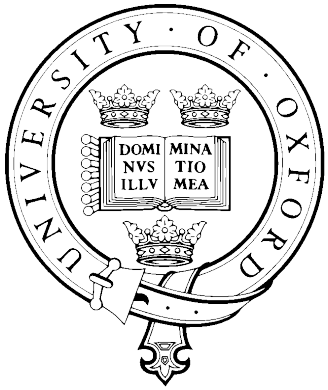


An integrative approach to understanding the fitness cost of rifampicin resistance in *Pseudomonas aeruginosa*

Qin Qi

Magdalen College, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2014

Abstract

Antibiotic resistance in bacteria is acquired through spontaneous chromosomal mutations or horizontal gene transfer. In the absence of antibiotics, resistant mutants generally show reduced fitness due to compromised growth rate, competitive ability and virulence compared to their antibiotic-sensitive ancestors. The focus of my research is to dissect the molecular underpinnings of the variations in the fitness cost of chromosomal antibiotic resistance using a systems-level approach. From an evolutionary perspective, my research aims are to understand how the fitness cost influences adaptation in resistant populations in an antibiotic-free environment. Using rifampicin resistance in *Pseudomonas aeruginosa* as a model, my work shows that most of the variation in the fitness cost of rifampicin resistance can be attributed to the direct effect of rifampicin resistance mutations on transcriptional efficiency. Through RNA-Seq transcriptome profiling, I demonstrate that global changes in gene expression levels associated with resistance mutations are surprisingly subtle, suggesting that the transcriptional regulatory network of *P. aeruginosa* is robust against compromised transcriptional efficiency. Using experimental evolution and whole-genome sequencing, my work reveals a systematic difference in the genetic basis of adaptation in mutants that were propagated in the absence of antibiotics. During compensatory adaptation, resistant mutants can recover the fitness cost of resistance by fixing second-site mutations that directly offset the deleterious effects of resistance mutations. Amongst resistant mutant populations with low fitness costs, general adaptation limits compensatory adaptation, which is most likely to be due to the rarity of compensatory mutations and clonal interference. Far from being the most ubiquitous mechanism in the evolution of resistance, compensatory adaptation is the exception that is more likely to be observed in resistant mutants with high fitness costs. In addition, I applied key elements of the integrative experimental approach developed in this work to dissect the molecular basis of the fitness cost associated with carriage of the pNUK73 small plasmid in *P. aeruginosa*, which carries the *rep* gene encoding a plasmid replication protein. My results confirmed that *rep* expression generates a significant fitness cost in *P. aeruginosa* and demonstrate how the molecular origins of the fitness cost of resistance can be dissected in a different biological context.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Dr Craig MacLean and Dr Gail Preston for all their guidance and encouragement throughout my time in Oxford. As highly dedicated mentors, Craig and Gail have always given me their full and unwavering support, especially in times of need. At the same time, they also gave me the intellectual freedom to pursue my research interests and formulate my own ideas. Indeed, I could not have wished for two better supervisors.

I would like to extend my gratitude to my collaborators Dr Macarena Toll-Riera and Dr Alvaro San Millan for their valuable contributions towards the successful completion of my doctoral research. In addition, I would like to thank all members of the MacLean and Preston research groups for their continuous support over the past three years. I am very grateful to Dr Ashleigh Griffin, Professor Kevin Foster, Dr Sebastian Shimeld, Professor Stuart West, Dr Ian Moore and members of their research teams from the Department of Zoology and Department of Plant Sciences for their encouragement and constructive feedback. I would also like to thank my external examiner Dr Stephen Diggle from the Centre for Biomolecular Sciences, University of Nottingham.

My parents have given me a lot of spiritual, moral and financial support throughout my life. I am very grateful for the values they instilled in me, without which I would not have had the determination to pursue a scientific career, and it is in this spirit that I would like to dedicate this thesis to my parents.

Contributions of Authors

This thesis features three manuscripts that have been prepared for publication in peer-reviewed scientific journals.

- I designed the experiments in Chapter 2 in collaboration with my supervisors Dr RC MacLean and Dr GM Preston. Unless otherwise mentioned, I performed the experiments and analyzed the data in this work. RNA-Seq and differential gene expression analysis were performed by research staff from the Wellcome Trust Centre for Human Genetics, University of Oxford. Dr Cedric Patthey assessed the integrity of RNA samples for RNA-Seq. I prepared the manuscript in collaboration with my supervisors.
- I planned the experiments in Chapter 3 in collaboration with my supervisors and Dr M Toll-Riera. With the exception of whole-genome sequencing of genomic DNA, which was carried out at the Wellcome Trust Centre for Human Genetics, University of Oxford, I performed all the experiments in this work. Dr M Toll-Riera analyzed the whole-genome sequencing data, while Dr RC MacLean and I jointly analyzed all remaining experimental data. All authors contributed to the preparation of the manuscript.
- Chapter 4 focuses on my contributions to the manuscript included in Appendix A, which is the product of a collaborative effort under the joint first authorship of Dr A San Millan and Dr M Toll-Riera. Dr A San Millan performed all the RNA extraction procedures and competitive fitness experiments, while Dr M Toll-Riera performed all the bioinformatic analyses in the entire work. I designed and performed the molecular cloning and transposon mutagenesis experiments for generating two mini-Tn7 insertion mutants of *P. aeruginosa*. In consultation with Dr A San Millan and Dr RC MacLean, I designed and performed all the experiments related to gene expression analysis by quantitative real-time PCR (qPCR). I also provided technical advice for RNA extraction and analyzed the qPCR data. All authors contributed to the preparation of the manuscript.

