_K8 plasmid. Laboratory-tested bacterial isolates that suffer large costs due to carriage of pOXA-48_K8 are genetically more closely related to clinical isolates that rarely carry pOXA-48_K8 [145].

One complication in interpreting gene-host association surveillance data is the potential for the relationship to be complicated by other adaptive traits. Genome-wide association studies found that a mupirocin resistance conferring *ileS* mutation has strong epistatic association with the virulence of *Staphylococcus aureus* isolates [138, 246]. Both the resistant *ileS* mutant and toxin production are costly. However, the resistant allele can reduce the production of costly toxins, and thus there is no fitness difference between the antibiotic susceptible and resistant bacteria. Yet, the fitness cost of the resistance is only seen in toxin-deficient cells [138]. The fitness tradeoff between the antibiotic resistance and the bacterial virulence complicates the association of the *ileS* resistance mutation with its host.

The association of antibiotic resistance with a fitness cost has inspired policies on restricting antibiotic usage to limit the spread of resistant bacteria [153, 247]. However, clinical studies revealed the restriction policy was not widely effective in reducing the prevalence of resistant pathogens [248-250]. In addition to compensatory evolution, host-dependent costs may contribute to the failure of the policy. It is plausible that an AR determinant can remain widespread in hosts that do not suffer an associated cost while becoming less common in hosts in which it confers a large cost. Taken together, host-dependent costs imply that there is a biological barrier to prevent the spread of antibiotic resistance determinants across bacterial hosts that will give rise to an unequal distribution of an AR determinant in bacterial populations.

The long-term effects of host-dependent costs on persistence of antibiotic resistance determinants in bacterial populations are not well studied. In this study, I showed that purifying selection and compensatory evolution on the variation in $bla_{TEM\text{-}116}$* cost can lead to contrasting evolutionary dynamics of the $bla_{TEM\text{-}116}$* plasmid. The rise of compensatory mutations altered a plasmid's dynamics in a host and may transform the host into an ARG

reservoir. Moreover, the chromosomal location of the compensatory mutations suggests that a host's responses to similar ARGs may have changed. The host-dependent cost can influence the persistence and evolution of an ARG in individual bacterial hosts. Nonetheless, it is not clear how strong this influence will be in determining the dynamics of antibiotic resistance in a bacterial community. A simulation model predicted that a costly plasmid can have higher persistence in a community when the cost's variability is higher between bacterial members [145]. The variation suggests that a costly resistance determinant can stably exist in bacterial populations even without antibiotic selection [145]. Complicating predictions of determinant dynamics, however, is that compensatory evolution in individual hosts can modify the overall size and variability of any cost. In summary, I argue that knowledge of the long-term persistence and evolution of resistance determinants in bacterial communities is required to predict the prevalence of antibiotic resistance in commensal bacteria [251, 252].

Due to stagnation of new drug development research on the effect of combinatorial antibiotic therapies is growing [253]. This approach places emphasis on the control of antibiotic resistant bacteria informed by knowledge of cross-resistance and collateral sensitivity of AR determinants [253]. Cross-resistance is found when a resistance allele is resistant to not only its cognate antibiotics but also other classes of drugs. For instance, efflux pumps can provide a broad-spectrum resistance to structurally different drugs [224, 254]. Collateral sensitivity may involve complex genetic interactions and physiological changes. Bacteria carrying the resistance allele become more susceptible to other non-cognate drugs. For instance, *E. coli* that was experimentally evolved in aminoglycoside environments not only developed resistance to the selected drug but also became more sensitive to other drugs, such as β-lactams (cell wall synthesis inhibitors), fluoroquinolones (DNA synthesis inhibitors), and tetracyclines (protein

synthesis inhibitors) [255]. The aminoglycoside resistance mutations affect genes that mediate proton-motive force (PMF). If PMF is reduced, it may compromise PMF-dependent efflux proteins and consequently increase the cellular sensitivity to other drugs [255]. In this study, I found bacteria carrying $bla_{TEM\text{-}116}$* have lower fitness in environments supplemented with kanamycin, even though they encoded a second ARG conferring resistance to that drug (Figure 2.6, Table 2.4 and Table 2.5). The kanamycin-aggravated cost of $bla_{TEM\text{-}116}$* indicates that kanamycin can enhance purifying selection against $bla_{TEM\text{-}116}$*. It is unclear what genetic and physiological mechanisms underlie the interaction of kanamycin and $bla_{TEM\text{-}116}$*, but future studies on the interaction may provide insights into the control of multidrug resistance.

The six prevalent ARGs tested in this study have a small fitness cost, yet the findings are in contrast to some studies. A recent study on the fitness costs of over 200 ARGs in *E. coli* showed that many ARGs can impose a fitness cost and some can result in a 30% fitness reduction [256]. Specially, their *aad*, *cat*, bla_{TEM}, $bla_{CTX\text{-}M}$ and bla_{SHV} variants showed 0~20% fitness cost. A possible explanation for the different magnitude of fitness estimates relative to those obtained here include key differences in the experimental designs. Both studies placed test ARGs on a synthetic non-transmissible plasmid, but the copy numbers of the plasmids differ. While the pmFP plasmid (with the pSC101 replication origin) has an average of 5 copies per cell, Porse et al. (2018) used the pZAT plasmid (p15A origin) that has an average number of 10 [257]. The slightly higher copy number of pZAT may increase the expression of its cargo gene. Moreover, in contrast to a native promoter used in this study, their ARGs were under a common non-native low/medium strength promoter. Therefore, the expression difference of the ARGs between the two studies are further intensified, which may partly explain the inconsistent results. Indeed, it has been shown that the cost of a plasmid-borne gene increases with its copy

number [258]. Another difference in the experimental design is the fitness measurement. While I adopted a fitness competition assay, which considers growth differences in all growth stages, they measured the maximal growth at the exponential growth phase. Different experimental designs may influence the expression level of an ARG and hence its fitness cost. Nonetheless, the differences in expression level could reflect another layer of selection that may act on ARGs. In nature, many factors may influence the expression level of an ARG, including its native promoter, any influence on transcription by an integron, the plasmid copy number, and a host's transcriptional machineries [259-262]. Hence, the fitness cost of an ARG in nature may be more variable than determined when many of these variables are held constant.

Overall, my results highlight the significance of an often overlooked effect of antibiotic resistance—the fitness cost—on the evolutionary dynamics of an ARG. Future studies can expand to test the influence of ARG costs on their dynamics in bacterial communities contributing to goals of the One Health plan.

Chapter 3. Antagonistic interaction with a prophage gene underlies the fitness cost of a β-lactamase

3.1. Abstract

The fitness cost of an antibiotic resistance gene (ARG) can determine its maintenance in a bacterial host population in the absence of antibiotic selection. Moreover, ARGs are frequently transferred between bacteria and costs vary across host backgrounds. Despite the importance of ARG-host interactions in determining ARG maintenance, the mechanisms underlying the fitness cost of an individual ARG and its variability between hosts are not well understood. I investigated the genetic basis of a host-dependent cost of a β-lactamase, BlaTEM-116* and identified mutations in a P1 phage-encoded gene, $relA_{P1}$, that can compensate this cost in experimentally evolved cells. Using genetic construction experiments I found that the phage gene $relA_{P1}$ is sufficient to recapitulate the cost of $bla_{TEM\text{-}116}$* in an alternative host in which it was initially neutral. To my knowledge, these findings represent the first demonstrated case of a genetic interaction between a phage gene and an ARG. The interaction between a phage gene and a plasmid-borne adaptive gene highlights the complexity of selective forces determining the maintenance and spread of ARGs in natural bacterial communities.

3.2. Introduction

The spread of antibiotic resistance in pathogenic bacteria, mainly attributed to selection imposed by frequent exposure to antibiotics, has become an increasing problem to public health [90, 91]. Yet, other mechanisms may contribute to the prevalence of antibiotic resistance [153]. In the absence of antibiotic selection, the genetic determinants of antibiotic resistance, antibiotic resistance mutations (ARMs) and antibiotic resistance genes (ARGs), can impose a fitness cost on bacteria [95, 263]. This cost can facilitate purifying selection against the resistant bacteria in the absence of antibiotic selection, but counterintuitively, it also enhances the selection for compensatory mutations to alleviate the cost [95, 153]. Indeed, compensatory evolution increases with the severity of the fitness defect of a mutation [264]. Moreover, there may be more gene targets for compensation, than for a simple reversion of the resistance mechanism, as a result of the usually pleiotropic effects of resistance mutations [245, 265]. As a result, compensatory evolution on antibiotic resistance can be more common than a reversion of resistance when an antibiotic treatment stops [95, 266].

Unlike ARMs that generally occur on genes native to the host genome, ARGs originate from other bacteria and are often not integrated into the core genetic network of the host [267]. Hence, the mechanisms causing their fitness cost are less understood, though candidate mechanisms that contribute to the cost of a horizontal gene transfer (HGT) event have been proposed. These include: compatibility of GC content and codon usage of an ARG with the recipient genome [129, 267], gene function being antagonistic in the context of a host cell [268], protein interactivity [126], toxicity to a new host [269], intrinsic costs of gene expression [130,

270], and the 5' UTR stability of mRNA [271, 272]. Many of the above sequence features have been examined in *Escherichia coli* for ~200 clinically prevalent ARGs [256]. That study found GC content, codon usage, gene length and 5' UTR stability of mRNA were poor predictors of ARG fitness cost. Instead, costs were best predicted by considering a combination of an ARG's phylogenetic origin and the functional category of its resistance mechanism [256]. For example, the fitness cost of cell-interacting ARGs, such as efflux proteins, correlated positively with the phylogenetic distance between the donor bacterium and *E. coli*. Yet, no such correlation was found for drug-interacting ARGs, such as drug modification enzymes [256].

Generic sequence and protein function features provide a mechanistic insight into the evolutionary success of a HGT event, yet detailed studies on individual ARGs can reveal gene-specific mechanisms underlying the fitness costs. For example, the cost of the tetracycline resistance operon *tetAR* is attributed to its unregulated expression in the absence of tetracycline [108, 135]. The signal peptides rather than the enzymatic activity of β-lactamases SME-1 and CTX-M-15 are the main cause of associated fitness costs, which may result from a disruption of cell membranes [134, 135]. By contrast, the expression of β-lactamases OXA and SFO-1 are accompanied by changes in peptidoglycan structure, suggesting residual DD-transpeptidase activity may underlie their fitness costs [24].

Mutations that compensate for the effect of a change in a focal gene are likely to occur at secondary sites in that gene or in genes that encode products that interact with the focal gene [245, 273]. Compensatory mutations to antibiotic resistance have risen repeatedly in laboratory

evolution experiments and are also found in clinical isolates [147, 185, 187, 274]. The repeated evolution implies that studying the mechanism of an ARG's cost may help to predict compensatory evolution. For instance, clinical bacterial isolates with rifampicin-resistance conferring mutations in the β subunit of RNA polymerase *rpoB* often carry compensatory mutations on other RNA polymerase subunits [147, 274]. Conversely, identifying compensatory mutations may uncover the sources of the cost of resistance. Studies on different types of ARG-carrying plasmids have repeatedly found compensatory mutations on DNA helicases, suggesting that DNA replication stresses induced by a plasmid could be a common attribute to a plasmid's fitness cost [185, 187]. Therefore, studying and dissecting the genetic basis of compensatory evolution in antibiotic resistance can explore the sources of the cost of resistance.

The fitness effect of a mutation can depend on its genetic background [137]. Such dependency can be especially strong in bacteria given their diversity of gene content [275]. For example, for a given *E. coli* strain, only around one half of its genes, the core genome, is shared among nearly all other *E. coli* isolates [276]. Moreover, genes in the core genome are polymorphic between isolates [277]. Therefore, it is not surprising that the fitness effect of a genetic change can vary between and within bacterial species. Variation in the fitness cost of antibiotic resistance can be consequential. It can result in an association of resistance mutations or ARGs with particular lineages of pathogenic bacteria and facilitate the maintenance of ARG-carrying plasmids. For instance, the metallo-β-lactamases BlaVIM-2 and BlaSPM-1 are frequently found in *Pseudomonas aeruginosa* but rarely in other enterobacteria [133]. The gene-host

association is likely due to these metallo-β-lactamases conferring a cost in most enterobateria but not in *P. aeruginosa* [133]. Similarly, the carriage of the carbapenemase-carrying plasmid pOXA-48_K8 has different fitness effects in enterobacterial isolates [145]. The plasmid can impose an up to a 20% fitness cost in some bacterial isolates, but can also confer a 20% fitness benefit in others [145]. Such variation is beneficial for the maintenance of the plasmid at a population level in a complex bacterial community [145]. Although it is presumably common to observe a variation in the fitness cost of antibiotic resistance, the genetic divergence between bacteria can make it a challenge to identify the genetic basis underlying the variation.

In **Chapter 2**, I found the β-lactamase $bla_{TEM\text{-}116}$* imposed a 10% fitness cost in the *E. coli* isolate M114 but was only slightly costly or nearly neutral in 11 other *Escherichia* isolates. I evolved M114 with a $bla_{TEM\text{-}116}$* plasmid in the absence of selection for the plasmid and found that population-level loss of the $bla_{TEM\text{-}116}$* plasmid was greatly reduced through time. The slowed plasmid loss was due to compensatory evolution to the fitness cost of $bla_{TEM\text{-}116}$*. In this study, I investigate the genetic basis of this compensation and demonstrate that this provides insight into the genetic basis of the host-dependent cost of $bla_{TEM\text{-}116}$*.

3.3. Materials and Methods

3.3.1. Bacterial strains and media.

Bacteria were cultured as described in **Chapter 2**. *E. coli* strains were maintained in DM medium supplemented with 250 μg/ml glucose at 37° C. The plasmids, pmFP and pmFP-bla$_{TEM-116}$*, were maintained in that medium supplemented with 50 μg/ml kanamycin.

3.3.2. Scarless genetic modification of relA$_{P1}$.

Genetic modification of the P1-like prophage gene *relA$_{P1}$* was accomplished using the portMAGE protocol [278]. Briefly, the portMAGE2 plasmid was transformed into target cells. The transformants were maintained at 30º C to prevent induction of the encoded cI857-repressed λ red recombinase enzymes. Overnight cultures of transformants were diluted 1:200 into 5 ml Lysogeny broth (LB) supplemented with 100 μg/ml ampicillin and grown for three hours. To induce the expression of the red recombination machinery and the transitive mutator phenotype (MutL E32K), these cells were induced at 42º C for 15 minutes with constant shaking. After induction the cultures were immediately placed on ice for 10 minutes then used to prepare electrocompetent cells by washing three times with 10 ml ice-cold Milli-Q® water. The cells were resuspended in 100 μL water and placed on ice before mixing with 1μL of 100 μM of an oligonucleotide designed to introduce focal mutations into the target cell. Electroporation was conducted in 1 mm gap electroporation cuvettes (BIO-RAD) with the setting 1.8 kV, 200 Ω, 25 μF. The cells were recovered in LB medium supplemented with

ampicillin at 30º C for 2 hours before either plating on LB agar plates supplemented with ampicillin or inoculation into 5 mL LB supplemented with ampicillin with overnight incubation for use in a subsequent portMAGE cycle. Four cycles of portMAGE were performed before genotyping the target gene. To cure the portMAGE2 plasmid, the cells were cultured at 42º C overnight on an LB agar plate and then streaked out and cultured on LB plates at 37º C. Colonies susceptible to ampicillin were selected for genotyping. The 5'end phosphorothioated (*) oligonucleotides for the gene editing were designed as suggested in Gallagher et al., 2014 [279]: MAGE_relA_anc_o1: C*T*TCGATCAGAAAGAGGATCAATATCCGTGGCGAAAAACAAAGATTGAGATA CAGTTAAGAACTCAATTGCAGCATGCTTGGGCTACTAG; MAGE_relA_evo_o1: C*T*TCGATCAGAAAGAGGATCAATATCCGTGGCGAAAAACAAAGAgTGAGATAC AGTTAAGAACTCAATTGCAGCATGCTTGGGCTACTAG; MAGE_relA_null_o1: C*T*TCGATCAGAAAGAGGATCAATATCCGTGGCGAtAAACAtAGATTtAGATAtAG TTAAGAACTCAATTGCAGCATGCTTGGGCTACTAG.

3.3.3. relA$_{P1}$ integration at attTn7 site.

To integrate *relA$_{P1}$* into REL606 genome, *relA$_{P1}$* was first cloned into the pGRG36 (ampicillin resistant) plasmid [280]. The resulting pGRG36-*relA$_{P1}$* plasmid was used to transform REL606 and transformants were selected on ampicillin containing LB agar plates at 30° C. To induce mini-Tn*7* transposition, cells were grown in LB supplemented with 0.1% arabinose at 30° C. To cure the plasmid, cells were streaked out and cultured at 42° C twice. The integration of *relA$_{P1}$* at *att*Tn*7* site was confirmed by PCR.

3.3.4. Fitness competition assay.

A flow cytometry-based competition assay protocol was used to estimate the fitness cost of a plasmid or a gene. The procedure was described in **Chapter 2**. Briefly, the cells were precultured in DM medium supplemented with 250 μg/ml glucose (and 50 μg/ml kanamycin if they carried a pmFP-based plasmid). The overnight cultures were then diluted 1:100 and preconditioned in fresh DM medium supplemented with 250 μg/ml glucose for two growth cycles. Strains were then separately mixed 1:1 with reference cells containing a pUA66-P_{rpsL}GFP plasmid (Day 0). The competition mixture was diluted 1:100 in fresh DM medium supplemented with 250 μg/ml glucose and cultured overnight (~6.7 generations; Day 1). The Day 0 and Day 1 samples were assayed using flow cytometry to distinguish test and reference cells. The proportion of GFP-positive and GFP-negative events were used to estimate the relative fitness of the cells [281].

3.3.5. Whole genome sequencing of evolved clones.

One plasmid-carrying clone was isolated from each M114 evolved line as described in **Chapter 2**. Genomic DNA of the evolved clones was isolated with Wizard® Genomic DNA Purification Kit (Promega) and short-read sequenced (HiSeq platform by The Microbial Genomic Sequencing Center, USA). Breseq software was used to call variants [282].

3.4. Results

3.4.1. Mutations on the prophage alleviate the cost of the blaTEM-116 plasmid*

In **Chapter 2**, I reported experiments in which I evolved six replicate lines of *E. coli* strain M114 carrying a control plasmid and nine replicate lines of M114 carrying a *bla*$_{TEM\text{-}116}$* plasmid for ~400 generations in antibiotic-free medium. I found that the fitness cost conferred by the *bla*$_{TEM\text{-}116}$* plasmid was greatly reduced in the evolved *bla*$_{TEM\text{-}116}$* plasmid-carrying clones, being reduced from 10% to nearly zero. In this chapter, I aimed to identity the genetic basis of the compensatory evolution. I reasoned that compensatory mutations that reduce the fitness cost of *bla*$_{TEM\text{-}116}$* are likely to be over-represented among the nine *bla*$_{TEM\text{-}116}$* plasmid-containing populations compared to the six control populations. To identify genetic changes in the evolved clones, I sequenced the genome of one plasmid-carrying clone isolated from each independently evolved line.

I found mutations in a native 90 kb P1-like prophage in M114 (hereafter, P1_M114; Figure 3.1) matched this criterium, occurring in all nine *bla*$_{TEM\text{-}116}$* plasmid clones, but in none of the

Table 3-1 Mutations found in the evolved plasmid-carrying clones.

v1	v2	v3	v4	v5	v6	b1	b2	b3	b4	b5	b6	b7	b8	b9	Gene	Description
A1bp	A72bp			A1bp											*GJAEAM04_00676*	O-antigen polymerase
			K36*												*rfaL*	O_antigen ligase
		P52S													*dut*	dUTPase
R562C	E82*	R62L		R562C	R562C				L92Q	Q280K				+1bp	*prkA*	Serine protein kinase
												A196V			*ydhP*	MFS transporter
						Q407*							S318P		*yhbH*	Probable sigma(54) modulation protein
										+2bp					*GJAEAM04_00573*	Phage tail protein
							H210Y								*GJAEAM04_02332*	Glycosyltransferase
													D58Y		*nfsA*	Nitroreductase
											D54bp				*ycgB*	Putative SpoVR family protein; belongs to the RpoS regulon
						Loss	Loss		Loss	Loss	Loss			Loss	*Phage P1_M114*	Phage P1_M114
								I179S				I179S	I179S		*relAP1*	Putative small ppGpp synthase

* Note that intergenic and synonymous mutations are not included in this table and none of the these mutations match the selection criteria for the compensatory mutations.