Contents

1. Introduction 6

1.1 Research background 6

1.2 The fitness cost of antibiotic resistance mutations 9

1.3 The fitness cost of rifampicin resistance 13

1.4 The opportunistic pathogen *Pseudomonas aeruginosa* 17

1.5 An experimental evolution approach for studying adaptive evolution of antibiotic resistant mutants 19

1.6 Research objectives and hypotheses 21

1.6.1 System-wide molecular mechanisms that contribute to the fitness cost of rifampicin resistance 21

1.6.2 Understanding the role of fitness cost of resistance in influencing compensatory adaptation 23

1.6.3 Dissecting the fitness cost of pNUK73 small plasmid carriage in *P. aeruginosa* 24

2. Linking system-wide impacts of RNA polymerase mutations to the fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* 28

3. Limits to compensatory adaptation in rifampicin resistant populations of *Pseudomonas aeruginosa* 73

4. Molecular mechanisms underpinning the fitness cost of pNUK73 small plasmid carriage in *Pseudomonas aeruginosa* 94

5. General Discussions 108

6. General Conclusions 117

7. Appendices 119

8. References 152

1. Introduction

1.1 Research background

Since the discovery of penicillin as the world's first antibiotic in 1929, the development of a wide range of antibacterial compounds has led to major improvements in public health over many decades [1-3]. Today, routine medical procedures such as surgeries, transplantations, cancer chemotherapy and clinical care of the critically ill all require the use of effective antibiotics [4]. However, the widespread misuse of antibiotics in both developing and developed countries has led to the spread of antibiotic resistance worldwide [5-7]. The rise of multi-drug resistant pathogenic bacteria is one of the greatest threats facing modern healthcare and has severely limited our repertoire of effective antimicrobials for treating bacterial infections [2, 8]. The present situation is exacerbated by the small pipelines of novel antibiotics and treatment strategies that are currently under development [6, 9, 10]. In light of this situation, it has become more imperative than ever before that existing antibiotics should be deployed in a more effective and strategic manner during clinical treatment to constrain the spread of antibiotic resistance [2, 11]. At the same time, intensive efforts are required to establish innovative and sustainable platforms for discovering novel antimicrobial compounds [12, 13]. However, these endeavors hinge on our ability to expand our existing knowledge of antibiotic resistance at the interface between various disciplines ranging from medicinal chemistry to evolutionary biology.

Antibiotics can be broadly classified into bactericidal drugs that induce bacterial cell death, and bacteriostatic drugs that arrest cell growth [1]. However, whether an antibiotic is bactericidal or bacteriostatic can depend on factors such as antibiotic concentration, bacterial species, growth environment and growth rate, rendering this distinction ambiguous under certain situations [14]. In general, antibiotics inhibit vital cellular processes such as DNA, RNA, protein, cell wall and folic acid synthesis by targeting essential gene products that are required for these cellular functions [1, 14-16]. As shown in Table 1.1, each family of antibiotics can interact directly with one or more specific

primary targets, but the effects of such interactions typically extend to multiple bacterial cellular pathways. While the biochemical aspects of these direct interactions between the antibiotics and their primary targets are very well characterized, antibiotic-mediated cellular response mechanisms that eventually lead to cell death are not very well understood due to the complexities of interactions between multiple affected cellular pathways [1].

Family of antimicrobial drug	Examples	Primary Target	Cellular pathways affected
Fluoroquinolones: DNA synthesis inhibitors	Nalidixic acid, ciprofloxacin	DNA gyrase, topoisomerase IV	DNA replication, SOS response, cell division, ATP generation, Krebs cycle, iron–sulfur (Fe–S) cluster synthesis, reactive-oxygen species (ROS) formation, envelope & redox-responsive two-component systems
Trimethoprim–sulfamethoxazole: DNA synthesis inhibitors	Co-trimoxazole	Tetrahydrofolic acid synthesis inhibitors	Nucleotide biosynthesis, DNA replication
Rifamycins: RNA synthesis inhibitors	Rifamycins, rifampicin	RNA polymerase	Transcription, DNA replication, SOS response
β-lactams: Cell wall synthesis inhibitors	Penicillin, ampicillin, carbapenems	Penicillin-binding proteins	Cell wall synthesis, cell division, autolysin activity, SOS response, Krebs cycle, Fe-S cluster synthesis, ROS formation, envelope & redox-responsive two-component systems
Glycopeptides & glycolipopeptides: Cell wall synthesis inhibitors	Vancomycin	Peptidoglycan units	Cell wall synthesis, transglycosylation, transpeptidation, autolysin activation
Lipopeptides: Cell wall synthesis inhibitors	Daptomycin, polymixin B	Cell membrane	Cell wall synthesis, envelope two-component systems
Aminoglycosides: Protein synthesis inhibitors	Gentamicin, streptomycin, kanamycin	30S ribosome	Translation (mistranslation by tRNA mismatching), ETC, SOS response, Krebs cycle, Fe–S cluster synthesis, ROS formation, envelope & redox-responsive two-component systems
Tetracyclines: Protein synthesis inhibitors	Tetracycline, doxycycline		Translation (inhibition of binding between aminoacyl tRNA and ribosome)
Macrolides: Protein synthesis inhibitors	Erythromycin	50S ribosome	Translation (inhibition of elongation and translocation), free tRNA depletion
Streptogramins: Protein synthesis inhibitors	Pristinamycin		Translation (inhibition of initiation, elongation and translocation), free tRNA depletion
Phenicols: Protein synthesis inhibitors	Chloramphenicol		Translation (inhibition of elongation)