1

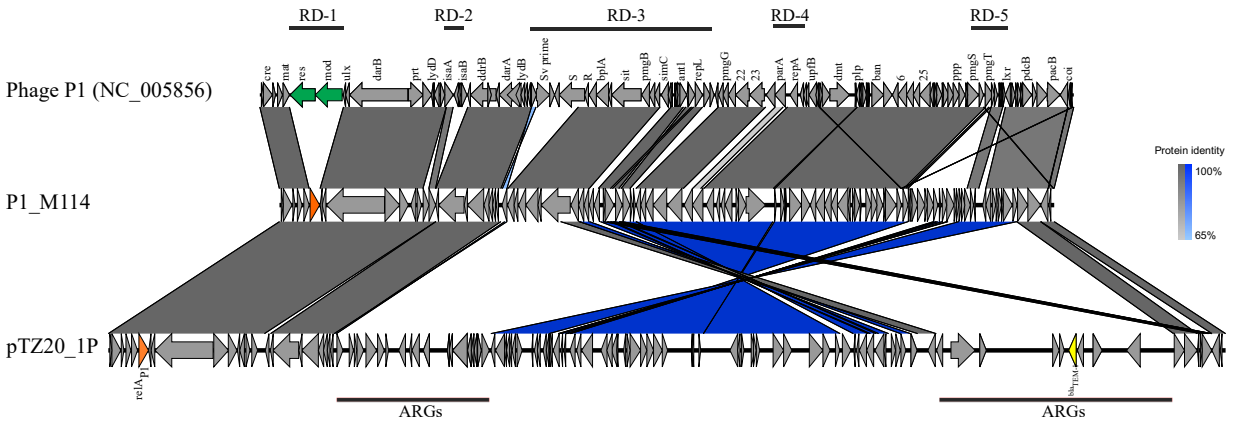

Figure 3.1. Genetic map and alignment of phage P1 variants.
Genome comparison between phage P1 variants. RD-1~5 are the region of sequence difference between the lab strain and the P1 variants [2]. The protein sequence identity between P1_M114 and other phage variants is colored by grey or blue gradient (inversion). *relA*$_{P1}$ (orange arrow) is present in P1_M114 and pTZ20_1P but absent in the laboratory phage strain (NCBI reference sequence: NC_005856). Green arrows indicate the type III restriction-modification system, *res* and *mod*, present in the type phage that flank the insertion site of *relA*$_{P1}$. pTZ20_1P has acquired multiple ARGs, including one *bla*$_{TEM-1}$ (indicated by yellow arrow).

control plasmid clones (Table 3.1). Six of the evolved *bla*$_{TEM-116}$* plasmid clones lost the entire prophage P1_M114, and the other three clones gained a mutation, I179S, in *relA,* a predicted phage-encoded ppGpp synthetase (Figure 3.2 and Table 3.1). *E. coli* has a full length *relA* encoded in its chromosome, so to make a distinction from the phage encoded *relA*, I hereafter refer to the phage-encoded *relA* as *relA*$_{P1}$ (Figure 3.3). The high parallelism of mutations in the prophage P1_M114 exclusively among *bla*$_{TEM-116}$* plasmid-containing populations suggests they may reduce the fitness cost of the *bla*$_{TEM-116}$* plasmid. To examine if the mutations on the prophage P1_M114 are compensatory, I first determined the cost of the *bla*$_{TEM-116}$* plasmid in the M114 ancestor with the I179S mutation in *relA*$_{P1}$ and compared it to the cost of the same plasmid in the WT M114 strain (Figure 3.4). The cost of the *bla*$_{TEM-116}$* plasmid was reduced from 16.1% (95% CI: 14.1%–18.1%]) in the wild type M114 to nearly zero in all of three independently constructed M114 *relA*$_{P1}$*/I179S* clones (transformant tf.1, 0.04% in the 95% CI: -1.98% –2.05% and 0.89% in the transformant tf.2, 95% CI: -1.12%–2.91%), indicating that the evolved *relA*$_{P1}$ allele was sufficient to compensate for the cost of the *bla*$_{TEM-116}$* plasmid (Fig. 3.4). To further verify the compensation, I reverted the *relA*$_{P1}$*/I179S* mutation to the ancestral genotype in one evolved clone, and the reversion recapitulated the majority of the fitness cost of the *bla*$_{TEM-116}$* plasmid in the evolved clone (Figure 3.5). These results indicate both that the *relA*$_{P1}$*/I179S* mutation was sufficient and, at least in the tested evolved line, necessary to compensate for the cost of the *bla*$_{TEM-116}$* plasmid.

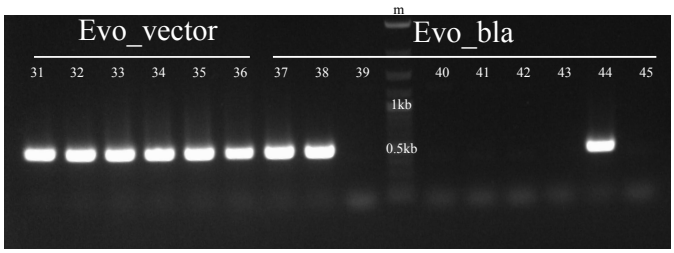

Figure 3.3. Presence of P1_M114 in evolved clones.

The presence of the phage P1_M114 in evolved clones was confirmed by PCR. Clone 31-36 (evolved with the control plasmid) and clone 37-45 (evolved with the *bla*$_{TEM-116}$* plasmid) indicate PCR results of clones isolated from each evolution line. Consistent with the WGS result, six evolved clones from the *bla*$_{TEM-116}$* plasmid populations lost the P1 phage. The lane labelled 'm' denotes a 1 Kb ladder with bands indicating fragments of 0.5 and 1 Kb highlighted. Note that clone IDs are different from that used in WGS (Table 3.1) as sample names were mistakenly swapped during sample preparation for genome sequencing.

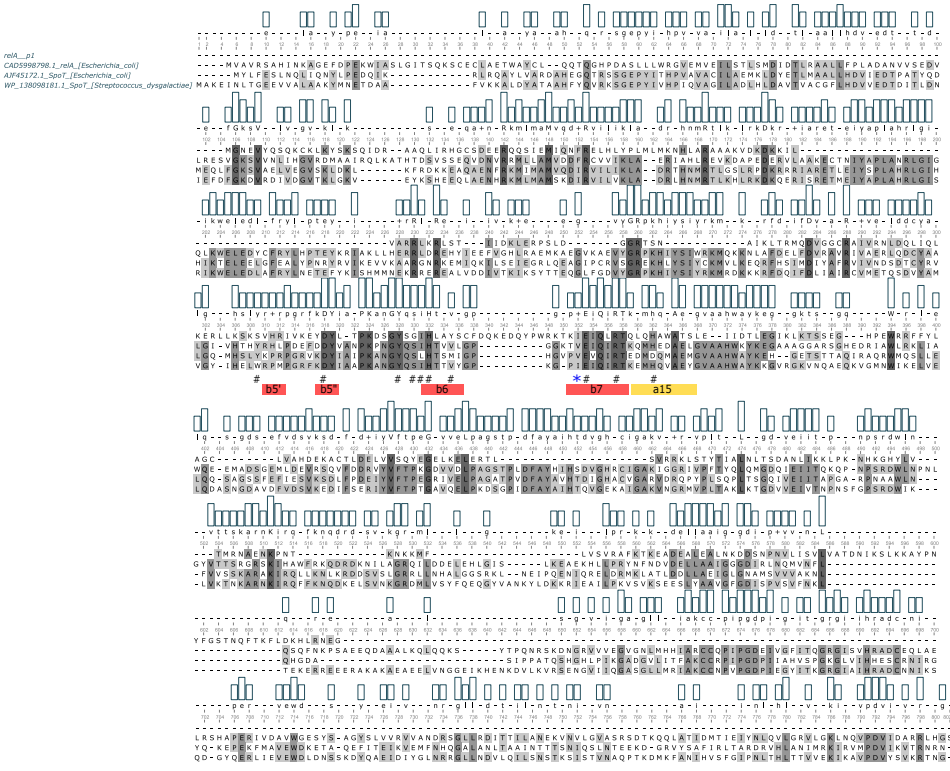

Figure 3.2. Protein alignment of RelA/SpoT homologs.

Protein homologs of RelA/SpoT from different species were aligned using CLUSTAL W 2.0. SpoT from *Streptococcus dysgalactiae* is a bifunctional protein with both ppGpp synthetase and hydrolase activity in the N-terminal domain. RelA from *E. coli* only has the synthetase activity, and SpoT from *E. coli* has both synthetase and hydrolase activity. The C-terminal region of the full length RelA/SpoT is a regulatory domain that binds to other proteins such as ribosomes. RelA$_{P1}$ from phage P1 aligns to the N-terminal region of the full length RelA/SpoT. The position of the I179S mutation is highlighted with a blue star. Amino acids kknown to be critical for the synthetase activity around I179 of RelA$_{P1}$ are marked with the hash symbol. The α helix and β sheet are highlighted with the yellow and red boxes, respectively [5].

The protein length of *relA*$_{P1}$ is shorter than the chromosomal ppGpp synthetase encoded by *relA* and the dual function ppGpp synthetase/hydrolase encoded by *spoT* (Figure 3.3). RelA$_{P1}$ lacks the regulatory C-terminal domain of RelA and SpoT, and only shares sequence similarity to their N-terminal enzymatic domains. Sequence alignments between the RelA/SpoT homologs (RSHs) and RelA$_{P1}$ showed that the I179 residue of RelA$_{P1}$ is located near residues critical for ppGpp synthetase activity, suggesting RelA$_{P1}$ has a functional ppGpp synthetase activity (Figure 3.3). To investigate the functional impacts of the I179S mutation on *relA*$_{P1}$, I introduced 4 stop codons at the start of *relA*$_{P1}$ in the ancestral M114 strain. This null allele conferred a similar reduction in cost of the *bla*$_{TEM\text{-}116}$* plasmid as did the evolved I179S allele,

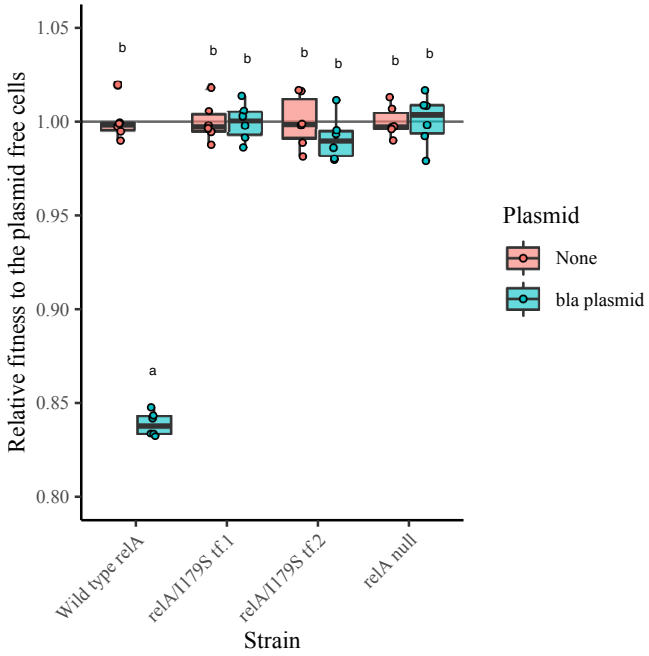

Figure 3.4. Compensatory mutation in *relA*$_{P1}$ reduces the cost of *bla*$_{TEM\text{-}116}$*.

I179S or null *relA*$_{P1}$ mutations were introduced into the M114 ancestor and their influence on the fitness effect of the *bla*$_{TEM\text{-}116}$* plasmid was examined. Two independent clones, tf.1 and tf.2, of *relA*/I179S replacement were obtained and included to rule out the influence of off-target mutations. Dots represent independent fitness estimates (n=5 or 6). Boxes indicate median (internal line) and inter-quartile fitness estimates for each host-plasmid combination. Boxes sharing the same letter are not significantly different (alpha = 0.05, Tukey-adjusted).

indicating that the I179S substitution may cause a substantial decrease or loss of RelA$_{P1}$ activity (Figure 3.4). Neither the I179S or the null *relA$_{P1}$* alleles affected cell fitness (Figure 3.6), suggesting the *relA$_{P1}$*/I179S was not selected in the evolved clones due to general adaptation to the lab environment but, rather, to specifically compensate the cost of the *bla$_{TEM-116}$** plasmid.

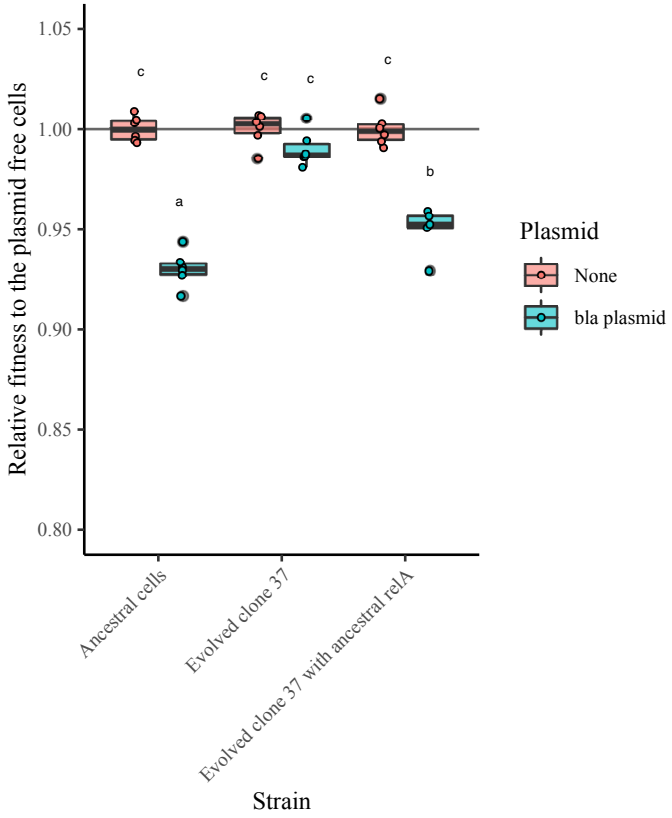

Figure 3.5. Reversion to the ancestral *relA$_{P1}$* recapitulated the cost of *bla$_{TEM-116}$ in evolved cells.**

The cost of the *bla$_{TEM-116}$** plasmid was greatly reduced in evolved clone 37, which contained the evolved *relA$_{P1}$/I179S* allele, compared to its ancestor. Replacement of the evolved *relA$_{P1}$/I179S* allele with the ancestral allele recapitulated the original cost of the *bla$_{TEM-116}$** plasmid. Dots represent independent fitness estimates (n=6). Boxes indicate median (internal line) and inter-quartile fitness estimates for each host-plasmid combination. Boxes sharing the same letter are not significantly different (alpha = 0.05, Tukey-adjusted).

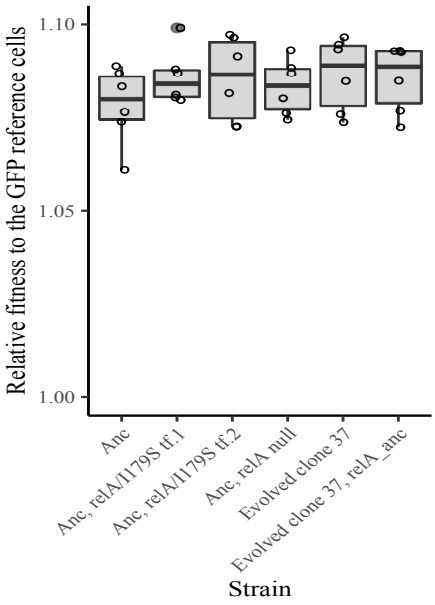

Figure 3.6. Mutations on *relA*$_{P1}$ did not change fitness in the absence of *bla*$_{TEM\text{-}116}$*.

Neither I179S or null *relA*$_{P1}$ alleles in the ancestral M114 strain nor the ancestral *relA*$_{P1}$ in the plasmid-cured evolved clone 37 influenced cell fitness. Relative fitness was estimated from the relative competitive fitness of test strains to a GFP-positive derivative of the M114 ancestor. Dots represent independent fitness estimates (n=5 or 6). Boxes indicate median (internal line) and inter-quartile fitness estimates for each host-plasmid combination. Analysis of the variation found no fitness difference between different cell types ($F_{5,30}$=0.62, P=0.68).

3.4.2. Expression of relA$_{P1}$ *results in the cost of the bla*$_{TEM\text{-}116}$* *plasmid*

That six evolved lines started with the *bla*$_{TEM\text{-}116}$* plasmid lost the entire P1_M114 prophage suggests this change can also compensate the cost of the *bla*$_{TEM\text{-}116}$* plasmid (Figure 3.2 and Table 3.1). Although the above results clearly demonstrate that the evolved *relA*$_{P1}$ gene can explain the compensation phenotype (Figure 3.4 and 3.5), it is possible that it is through an interaction with *relA*$_{P1}$ and other P1_M114 encoded genes. To test this possibility, I first attempted to cure the prophage P1_M114 from the ancestral strain. Despite repeated attempts, including protocols exposing cells to chemical or physical stresses (ethidium bromide and UV

light), I was unable to isolate any prophage cured clones. For this reason, to isolate the influence of *relA*$_{P1}$ in determining the cost of the *bla*$_{TEM\text{-}116}$* plasmid, I instead introduced *relA*$_{P1}$ into the lab strain REL606, which does not naturally have the prophage P1_M114 or *relA*$_{P1}$ and has no cost when bearing the *bla*$_{TEM\text{-}116}$* plasmid. If *relA*$_{P1}$ is the causal determinant of the cost of the bla$_{TEM\text{-}116}$* plasmid, I expected that expressing *relA*$_{P1}$ alone would incur a cost of the plasmid similar to that in M114. By contrast, if other prophage P1_M114 encoded genes or M114-specific genes are involved, I expected that expressing *relA*$_{P1}$ alone in REL606 would have only little or perhaps no influence on the fitness effect of *bla*$_{TEM\text{-}116}$*.

I integrated *relA*$_{P1}$ into the *attTn7* site in the REL606 chromosome and measured the cost of *bla*$_{TEM\text{-}116}$*. I found that *bla*$_{TEM\text{-}116}$* confers a 10% fitness cost in REL606 Tn7:: *relA*$_{P1}$ but had no effects in the wild type REL606 (Figure 3.7). This result indicates that *relA*$_{P1}$ is sufficient to cause *bla*$_{TEM\text{-}116}$* to confer a cost in a strain in which it is otherwise neutral. Taken together, the compensated cost of *bla*$_{TEM\text{-}116}$* in M114 with a *relA*$_{P1}$ null mutant (Figure 3.4) and the emergent cost of *bla*$_{TEM\text{-}116}$* in REL606 Tn7:: *relA*$_{P1}$ (Figure 3.7) indicates the cost of *bla*$_{TEM\text{-}116}$* is a result of the antagonistic interaction between *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$*. The *relA*$_{P1}$ dependent cost of *bla*$_{TEM\text{-}116}$* may explain why *bla*$_{TEM\text{-}116}$* imposes a cost in M114, but not the other 11 *E. coli* isolates that lack P1_M114 and *relA*$_{P1}$ (**Chapter 2**).

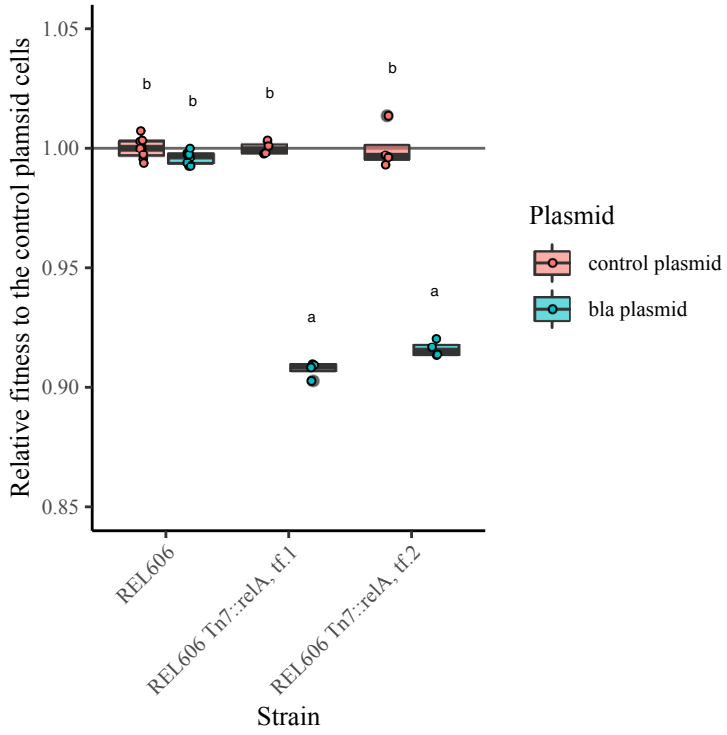

Figure 3.7. Expression of *relA*$_{P1}$ in REL606 reproduced a fitness cost of *bla*$_{TEM\text{-}116}$*.

Carriage of the *bla*$_{TEM\text{-}116}$* plasmid yielded a similar fitness effect to the control plasmid in the wild type REL606, indicating *bla*$_{TEM\text{-}116}$* does not confer a cost. By contrast, introduction of *relA*$_{P1}$ at the chromosomal *attTn7* site resulted in a 10% fitness cost of *bla*$_{TEM\text{-}116}$*. Dots represent independent fitness estimates (n=5). Boxes indicate median (internal line) and inter-quartile fitness estimates for each host-plasmid combination. Boxes sharing the same letter are not significantly different (alpha = 0.05, Tukey-adjusted).

3.4.3. The prevalence of the relAP1 encoded P1 prophage.

Despite *relA*$_{P1}$ being encoded by the prophage P1_M114, it is not ubiquitous across all P1 prophages (Figure 3.1 and Figure 3.8). For instance, the laboratory bacteriophage P1 strain (P1 mod749::IS*5* c1.100 GenBank accession: NC_005856) does not have *relA*$_{P1}$, but the gene is found in another P1 strain, pTZ20_1P (GenBank accession: MN510447). Synteny and sequence comparison of the regions surrounding *relA*$_{P1}$ in P1_M114 and pTZ20_1P shows a

high similarity despite a difference in the gene content of the rest of the genome (Figure 3.1 and Figure 3.8). The *relA*$_{P1}$ gene is located in the region of difference 1 (RD-1) in P1_M114 and pTZ20_1P, flanked by the genes *mat* and *lxc* (a phage maturation protein and a modulator of C1-mediated repression, respectively) [2]. RD-1 is a variable region and encodes various genes in different P1 strains (Figure 3.8) [2]. The low GC content in RD-1 compared to the rest of the phage genome suggests it comprises recently acquired genes [2]. For instance, the RD-1 in the lab strain, P1 mod749::IS*5* c1.100, has recently acquired *res* and *mod*, a type III restriction-modification system. If functional, this system can restrict the entry of foreign lysogens although a recent IS*5* insertion has disrupted the *mod* gene [283]. The functions of other acquired genes are less clear but may contribute to niche specialization of different strains along with gene gain and loss in other RD regions [2].