Table 1.1: A summary of major families of currently available antibiotics, their primary drug targets and bacterial cellular pathways that are adversely affected by these antibiotic drug-target interactions.

Table adapted from Kohanski *et al* (2010)

1.2 The fitness cost of antibiotic resistance mutations

Bacteria can acquire antibiotic resistance via two main mechanisms - spontaneous chromosomal mutations or horizontal gene transfer [3, 15, 17]. In the presence of antibiotics, bacterial populations evolve resistance to antibiotics through adaptation by natural selection [3]. Chromosomal antibiotic resistance mutations can alter the biochemical structures of primary antibiotic targets, such that antimicrobial molecules can no longer bind to these gene products [15, 18]. Other common resistance mechanisms include upregulation of drug efflux pumps [15, 19], inactivation of antibiotics via biochemical modification of antimicrobial molecules [15, 18], as well as gene duplication and amplification (GDA) mechanisms [15, 20]. All these resistance mechanisms are known to carry fitness costs, which are typically expressed in terms of reduced growth rate, competitive ability and virulence in resistant mutants relative to their antibiotic-sensitive ancestors [3, 15, 21-24]. In a laboratory setting, the fitness cost of resistance can be quantified *in vitro* and *in vivo* using the three main experimental approaches as illustrated in Figure 1.1 [15].

Chromosomal antibiotic resistance mutations in genes encoding essential proteins disrupt normal physiological processes [25] and generate a range of fitness costs, which have been observed in a variety of different model bacterial species *in vivo* and *in vitro* [23]. Similarly, plasmid-conferred antibiotic resistance is known to carry fitness costs due to the additional metabolic burden that is associated with plasmid carriage [3, 25], as well as intragenomic conflicts between plasmid and chromosomal genes [3, 26-28]. The fitness cost of resistance is expected to play a crucial role in influencing the evolutionary dynamics of resistance, including the rate of development of resistance and compensatory adaptation in resistant mutant populations [3, 15].

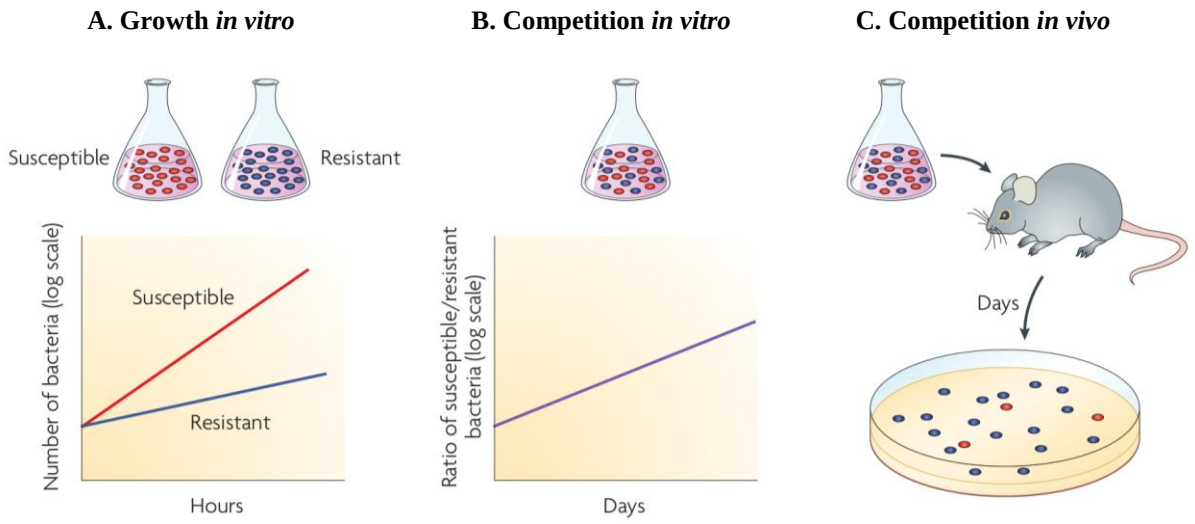


Figure 1.1: Three experimental approaches for quantifying the fitness cost of antibiotic resistance using *in vitro* and *in vivo* models

Figure from Andersson & Hughes (2010)

Panel A: The difference between the *in vitro* growth rates of an antibiotic resistant mutant and an antibiotic susceptible ancestral strain during the exponential phase can be used as a proxy measurement for the fitness cost of resistance. However, such exponential growth rate measurements obtained from individual bacterial cultures do not take into consideration any differences in competitive fitness during the lag and stationary phases of bacterial growth.

Panel B: In a pairwise *in vitro* competition experiment, an antibiotic resistant mutant is typically co-cultured with an antibiotic susceptible ancestral strain that has been tagged with a reporter gene (fluorescent protein genes, β-galactosidase encoded by *lacZ* etc.). The ratio of the cell number of the two different strains is determined before subjecting the co-culture to a bottleneck and overnight incubation each day. The difference in the competitive fitness of the two co-cultured strains should result in an altered ratio of the cell number of the two strains after each day.

Panel C: Pairwise competition experiments can be performed in an *in vivo* model. Animal subjects are injected with a co-culture of a resistant mutant and a susceptible ancestral strain. The ratio of the cell number of the two different strains isolated from the infected animal subjects is expected to change after a few days due to the cost of resistance in the resistant mutant strain.

In the context of chromosomal antibiotic resistance, our current understanding of the molecular underpinnings of the cost of resistance often centers on the genetic and biochemical aspects of how resistance mutations interfere with vital cellular functions of essential proteins that have undergone structural changes as a result of these deleterious mutations [3]. In an extensively characterized example of chromosomal antibiotic resistance, the fitness effect of streptomycin resistance in *Salmonella typhimurium* has been shown to be a consequence of reductions in ribosomal activity by compromising translation elongation rate due to mutations in the ribosomal protein gene *rpsL* [29, 30]. In addition to the direct effects of *rpsL* mutations on elongation rate, Paulander *et al* have also elucidated the pleiotropic effects of *rpsL* mutations on the activation of the carbon starvation response. While *rpsL* mutants grow slower than the wild-type in a rich growth medium, the reverse trend was observed in growth media with poorer carbon sources due to the failure of *rpsL* mutants to activate the stress-inducible sigma factor RpoS and the RpoS-mediated general stress response [29]. This observation demonstrates that the cost of resistance can be strongly influenced by the environment in which it is measured, which has also been described in other models of chromosomal antibiotic resistance [15, 24, 31, 32].

A large number of laboratory-based *in vitro* growth rate and competitive fitness assays have demonstrated that even within the same environment, different resistance mutations in the same chromosomal genes that confer resistance to the same antibiotics can be associated with very different fitness costs [29, 33-37]. This is an important research topic in antibiotic resistance and microbial evolution, because understanding the molecular underpinnings of how resistance mutations influence bacterial phenotype and translate into variations in mutant fitness would allow us to predict the fitness cost of resistance with a view to developing novel treatment strategies that generate weaker selection for resistance [3]. However, our current understanding of the molecular underpinnings which contribute to these variations in costs of resistance is limited [3]. In two notable exceptions, the molecular mechanisms underlying variations in the fitness effects of mutations that confer resistance to fusidic acid and peptide deformylase inhibitors (PDFIs) have been successfully dissected and were shown to be due to variations in protein translation rate [36, 37]. More specifically, fusidic acid

resistance mutations in the *fusA* gene, which encodes elongation factor G (EF-G), result in compromised translation elongation rate [36], while PDFI resistance mutations in the *fmt* gene, which encodes formyl-methionine-transferase, have deleterious effects on translation initiation rate [37]. In both examples, resistance mutations that result in the highest impairments in enzymatic functions are associated with the highest fitness costs.

A systems biology approach is highly suitable for dissecting the molecular origins of the cost of resistance, because resistance mutations are expected to have direct impacts on the enzymatic activities of essential proteins that have acquired these deleterious mutations, as well as genome-wide indirect effects on other cellular pathways and networks that interact with these mutant enzymes. For example, in addition to the direct impacts of fusidic acid resistance mutations on translation elongation rate, *fusA* mutations also have the pleiotropic effects of perturbing cellular levels of the transcriptional regulator molecule guanosine tetraphosphate (ppGpp) during both normal growth and starvation [36]. Since the ppGpp-mediated stringent response is a genome-wide stress response to a variety of physiological stress and nutritional starvation [38], fusidic acid resistant mutants are expected to show global changes in their gene expression patterns compared to the wild-type strain when grown in minimal medium. A systems-level approach allows such genome-wide changes in transcriptome patterns in resistant mutants to be captured. This could in turn help to pinpoint specific, pleiotropic transcriptional changes that may have contributed to deleterious phenotypic consequences and retarded growth rate.