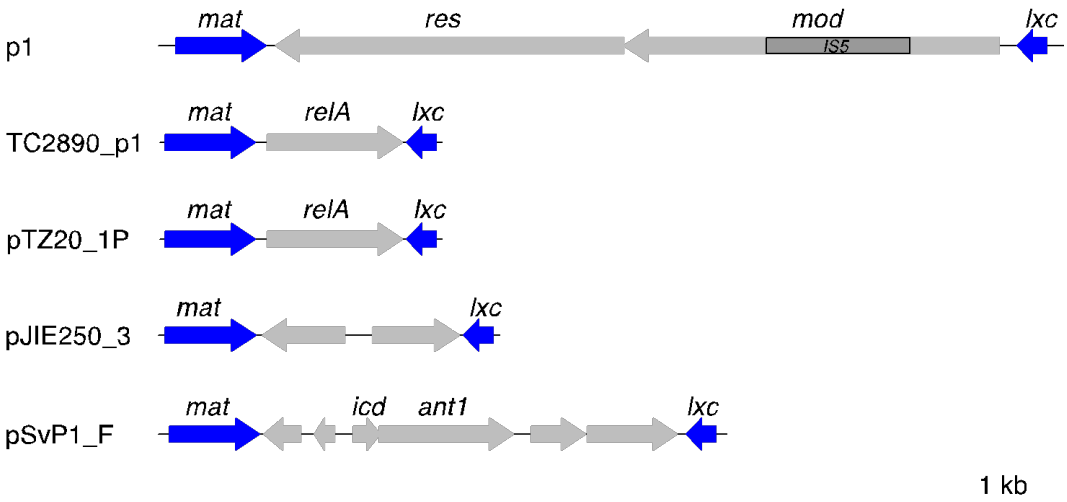

Figure 3.8. Synteny of *relA*$_{P1}$ region RD-1 shows a variable gene repertoire.

A close-up of the RD-1 of phage P1 in Figure 1. The boundary (*mat* and *lxc* gene) of RD-1 is intact in the five P1 variants, but the genes present between these genes is variable [3].

P1 is a temperate phage that possess both the life styles of a phage and a plasmid . Unlike other temperate phages that can enter a lysogeny state by integrating into a bacterial host's genome, P1 can exist as a extrachromosomal plasmid [89]. To survey the incidence of $relA_{P1}$ in P1 prophages, I identified genome accessions in the NCBI database that contain features of P1 phage (the phage replication protein, *repL*) and plasmid (the P1 plasmid replication protein, *repA*) lifestyles using BLASTn [2, 284]. Using a cutoff of 90% query coverage and 90% sequence identity, 270 accessions matched *repL*, 273 matched *repA*, and 141 matched both. Using the sequence of $relA_{P1}$ and the entire *mat-*$relA_{P1}$*-lxc* region with the same cutoffs, I found 31 accessions contain $relA_{P1}$ and all the 31 accessions matched the larger *mat-*$relA_{P1}$*-lxc* region. This strong genetic linkage occurring in an otherwise variable region suggests a recent acquisition of $relA_{P1}$ in the *mat-*$relA_{P1}$*-lxc* region. Comparing accessions matching $relA_{P1}$ with those matching *repL* and *repA* (i.e., P1 phage features), I found $relA_{P1}$ sequence in 4% of accessions with both *repL* and *repA*, indicating a substantial number of P1 phages carry $relA_{P1}$. The function of $relA_{P1}$ to the P1 phage still needs further investigation.

3.5. Discussion

Many studies have investigated the cost of ARG-encoding plasmids, but little is known about the cost of individual ARGs, let alone about the mechanistic basis of such costs [256]. Here, I investigated the genetic basis of a host-dependent cost of $bla_{TEM\text{-}116}$* and of compensation to this cost. This ARG was previously demonstrated to decrease the cellular fitness by 10% in *Escherichia* strain M114 but to have no fitness effect in another 11 tested strains (**Chapter 2**). I found that mutations on a native P1-like bacteriophage element in M114, resulting in either

loss of the element or loss of function of the phage encoded gene, *relA*$_{P1}$, occurred in evolved clones that have alleviated the cost of *bla*$_{TEM-116}$*. Follow-up reverse genetics experiments confirmed that an evolved *relA*$_{P1}$ mutation can compensate the cost of *bla*$_{TEM-116}$* in the M114 ancestral strain. Moreover, introducing *relA*$_{P1}$ in a naïve host, REL606, in which *bla*$_{TEM-116}$* was initially neutral, recapitulates the cost of *bla*$_{TEM-116}$* to a level comparable to that in M114. This indicates that the cost of *bla*$_{TEM-116}$* is the result of an antagonistic interaction between phage encoded *relA*$_{P1}$ and *bla*$_{TEM-116}$*. Evidently, the antagonism between potentially horizontally mobile elements acts antagonistically to the maintenance of *bla*$_{TEM-116}$* in natural bacterial populations. This result highlights the potential influence of bacterial accessory genomes on the maintenance and evolution of ARGs.

In this study, the cost of *bla*$_{TEM-116}$* depends on the presence of the phage encoded *relA*$_{P1}$. Protein sequence alignment indicate that RelA$_{P1}$ has similarity to the enzymatic domain of the chromosomal *relA/spoT* genes but lacks the regulatory domain that those genes encode (Figure 3.3). During nutrient starvation (e.g., amino acid, glucose and fatty acids), the ppGpp level is abruptly increased via the activation of the ppGpp synthetase activity of RelA or SpoT [285]. ppGpp regulates transcriptional machineries to regulate various genes involved in cell physiology, resulting in growth arrest as part of a so-called stringent response [285]. In addition, the basal level of ppGpp can coordinate various cell processes during exponential growth [286]. In the absence of a regulatory domain, the prophage RelA$_{P1}$ is likely constitutively active and may increase the basal level of ppGpp. The ppGpp synthetase activity of the full-length RelA is positively regulated by its product ppGpp [287]. This positive feedback ensures a rapid

accumulation of ppGpp when the stringent response is induced. Therefore, $RelA_{P1}$ may increase the basal ppGpp level and hypersensitize cells to mild stresses. For instance, a compromised cell wall or membrane components caused by the signal peptide or the DD-transpeptidase activity of BlaTEM-116* that may be tolerated in the wild type cells but may induce the stringent response in $RelA_{P1}$-hypersensitized cells [24, 133]. The hypersensitivity may explain why the fitness cost of $bla_{TEM\text{-}116}$* is only observed in the presence of $relA_{P1}$. This hypothesis will be examined in future experiments aiming to determine the molecular basis of the epistatic $relA_{P1}$ and $bla_{TEM\text{-}116}$* interaction.

The function of $relA_{P1}$ in the context of the prophage P1_M114 is unclear. A candidate adaptive mechanism for $relA_{P1}$ is as a toxin or component of a toxin-antitoxin module, which small ppGpp synthetase and hydrolase enzymes can function as [288]. For example. a homolog of ppGpp synthase, Tas1, which pyrophosphorylates adenosine nucleotides to produce ppApp, can be delivered via a type VI secretion system to neighbouring cells to cause cell death [289]. Arguing against this possibility, expression of $relA_{P1}$ in a naïve host did not decrease cell fitness, suggesting either that $RelA_{P1}$ is unlikely to be a toxin or that the host already encoded a cognate antitoxin (Fig. 3.9). Alternatively, $RelA_{P1}$ might increase the basal levels of ppGpp, causing the lysogenized cells to become more sensitive to nutrient starvation and therefore to induce the stringent response before nutrients completely run out. Prophage conferring a hypersensitized stringent response on host cells might gain a selective advantage by an early induction of the lytic cycle in harsh environments. Moreover, the elevated basal level of ppGpp may increase the proportion of persister cells induced by stochastic activation of the stringent

response [290-293]. Increased production of persister cells may increase the survival of the host population (and hence the temperate phage) in fluctuating environments.

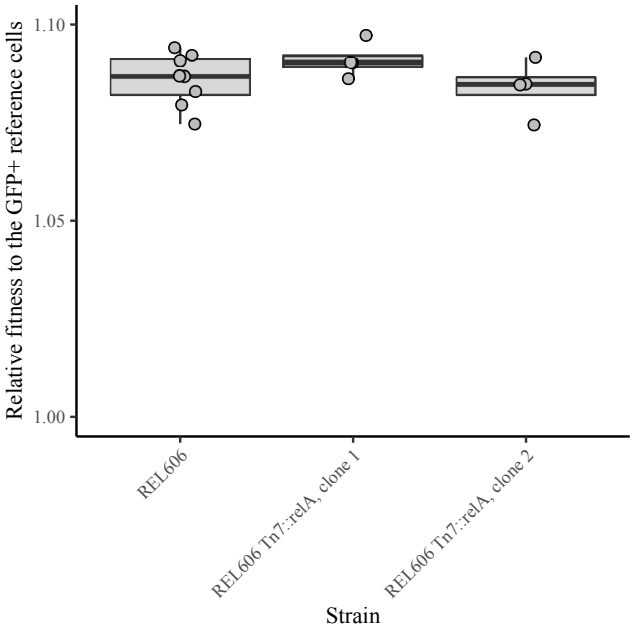

Figure 3.9. *relA$_{P1}$* did not influence the cellular growth.

The fitness of plasmid-free REL606 (WT) and two independently constructed derivatives having the *relA$_{P1}$* gene at the chromosomal *attTn7* was measured by competition with the wild type GFP-positive REL606. Dots represent independent fitness estimates (n= 4 or 8). Boxes indicate median (internal line) and inter-quartile fitness estimates for each host-plasmid combination. An analysis of variance test found no growth difference between different cell types ($F_{2,13}$=1.37, p=0.29).

The genetic interaction between *relA$_{P1}$* and *bla$_{TEM-116}$** may be incidental to the selection of both genes. However, β-lactamases share some sequence homology with phage tails [88], which may be involved in regenerating cell wall and membrane components of a host during infection [294]. It is tempting to speculate that the observed antagonism between *relA$_{P1}$* and *bla$_{TEM-116}$** may originate from a historical conflict of interest. In any case, studying the interaction dynamics of *relA$_{P1}$* and *bla$_{TEM-116}$** may have practical applications in developing

approaches to reduce the dissemination of $bla_{TEM\text{-}116}$*. Finally, I note that one P1 strain, pTZ20-1P, encodes both $relA_{P1}$ and $bla_{TEM\text{-}1}$ (Figure 3.1) [2]. It would be interesting to investigate if the antagonism between these genes exists in this prophage and if not, how it has been resolved.

In summary, I have identified the genetic basis of a host-dependent cost of $bla_{TEM\text{-}116}$* as originating from an antagonistic interaction with a phage gene $relA_{P1}$. It is unclear if such antagonistic interactions are common, but a survey of the fitness cost of ~ 200 ARGs has shown that cell-interacting ARGs are more likely to incur a cost [256], emphasizing the importance of the compatibility between a HGT gene and its host. Future studies on the identification of the genetic antagonism between an ARG and its host and on the significance of the antagonistic interactions on the dissemination and maintenance of ARGs in bacterial populations may provide a comprehensive view on the flow of ARGs between bacteria in nature.

Chapter 4. Investigation of the dependency of genetic variation in *bla*$_{TEM-116}$* on its fitness cost

4.1. Abstract

The rise of new antibiotic resistance from promiscuous antibiotic modifying enzymes represents a newly recognized challenge to combatting antibiotic resistance. With as little as one amino acid change, a promiscuous enzyme can increase its specificity to a new antibiotic substrate and expand its resistance spectrum. In natural environments antibiotic resistance genes (ARGs) are often subject to a relaxation of selection for their resistance phenotype and, concomitantly, to purifying selection if they impose any fitness cost. Complicating any effort to determine the net effect of these factors, the background genotype can influence the effect of any mutations in an ARG. For example, costs may tend to increase when tested in fitter genetic backgrounds. These factors can influence the chance that an ARG is maintained in a population and suggests that the degree of standing genetic variation of an ARG may depend on the fitness cost of accessible mutational variants. Cryptic genetic variation has been shown to facilitate the gain of a new phenotype of a protein. One implication is the potential for the fitness cost of an ARG to influence its evolution to a new resistance phenotype via accumulation of genetic variation. In this chapter, I investigate if the fitness cost of *bla*$_{TEM-116}$* in different hosts can affect maintenance of introduced genetic variation. I introduced a mutant library of *bla*$_{TEM-116}$* into two *E. coli* strains, one suffers a 10% fitness reduction when carrying *bla*$_{TEM-116}$* and the other does not. After 200 generations of selection in the absence of any β-

lactam antibiotics, I found that *bla*$_{TEM\text{-}116}$* diversity remained much higher in the costly host strain: 80% of the starting *bla*$_{TEM\text{-}116}$* mutations were well preserved in the host with a *bla*$_{TEM\text{-}116}$* cost, but nearly 95% of the starting mutations were lost in the cost-free host. In many cases these losses did not occur in parallel across the replicate lines of the cost-free host, and moreover, the frequency and the frequency changes of the selected mutations have little or no correlation between the lines, suggesting the loss of *bla*$_{TEM\text{-}116}$* diversity in the cost-free host was due to selection on linked chromosomal mutations. I discuss results and propose modifications on the experimental designs for future experiments.

4.2. Introduction

Antibiotic resistance genes (ARGs) have contributed to the prevalence of antibiotic resistance in pathogenic bacteria. ARGs not only provide immediate resistance to cognate antibiotics but also represent a gene pool that can readily evolve resistance to new antibiotics [295]. Many antibiotic modifying enzymes have a spectrum of specificity to a range of related substrates [296, 297]. This often-narrow specificity spectrum can limit the cross-resistance to similar antibiotics derived from the same lead compounds [298]. However, this specificity can be modified with a few changes of amino acids on an antibiotic modifying enzyme and undermine the efficacy of alternative antibiotics. For instance, the predominant β-lactamase, TEM-1 (BlaTEM-1), has a low specificity to the cefotaxime class of β-lactam drugs, and hence, the effectiveness of cefotaxime is not affected by BlaTEM-1. Nonetheless, with only a few amino acid substitutions, the activity of BlaTEM-1 to cefotaxime can increase ~100,000-fold to confer clinical levels of resistance [57]. So far, over 170 BlaTEM variants with different substrate specificity have been identified with each variant differing by a few amino acids [58].Yet, in the past 40 years, over 60% of developed antibiotics were derived from known antibiotics [299, 300]. The rise of new resistance from an ARG can undermine the effectiveness and development of new antibiotics. Hence, it is important to understand how new resistance rises from an ARG.

The evolvability of a protein, defined as its ability to develop new phenotypes [301], can be affected by a range of factors. For example, the stability of a protein can affect its enzymatic

activity and its ability to gain new functions. Around half of random single mutations can destabilize protein structure [302, 303], and the stability can be further deteriorated by negative interactions between destabilizing mutations [304]. Thus, an increase in the intrinsic stability of a protein product can increase a protein's tolerance to mutations, including neo-functional mutations, and consequently facilitate its evolvability [302, 305].

The relationship between protein stability and evolvability can also be influenced by the selection strength on a protein's original phenotype. For example, selection for a strong signal of yellow fluorescence protein (YFP) concomitantly selected for protein stabilizing mutations [301]. These stabilizing mutations increased the protein's tolerance to mutations that would otherwise destabilize protein structure but emit green fluorescence, the new phenotype. Therefore, strong positive selection for yellow fluorescence increases YFP's evolvability to green fluorescence [301]. However, this selection strength-evolvability relationship may not be generalizable to all proteins. Strong selection for ampicillin resistance conferred by BlaTEM-1 purges cefotaxime resistance mutations that are neutral in a low ampicillin concentration but deleterious in a high concentration, decreasing BlaTEM-1's evolvability [306].

Another mechanism to increase a protein's evolvability is via an accumulation of cryptic genetic variation (CGV) —standing genetic variants that have neutral or nearly neutral fitness effect in the current environment but possess new phenotypes that can be unmasked following an environmental change [307]. CGV allows the exploration of otherwise inaccessible genotypes and corresponding protein sequence space, potentially facilitating the evolution of

new phenotypes [308]. These mechanisms govern the evolution of protein and thus may underlie the rise of new resistance from an ARG.

In the absence of antibiotic selection, the carriage and expression of an ARG can incur a fitness cost, and the cost can influence its maintenance in a population [309]. The cost can lead to purifying selection and disfavour maintenance of the ARG in a bacterial population [95, 310], yet it also facilitates compensatory evolution to alleviate the cost. The rise of compensatory mutations depends on the degree of the deleterious genotype, where a more deleterious genotype is more likely to be compensated [245, 264]. This dependency is a result of pervasive epistatic interactions between mutations, which can lead to a generic dependence of a mutation's fitness effect based on the background genotype fitness [311-314]. A beneficial mutation will have a larger fitness effect in a less fit background [315], and hence, compensatory evolution is more likely to be selected. On the contrary, a deleterious mutation will impose a larger cost when it occurs in a fitter background [312].

The dependence between the effect of new mutations and the mean contribution of an ARG to the fitness of a host genotype may not only influence selections for compensatory mutations but also purifying selection against deleterious mutations. In natural environments, positive selection for the resistance of an ARG can be weak and sporadic [11], and thus an ARG may frequently experience a relaxation of antibiotic selection. As a result, the fitness cost may play an important role on the evolution of an ARG. Yet, it is not known the extent of an ARG's evolution hinges on its fitness cost, and whether the fitness cost can affect an ARG's evolution to new functions.

In Chapter 2, I found the β-lactamase, *bla*$_{TEM\text{-}116}$*, can impose a fitness cost on bacterial hosts, but the cost is host-dependent. While in the lab strain REL606, it does not incur a measurable cost in an antibiotic-free environment, it causes a 10% fitness reduction in the wild isolate M114. This host-dependent cost serves an ideal model to examine if the level of an ARG's cost can influence its molecular evolution and perhaps its ability to evolve a new function. In this chapter, I focus on whether the cost conferred by *bla*$_{TEM\text{-}116}$* can influence the maintenance of an initially introduced amount of genetic variation. To do this, I evolved lines containing diverse variant libraries of *bla*$_{TEM\text{-}116}$* in either REL606 and M114 hosts in the absence of positive selection for the β-lactam resistance. While I found that the genetic pools of the *bla*$_{TEM\text{-}116}$* mutation library differ between the hosts, the results remain inconclusive and require further experimental verification. I discuss the interpretations on the results and proposed experiments to confirm the findings.

4.3. Materials and methods

4.3.1. Construction of the bla$_{TEM\text{-}116}$** library.*

The *bla*$_{TEM\text{-}116}$* mutation library was constructed using error-prone PCR based on the Diversity PCR Random Mutagenesis kit protocol (Clontech). Two mutagenesis conditions were used to generate a predicted mutation density of 2 and 4 mutations per kilobase across the *bla*$_{TEM\text{-}116}$* ORF and promoter. The oligo pair pUA66_cm_F: ccgacgtctaagaaaccatgaattcttctggtcggcgcatagctg and pUA66_cm_R: cctcgtgatacgcctattgcaagcagcagattacgcg were used to generate the error-prone PCR fragments

from the plasmid pmFP- *bla*$_{TEM-116}$*. Because protein expression, in addition to protein sequence, can play a significant role in protein evolution, I included flanking sequence 72 bp upstream and 148 bp downstream of the coding region to give a total insert length of 1081bp [316, 317]. The recipient plasmid backbone was amplified by the oligo pair pUA66_F: aataggcgtatcacgagg and pUA66_EcoRI_R: gaattcatggtttcttagacgtcgg using a high-fidelity DNA polymerase (Hi-fidelity Q5 master mix, NEB). PCR fragments were combined using the NEBuilder® HiFi DNA Assembly kit (NEB) and used to transform the lab strain DH5α. Following overnight incubation, around 14,000 colonies grown on LB supplemented with kanamycin (50 μg/ml) plates were collected by scraping from plates and pooled. This cell suspension represented the master *bla*$_{TEM-116}$* mutation library used in subsequent manipulations.

4.3.2. Short-read sequencing and SNP calling.

The *bla*$_{TEM-116}$* mutation library was isolated from host cells using the QIAprep Spin Miniprep Kit (QIAGEN) and sequenced by Illumina sequencing (HiSeq2500; The Microbial Genomic Sequencing Center, USA). The sequencing results were analysed using *breseq* with POLYMORPHISM mode at default settings except the frequency cutoff for detecting a polymorphism was set to levels as indicated in Results [282].