Some resistant mutants can recover their compromised fitness by fixing compensatory mutations (also known as suppressor mutations) without compromising their resistance phenotype [3, 15, 25, 39]. Compensatory mutations are typically second-site mutations that specifically offset the deleterious effects of resistance mutations through epistatic interactions with the original resistance mutations [40-44]. Understanding how compensatory mutations offset the deleterious fitness effects of resistance mutations can provide additional useful insight into the molecular origins of the fitness cost of resistance. In general, there are two types of compensatory mutations. Intragenic compensatory mutations restore the structure and function of mutant gene products, while extragenic compensatory

mutations restore the structure and function of mutant multi-subunit enzymes or organelles [42]. A variety of intragenic and extragenic compensatory mutations have been observed in a wide range of *in vitro* and *in vivo* studies [15]. Compensatory mutations tend to directly or indirectly restore the enzymatic efficiency or functions of mutant gene products [15]. However, it has also been noted that it can be difficult for compensatory mutations to completely ameliorate the fitness defects associated with pleiotropic effects of resistance mutations that affect multiple cellular processes [39, 45]. A systems biology approach should enable us to quantify the extent to which compensatory mutations can ameliorate the direct effects of resistance mutations on the enzymatic activities of affected gene products, and offset the pleiotropic fitness costs of resistance mutations at a genome-wide level.

1.3 The fitness cost of rifampicin resistance

Rifampicin resistance is a useful model for studying variations in the fitness costs of chromosomal antibiotic resistance, because rifampicin resistance mutations are associated with a range of different fitness costs [17, 46-48]. First synthesized in 1959, rifampicin (also known as rifampin) is a semi-synthetic derivative of rifamycin B, which was originally isolated from the Gram-positive actinomycete *Amycolatopsis mediterranei* S699 [49, 50]. Rifampicin is one of the most potent and broad-spectrum antibiotics that is currently used to treat a range of diseases such as tuberculosis, leprosy, brucellosis and AIDS-related mycobacterial infections [51-54]. Depending on its local concentration, rifampicin can be either bacteriostatic or bactericidal [55-57].

Rifampicin inhibits transcription by targeting a small, but highly conserved pocket located within the DNA/RNA channel of the RNA polymerase (RNAP) β-subunit, which is encoded by *rpoB* [58, 59]. Once rifampicin binds to this pocket, nascent RNA cannot elongate beyond 2-3 nucleotides in length as its path is sterically blocked (Figure 1.2) [51]. Non-synonymous mutations in residues that make direct contact with rifampicin or are in the vicinity of the binding site confer resistance by altering the three-dimensional structure of the rifampicin-binding pocket in the highly conserved DNA/RNA channel and decreasing the binding affinity between rifampicin and RNAP [60].

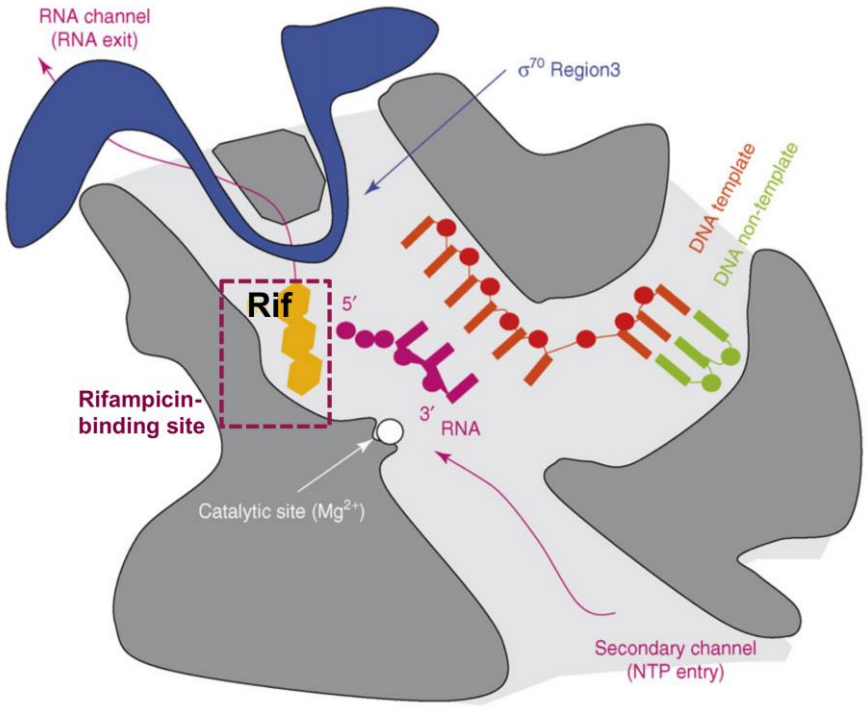
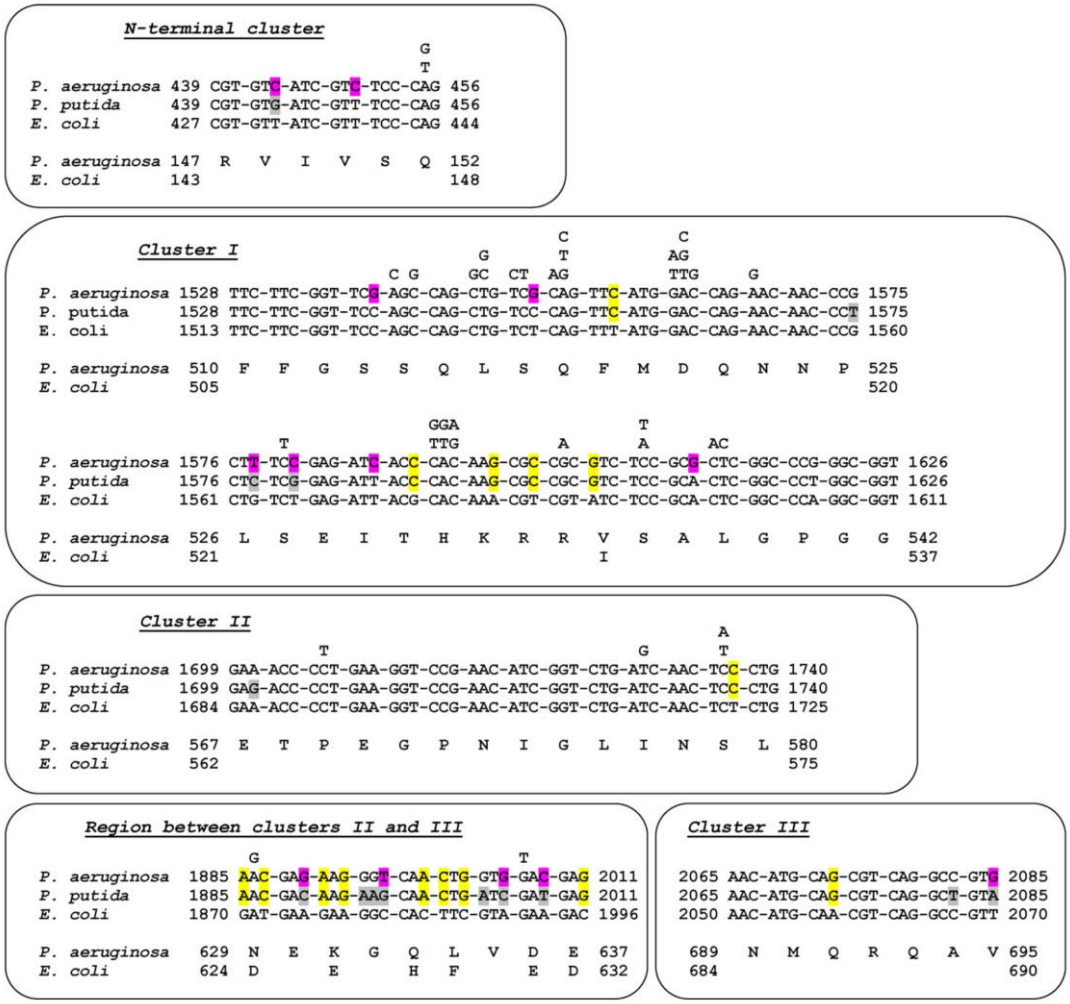