4.3.3. Experimental evolution of the bla$_{TEM-116}$ library.*

The *bla$_{TEM-116}$** mutation library was isolated from DH5α and used to transform the lab strain REL606 and the natural isolate M114. To reduce the chance of losing diversity, over ten times the number of transformants present in the original library were collected for each strain. To set up the experimental evolution, REL606 and M114 library populations were first cultured in LB supplemented with 50 μg/ml kanamycin overnight at 37°C. The overnight cultures were diluted 1:100, diluted in three tubes of 10 ml Davis minimal broth supplemented with 1000 μg/ml glucose and 50 μg/ml kanamycin and cultured overnight [318]. The overnight cultures were diluted 1:100 to fresh medium and again incubated overnight before being used to start the experimental evolution lines. The evolution experiment was conducted in 10 ml Davis minimal broth supplemented with 1000 μg/ml glucose and 50 μg/ml kanamycin with 1:100 daily transfer to fresh medium for 30 Days. The kanamycin present in the evolution environment ensures maintenance of the plasmid, while not representing any direct selection on the *bla$_{TEM-116}$** gene. Samples of the three replicate lines evolved for each strain were frozen at Day 0 and Day 30. Analysis of these populations were carried out as described in the text.

4.4. Results

4.4.1. Assessment of the random mutation library of bla$_{TEM-116}$.*

To study the influence of fitness cost on the genetic variation of *bla$_{TEM-116}$**, I constructed a random mutation library of *bla$_{TEM-116}$** (Material and Methods). To characterize this library, I

used short-read sequencing to sequence both the *bla*$_{TEM\text{-}116}$* insert and the plasmid vector backbone. Mutations in the plasmid backbone, which was not subject to error-prone PCR, were used to estimate a baseline error rate of the sequencing and basecalling above which *bla*$_{TEM\text{-}116}$* mutations could be identified. Given a mean read depth of 84,366 across the 4.7kb *bla*$_{TEM\text{-}116}$* plasmid, if each transformant in the original library (14,000 colonies) carried a unique mutational genotype, the expected mean frequency of each SNP is 7×10^{-5}. Against this expectation, three conservative frequency thresholds—5×10^{-4}, 1×10^{-4}, and 5×10^{-5}—were used to call variants. The signal-to-noise ratio at each threshold was assessed by comparing mutations in the mutagenized insert to the plasmid backbone.

I found an enrichment of mutations in the mutagenized *bla*$_{TEM\text{-}116}$* region at 46-61-fold excess relative to the backbone (Figure 4.1 and Table 4.1). The number of mutations identified across the backbone do not differ significantly at different polymorphism threshold cutoffs (31, 32, and 34 mutations at cutoffs of 5×10^{-4}, 1×10^{-4}, and 5×10^{-5}, respectively). This indicates that the number of mutations called due to errors in sequencing, read mapping, and variant calling is

Table 4-1. Number of mutations found in the DH5a library under different frequency cutoffs.

5×10^{-4}	mutagenized*	1284	529	46
	non-mutagenized§	3456	31	
1×10^{-4}	mutagenized	1284	727	61
	non-mutagenized	3456	32	
5×10^{-5}	mutagenized	1284	720	57
	non-mutagenized	3456	34	

§ Vector backbone not subject to deliberate mutagenesis.
* Region containing *bla* $_{TEM\text{-}116}$* and flanking regions.

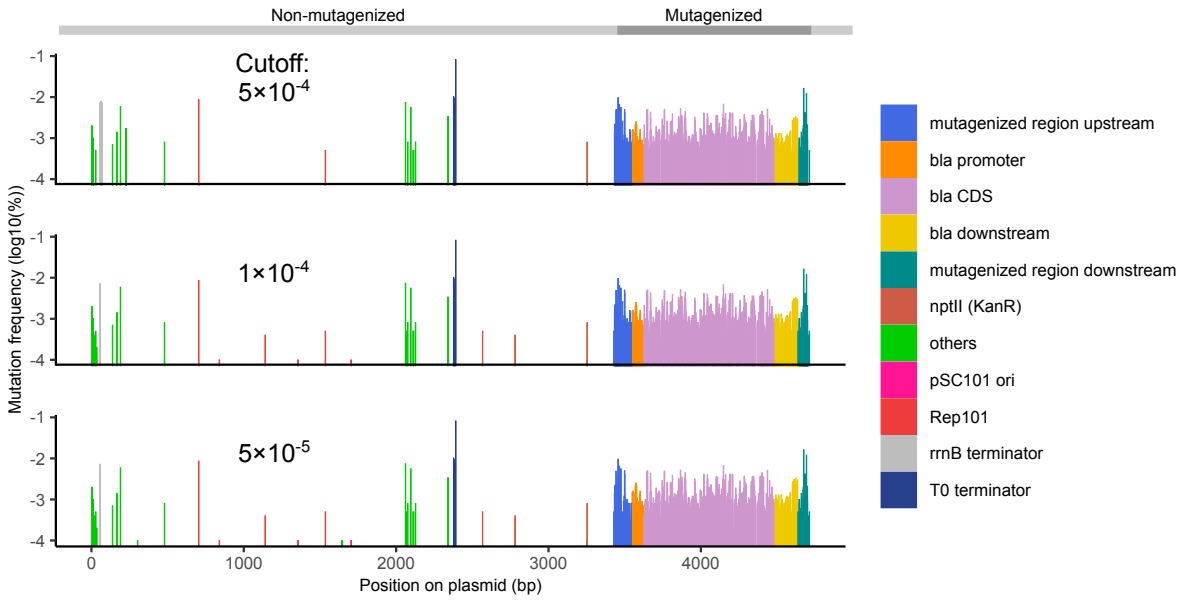

Figure 4.1. Mutations identified at different frequency cutoffs.

The *bla*$_{TEM\text{-}116}$* library constructed in DH5α was short-read sequenced. The POLYMORPHISM mode of breseq was used to call variants with three frequency cutoffs, 5×10^{-4}, 1×10^{-4} and 5×10^{-5}. Mutations found were plotted along the plasmid map and colored according to the gene function or sequence features. The region identified as being mutagenized includes the *bla*$_{TEM\text{-}1\ gene}$ and flanking regions. The non-mutagenized region includes the vector backbone.

low. By contrast, the number of mutations in the mutagenized region increase from 529 (5×10^{-4} cutoff) to 727 (1×10^{-4} cutoff) but stalled at the cutoff of 5×10^{-5} (720 mutations). Given that the number of mutagenized mutations plateaued at the 1×10^{-4} cutoff, I used this cutoff for further analysis. To further reduce the chance of incorporating spurious errors, I removed mutations that were identified at the 1×10^{-4} cutoff but not at the 5×10^{-5} cutoff. Applying this criterion removed 19 out of 759 mutations occurring in the mutagenized region that includes both the *bla*$_{TEM\text{-}116}$* sequence and the vector sequences that flank the *bla*$_{TEM\text{-}116}$* sequence. The resulting mutation set comprises 591 mutations occurring at 586 sites of the *bla*$_{TEM\text{-}116}$* sequence with 75% of mutations in the coding region being either non-synonymous or non-sense mutations (Table 4.2).

Table 4-2. Mutation number and types of the DH5a library under the 1×10-4 cutoff and with dubious mutations removed.

	No. of mutations	No. of mutation types
bla $_{TEM-116}$ * gene (including regulatory regions)	591 (586 sites)	-
bla $_{TEM-116}$ * CDS	-	Non-sense: 40
	-	Synonymous: 117
	-	Non-synonymous: 312

4.4.2. Influence of a host's genetic background on the genetic pool of the bla$_{TEM-116}$ library.*

The fitness effect of a mutation can depend on the basal fitness of its genetic background, where in a fitter background a beneficial mutation has a less effect but a deleterious mutation has a larger effect [313, 314]. To study the influence of fitness cost on the molecular evolution of *bla$_{TEM-116}$**, I examined the maintenance of genetic variation of *bla$_{TEM-116}$** in different genetic background. I transformed the *bla$_{TEM-116}$** library into two *E. coli* strains, REL606 and M114. The ancestral *bla$_{TEM-116}$** imposes a 10% fitness cost on M114, but no measurable cost on REL606. For each strain, the resulting libraries were separated into three independent lines to set up the experimental evolution. The founder lines contain a considerable coverage of the original library and have a high overlap between the two hosts (Table 4.3). Of the mutations identified in the original library, 73% and 59% were in common across the three founding lines started with REL606 and M114, respectively. The REL606 libraries share 76% of mutations with the M114 libraries, and 95% of mutations in the M114 libraries can be found in the

REL606 libraries. The mutations are broadly distributed across the mutagenized region and frequencies are highly correlated across the independent lines (Figure 4.2 and Figure 4.3A). The independent lines were evolved by serial transfer for ~200 generations (30 daily transfers) in medium containing kanamycin to select plasmid-carrying cells but not the β-lactam resistance of $bla_{TEM\text{-}116}$*. The libraries of the selected populations from Day 30 were sequenced and mutations were identified using the same methods and procedures used to characterize ancestral libraries.

Table 4-3. Overlap of $bla_{TEM\text{-}116*}$ library variation in different hosts.

	Host strain	
	REL606	M114
Mutations also found in the original library (DH5α)	549/591 (93%)	530/591 (89%)
Mutations shared among the parallel lines	431/591 (73%)	348/591 (59%)
Parallel mutations found in other host	330/431 (76%)	330/348 (95%)

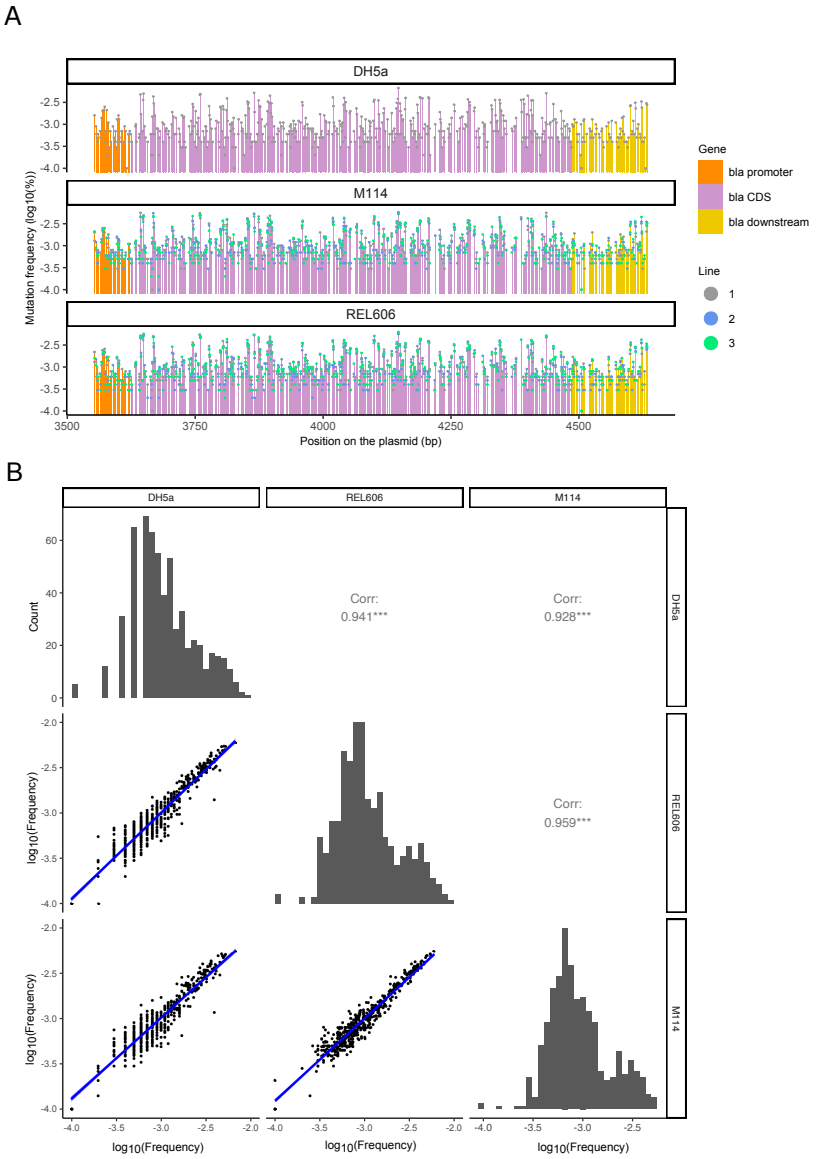

Figure 4.2. Frequency and location distribution of mutations in the ancestral *bla*$_{TEM-116}$* library in different host strains.

(A) Frequency and distribution of *bla*$_{TEM-116}$* mutations. The library was constructed in DH5α and then transformed into REL606 and M114. The libraries of REL606 and M114 were then used to establish three independent lines for the experimental evolution. The independent lines sampled on Day 0 of the evolution experiment were sequenced. Bars indicate sites in the *bla*$_{TEM-116}$* gene and flanking regions where a mutation was identified. They are colored based on the sequence features. Color of dots indicates different lines. **(B)** Correlation of the mutation frequencies between different libraries. The mutations of the independent lines of REL606 and M114 are pooled together to present the initial library diversities. The blue lines represent the linear regression lines. Pearson correlations are reported where "***" indicates a p-value < 0.001. The histograms represent the distribution of mutation frequencies in each library.

If the initial level of the *bla*$_{TEM\text{-}116}$* gene's cost systematically influences the relative fitness effect of genetic variants, the selected mutation pool should differ between the evolved REL606 and M114 libraries. To examine this, I identified the number of unique variant SNPs present in each selected line at Day 30 (Figure 4.3 and 4.4). The M114 lines retained more variation than the REL606 lines (Figure 4.3A). Nearly 78% of the starting mutations (around 400 mutations) were retained in the M114 lines compared to only 26% (127-140 mutations) across the REL606 lines (Figure 4.3A). Moreover, the number of the mutations shared across replicate lines decreased even more in the REL606 lines after the selection (Figure 4.3B and Figure 4.4). Whereas 80.3% (282/351) of the pool of the starting mutations that were shared across the M114 lines were still detected in all lines on Day 30, only 5.3% (23 of 432) remained in the REL606 lines.

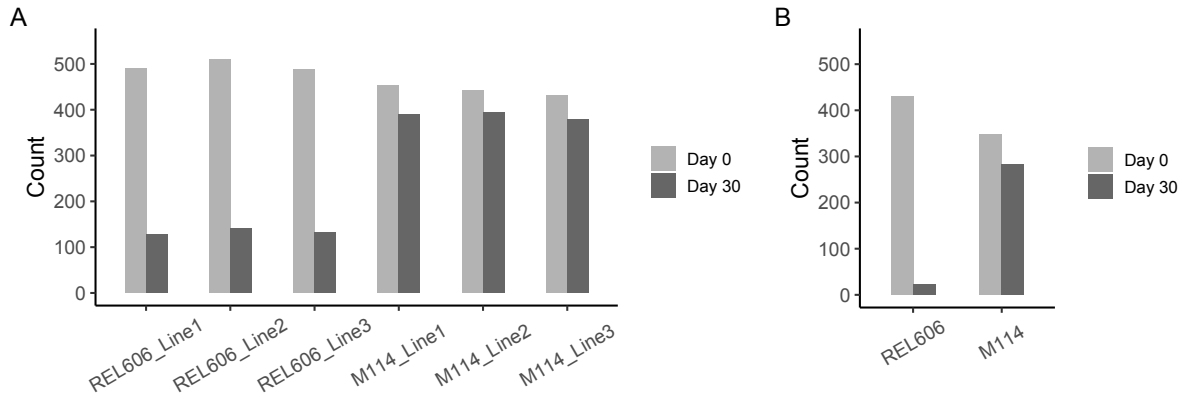

Figure 4.3. Changes of the blaTEM-116* library size in the two *E. coli* strains.

(A) The *bla*$_{TEM\text{-}116}$* library size in independent lines at Day 0 and Day 30. The master *bla*$_{TEM\text{-}116}$* library was transformed into REL606 and M114 and divided into three independent lines. Each line was selected for ~200 generations in medium with antibiotic selection for the plasmid but not for the β-lactam resistance. The *bla*$_{TEM\text{-}116}$* mutations in each line at Day 0 and Day 30 were identified and counted. **(B)** The number of *bla*$_{TEM\text{-}116}$* variants shared across three independent lines of each strain at Day 0 and Day 30.

The smaller number of the *bla*$_{TEM\text{-}116}$* variants remaining in the selected REL606 libraries is consistent with the action of stronger selection for particular variants in that host background. If so, I expected specific mutations to have similar fates in the replicate lines. Yet, comparisons on the mutation frequencies on Day 30 and their fold changes between the independent lines show a higher divergence among REL606 than M114 lines (Figure 4.5). The frequencies of the mutations between the lines initially have high correlations (Figure 4.5A), but the correlations decay more quickly in the REL606 lines (Figure 4.5B). I note that this measure of change is conservative because it only considers mutations that remain identified as present in each pair of lines. Similarly, the fold-change in each mutation's frequency was correlated across the M114 lines, but only between Line 1 and Line 3 of REL606 (Figure 4.5B and 4.5C). The loss of the mutation frequency correlations among the REL606 lines is consistent with mutations were under quantitatively, or even qualitatively different selection pressures. One possible explanation for this is that background selection of chromosomal mutations rather than selection on *bla*$_{TEM\text{-}116}$* mutations was dominant. Further experiments will need to verify the fitness effects of the selected REL606 libraries.

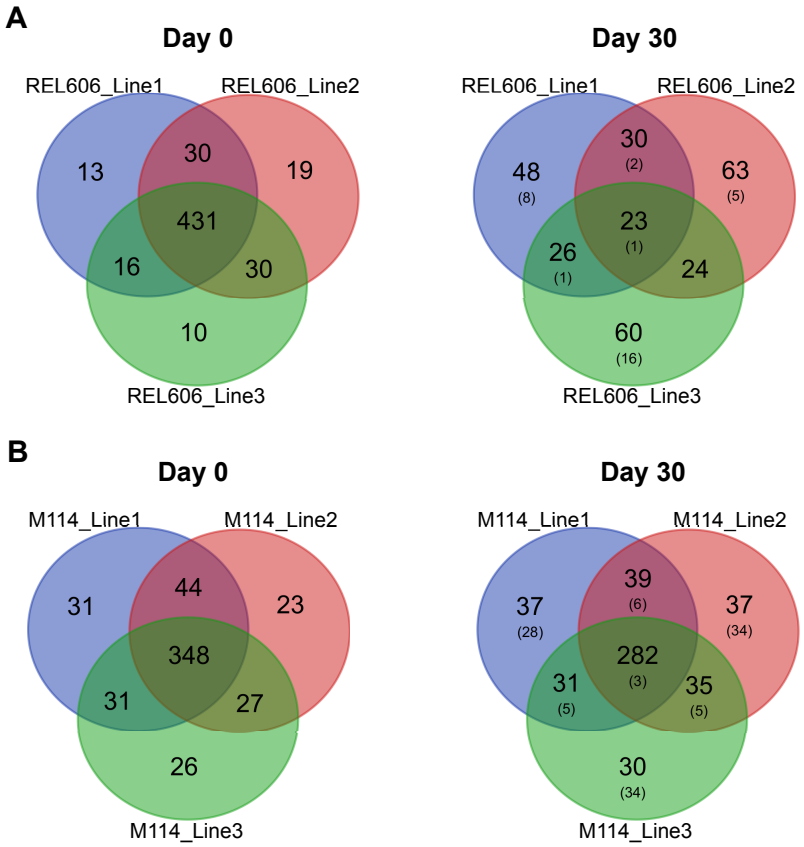

Figure 4.4. Overlapping blaTEM-116* mutations between replicate lines.

(A) The libraries of the REL606 lines. (B) The libraries of the M114 lines. The Venn diagrams show the number of the *bla*$_{TEM-116}$* mutations overlapped between the independent lines. The numbers in parentheses represent the number of the mutations in a category that were not detected on Day 0 but were present on Day 30.

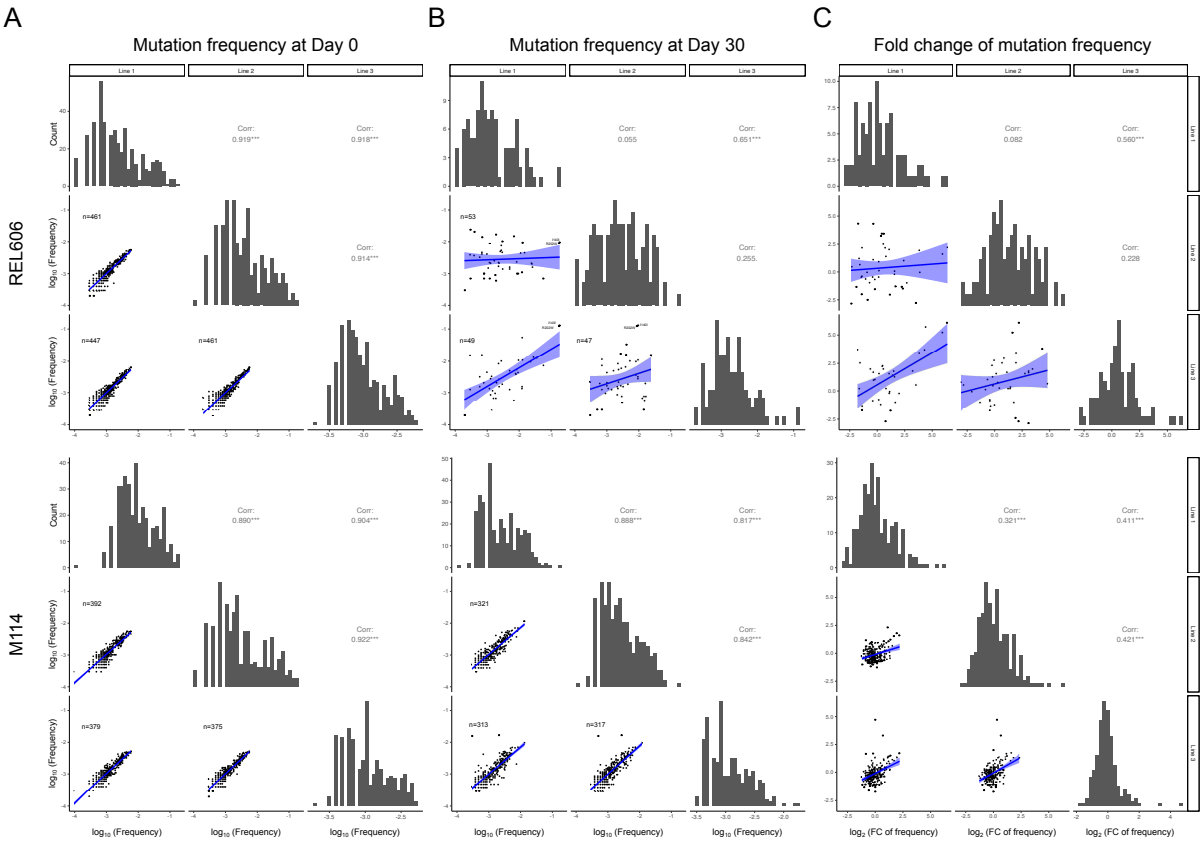

Figure 4.5. Frequency changes of *bla*$_{TEM\text{-}116}$* library variants after selection.

(A) Correlations of mutation frequencies between the three independent lines in REL606 and M114 at Day 0. **(B)** Correlations of mutation frequencies between the three independent lines at Day 30. **(C)** Correlations of frequency changes between the three independent lines in REL606 and M114. The frequency change was the log2 value of mutation frequency on Day 0 over Day 30. Blue lines and shaded regions indicate best fit and 95% CI estimated by linear models, respectively.

4.5. Discussion

Evolution of antibiotic modifying enzymes can quickly make ineffective new antibiotics with the same chemical core structure to those previously targeted. One mechanism that may influence the rise of a new resistance in an ARG is via an accumulation of CGV. Several factors

can affect the accumulation of CGV including the basal fitness of a genetic background that can shape the fitness effect of a newly arising mutations. In this chapter, I used experimental evolution to investigate if the fitness cost of the β-lactamase, BlaTEM-116*, can affect maintenance of a pool of genetic variation. After 30 days of serial transfer without selection for β-lactam resistance, the genetic variation of *bla*$_{TEM\text{-}116}$* was smaller in an *E. coli* host without a *bla*$_{TEM\text{-}116}$* cost. Yet, the current experimental results and designs fail to distinguish two possible mechanisms that can lead to the reduced genetic variation of *bla*$_{TEM\text{-}116}$*—selections on the *bla*$_{TEM\text{-}116}$* variants and selections on the chromosomal mutations. Below, I discuss limitations of the experimental designs and propose modifications to improve future experiments.

In this chapter, I expressed a *bla*$_{TEM\text{-}116}$* library comprised of randomly generated variants in two bacterial hosts and imposed a simple selection which is expected to result in increasing frequency of any variants conferring a fitness benefit relative to the mean of the entire library pool. While random genetic drift stochastically affects the frequency of mutations when they are rare, the starting mutation frequencies in my experiment were similar across the two host strains (Figure 4.2B), allowing a meaningful comparison of library dynamics on the basis of host identity. The percentage of shared mutations in the REL606 lines has decreased from 85~89% on Day 0 to 16~18% on Day 30 in contrast to the relatively stable numbers in the M114 lines (from 77~80% to 72~75%. Figure 4.4). Moreover, the frequency of mutations and their fold change are less similar among REL606 lines (Pearson's r=0.05-0.65 and r= 0.08-0.56) than among M114 lines (r=0.82-0.89 and r= 0.32-0.42). The greater loss of shared

mutations in REL606 lines suggests either the purifying selection on the REL606 *bla*$_{TEM\text{-}116}$* library was stronger than in M114 or other evolutionary mechanisms dominated the selection process.

It has been shown that the fitness distribution of a whole-gene saturation *bla*$_{TEM\text{-}1}$ library in the *E. coli* lab strain DH10B in the absence of β-lactam is normally distributed with a mean fitness effect of -0.9% and a variance of 7% [306]. In that study, no mutations with a substantially different fitness effect relative to the WT gene were identified, which raises the possibility that my libraries did not contain variants with different fitness effects. If so, *bla*$_{TEM\text{-}116}$* library dynamics in my experiment might be determined mostly be genetic drift. Suggesting against this, my finding of a strong background dependence for the WT allele demonstrates that library fitness distributions cannot be simply extrapolated across host strains. Moreover, my library contained alleles with multiple mutations in contrast to the single amino acid substitutions comprising the library of Stiffler et al. (2015).

Another explanation to the uncorrelated loss of *bla*$_{TEM\text{-}116}$* mutations among REL606 lines is the action of selection on newly arising chromosomal mutations. The starting libraries of REL606 and M114 comprised a collection of over 140,000 independent transformants with approximately ~200 generations of independent propagation prior to the beginning of the evolution experiment. Chromosomal mutations occurring during propagation and/or that arose during the evolution experiment may obscure the effect of selection on the *bla*$_{TEM\text{-}116}$* mutations and drive uncorrelated changes of *bla*$_{TEM\text{-}116}$* mutations among the replicate lines.

This possibility could be tested by measuring fitness changes of the selected library in a clean ancestral background to break the linkage of chromosomal and $bla_{TEM\text{-}116}$* mutations. I intend to do this experiment in the near future.

There are several aspects of my experiment that could be changed to more effectively test for a role of mean fitness effect on evolvability of a gene. First, mutations are the source of new phenotypes and it is critical to maximize the mutation diversity to capture the mutations with a desired phenotype. In this study, the $bla_{TEM\text{-}116}$* library size (312 non-synonymous substitutions) is small even compared to the size of the whole-gene saturation $bla_{TEM\text{-}1}$ library used in Stiffler et al., 2015 (4,997 single amino acid substitutions in each amino acid site). To capture mutations that confer a range of fitness effects, the library size can be increased. Second, the selection on the fitness cost of the $bla_{TEM\text{-}116}$* mutations is purely based on bacterial growth without any strong environmental stress. Any mutations, including chromosomal mutations, that improve cellular growth can be selected. Thus, this selection is prone to the confounding effect of background selection. It is necessary either to remove this effect, for example by using ancestral strains that are pre-adapted to the selection environment, or to verify the fitness effects of the mutations of the selected library after the selection. Below, I propose strategies to increase library diversity and to optimize the selection processes.

Library construction and identification. Systematic study of protein evolutionary potential by natural selection requires a large and random mutational library. Error-prone PCR based library construction is a useful technique, but it suffers from limitations, such as the transition/transversion mutational bias of a DNA polymerase and the inheritance of mutations

from the previous PCR cycle, leading to mutation enrichment [319]. Instead of using an error-prone PCR based library, deep mutational scanning approaches that utilize synthesized oligonucleotides to achieve mutation saturation and eliminate mutational bias have been widely used in directed protein evolution [320, 321]. For instance, the experimental approach, PFunkle, has been proven to generate a single codon replacement library for bla_{TEM-1} consisting of 97% of possible amino acid substitutions (287 codons × 63 possible codon substitutions) [322]. In addition to error-prone PCR, synthetic libraries are good alternatives to generate a random mutation library.

In this study, I used short-read sequencing on pooled samples to estimate the frequency of mutations. However, short-read sequencing cannot distinguish linkage between mutations on the DNA fragments longer than ~300bp. A better approach is to generate a mutation library with random barcodes. The original library can be sequenced by long-read sequencing, such as single-molecule real-time (SMRT) sequencing [323], to associate the barcodes to individual variants. Later, the selected libraries can be sequenced by short read sequencing to monitor the frequency of individual barcodes.

Optimization of the selection processes. To decrease the effect of unwanted selection on chromosomal mutations and increase the effect of selection for $bla_{TEM-116}^{*}$ mutations with a weak fitness effect, I plan to retransform the selected library into the ancestral background and include an additional mutagenesis on the selected library after each selection step but before the retransformation. A "cleaning" step can be added in the evolution-selection cycle to remove unwanted chromosomal mutations as practiced in Zheng et al., 2019 [308]. To further decrease

confounding factors due to the genetic divergence between REL606 and M114, I will conduct the evolution in nearly isogenic hosts. In Chapter 3, I found the interaction between $bla_{TEM\text{-}116}$* and a phage gene $relA_{P1}$ is the main cause of the $bla_{TEM\text{-}116}$*'s fitness cost. Hence, the directed protein evolution can be conducted in isogenic cells, for instance, REL606 with or without $relA_{P1}$, to reduce the influences other than the $bla_{TEM\text{-}116}$*- $relA_{P1}$ interaction, for instance, the difference in the spontaneous mutation rates between bacterial hosts.

ARGs are frequently transferred between bacteria and subject to fluctuating selections for the antibiotic resistance. It is less known how the relaxation of antibiotic selection and the host genetic background on the evolution of an ARG. Although this study cannot be completed with this doctoral study, the preliminary results yielded an invaluable guideline to optimize the experimental procedures. The proposed experimental modifications should be able to reduce the aforementioned complications and to proceed the research.

Chapter 5. Conclusion

5.1. Overview

The continuing rise of antibiotic resistance and its rapid spread in pathogenic bacteria have outpaced the development of new drugs and exhausted the list of existing antibiotics [8]. Combating antibiotic resistance in pathogens requires a holistic approach that looks beyond the health of humans [11]. For instance, while the rise and spread of antibiotic resistance in environmental bacteria do not directly threaten human health, the gene flow between environmental bacteria and human pathogens may increase the chance of a gain of antibiotic resistance in pathogenic bacteria [324]. Given that antibiotic resistance exists in bacteria long before the discovery and application of antibiotics and is also commonly found in remote areas with little human activities [79, 155, 160], knowledge of the evolution and spread of antibiotic resistance in bacteria found in natural environments and environments with little antibiotic exposure can provide insight into a broader set of selective pressures that promote the resistance in pathogenic bacteria. In this doctoral study, I aim to close to the gap in our knowledge of the antibiotic resistance between clinical settings and natural environments. More specifically, my study focuses on the maintenance and evolution of antibiotic resistant genes (ARGs) in bacteria without antibiotic selection.

Chapter 2 quantifies the variation of the fitness cost of six ARGs in twelve *Escherichia* isolates and investigates the evolutionary consequence of the variation that was found. My results show that, while on average, the six ARGs impose little fitness cost on bacterial hosts, the interaction

between an ARG and its bacterial host underlies the considerable host-specific variation of the fitness cost. Using an experimental evolution, I find that these interactions can lead to evolutionary dynamics of an ARG-carrying plasmid that vary in a host-dependent manner. This finding highlights that the spread and evolution of antibiotic resistance may not be uniform across strains. **Chapter 3** exploits whole-genome sequencing on the experimentally evolved lines from **Chapter 2** to identify the genetic basis underlying the host-dependent cost of a β-lactamase. I find a negative genetic interaction between a phage-encoded gene and the β-lactamase leads to a host-dependent cost of the β-lactamase. In **Chapter 4**, I evaluate the importance of bacterial hosts on the genetic variation of a β-lactamase. To study if the host-dependent cost of the β-lactamase can influence its genetic variation, I delivered a random mutagenized β-lactamase library into two different hosts and selected the libraries in the absence of β-lactam. Although the resulting genetic variation differ between the two hosts, the results remain inconclusive and need a further confirmation.

Taken together, my findings demonstrate that even without antibiotic selection, natural selection continues to shape the evolution of antibiotic resistance and its maintenance in bacteria populations. It emphasizes that a comprehensive understanding of antibiotic resistance in natural environments is required to contain and combat antibiotic resistance in pathogens. Below, I discuss the implications of my findings and some of the limitations of my work. I finish with suggested directions for future studies.

5.2. The generality of the fitness cost of an ARG and the distribution of the cost in bacteria

An ARG is not wired into the genetic network of a bacterial genome in contrast to an antibiotic resistant mutation (ARM) which often occurs in genes essential for bacterial growth. As a result, the fitness cost of an ARG is expected to be smaller than the cost of an ARM. Indeed, a meta-analysis of studies that estimated the fitness cost of resistance found that on average the fitness cost of ARMs (20% of fitness reduction) is larger than that of entire ARG-encoding plasmids (10% of fitness reduction) [117]. The smaller cost of ARG- encoding plasmids suggests that the cost of individual ARGs may be also smaller. Consistent with this expectation, **Chapter 2** shows that the average fitness costs of four tested ARGs among different *Escherichia* strains are small (1.1-3.3%) and two other ARGs do not impose a measurable cost.

The sometimes small cost of an ARG suggests that purifying selection against its carriage by a host might be weak, raising the possibility that ARGs might stably persist in bacterial populations in the absence of antibiotic selection. Nevertheless, epidemiological surveillance finds a negative correlation between the antibiotic use and the prevalence of antibiotic resistance in pathogens [325]. The correlation between the antibiotic restriction and the reduced resistance prevalence suggests that the biological fitness cost of antibiotic resistance is sufficient to play a role in its success. Therefore, an understanding on the physiological cost of antibiotic resistance may provide a mechanistic explanation on the prevalence of antibiotic resistance in bacteria.

One major limitation of **Chapter 2** is that the results are drawn from a small subset of genes and bacterial strains. The small sample size prevents a generalization of the findings to other ARGs and bacterial hosts. Moreover, the six candidate genes are selected based on their high prevalence in published *Escherichia coli* genomes. The high prevalence of an ARG could partly be due to its small fitness cost in bacteria. Hence, the results drawn from the six candidate genes may not reflect the fitness cost of a randomly chosen ARG. In addition, the small sample size of the *Escherichia* strains restricts the statistical power to examine the statistical association between the variation of the fitness cost of an ARG in bacterial strains and the phylogenetic relationship of the bacterial hosts. Besides, a lack of knowledge on the historical exposure of the selected bacteria to antibiotics makes it challenging to determine if prior compensatory evolution has shaped the response of a strain to the carriage of an ARG. Future studies could increase the sample sizes of ARGs and bacterial strains/species to explore the generality of the fitness cost of an ARG and its variability between bacterial hosts.

While **Chapter 2** only considers the fitness cost of individual ARGs, multiple drug resistant (MDR) bacteria are common in clinics and in nature. It is not possible to simply extrapolate the effects of individual ARGs to their effects in combination, predicting the fitness cost of MDR and its relation to the MDR prevalence. Future experimental work on the cost of multiple ARGs may reveal the influence of fitness cost on the spread of MDR in bacteria and on the formation of MDR [326, 327].

5.3. The variation of the fitness cost of an ARG on the evolution and maintenance of antibiotic resistance in a microbial community

Antibiotics are used to kill or inhibit the growth of pathogenic bacteria, yet unavoidably, they also reduce the microbial diversity of normal microbial flora and increase the prevalence and persistence of antibiotic resistance in human microbiomes [328]. The abundance of ARGs is increased in human gut microbiomes after antibiotic treatments, and this change can persist months and even years after the treatments [329]. Even though commensal bacteria do not cause bacterial infections (except opportunistic bacteria), they are potential reservoirs of ARGs and may transfer the carried ARGs to incoming pathogens [330, 331]. Given that ARGs can persist in a microbiome long after an antibiotic treatment, knowledge of the ecological and evolutionary mechanisms that allow the persistence of resistance without constant antibiotic selection may facilitate the invention of intervening strategies to clear ARGs or antibiotic resistant bacteria from microbiomes.

ARGs have undergone frequent transfer between bacterial hosts, yet it is less known if they impose different costs across hosts and, even if so, whether this has any biological consequence. In fact, even the variation of the fitness effect of a horizontal gene transfer (HGT) gene among bacteria is generally unknown, despite theoretical and simulation studies show that the distribution of the fitness effect of a HGT among bacteria can influence its frequency among bacterial populations and can shape the size and diversity of bacterial accessory genomes [332]. So far, researches that are most relevant to the variation of an ARG's cost are studies that measure the fitness cost of plasmids in different bacterial hosts [145, 194]. These studies found

that not all bacteria suffer a fitness reduction from the plasmid carriage. Instead, some bacterial hosts can benefit from the carriage even without a positive selection for the plasmid-borne adaptive traits. Although a plasmid can impose a fitness cost on some hosts and the cost disfavours the plasmid carriage, it has been shown that the variability in the fitness cost of a plasmid, on the contrary, can facilitate its persistence in a bacterial community [145, 194].

Chapter 2 investigates the short-term and long-term consequences of the variation of the cost of the β-lactamase $bla_{TEM\text{-}116}$* in clonal populations. I found the ecological and evolutionary dynamics of a $bla_{TEM\text{-}116}$* plasmid can depend on the level of the fitness cost of $bla_{TEM\text{-}116}$*. In the absence of positive selection for the plasmid, the $bla_{TEM\text{-}116}$* plasmid quickly declined in frequency in a clonal bacterial population in which the host initially suffers a large fitness reduction from the plasmid carriage. In time, however, compensatory mutations occur in the remaining plasmid-carriers, slow down the plasmid loss, and change the plasmid dynamics. In contrast, the bacterial hosts that suffer zero or little fitness reduction from the $bla_{TEM\text{-}116}$* plasmid carriage, have slower plasmid loss yet have no changes in the plasmid dynamics in the long-term. The findings in **Chapter 2** indicates that the variation of an ARG's cost can shape the evolutionary dynamics of the ARG in a bacterial population in a host-dependent manner. The results suggest that the variation of an ARG's cost among hosts may also change the evolutionary dynamics of the ARG in a bacterial community. A follow-up study can investigate if the dynamics of an ARG within clonal populations can predict its short-term and long-term persistence in a mixed population and if the variability of the fitness cost of an ARG can influence its long-term persistence in a bacterial community.

5.4. The origin, significance and application of genetic conflict between a phage and beta-lactamase

Genetic interactions between accessory genes can shape the bacterial pangenome [333, 334]. For instance, negative genetic interactions between accessory genes can select against their co-occurrence in the same bacterial genome, and, in contrast, positive interactions between accessory genes can promote their co-occurrence [335]. It has been shown that in 40 *Pseudomonas* species, around 87% of common accessory genes are involved in more than one gene association/dissociation relationship [333]. Not yet clear is the evolutionary process through which a negative interaction between accessory genes leads to a particular pattern of gene avoidance in the same genome. **Chapter 3** presents an example of the negative interaction between two accessory genes—the P1 phage-encoded $relA_{P1}$ and the β-lactamase $bla_{TEM\text{-}116}$*, and this gene pair can serve as a model to study the interaction dynamics of negatively interacting accessory genes in a bacterial population. Because of the medical importance of β-lactamases, knowledge of the molecular mechanisms underlying the negative interaction and of the interaction dynamics of the phage gene and $bla_{TEM\text{-}116}$* may give insight on the prevalence and distribution of the β-lactamase in bacterial genomes. Furthermore, the negative genetic interaction may encourage the development of bio-control strategies to facilitate the elimination of the β-lactamase in bacterial populations. Below, I discuss potential research directions derived from **Chapter 3**.

The molecular mechanism of the negative interaction between *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$* remains unclear. The potential biological functions of *relA*$_{P1}$ to the phage and the possible mechanistic causes of the negative interaction between *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$* are discussed in **Chapter 3** and can be further experientially examined. Additionally, one can look into the interaction between *relA*$_{P1}$ and other non-type A β-lactamases that are not examined in this thesis, such as the type C and D serine β-lactamases and the type B metallo-β-lactamases [336]. The four types of β-lactamases differ in their substrate spectrum, protein sequences and the mechanisms of β-lactam hydrolysis [336]. A evaluation on the interaction between *relA*$_{P1}$ and other types of β-lactamases may narrow down the mechanistic causes of the negative interaction of *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$*. Moreover, it can serve as prior knowledge on examining the efficacy and the general application of a *relA*$_{P1}$-based bio-control application, which is discussed below.

The negative interaction between *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$* may provide a good system in which to study if a specific negative genetic interaction leads to gene avoidance in bacterial genomes. Tracking the temporal and evolutionary dynamics of the two genes in a bacterial population may reveal the influence of the gene antagonism on their distribution in bacterial genomes. In addition, the low incidence of *relA*$_{P1}$ in P1 variants suggests that *relA*$_{P1}$ is likely non-essential to P1 reproduction, but it may be conditionally beneficial to the phage. For instance, the expression of the RelA homolog in the Phrann prophage is induced in mycobacteria only when the bacterial host is infected by lytic phages [3]. The Phrann-encoded RelA is served as a viral defence system to protect the hosts and ensure the lysogeny. It is possible that the negative interaction between *relA*$_{P1}$ and *bla*$_{TEM\text{-}116}$* may underlie a conflict between the phage and the

β-lactam resistant bacteria. As seen in **Chapter 3**, all the examined evolved *bla*$_{TEM-116}$*-carrying cells with compensatory mutations have either lost the phage or gained a loss-of-function *relA*$_{P1}$. Exploring the conditions, such as the levels of nutrients that influence the induction of the phage lytic cycle and maintenance of the lysogeny or the concentrations of β-lactam that influence the fitness effect of *bla*$_{TEM-116}$*, that favor the stable existence of either the phage or the *bla*$_{TEM-116}$* or their co-existence in a bacterial host may reveal ecological conditions that influence their distribution in bacterial genomes [335].