Figure 1.2: DNA/RNA channel of wild-type RNAP. Rifampicin (Rif) binds to a highly-conserved rifampicin-binding pocket and inhibits transcription [55].

Figure adapted from Villain-Guillot *et al* (2007)

The vast majority of rifampicin resistance mutations are single base substitutions that can be detected in Clusters I, II and III of the central rifampicin resistance region (Figure 1.3), but a small number of resistance mutations can also be found in the N-terminal cluster [17]. At a phenotypic level, *rpoB* mutations confer different levels of resistance to rifampicin, which were broadly described as strong (full resistance) or weak (partial resistance) in *P. aeruginosa* by Jatsenko *et al* [17]. However, there is no correlation between the level of rifampicin resistance and the fitness cost, at least in *Staphylococcus aureus* [48] and *Pseudomonas aeruginosa* (MacLean, unpublished data).



Key:

Pink - nucleotides unique to the *Pseudomonas aeruginosa* PAO1reference strain
Grey - nucleotides unique to the *Pseudomonas putida* PaW85 reference strain
Yellow - nucleotides identical in both *Pseudomonads*, but not in *E. coli*

Figure 1.3: Rifampicin-resistant *rpoB* mutations are annotated above the nucleotide sequences of the central rifampicin resistance region in the *rpoB* gene from three Gram-negative species

Figure from Jatsenko *et al* (2009)

Classic experiments involving rifampicin resistant mutants of *E. coli* showed that various *rpoB* mutations have adverse impacts on the enzymatic activity of RNAP characterized by altering transcription initiation rate, elongation rate and termination phenotypes [61-63]. A range of fitness costs associated with different *rpoB* mutations have been observed in at least four different bacterial species [21, 46, 48, 64, 65]. *rpoB* mutations are thought to generate a direct fitness cost by

compromising RNAP enzymatic activity. In support of this idea, Brandis *et al* established that one *rpoB* mutant of *Salmonella enterica* with reduced growth rate showed a reduction in transcriptional efficiency relative to the wild-type [66]. Using a semi-quantitative β-galactosidase reporter gene assay, Reynolds demonstrated that three costly rifampicin resistant mutants of *E. coli* showed reductions in transcriptional efficiency, and reported a significant linear correlation between mutant fitness and transcriptional efficiency [46]. However, both studies are highly limited in terms of the number of different *rpoB* mutations that were phenotypically assayed. Given the semi-quantitative nature of transcriptional efficiency measurements and the small number of *rpoB* mutations that were characterized in Reynolds' study, further investigation is required to establish the validity and the causality of the observed linear correlation.