The stagnation of new antibiotic development and the emergence of new antibiotic resistance have driven the development of new strategies designed to either slow the spread or prevent the rise of antibiotic resistance in bacteria [337, 338]. For instance, the cross-resistance between antibiotics—i.e. the sensitivity of Drug A-resistant bacterium to Drug B can influence the rise of the resistance to Drug B and vice versa—may be used to contain antibiotic resistance in the long-term by following alternating antibiotic treatments [20, 339, 340]. Other approaches against pathogens and bacteria, such as the exploitation of bacteriophages on controlling food-borne pathogens and environmental bacteria in wastewater plants or as alternative bactericidal agents, can also have a long-term influence on reducing antibiotic resistance in bacteria [341-343]. Given the negative genetic interaction of *relA*$_{P1}$ with *bla*$_{TEM-116}$* leading to an evolutionary incompatibility (as seen in **Chapter 3**), *relA*$_{P1}$ may be exploited as a bio-control agent to facilitate the purge of the *bla*$_{TEM}$ in a bacterial population in the absence of β-lactams. The information from the dynamics of *relA*$_{P1}$ and *bla*$_{TEM-116}$* in a population and the genetic

interaction of *relA*$_{P1}$ and other types of β-lactamases as proposed above can be exploited to evaluate the possibility and the effectiveness of the biocontrol application.

5.5. The influence of fitness cost on the evolution of the TEM beta-lactamase

Conventional wisdom in protein evolution is that a preservation of the original function of a protein constrains its ability to gain new functions [344]. It is assumed that the original function of a protein is critical for the organismal survival, and consequently, purifying selection for the original function constrains the functional divergence of a protein. To explain the diversity of modern proteins, gene duplication that allows the preservation of the original function in one gene copy yet functional divergence in another has been proposed as a mechanism of protein evolution and diversification [345]. Yet, for a conditionally adaptive gene, such as an ARG, its function only contributes to the organismal fitness in certain conditions, i.e. antibiotic containing environments. Thus, purifying selection for the functionality of a conditionally adaptive ARG fluctuates with the concentration of the cognate antibiotic in environments. It is less studied how a protein evolves and diversifies when selection for its original function fluctuates or is often relaxed.

Chapter 4 investigates if a relaxation of selection on the β-lactam resistance phenotype of the β-lactamase *bla*$_{TEM-116}$* can influence the maintenance of its genetic variation. I exploit a combination of random mutagenesis to create a library of gene variants and a natural selection experiment to compare the genetic variation of the *bla*$_{TEM-116}$* library in two bacterial hosts that have contrasting fitness costs due to *bla*$_{TEM-116}$*. Although the time constraints of this

doctoral study prevents me from verifying the results and optimizing the experiments, I have proposed future experiments to address these issues in **Chapter 4**. β-lactamases are known for their promiscuous activity to the β-lactam drugs and β-lactamase inhibitors [336]. Hence, the study in **Chapter 4** can be extended to examine the influence of the fitness cost of $bla_{TEM\text{-}116}$* on its cryptic genetic variation— a subset of genetic variation that possesses new functions, i.e. resistance to other classes of β -lactams [307]. Investigating the cryptic genetic variation of $bla_{TEM\text{-}116}$* can reveal the effect of a relaxation of selection for the β-lactam resistance of $bla_{TEM\text{-}116}$* on its evolvability.

5.6. Conclusion

The goal in my doctoral work is to understand antibiotic resistance in bacteria from an evolutionary perspective. To this end, I measure the fitness costs of ARGs and determine how they are influenced by different bacterial hosts. **Chapter 2** quantifies the variation of the cost and studies the evolutionary consequences of the variation. **Chapter 3** identifies the genetic basis underlying the variation of the fitness cost of an ARG. **Chapter 4** investigates the fitness cost on the molecular evolution of an ARG. These studies are by no means comprehensive to understand how antibiotic resistance evolves and persists in bacteria, yet they demonstrate clear evidences for the ARG-host interactions that must be understood to predict the likely success of ARGs in natural populations. In addition, I uncover findings that are important to investigate further, for example, the extent to which the negative genetic interaction between a phage gene

and a β-lactamase uncovered in **Chapter 3** could be a genetic barrier that constrains the spread of β-lactamase in bacterial populations.

1

Reference

1. Wang, Y., et al., *Benefit of transferred mutations is better predicted by the fitness of recipients than by their ecological or genetic relatedness.* Proc Natl Acad Sci U S A, 2016. **113**(18): p. 5047-52.
2. Venturini, C., et al., *Diversity of P1 phage-like elements in multidrug resistant Escherichia coli.* Sci Rep, 2019. **9**(1): p. 18861.
3. Dedrick, R.M., et al., *Prophage-mediated defence against viral attack and viral counter-defence.* Nat Microbiol, 2017. **2**: p. 16251.
4. Partridge, S.R., et al., *Mobile Genetic Elements Associated with Antimicrobial Resistance.* Clin Microbiol Rev, 2018. **31**(4).
5. Hogg, T., et al., *Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response [corrected].* Cell, 2004. **117**(1): p. 57-68.
6. Djordjevic, S.P. and B.S. Morgan, *A One Health genomic approach to antimicrobial resistance is essential for generating relevant data for a holistic assessment of the biggest threat to public health.* Microbiology Australia, 2019. **40**(2): p. 73-76.
7. Hutchings, M.I., A.W. Truman, and B. Wilkinson, *Antibiotics: past, present and future.* Curr Opin Microbiol, 2019. **51**: p. 72-80.
8. Butler, M.S., et al., *Analysis of the Clinical Pipeline of Treatments for Drug-Resistant Bacterial Infections: Despite Progress, More Action Is Needed.* Antimicrob Agents Chemother, 2022. **66**(3): p. e0199121.
9. Mackenzie, J.S. and M. Jeggo, *The One Health Approach-Why Is It So Important?* Trop Med Infect Dis, 2019. **4**(2).
10. Ghai, R.R., et al., *A generalizable one health framework for the control of zoonotic diseases.* Sci Rep, 2022. **12**(1): p. 8588.
11. Larsson, D.G.J. and C.F. Flach, *Antibiotic resistance in the environment.* Nat Rev Microbiol, 2022. **20**(5): p. 257-269.
12. Blair, J.M., et al., *Molecular mechanisms of antibiotic resistance.* Nat Rev Microbiol, 2015. **13**(1): p. 42-51.
13. Campbell, E.A., et al., *Structural mechanism for rifampicin inhibition of bacterial rna polymerase.* Cell, 2001. **104**(6): p. 901-12.
14. Ramaswamy, S. and J.M. Musser, *Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update.* Tuber Lung Dis, 1998. **79**(1): p. 3-29.

15. Vedithi, S.C., et al., *Structural Implications of Mutations Conferring Rifampin Resistance in Mycobacterium leprae.* Sci Rep, 2018. **8**(1): p. 5016.

16. Shaheen, A., et al., *Mutational Diversity in the Quinolone Resistance-Determining Regions of Type-II Topoisomerases of Salmonella Serovars.* Antibiotics (Basel), 2021. **10**(12).

17. Wohlkonig, A., et al., *Structural basis of quinolone inhibition of type IIA topoisomerases and target-mediated resistance.* Nat Struct Mol Biol, 2010. **17**(9): p. 1152-3.

18. Ning, Q., et al., *Predicting rifampicin resistance mutations in bacterial RNA polymerase subunit beta based on majority consensus.* BMC Bioinformatics, 2021. **22**(1): p. 210.

19. Portelli, S., et al., *Prediction of rifampicin resistance beyond the RRDR using structure-based machine learning approaches.* Sci Rep, 2020. **10**(1): p. 18120.

20. Yelin, I., et al., *Personal clinical history predicts antibiotic resistance of urinary tract infections.* Nat Med, 2019. **25**(7): p. 1143-1152.

21. Choi, U. and C.R. Lee, *Distinct Roles of Outer Membrane Porins in Antibiotic Resistance and Membrane Integrity in Escherichia coli.* Front Microbiol, 2019. **10**: p. 953.

22. Yoshida, T., et al., *Transcription regulation of ompF and ompC by a single transcription factor, OmpR.* J Biol Chem, 2006. **281**(25): p. 17114-17123.

23. Lou, H., et al., *Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant E. coli.* PLoS One, 2011. **6**(10): p. e25825.

24. Fernandez, A., et al., *Expression of OXA-type and SFO-1 beta-lactamases induces changes in peptidoglycan composition and affects bacterial fitness.* Antimicrob Agents Chemother, 2012. **56**(4): p. 1877-84.

25. Ma, Z., et al., *A MexR Mutation Which Confers Aztreonam Resistance to Pseudomonas aeruginosa.* Front Microbiol, 2021. **12**: p. 659808.

26. Poole, K., et al., *Expression of the multidrug resistance operon mexA-mexB-oprM in Pseudomonas aeruginosa: mexR encodes a regulator of operon expression.* Antimicrob Agents Chemother, 1996. **40**(9): p. 2021-8.

27. Vettoretti, L., et al., *Efflux unbalance in Pseudomonas aeruginosa isolates from cystic fibrosis patients.* Antimicrob Agents Chemother, 2009. **53**(5): p. 1987-97.

28. Poole, K., *Efflux pumps as antimicrobial resistance mechanisms.* Annals of Medicine, 2007. **39**(3): p. 162-176.

29. Novozhilov, A.S., G.P. Karev, and E.V. Koonin, *Mathematical modeling of evolution of horizontally transferred genes.* Mol Biol Evol, 2005. **22**(8): p. 1721-32.

30. Woods, L.C., et al., *Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation.* Proc Natl Acad Sci U S A, 2020. **117**(43): p. 26868-26875.

31. Andersson, D.I. and D. Hughes, *Evolution of antibiotic resistance at non-lethal drug concentrations.* Drug Resist Updat, 2012. **15**(3): p. 162-72.
32. Kohanski, M.A., et al., *A common mechanism of cellular death induced by bactericidal antibiotics.* Cell, 2007. **130**(5): p. 797-810.
33. Gutierrez, A., et al., *beta-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity.* Nat Commun, 2013. **4**: p. 1610.
34. Kohanski, M.A., M.A. DePristo, and J.J. Collins, *Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis.* Mol Cell, 2010. **37**(3): p. 311-20.
35. Sniegowski, P.D., P.J. Gerrish, and R.E. Lenski, *Evolution of high mutation rates in experimental populations of E. coli.* Nature, 1997. **387**(6634): p. 703-5.
36. Giraud, A., et al., *The rise and fall of mutator bacteria.* Curr Opin Microbiol, 2001. **4**(5): p. 582-5.
37. Oliver, A., et al., *High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection.* Science, 2000. **288**(5469): p. 1251-4.
38. Komp Lindgren, P., et al., *Prevalence of hypermutators among clinical Acinetobacter baumannii isolates.* J Antimicrob Chemother, 2016. **71**(3): p. 661-5.
39. Denamur, E., et al., *High frequency of mutator strains among human uropathogenic Escherichia coli isolates.* J Bacteriol, 2002. **184**(2): p. 605-9.
40. Schaenzer, A.J. and G.D. Wright, *Antibiotic Resistance by Enzymatic Modification of Antibiotic Targets.* Trends Mol Med, 2020. **26**(8): p. 768-782.
41. Skinner, R., E. Cundliffe, and F.J. Schmidt, *Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics.* J Biol Chem, 1983. **258**(20): p. 12702-6.
42. Long, K.S., et al., *The Cfr rRNA methyltransferase confers resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A antibiotics.* Antimicrob Agents Chemother, 2006. **50**(7): p. 2500-5.
43. Donhofer, A., et al., *Structural basis for TetM-mediated tetracycline resistance.* Proc Natl Acad Sci U S A, 2012. **109**(42): p. 16900-5.
44. Trieber, C.A., et al., *Ribosomal protection from tetracycline mediated by Tet(O): Tet(O) interaction with ribosomes is GTP-dependent.* Biol Chem, 1998. **379**(7): p. 847-55.
45. Vetting, M.W., et al., *Structure of QnrB1, a plasmid-mediated fluoroquinolone resistance factor.* J Biol Chem, 2011. **286**(28): p. 25265-73.
46. Hooper, D.C. and G.A. Jacoby, *Mechanisms of drug resistance: quinolone resistance.* Ann N Y Acad Sci, 2015. **1354**: p. 12-31.
47. Burchall, J.J., L.P. Elwell, and M.E. Fling, *Molecular mechanisms of resistance to trimethoprim.* Rev Infect Dis, 1982. **4**(2): p. 246-54.

48. Silver, S., *Bacterial resistances to toxic metal ions--a review.* Gene, 1996. **179**(1): p. 9-19.

49. Kermani, A.A., et al., *Guanidinium export is the primal function of SMR family transporters.* Proc Natl Acad Sci U S A, 2018. **115**(12): p. 3060-3065.

50. Bay, D.C. and R.J. Turner, *Diversity and evolution of the small multidrug resistance protein family.* BMC Evol Biol, 2009. **9**: p. 140.

51. Li, X.Z. and H. Nikaido, *Efflux-mediated drug resistance in bacteria: an update.* Drugs, 2009. **69**(12): p. 1555-623.

52. Thaker, M., P. Spanogiannopoulos, and G.D. Wright, *The tetracycline resistome.* Cell Mol Life Sci, 2010. **67**(3): p. 419-31.

53. Webber, M.A. and L.J. Piddock, *The importance of efflux pumps in bacterial antibiotic resistance.* J Antimicrob Chemother, 2003. **51**(1): p. 9-11.

54. Saenger, W., et al., *The Tetracycline Repressor-A Paradigm for a Biological Switch.* Angew Chem Int Ed Engl, 2000. **39**(12): p. 2042-2052.

55. De Pascale, G. and G.D. Wright, *Antibiotic resistance by enzyme inactivation: from mechanisms to solutions.* Chembiochem, 2010. **11**(10): p. 1325-34.

56. Zarate, S.G., et al., *Overcoming Aminoglycoside Enzymatic Resistance: Design of Novel Antibiotics and Inhibitors.* Molecules, 2018. **23**(2).

57. Weinreich, D.M., et al., *Darwinian evolution can follow only very few mutational paths to fitter proteins.* Science, 2006. **312**(5770): p. 111-4.

58. Salverda, M.L., J.A. De Visser, and M. Barlow, *Natural evolution of TEM-1 beta-lactamase: experimental reconstruction and clinical relevance.* FEMS Microbiol Rev, 2010. **34**(6): p. 1015-36.

59. Gniadkowski, M., *Evolution of extended-spectrum beta-lactamases by mutation.* Clin Microbiol Infect, 2008. **14 Suppl 1**: p. 11-32.

60. Wang, M., et al., *Plasmid-mediated quinolone resistance in clinical isolates of Escherichia coli from Shanghai, China.* Antimicrob Agents Chemother, 2003. **47**(7): p. 2242-8.

61. Robicsek, A., et al., *Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase.* Nat Med, 2006. **12**(1): p. 83-8.

62. Marshall, C.G., et al., *Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms.* Antimicrob Agents Chemother, 1998. **42**(9): p. 2215-20.

63. Patel, R., et al., *The biopesticide Paenibacillus popilliae has a vancomycin resistance gene cluster homologous to the enterococcal VanA vancomycin resistance gene cluster.* Antimicrob Agents Chemother, 2000. **44**(3): p. 705-9.

64. Poirel, L., et al., *Origin of plasmid-mediated quinolone resistance determinant QnrA.* Antimicrob Agents Chemother, 2005. **49**(8): p. 3523-5.

65. Sanchez, M.B., et al., *Predictive analysis of transmissible quinolone resistance indicates Stenotrophomonas maltophilia as a potential source of a novel family of Qnr determinants.* BMC Microbiol, 2008. **8**: p. 148.

66. Poirel, L., et al., *Vibrionaceae as a possible source of Qnr-like quinolone resistance determinants.* J Antimicrob Chemother, 2005. **56**(6): p. 1118-21.

67. Humeniuk, C., et al., *Beta-lactamases of Kluyvera ascorbata, probable progenitors of some plasmid-encoded CTX-M types.* Antimicrob Agents Chemother, 2002. **46**(9): p. 3045-9.

68. Lartigue, M.F., et al., *In vitro analysis of ISEcp1B-mediated mobilization of naturally occurring beta-lactamase gene blaCTX-M of Kluyvera ascorbata.* Antimicrob Agents Chemother, 2006. **50**(4): p. 1282-6.

69. Currie, C.R., U.G. Mueller, and D. Malloch, *The agricultural pathology of ant fungus gardens.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7998-8002.

70. Cundliffe, E. and A.L. Demain, *Avoidance of suicide in antibiotic-producing microbes.* J Ind Microbiol Biotechnol, 2010. **37**(7): p. 643-72.

71. Marshall, C.G., et al., *D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB.* Proc Natl Acad Sci U S A, 1997. **94**(12): p. 6480-3.

72. Graham, M.Y. and B. Weisblum, *23S ribosomal ribonucleic acid of macrolide-producing streptomycetes contains methylated adenine.* J Bacteriol, 1979. **137**(3): p. 1464-7.

73. Macinga, D.R. and P.N. Rather, *The chromosomal 2'-N-acetyltransferase of Providencia stuartii: physiological functions and genetic regulation.* Front Biosci, 1999. **4**: p. D132-40.

74. Pramer, D. and R.L. Starkey, *Decomposition of streptomycin.* Science, 1951. **113**(2927): p. 127.

75. Abd-El-Malek, Y., M. Monib, and A. Hazem, *Chloramphenicol, a simultaneous carbon and nitrogen source for a Streptomyces sp. from Egyptain soil.* Nature, 1961. **189**: p. 775-6.

76. Dantas, G., et al., *Bacteria subsisting on antibiotics.* Science, 2008. **320**(5872): p. 100-3.

77. Walsh, F., S.G. Amyes, and B. Duffy, *Challenging the concept of bacteria subsisting on antibiotics.* Int J Antimicrob Agents, 2013. **41**(6): p. 558-63.

78. Riesenfeld, C.S., R.M. Goodman, and J. Handelsman, *Uncultured soil bacteria are a reservoir of new antibiotic resistance genes.* Environ Microbiol, 2004. **6**(9): p. 981-9.

79. Allen, H.K., et al., *Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil.* ISME J, 2009. **3**(2): p. 243-51.

80. Kim, D.W. and C.J. Cha, *Antibiotic resistome from the One-Health perspective: understanding and controlling antimicrobial resistance transmission.* Exp Mol Med, 2021. **53**(3): p. 301-309.

81. Lipszyc, A., M. Szuplewska, and D. Bartosik, *How Do Transposable Elements Activate Expression of Transcriptionally Silent Antibiotic Resistance Genes?* Int J Mol Sci, 2022. **23**(15).

82. Khedkar, S., et al., *Landscape of mobile genetic elements and their antibiotic resistance cargo in prokaryotic genomes.* Nucleic Acids Res, 2022. **50**(6): p. 3155-3168.

83. Mazel, D., *Integrons: agents of bacterial evolution.* Nat Rev Microbiol, 2006. **4**(8): p. 608-20.

84. Chen, C.Y., et al., *Comparative genome analysis of Vibrio vulnificus, a marine pathogen.* Genome Res, 2003. **13**(12): p. 2577-87.

85. Graf, F.E., et al., *Inhibiting conjugation as a tool in the fight against antibiotic resistance.* Drug Development Research, 2019. **80**(1): p. 19-23.

86. Palencia-Gandara, C., et al., *Conjugation Inhibitors Effectively Prevent Plasmid Transmission in Natural Environments.* mBio, 2021. **12**(4): p. e0127721.

87. Volkova, V.V., et al., *Modeling the infection dynamics of bacteriophages in enteric Escherichia coli: estimating the contribution of transduction to antimicrobial gene spread.* Appl Environ Microbiol, 2014. **80**(14): p. 4350-62.

88. Enault, F., et al., *Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses.* ISME J, 2017. **11**(1): p. 237-247.

89. Pfeifer, E., R.A. Bonnin, and E.P.C. Rocha, *Phage-Plasmids Spread Antibiotic Resistance Genes through Infection and Lysogenic Conversion.* mBio, 2022: p. e0185122.

90. Llor, C. and L. Bjerrum, *Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem.* Ther Adv Drug Saf, 2014. **5**(6): p. 229-41.

91. Tan, S.Y., et al., *Correlation between antibiotic consumption and the occurrence of multidrug-resistant organisms in a Malaysian tertiary hospital: a 3-year observational study.* Sci Rep, 2022. **12**(1): p. 3106.

92. Davies, J. and D. Davies, *Origins and evolution of antibiotic resistance.* Microbiol Mol Biol Rev, 2010. **74**(3): p. 417-33.

93. Olesen, S.W., et al., *The distribution of antibiotic use and its association with antibiotic resistance.* Elife, 2018. **7**.

94. Goossens, H., et al., *Outpatient antibiotic use in Europe and association with resistance: a cross-national database study.* Lancet, 2005. **365**(9459): p. 579-87.

95. Andersson, D.I. and D. Hughes, *Antibiotic resistance and its cost: is it possible to reverse resistance?* Nat Rev Microbiol, 2010. **8**(4): p. 260-71.

96. McGowan, J.E., Jr. and D.N. Gerding, *Does antibiotic restriction prevent resistance?* New Horiz, 1996. **4**(3): p. 370-6.

97. Livermore, D., *Can better prescribing turn the tide of resistance?* Nat Rev Microbiol, 2004. **2**(1): p. 73-8.

98. Marchant, R., et al., *What are high-temperature bacteria doing in cold environments?* Trends Microbiol, 2002. **10**(3): p. 120-1.

99. Stearns, S.C., *Life history evolution: successes, limitations, and prospects.* Naturwissenschaften, 2000. **87**(11): p. 476-86.

100. Garland, T., Jr., C.J. Downs, and A.R. Ives, *Trade-Offs (and Constraints) in Organismal Biology.* Physiol Biochem Zool, 2022. **95**(1): p. 82-112.

101. Purrington, C.B., *Costs of resistance.* Curr Opin Plant Biol, 2000. **3**(4): p. 305-8.

102. Coustau, C., C. Chevillon, and R. ffrench-Constant, *Resistance to xenobiotics and parasites: can we count the cost?* Trends Ecol Evol, 2000. **15**(9): p. 378-383.

103. Simms, E.L. and M.D. Rausher, *The Evolution of Resistance to Herbivory in Ipomoea Purpurea. Ii. Natural Selection by Insects and Costs of Resistance.* Evolution, 1989. **43**(3): p. 573-585.

104. Lenormand, T., N. Harmand, and R. Gallet, *Cost of resistance: an unreasonably expensive concept.* Rethinking Ecology, 2018. **3**: p. 51-70.

105. Lenski, R.E., *Experimental Studies of Pleiotropy and Epistasis in Escherichia Coli. I. Variation in Competitive Fitness among Mutants Resistant to Virus T4.* Evolution, 1988. **42**(3): p. 425-432.

106. Lenski, R.E., *Bacterial evolution and the cost of antibiotic resistance.* Int Microbiol, 1998. **1**(4): p. 265-70.

107. Schrag, S.J., V. Perrot, and B.R. Levin, *Adaptation to the fitness costs of antibiotic resistance in Escherichia coli.* Proc Biol Sci, 1997. **264**(1386): p. 1287-91.

108. Nguyen, T.N., et al., *Effects of carriage and expression of the Tn10 tetracycline-resistance operon on the fitness of Escherichia coli K12.* Mol Biol Evol, 1989. **6**(3): p. 213-25.

109. Melnyk, A.H., A. Wong, and R. Kassen, *The fitness costs of antibiotic resistance mutations.* Evol Appl, 2015. **8**(3): p. 273-83.

110. Bauer, C.R., S. Li, and M.L. Siegal, *Essential gene disruptions reveal complex relationships between phenotypic robustness, pleiotropy, and fitness.* Mol Syst Biol, 2015. **11**(1): p. 773.

111. Cooper, T.F., E.A. Ostrowski, and M. Travisano, *A negative relationship between mutation pleiotropy and fitness effect in yeast.* Evolution, 2007. **61**(6): p. 1495-9.

112. Orr, H.A., *The distribution of fitness effects among beneficial mutations.* Genetics, 2003. **163**(4): p. 1519-26.

113. Wagner, G.P. and J. Zhang, *The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms.* Nat Rev Genet, 2011. **12**(3): p. 204-13.

114. Qi, Q., G.M. Preston, and R.C. MacLean, *Linking system-wide impacts of RNA polymerase mutations to the fitness cost of rifampin resistance in Pseudomonas aeruginosa.* mBio, 2014. **5**(6): p. e01562.

115. Murren, C.J., et al., *Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity.* Heredity (Edinb), 2015. **115**(4): p. 293-301.

116. Yamamoto, N., et al., *Update on the Keio collection of Escherichia coli single-gene deletion mutants.* Mol Syst Biol, 2009. **5**: p. 335.

117. Vogwill, T. and R.C. MacLean, *The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach.* Evol Appl, 2015. **8**(3): p. 284-95.

118. Paulander, W., S. Maisnier-Patin, and D.I. Andersson, *Multiple mechanisms to ameliorate the fitness burden of mupirocin resistance in Salmonella typhimurium.* Mol Microbiol, 2007. **64**(4): p. 1038-48.

119. Dutta, D., et al., *Linking RNA polymerase backtracking to genome instability in E. coli.* Cell, 2011. **146**(4): p. 533-43.

120. Proshkin, S., et al., *Cooperation between translating ribosomes and RNA polymerase in transcription elongation.* Science, 2010. **328**(5977): p. 504-8.

121. Lubelski, J., W.N. Konings, and A.J. Driessen, *Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria.* Microbiol Mol Biol Rev, 2007. **71**(3): p. 463-76.

122. Olivares, J., C. Alvarez-Ortega, and J.L. Martinez, *Metabolic compensation of fitness costs associated with overexpression of the multidrug efflux pump MexEF-OprN in Pseudomonas aeruginosa.* Antimicrob Agents Chemother, 2014. **58**(7): p. 3904-13.

123. Olivares Pacheco, J., et al., *Metabolic Compensation of Fitness Costs Is a General Outcome for Antibiotic-Resistant Pseudomonas aeruginosa Mutants Overexpressing Efflux Pumps.* mBio, 2017. **8**(4).

124. Olivares, J., et al., *Overproduction of the multidrug efflux pump MexEF-OprN does not impair Pseudomonas aeruginosa fitness in competition tests, but produces specific changes in bacterial regulatory networks.* Environ Microbiol, 2012. **14**(8): p. 1968-81.

125. Zhou, H., J.F. Beltran, and I.L. Brito, *Functions predict horizontal gene transfer and the emergence of antibiotic resistance.* Sci Adv, 2021. **7**(43): p. eabj5056.

126. Jain, R., M.C. Rivera, and J.A. Lake, *Horizontal gene transfer among genomes: the complexity hypothesis.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3801-6.

127. Porse, A., et al., *Dominant resistance and negative epistasis can limit the co-selection of de novo resistance mutations and antibiotic resistance genes.* Nat Commun, 2020. **11**(1): p. 1199.

128. Medrano-Soto, A., et al., *Successful lateral transfer requires codon usage compatibility between foreign genes and recipient genomes.* Mol Biol Evol, 2004. **21**(10): p. 1884-94.

129. Popa, O., et al., *Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes.* Genome Res, 2011. **21**(4): p. 599-609.

130. Park, C. and J. Zhang, *High expression hampers horizontal gene transfer.* Genome Biol Evol, 2012. **4**(4): p. 523-32.

131. Drummond, D.A. and C.O. Wilke, *Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution.* Cell, 2008. **134**(2): p. 341-52.

132. Geiler-Samerotte, K.A., et al., *Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast.* Proc Natl Acad Sci U S A, 2011. **108**(2): p. 680-5.

133. Lopez, C., et al., *Protein determinants of dissemination and host specificity of metallo-beta-lactamases.* Nat Commun, 2019. **10**(1): p. 3617.

134. Marciano, D.C., O.Y. Karkouti, and T. Palzkill, *A fitness cost associated with the antibiotic resistance enzyme SME-1 beta-lactamase.* Genetics, 2007. **176**(4): p. 2381-92.

135. Rajer, F. and L. Sandegren, *The Role of Antibiotic Resistance Genes in the Fitness Cost of Multiresistance Plasmids.* mBio, 2022: p. e0355221.

136. Kadonaga, J.T., et al., *The role of the beta-lactamase signal sequence in the secretion of proteins by Escherichia coli.* J Biol Chem, 1984. **259**(4): p. 2149-54.

137. Goldstein, I. and I.M. Ehrenreich, *The complex role of genetic background in shaping the effects of spontaneous and induced mutations.* Yeast, 2021. **38**(3): p. 187-196.

138. Yokoyama, M., et al., *Epistasis analysis uncovers hidden antibiotic resistance-associated fitness costs hampering the evolution of MRSA.* Genome Biol, 2018. **19**(1): p. 94.

139. Paulander, W., S. Maisnier-Patin, and D.I. Andersson, *The fitness cost of streptomycin resistance depends on rpsL mutation, carbon source and RpoS (sigmaS).* Genetics, 2009. **183**(2): p. 539-46, 1SI-2SI.

140. Lin, W., et al., *Reduction of the fitness cost of antibiotic resistance caused by chromosomal mutations under poor nutrient conditions.* Environ Int, 2018. **120**: p. 63-71.

141. Gifford, D.R., E. Moss, and R.C. MacLean, *Environmental variation alters the fitness effects of rifampicin resistance mutations in Pseudomonas aeruginosa.* Evolution, 2016. **70**(3): p. 725-30.

142. Maharjan, R. and T. Ferenci, *The fitness costs and benefits of antibiotic resistance in drug-free microenvironments encountered in the human body.* Environ Microbiol Rep, 2017. **9**(5): p. 635-641.

143. Enne, V.I., et al., *Assessment of the fitness impacts on Escherichia coli of acquisition of antibiotic resistance genes encoded by different types of genetic element.* J Antimicrob Chemother, 2005. **56**(3): p. 544-51.

144. Jordan, J.A., R.E. Lenski, and K.J. Card *Idiosyncratic Fitness Costs of Ampicillin-Resistant Mutants Derived from a Long-Term Experiment with Escherichia coli*. Antibiotics, 2022. **11**, DOI: 10.3390/antibiotics11030347.

145. Alonso-Del Valle, A., et al., *Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities.* Nat Commun, 2021. **12**(1): p. 2653.

146. Vogwill, T., et al., *Testing the role of genetic background in parallel evolution using the comparative experimental evolution of antibiotic resistance.* Mol Biol Evol, 2014. **31**(12): p. 3314-23.

147. Comas, I., et al., *Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes.* Nat Genet, 2011. **44**(1): p. 106-10.

148. Brandis, G., et al., *Fitness-compensatory mutations in rifampicin-resistant RNA polymerase.* Mol Microbiol, 2012. **85**(1): p. 142-51.

149. Jin, D.J. and C.L. Turnbough, Jr., *An Escherichia coli RNA polymerase defective in transcription due to its overproduction of abortive initiation products.* J Mol Biol, 1994. **236**(1): p. 72-80.

150. Pennings, P.S., C.B. Ogbunugafor, and R. Hershberg, *Reversion is most likely under high mutation supply when compensatory mutations do not fully restore fitness costs.* G3 (Bethesda), 2022. **12**(9).

151. Maisnier-Patin, S., et al., *Compensatory adaptation to the deleterious effect of antibiotic resistance in Salmonella typhimurium.* Mol Microbiol, 2002. **46**(2): p. 355-66.

152. Filteau, M., et al., *Evolutionary rescue by compensatory mutations is constrained by genomic and environmental backgrounds.* Mol Syst Biol, 2015. **11**(10): p. 832.

153. Bjorkman, J. and D.I. Andersson, *The cost of antibiotic resistance from a bacterial perspective.* Drug Resist Updat, 2000. **3**(4): p. 237-245.

154. Steinchen, W., V. Zegarra, and G. Bange, *(p)ppGpp: Magic Modulators of Bacterial Physiology and Metabolism.* Front Microbiol, 2020. **11**: p. 2072.

155. D'Costa, V.M., et al., *Antibiotic resistance is ancient.* Nature, 2011. **477**(7365): p. 457-61.

156. Santiago-Rodriguez, T.M., et al., *Gut Microbiome of an 11th Century A.D. Pre-Columbian Andean Mummy.* PLoS One, 2015. **10**(9): p. e0138135.

157. Warinner, C., et al., *Pathogens and host immunity in the ancient human oral cavity.* Nat Genet, 2014. **46**(4): p. 336-44.

158. Perron, G.G., et al., *Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics.* PLoS One, 2015. **10**(3): p. e0069533.

159. McCann, C.M., et al., *Understanding drivers of antibiotic resistance genes in High Arctic soil ecosystems.* Environ Int, 2019. **125**: p. 497-504.

160. Van Goethem, M.W., et al., *A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils.* Microbiome, 2018. **6**(1): p. 40.

161. Bhullar, K., et al., *Antibiotic resistance is prevalent in an isolated cave microbiome.* PLoS One, 2012. **7**(4): p. e34953.

162. McConnell, M.M., et al., *Sources of Antibiotic Resistance Genes in a Rural River System.* J Environ Qual, 2018. **47**(5): p. 997-1005.

163. Karkman, A., et al., *Antibiotic-Resistance Genes in Waste Water.* Trends Microbiol, 2018. **26**(3): p. 220-228.

164. Knapp, C.W., et al., *Evidence of increasing antibiotic resistance gene abundances in archived soils since 1940.* Environ Sci Technol, 2010. **44**(2): p. 580-7.

165. Pawlowski, A.C., et al., *A diverse intrinsic antibiotic resistome from a cave bacterium.* Nat Commun, 2016. **7**: p. 13803.

166. Forsberg, K.J., et al., *The shared antibiotic resistome of soil bacteria and human pathogens.* Science, 2012. **337**(6098): p. 1107-11.

167. Smillie, C.S., et al., *Ecology drives a global network of gene exchange connecting the human microbiome.* Nature, 2011. **480**(7376): p. 241-4.

168. Sommer, M.O.A., G. Dantas, and G.M. Church, *Functional characterization of the antibiotic resistance reservoir in the human microflora.* Science, 2009. **325**(5944): p. 1128-1131.

169. Groussin, M., et al., *Elevated rates of horizontal gene transfer in the industrialized human microbiome.* Cell, 2021. **184**(8): p. 2053-2067 e18.

170. Van Boeckel, T.P., et al., *Reducing antimicrobial use in food animals.* Science, 2017. **357**(6358): p. 1350-1352.

171. Ebmeyer, S., E. Kristiansson, and D.G.J. Larsson, *A framework for identifying the recent origins of mobile antibiotic resistance genes.* Commun Biol, 2021. **4**(1): p. 8.

172. Andersson, D.I. and D. Hughes, *Persistence of antibiotic resistance in bacterial populations.* FEMS Microbiol Rev, 2011. **35**(5): p. 901-11.

173. Harrison, E., et al., *Parallel compensatory evolution stabilizes plasmids across the parasitism-mutualism continuum.* Curr Biol, 2015. **25**(15): p. 2034-9.

174. Hulter, N., et al., *An evolutionary perspective on plasmid lifestyle modes.* Curr Opin Microbiol, 2017. **38**: p. 74-80.

175. Sorensen, S.J., et al., *Studying plasmid horizontal transfer in situ: a critical review.* Nat Rev Microbiol, 2005. **3**(9): p. 700-10.

176. Ebersbach, G. and K. Gerdes, *Plasmid segregation mechanisms.* Annu Rev Genet, 2005. **39**: p. 453-79.

177. Gerdes, K., J.S. Jacobsen, and T. Franch, *Plasmid stabilization by post-segregational killing.* Genet Eng (N Y), 1997. **19**: p. 49-61.

178. Coluzzi, C., et al., *Evolution of Plasmid Mobility: Origin and Fate of Conjugative and Nonconjugative Plasmids.* Mol Biol Evol, 2022. **39**(6).

179. Pal, C., et al., *Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential.* BMC Genomics, 2015. **16**: p. 964.

180. Darmancier, H., et al., *Are Virulence and Antibiotic Resistance Genes Linked? A Comprehensive Analysis of Bacterial Chromosomes and Plasmids.* Antibiotics (Basel), 2022. **11**(6).

181. Turton, J., et al., *Hybrid Resistance and Virulence Plasmids in "High-Risk" Clones of Klebsiella pneumoniae, Including Those Carrying blaNDM-5.* Microorganisms, 2019. **7**(9).

182. San Millan, A. and R.C. MacLean, *Fitness Costs of Plasmids: a Limit to Plasmid Transmission.* Microbiol Spectr, 2017. **5**(5).

183. Hall, J.P.J., et al., *Plasmid fitness costs are caused by specific genetic conflicts enabling resolution by compensatory mutation.* PLoS Biol, 2021. **19**(10): p. e3001225.

184. Ingmer, H., C. Miller, and S.N. Cohen, *The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response.* Mol Microbiol, 2001. **42**(2): p. 519-26.

185. San Millan, A., et al., *Interactions between horizontally acquired genes create a fitness cost in Pseudomonas aeruginosa.* Nat Commun, 2015. **6**: p. 6845.

186. Yano, H., et al., *Evolved plasmid-host interactions reduce plasmid interference cost.* Mol Microbiol, 2016. **101**(5): p. 743-56.

187. Loftie-Eaton, W., et al., *Compensatory mutations improve general permissiveness to antibiotic resistance plasmids.* Nat Ecol Evol, 2017. **1**(9): p. 1354-1363.

188. Sota, M., et al., *Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein.* ISME J, 2010. **4**(12): p. 1568-80.

189. Loftie-Eaton, W., et al., *Evolutionary Paths That Expand Plasmid Host-Range: Implications for Spread of Antibiotic Resistance.* Mol Biol Evol, 2016. **33**(4): p. 885-97.

190. Harrison, E., et al., *Rapid compensatory evolution promotes the survival of conjugative plasmids.* Mob Genet Elements, 2016. **6**(3): p. e1179074.

191. San Millan, A., et al., *Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids.* Nat Commun, 2014. **5**: p. 5208.

192. Hall, J.P., et al., *Source-sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities.* Proc Natl Acad Sci U S A, 2016. **113**(29): p. 8260-5.

193. Kottara, A., et al., *The dilution effect limits plasmid horizontal transmission in multispecies bacterial communities.* Microbiology (Reading), 2021. **167**(9).

194. Li, L., et al., *Plasmids persist in a microbial community by providing fitness benefit to multiple phylotypes.* ISME J, 2020. **14**(5): p. 1170-1181.

195. Modi, R.I., et al., *Genetic changes accompanying increased fitness in evolving populations of Escherichia coli.* Genetics, 1992. **130**(2): p. 241-9.

196. DelaFuente, J., et al., *Within-patient evolution of plasmid-mediated antimicrobial resistance.* bioRxiv, 2022: p. 2022.05.31.493991.

197. De Gelder, L., et al., *Combining mathematical models and statistical methods to understand and predict the dynamics of antibiotic-sensitive mutants in a population of resistant bacteria during experimental evolution.* Genetics, 2004. **168**(3): p. 1131-44.

198. Modi, R.I., et al., *Plasmid macro-evolution: selection of deletions during adaptation in a nutrient-limited environment.* Genetica, 1991. **84**(3): p. 195-202.

199. Sheppard, A.E., et al., *Nested Russian Doll-Like Genetic Mobility Drives Rapid Dissemination of the Carbapenem Resistance Gene blaKPC.* Antimicrob Agents Chemother, 2016. **60**(6): p. 3767-78.

200. Cuzon, G., T. Naas, and P. Nordmann, *Functional characterization of Tn4401, a Tn3-based transposon involved in blaKPC gene mobilization.* Antimicrob Agents Chemother, 2011. **55**(11): p. 5370-3.

201. Che, Y., et al., *Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes.* Proc Natl Acad Sci U S A, 2021. **118**(6).

202. David, S., et al., *Integrated chromosomal and plasmid sequence analyses reveal diverse modes of carbapenemase gene spread among Klebsiella pneumoniae.* Proc Natl Acad Sci U S A, 2020. **117**(40): p. 25043-25054.

203. Schwarz, S., et al., *Plasmid-Mediated Antimicrobial Resistance in Staphylococci and Other Firmicutes.* Microbiol Spectr, 2014. **2**(6).

204. Desmet, S., et al., *Antibiotic Resistance Plasmids Cointegrated into a Megaplasmid Harboring the blaOXA-427 Carbapenemase Gene.* Antimicrob Agents Chemother, 2018. **62**(3).

205. Antimicrobial Resistance, C., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* Lancet, 2022. **399**(10325): p. 629-655.

206. Aminov, R.I., *A brief history of the antibiotic era: lessons learned and challenges for the future.* Front Microbiol, 2010. **1**: p. 134.

207. Kieffer, N., S. Ebmeyer, and D.G.J. Larsson, *The Class A Carbapenemases BKC-1 and GPC-1 Both Originate from the Bacterial Genus Shinella.* Antimicrob Agents Chemother, 2020. **64**(12).

208. Allen, H.K., et al., *Call of the wild: antibiotic resistance genes in natural environments.* Nat Rev Microbiol, 2010. **8**(4): p. 251-9.

209. Shapiro, J.T., et al., *Metapopulation ecology links antibiotic resistance, consumption, and patient transfers in a network of hospital wards.* Elife, 2020. **9**.

210. Zainab, S.M., et al., *Antibiotics and antibiotic resistant genes (ARGs) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks.* Water Res, 2020. **187**: p. 116455.

211. Gullberg, E., et al., *Selection of resistant bacteria at very low antibiotic concentrations.* PLoS Pathog, 2011. **7**(7): p. e1002158.

212. Bengtsson-Palme, J., V. Jonsson, and S. Hess, *What Is the Role of the Environment in the Emergence of Novel Antibiotic Resistance Genes? A Modeling Approach.* Environ Sci Technol, 2021. **55**(23): p. 15734-15743.

213. Huijben, S., et al., *The impact of within-host ecology on the fitness of a drug-resistant parasite.* Evol Med Public Health, 2018. **2018**(1): p. 127-137.

214. Sundqvist, M., *Reversibility of antibiotic resistance.* Ups J Med Sci, 2014. **119**(2): p. 142-8.

215. Letten, A.D., A.R. Hall, and J.M. Levine, *Using ecological coexistence theory to understand antibiotic resistance and microbial competition.* Nat Ecol Evol, 2021. **5**(4): p. 431-441.

216. Jones, M.E., et al., *Apparent plateau in beta-lactamase production among clinical isolates of Haemophilus influenzae and Moraxella catarrhalis in the United States: results from the LIBRA Surveillance initiative.* Int J Antimicrob Agents, 2002. **19**(2): p. 119-23.

217. Lehtinen, S., et al., *Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage.* Proc Natl Acad Sci U S A, 2017. **114**(5): p. 1075-1080.

218. Hall, A.R., et al., *Costs of antibiotic resistance - separating trait effects and selective effects.* Evol Appl, 2015. **8**(3): p. 261-72.

219. Bjorkman, J., et al., *Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance.* Science, 2000. **287**(5457): p. 1479-82.