Pleiotropy is defined as the phenomenon in which a single gene locus affects multiple apparently unrelated phenotypic traits, and is often identified as single mutations that affect multiple wild-type traits [67]. Using *Saccharomyces cerevisiae* as a model organism, Cooper *et al* (2007) demonstrated that deletion mutations in genes with more pleiotropic effects tend to have greater deleterious effects on fitness [68]. Since the RNAP holoenzyme interacts with every promoter of every transcribed gene or operon during transcription initiation [32], *rpoB* mutations are thought to be global regulatory mutations that change the transcriptional profiles of mutants by altering the promoter selectivity of mutant RNAP [69]. Given the highly pleiotropic nature of mutations in RNAP and the diversity of transcriptional regulatory mechanisms [70-72], *rpoB* mutations are expected to generate an additional fitness cost that is independent from that associated with the direct effects of *rpoB* mutations on RNAP activity. For example, rifampicin resistance mutations were shown to have extensive pleiotropic effects ranging from alterations in carbon catabolism in *rpoB* mutants *E. coli* [73], *B. subtilis* [32] and *P. aeruginosa* [65], to changes in lipid metabolism in *M. tuberculosis* [74]. However, these pleiotropic effects have not been studied in relation to their fitness effects on *rpoB* mutants.

1.4 The opportunistic pathogen *Pseudomonas aeruginosa*

P. aeruginosa is a versatile and ubiquitous bacterium that has gained clinical relevance as a significant opportunistic pathogen amongst patients suffering from conditions such as cystic fibrosis, burn injuries and immunodeficiency [75, 76]. The PAO1 reference strain of *P. aeruginosa* is the most commonly used strain in laboratory research [77, 78] and was originally isolated from the wound of a patient in 1954 [78]. The complete genome sequence of PAO1 was the first *P. aeruginosa* genome to be annotated [76]. It is thought that the relatively large size (6.3 Mb) and complexity of the PAO1 genome reflect evolutionary adaptations which allow *P. aeruginosa* to thrive in a variety of ecological niches, ranging from soil water to animal and plant tissues [76]. Another commonly used *P. aeruginosa* reference strain is the clinically derived PA14 strain, which is the most common clonal group worldwide [79] and shows a higher level of virulence compared to PAO1 [80].

Like many Gram-negative opportunistic pathogenic bacteria, *P. aeruginosa* utilizes quorum sensing via *N*-acylhomoserine lactone (AHL) signaling to coordinate population behavior during colonization of animal and plant hosts [81]. AHLs are diffusible autoinducers that accumulate in the extracellular environment during bacterial growth [82]. In *P.* aeruginosa, two different AHL autoinducers are synthesized from their precursors by a synthase enzyme (LasI or RhlI) and interact with their cognate transcriptional regulators (LasR or RhlR) to activate the expression of quorum sensing genes [83, 84]. Quorum sensing allows *P. aeruginosa* to sense if a critical population density has been reached in the host tissue [85]. Once AHL molecules accumulate to a threshold concentration, the transcription of quorum sensing regulated genes is activated [84]. It has been estimated that 6-10% of the *P. aeruginosa* genome is regulated by the LasR-LasI and RhlR-RhlI systems [86].

A key quorum sensing regulated trait in *P. aeruginosa* is the secretion of virulence factors that are responsible for damaging host tissues during the initiation of infection [85, 87]. Examples of these virulence factors include elastase, alkaline phosphatase and an extracellular protease, which are encoded by *lasB, aprA* and *lasA* respectively [84]. Quorum sensing also regulates complex cooperative behaviors such as biofilm formation [88-90]. Biofilms are formed when bacteria

aggregate within a surfaced-attached extracellular hydrated polymeric matrix synthesized by themselves [91]. *P. aeruginosa* biofilms present a significant challenge to the medical community due to their increased resistance to antimicrobials, disinfectants and actions of the host immune system [91, 92]. Biofilms are also responsible for persistent and chronic infections in cystic fibrosis patients, as well as amongst patients with indwelling medical devices such as catheters [91, 93, 94].

P. aeruginosa shows intrinsic resistance to many front-line antibiotics and disinfectants, which is mainly due to its low outer-membrane permeability to antimicrobials and the presence of multidrug efflux pumps [95, 96]. In addition, the *P. aeruginosa* genome encodes a periplasmic, AmpC-like inducible β-lactamase that can hydrolyze β-lactams such as penicillin [96]. Table 1.2 shows a list of common antibiotics that are currently used to treat *P. aeruginosa* infections. Each class of antibiotics offers certain advantages and disadvantages. Even though rifampicin is rarely used as an antipseudomonal drug [96, 97], it is conceivable that future clinical treatments of *P. aeruginosa* infections may involve the use of cocktails of rifampicin and other antibiotics if resistance against currently used antibiotics continues to develop and spread.