220. Rozen, D.E., et al., *Fitness costs of fluoroquinolone resistance in Streptococcus pneumoniae.* Antimicrob Agents Chemother, 2007. **51**(2): p. 412-6.

221. Kassen, R. and T. Bataillon, *Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria.* Nat Genet, 2006. **38**(4): p. 484-8.

222. Melnyk, A.H., et al., *Evolution of Cost-Free Resistance under Fluctuating Drug Selection in Pseudomonas aeruginosa.* mSphere, 2017. **2**(4).

223. Vogwill, T., M. Kojadinovic, and R.C. MacLean, *Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of Pseudomonas.* Proc Biol Sci, 2016. **283**(1830).

224. Alekshun, M.N. and S.B. Levy, *Molecular mechanisms of antibacterial multidrug resistance.* Cell, 2007. **128**(6): p. 1037-50.

225. Colquhoun, J.M., et al., *OXA-23 beta-Lactamase Overexpression in Acinetobacter baumannii Drives Physiological Changes Resulting in New Genetic Vulnerabilities.* mBio, 2021. **12**(6): p. e0313721.

226. Lenski, R.E., S.C. Simpson, and T.T. Nguyen, *Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness.* J Bacteriol, 1994. **176**(11): p. 3140-7.

227. Knopp, M. and D.I. Andersson, *Predictable Phenotypes of Antibiotic Resistance Mutations.* mBio, 2018. **9**(3).

228. Peng, F., et al., *Effects of Beneficial Mutations in pykF Gene Vary over Time and across Replicate Populations in a Long-Term Experiment with Bacteria.* Mol Biol Evol, 2018. **35**(1): p. 202-210.

229. Trindade, S., et al., *Positive epistasis drives the acquisition of multidrug resistance.* PLoS Genet, 2009. **5**(7): p. e1000578.

230. Silva, R.F., et al., *Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations.* PLoS Genet, 2011. **7**(7): p. e1002181.

231. Moura de Sousa, J., et al., *Multidrug-resistant bacteria compensate for the epistasis between resistances.* PLoS Biol, 2017. **15**(4): p. e2001741.

232. Alcock, B.P., et al., *CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database.* Nucleic Acids Res, 2020. **48**(D1): p. D517-D525.

233. Zaslaver, A., et al., *A comprehensive library of fluorescent transcriptional reporters for Escherichia coli.* Nat Methods, 2006. **3**(8): p. 623-8.

234. Lenski, R.E., et al., *Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations.* The American Naturalist, 1991. **138**(6): p. 1315-1341.

235. Turner, P.E., V.S. Cooper, and R.E. Lenski, *Tradeoff between Horizontal and Vertical Modes of Transmission in Bacterial Plasmids.* Evolution, 1998. **52**(2): p. 315-329.

236. Lee, C., et al., *Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli.* J Biotechnol, 2006. **123**(3): p. 273-80.

237. Revell, L.J., et al., *A new phylogenetic method for identifying exceptional phenotypic diversification.* Evolution, 2012. **66**(1): p. 135-46.

238. Team, R.C.D., *R: A language and environment for statistical computing.* 2021, R Foundation for Statistical Computing, Vienna, Austria.

239. Ellis B, H.P., Hahne F, Le Meur N, Gopalakrishnan N, Spidlen J, Jiang M, Finak G, *flowCore: flowCore: Basic structures for flow cytometry data.* 2022.

240. Kuznetsova, A., P.B. Brockhoff, and R.H.B. Christensen, *lmerTest Package: Tests in Linear Mixed Effects Models.* Journal of Statistical Software, 2017. **82**(13): p. 1 - 26.

241. Bates, D., et al., *Fitting Linear Mixed-Effects Models Using lme4.* Journal of Statistical Software, 2015. **67**(1): p. 1 - 48.

242. Papadovasilaki, M., et al., *Biophysical and enzymatic properties of aminoglycoside adenylyltransferase AadA6 from Pseudomonas aeruginosa.* Biochem Biophys Rep, 2015. **4**: p. 152-157.

243. Dhar, R., T. Bergmiller, and A. Wagner, *Increased gene dosage plays a predominant role in the initial stages of evolution of duplicate TEM-1 beta lactamase genes.* Evolution, 2014. **68**(6): p. 1775-91.

244. Pagel, M., *Inferring the historical patterns of biological evolution.* Nature, 1999. **401**(6756): p. 877-84.

245. Szamecz, B., et al., *The genomic landscape of compensatory evolution.* PLoS Biol, 2014. **12**(8): p. e1001935.

246. Laabei, M., et al., *Predicting the virulence of MRSA from its genome sequence.* Genome Res, 2014. **24**(5): p. 839-49.

247. Enne, V.I., *Reducing antimicrobial resistance in the community by restricting prescribing: can it be done?* J Antimicrob Chemother, 2010. **65**(2): p. 179-82.

248. Enne, V.I., et al., *Persistence of sulphonamide resistance in Escherichia coli in the UK despite national prescribing restriction.* Lancet, 2001. **357**(9265): p. 1325-8.

249. Gottesman, B.S., et al., *Impact of quinolone restriction on resistance patterns of Escherichia coli isolated from urine by culture in a community setting.* Clin Infect Dis, 2009. **49**(6): p. 869-75.

250. Dagan, R., et al., *Seasonality of antibiotic-resistant streptococcus pneumoniae that causes acute otitis media: a clue for an antibiotic-restriction policy?* J Infect Dis, 2008. **197**(8): p. 1094-102.

251. Nji, E., et al., *High prevalence of antibiotic resistance in commensal Escherichia coli from healthy human sources in community settings.* Sci Rep, 2021. **11**(1): p. 3372.

252. Nkansa-Gyamfi, N.A., et al., *Prevalence of multidrug-, extensive drug-, and pandrug-resistant commensal Escherichia coli isolated from healthy humans in community settings in low- and middle-income countries: a systematic review and meta-analysis.* Glob Health Action, 2019. **12**(sup1): p. 1815272.

253. Roemhild, R., T. Bollenbach, and D.I. Andersson, *The physiology and genetics of bacterial responses to antibiotic combinations.* Nat Rev Microbiol, 2022. **20**(8): p. 478-490.

254. Sodani, K., et al., *Multidrug resistance associated proteins in multidrug resistance.* Chin J Cancer, 2012. **31**(2): p. 58-72.

255. Lazar, V., et al., *Bacterial evolution of antibiotic hypersensitivity.* Mol Syst Biol, 2013. **9**: p. 700.

256. Porse, A., et al., *Biochemical mechanisms determine the functional compatibility of heterologous genes.* Nat Commun, 2018. **9**(1): p. 522.

257. Shao, B., et al., *Single-cell measurement of plasmid copy number and promoter activity.* Nat Commun, 2021. **12**(1): p. 1475.

258. Adler, M., et al., *High fitness costs and instability of gene duplications reduce rates of evolution of new genes by duplication-divergence mechanisms.* Mol Biol Evol, 2014. **31**(6): p. 1526-35.

259. Hall, J.P., et al., *Environmentally co-occurring mercury resistance plasmids are genetically and phenotypically diverse and confer variable context-dependent fitness effects.* Environ Microbiol, 2015. **17**(12): p. 5008-22.

260. Callens, M., C. Scornavacca, and S. Bedhomme, *Evolutionary responses to codon usage of horizontally transferred genes in Pseudomonas aeruginosa: gene retention, amelioration and compensatory evolution.* Microb Genom, 2021. **7**(6).

261. Souque, C., J.A. Escudero, and R.C. MacLean, *Integron activity accelerates the evolution of antibiotic resistance.* Elife, 2021. **10**.

262. Dimitriu, T., A.C. Matthews, and A. Buckling, *Increased copy number couples the evolution of plasmid horizontal transmission and plasmid-encoded antibiotic resistance.* Proc Natl Acad Sci U S A, 2021. **118**(31).

263. Ffrench-Constant, R.H. and C. Bass, *Does resistance really carry a fitness cost?* Curr Opin Insect Sci, 2017. **21**: p. 39-46.

264. Moore, F.B., D.E. Rozen, and R.E. Lenski, *Pervasive compensatory adaptation in Escherichia coli.* Proc Biol Sci, 2000. **267**(1442): p. 515-22.

265. Poon, A. and L. Chao, *The rate of compensatory mutation in the DNA bacteriophage phiX174.* Genetics, 2005. **170**(3): p. 989-99.

266. Dunai, A., et al., *Rapid decline of bacterial drug-resistance in an antibiotic-free environment through phenotypic reversion.* Elife, 2019. **8**.

267. Moradigaravand, D., et al., *Prediction of antibiotic resistance in Escherichia coli from large-scale pan-genome data.* PLoS Comput Biol, 2018. **14**(12): p. e1006258.

268. Rivera, M.C., et al., *Genomic evidence for two functionally distinct gene classes.* Proc Natl Acad Sci U S A, 1998. **95**(11): p. 6239-44.

269. Sorek, R., et al., *Genome-wide experimental determination of barriers to horizontal gene transfer.* Science, 2007. **318**(5855): p. 1449-52.

270. Dekel, E. and U. Alon, *Optimality and evolutionary tuning of the expression level of a protein.* Nature, 2005. **436**(7050): p. 588-92.

271. Goodman, D.B., G.M. Church, and S. Kosuri, *Causes and effects of N-terminal codon bias in bacterial genes.* Science, 2013. **342**(6157): p. 475-9.

272. Kudla, G., et al., *Coding-sequence determinants of gene expression in Escherichia coli.* Science, 2009. **324**(5924): p. 255-8.

273. Davis, B.H., A.F. Poon, and M.C. Whitlock, *Compensatory mutations are repeatable and clustered within proteins.* Proc Biol Sci, 2009. **276**(1663): p. 1823-7.

274. Merker, M., et al., *Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia.* Elife, 2018. **7**.

275. McInerney, J.O., A. McNally, and M.J. O'Connell, *Why prokaryotes have pangenomes.* Nat Microbiol, 2017. **2**: p. 17040.

276. Abram, K., et al., *Mash-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups.* Commun Biol, 2021. **4**(1): p. 117.

277. Vigue, L., et al., *Deciphering polymorphism in 61,157 Escherichia coli genomes via epistatic sequence landscapes.* Nat Commun, 2022. **13**(1): p. 4030.

278. Nyerges, A., et al., *A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species.* Proc Natl Acad Sci U S A, 2016. **113**(9): p. 2502-7.

279. Gallagher, R.R., et al., *Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA.* Nat Protoc, 2014. **9**(10): p. 2301-16.

280. McKenzie, G.J. and N.L. Craig, *Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn7 without need for selection of the insertion event.* BMC Microbiol, 2006. **6**: p. 39.

281. Satterwhite, R.S. and T.F. Cooper, *Constraints on adaptation of Escherichia coli to mixed-resource environments increase over time.* Evolution, 2015. **69**(8): p. 2067-78.

282. Deatherage, D.E. and J.E. Barrick, *Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq.* Methods Mol Biol, 2014. **1151**: p. 165-88.

283. Lobocka, M.B., et al., *Genome of bacteriophage P1.* J Bacteriol, 2004. **186**(21): p. 7032-68.

284. Johnson, M., et al., *NCBI BLAST: a better web interface.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W5-9.

285. Chau, N.Y.E., et al., *Emerging and divergent roles of pyrophosphorylated nucleotides in bacterial physiology and pathogenesis.* PLoS Pathog, 2021. **17**(5): p. e1009532.

286. Fernandez-Coll, L. and M. Cashel, *Possible Roles for Basal Levels of (p)ppGpp: Growth Efficiency Vs. Surviving Stress.* Front Microbiol, 2020. **11**: p. 592718.

287. Shyp, V., et al., *Positive allosteric feedback regulation of the stringent response enzyme RelA by its product.* EMBO Rep, 2012. **13**(9): p. 835-9.

288. Jimmy, S., et al., *A widespread toxin-antitoxin system exploiting growth control via alarmone signaling.* Proc Natl Acad Sci U S A, 2020. **117**(19): p. 10500-10510.

289. Ahmad, S., et al., *An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp.* Nature, 2019. **575**(7784): p. 674-678.

290. Svenningsen, M.S., et al., *Existence of log-phase Escherichia coli persisters and lasting memory of a starvation pulse.* Life Sci Alliance, 2022. **5**(2).

291. Balaban, N.Q., et al., *Definitions and guidelines for research on antibiotic persistence.* Nat Rev Microbiol, 2019. **17**(7): p. 441-448.

292. Balaban, N.Q., et al., *Bacterial persistence as a phenotypic switch.* Science, 2004. **305**(5690): p. 1622-5.

293. Pacios, O., et al., *(p)ppGpp and Its Role in Bacterial Persistence: New Challenges.* Antimicrob Agents Chemother, 2020. **64**(10).

294. Dunne, M., et al., *Molecular Basis of Bacterial Host Interactions by Gram-Positive Targeting Bacteriophages.* Viruses, 2018. **10**(8).

295. Woodford, N. and M.J. Ellington, *The emergence of antibiotic resistance by mutation.* Clin Microbiol Infect, 2007. **13**(1): p. 5-18.

296. Fong, D.H. and A.M. Berghuis, *Substrate promiscuity of an aminoglycoside antibiotic resistance enzyme via target mimicry.* EMBO J, 2002. **21**(10): p. 2323-31.

297. Palzkill, T., *Structural and Mechanistic Basis for Extended-Spectrum Drug-Resistance Mutations in Altering the Specificity of TEM, CTX-M, and KPC beta-lactamases.* Front Mol Biosci, 2018. **5**: p. 16.

298. Baquero, F., et al., *Allogenous Selection of Mutational Collateral Resistance: Old Drugs Select for New Resistance Within Antibiotic Families.* Front Microbiol, 2021. **12**: p. 757833.

299. Atanasov, A.G., et al., *Natural products in drug discovery: advances and opportunities.* Nat Rev Drug Discov, 2021. **20**(3): p. 200-216.

300. Miethke, M., et al., *Towards the sustainable discovery and development of new antibiotics.* Nat Rev Chem, 2021. **5**(10): p. 726-749.

301. Zheng, J., N. Guo, and A. Wagner, *Selection enhances protein evolvability by increasing mutational robustness and foldability.* Science, 2020. **370**(6521).

302. Bloom, J.D., et al., *Protein stability promotes evolvability.* Proc Natl Acad Sci U S A, 2006. **103**(15): p. 5869-74.

303. Rennell, D., et al., *Systematic mutation of bacteriophage T4 lysozyme.* J Mol Biol, 1991. **222**(1): p. 67-88.

304. Bershtein, S., et al., *Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein.* Nature, 2006. **444**(7121): p. 929-32.

305. Tokuriki, N. and D.S. Tawfik, *Stability effects of mutations and protein evolvability.* Curr Opin Struct Biol, 2009. **19**(5): p. 596-604.

306. Stiffler, M.A., D.R. Hekstra, and R. Ranganathan, *Evolvability as a function of purifying selection in TEM-1 beta-lactamase.* Cell, 2015. **160**(5): p. 882-892.

307. Paaby, A.B. and M.V. Rockman, *Cryptic genetic variation: evolution's hidden substrate.* Nat Rev Genet, 2014. **15**(4): p. 247-58.

308. Zheng, J., J.L. Payne, and A. Wagner, *Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks.* Science, 2019. **365**(6451): p. 347-353.

309. Baltrus, D.A., *Exploring the costs of horizontal gene transfer.* Trends Ecol Evol, 2013. **28**(8): p. 489-95.

310. Zwanzig, M., *The ecology of plasmid-coded antibiotic resistance: a basic framework for experimental research and modeling.* Comput Struct Biotechnol J, 2021. **19**: p. 586-599.

311. Orr, H.A., *The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution.* Evolution, 1998. **52**(4): p. 935-949.

312. Johnson, M.S., et al., *Higher-fitness yeast genotypes are less robust to deleterious mutations.* Science, 2019. **366**(6464): p. 490-493.

313. Lyons, D.M., et al., *Idiosyncratic epistasis creates universals in mutational effects and evolutionary trajectories.* Nat Ecol Evol, 2020. **4**(12): p. 1685-1693.

314. Reddy, G. and M.M. Desai, *Global epistasis emerges from a generic model of a complex trait.* Elife, 2021. **10**.

315. Chou, H.H., et al., *Diminishing returns epistasis among beneficial mutations decelerates adaptation.* Science, 2011. **332**(6034): p. 1190-2.

316. Emory, S.A., P. Bouvet, and J.G. Belasco, *A 5'-terminal stem-loop structure can stabilize mRNA in Escherichia coli.* Genes Dev, 1992. **6**(1): p. 135-48.

317. Ren, G.X., X.P. Guo, and Y.C. Sun, *Regulatory 3' Untranslated Regions of Bacterial mRNAs.* Front Microbiol, 2017. **8**: p. 1276.

318. Carlton, B.C., and B. J. Brown., *Gene mutation. Pp. 222-242 in P. Gerhardt (ed.), Manual of methods for general bacteriology.* American Society for Microbiology, Washington, D.C., 1981.

319. Wong, T.S., et al., *A statistical analysis of random mutagenesis methods used for directed protein evolution.* J Mol Biol, 2006. **355**(4): p. 858-71.

320. Kuiper, B.P., R.C. Prins, and S. Billerbeck, *Oligo Pools as an Affordable Source of Synthetic DNA for Cost-Effective Library Construction in Protein- and Metabolic Pathway Engineering.* Chembiochem, 2022. **23**(7): p. e202100507.

321. Macdonald, C.B., et al., *Deep Insertion, Deletion, and Missense Mutation Libraries for Exploring Protein Variation in Evolution, Disease, and Biology.* bioRxiv, 2022: p. 2022.07.26.501589.

322. Firnberg, E. and M. Ostermeier, *PFunkel: efficient, expansive, user-defined mutagenesis.* PLoS One, 2012. **7**(12): p. e52031.

323. Wenger, A.M., et al., *Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome.* Nat Biotechnol, 2019. **37**(10): p. 1155-1162.

324. He, Y., et al., *Antibiotic resistance genes from livestock waste: occurrence, dissemination, and treatment.* npj Clean Water, 2020. **3**(1): p. 4.

325. Blanquart, F., S. Lehtinen, and C. Fraser, *An evolutionary model to predict the frequency of antibiotic resistance under seasonal antibiotic use, and an application to Streptococcus pneumoniae.* Proc Biol Sci, 2017. **284**(1855).

326. Pecerska, J., et al., *Quantifying transmission fitness costs of multi-drug resistant tuberculosis.* Epidemics, 2021. **36**: p. 100471.

327. Kim, M., et al., *Gain and loss of antibiotic resistant genes in multidrug resistant bacteria: One Health perspective.* J Microbiol, 2021. **59**(6): p. 535-545.

328. Francino, M.P., *Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances.* Front Microbiol, 2015. **6**: p. 1543.

329. Anthony, W.E., et al., *Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults.* Cell Rep, 2022. **39**(2): p. 110649.

330. San Millan, A., *Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context.* Trends Microbiol, 2018. **26**(12): p. 978-985.

331. Salyers, A.A., A. Gupta, and Y. Wang, *Human intestinal bacteria as reservoirs for antibiotic resistance genes.* Trends in Microbiology, 2004. **12**(9): p. 412-416.

332. Domingo-Sananes, M.R. and J.O. McInerney, *Selection-based model of prokaryote pangenomes.* bioRxiv, 2019: p. 782573.

333. Whelan, F.J., R.J. Hall, and J.O. McInerney, *Evidence for Selection in the Abundant Accessory Gene Content of a Prokaryote Pangenome.* Mol Biol Evol, 2021. **38**(9): p. 3697-3708.

334. Whelan, F.J., M. Rusilowicz, and J.O. McInerney, *Coinfinder: detecting significant associations and dissociations in pangenomes.* Microb Genom, 2020. **6**(3).

335. Domingo-Sananes, M.R. and J.O. McInerney, *Mechanisms That Shape Microbial Pangenomes.* Trends Microbiol, 2021. **29**(6): p. 493-503.

336. Bush, K. and G.A. Jacoby, *Updated functional classification of beta-lactamases.* Antimicrob Agents Chemother, 2010. **54**(3): p. 969-76.

337. Baym, M., L.K. Stone, and R. Kishony, *Multidrug evolutionary strategies to reverse antibiotic resistance.* Science, 2016. **351**(6268): p. aad3292.

338. Bell, G. and C. MacLean, *The Search for 'Evolution-Proof' Antibiotics.* Trends Microbiol, 2018. **26**(6): p. 471-483.

339. Kim, S., T.D. Lieberman, and R. Kishony, *Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance.* Proc Natl Acad Sci U S A, 2014. **111**(40): p. 14494-9.

340. Munck, C., et al., *Prediction of resistance development against drug combinations by collateral responses to component drugs.* Sci Transl Med, 2014. **6**(262): p. 262ra156.

341. Jassim, S.A., R.G. Limoges, and H. El-Cheikh, *Bacteriophage biocontrol in wastewater treatment.* World J Microbiol Biotechnol, 2016. **32**(4): p. 70.

342. Jassim, S. and R. Limoges, *Bacteriophages: Practical Applications for Nature's Biocontrol*. 2017.

343. Jassim, S.A. and R.G. Limoges, *Natural solution to antibiotic resistance: bacteriophages 'The Living Drugs'.* World J Microbiol Biotechnol, 2014. **30**(8): p. 2153-70.

344. Konate, M.M., et al., *Molecular function limits divergent protein evolution on planetary timescales.* Elife, 2019. **8**.

345. Ohno, S., U. Wolf, and N.B. Atkin, *Evolution from fish to mammals by gene duplication.* Hereditas, 1968. **59**(1): p. 169-87.