Class of antibiotics	Examples	Advantages	Disadvantages
Penicillin	Carbenicillin, ticarcillin	Synergistic with aminoglycosides against *P. aeruginosa*	May induce β-lactamases in *P. aeruginosa*
Cephalosporin	Ceftazidime, cefoperazone	Can be used as a single agent against *P. aeruginosa*	
Aminoglycoside	Gentamicin, amikacin, tobramycin	Synergistic with β-lactams against *P. aeruginosa*	Narrow therapeutic/toxic ratio, penetrate poorly into cerebrospinal fluid
Quinolone	Ciprofloxacin	Can be taken orally	Not suitable for children under the age of 16
Polymyxin	Colistin	Very active, little resistance development	Possible toxicity concerns, used largely in cystic fibrosis patients
Carbapenem	Meropenem, imipenem,	Very broad-spectrum activity against Gram-negative bacteria	May induce β-lactamases, rapid development of resistance

Table 1.2: A summary of common antibiotics that are currently used to treat *P. aeruginosa* infections.

Table adapted from Hancock & Speert (2000)

As a leading cause of nosocomial infections [98, 99], *P. aeruginosa* is often considered a clinically relevant and useful pathogen for studying antibiotic resistance. To date, the genome of *P. aeruginosa* has been extensively characterized and annotated [76, 100, 101]. In addition, a large body of work focusing on the genotypic, phenotypic and pathogenic traits of *P. aeruginosa* is available in scientific literature, making this species a tractable model organism for laboratory experimentation. In the context of this study, *P. aeruginosa* PAO1 is used as a model organism for studying rifampicin resistance, and it is hoped that the conclusions of this study may be extrapolated to *Mycobacterium tuberculosis*, the causative agent of most cases of tuberculosis [102]. A similar approach involving the use of other model bacterial species to infer the genotypic and phenotypic effects of rifampicin resistance on *M. tuberculosis* has been adopted in several other recent studies [66, 103, 104].

1.5 An experimental evolution approach for studying adaptive evolution of antibiotic resistant mutants

A secondary theme of my doctoral research centers on understanding how the fitness costs of chromosomal antibiotic resistance influence the adaptation of resistant mutants in an antibiotic-free environment. Due to the cost of resistance, bacterial pathogens that carry resistance mutations are expected to be under negative selection in the absence of antibiotics, e.g. when patients stop antibiotic use, or when pathogens transmit between patients via antibiotic-free environments. Therefore, resistance is only likely to persist in the long term if resistant lineages can evolve increased fitness to offset the cost of resistance. The classic way to study the adaptation of resistant mutants is to subject them to sequential serial passage in the absence of antibiotics during *in vitro* or *in vivo* experimental evolution assays. During adaptive evolution, the average fitness of mutants is expected to rise over time due to selection for compensatory mutations [3, 42] and/or generally beneficial mutations [43]. Compensatory adaptation has been defined in several ways, but in general, compensatory mutations are defined as second-site deleterious mutations that fully or partially restore the deleterious effects of the original resistance mutations through epistatic interactions [40-44]. In contrast, generally

beneficial mutations are defined as mutations that confer a fitness advantage on mutants regardless of the genetic background in which they arise during general adaptation [43, 105]. Reversion mutations that result in the loss of resistance mutations are also possible during the adaptive evolution of resistance, but they are expected to be extremely rare events due to their extremely low frequencies [106] and tend to be lost subsequently due to genetic drift and clonal interference [107].

During compensatory adaptation, the antibiotic resistance phenotype is often maintained even though resistant mutants were serially passaged in an antibiotic-free environment [15, 46, 108, 109]. At a molecular level, there are currently many specific examples of compensatory mutations that have been well characterized in literature across various models of chromosomal antibiotic resistance in various bacterial species [42, 104, 109, 110]. In *Salmonella typhimurium*, a streptomycin resistant mutant carrying the point mutation K42N in the ribosomal gene *rpsL* was found to have acquired intragenic, as well as extragenic mutations in 3 additional functionally related ribosomal genes (*rpsD*, *rpsE* and *rplS*) during compensatory adaptation [42, 109, 111]. Likewise, previous work has shown that rifampicin resistant bacteria can adapt to the cost of resistance by fixing second-site mutations in both within the *rpoB* gene, as well as in *rpoA* and *rpoC*, which code for two other subunits of the RNAP core enzyme [3, 112].

However, the relative contribution of compensatory adaptation towards the adaptive evolution of antibiotic resistant mutants is not well understood. To date, a large number of experimental evolution studies aimed at understanding adaptation in resistant mutants have reported high frequencies and large varieties of compensatory mutations that can be found in endpoint populations of resistant mutants that were serially passaged in antibiotic-free environments [108, 109, 113-116]. However, in a study using *Mycobacterium smegmatis* strains carrying chromosomal mutations conferring resistance to three different classes of antibiotics that inhibit protein synthesis, Sander *et al* (2002) showed that compensatory adaptation was not observed in mutants which carry low-cost or no-cost mutations [117]. I aim to address such discrepancies in literature by initiating my *in vitro* laboratory evolution experiments using a range of resistant mutants carrying *rpoB* mutations that are associated with different fitness costs.

The combination of laboratory experimental evolution and whole-genome sequencing is a powerful technique that can be applied to study the adaptive evolution of resistant mutants under controlled laboratory conditions [118]. In the context of experimental evolution, a selection experiment involves introducing clonal groups of an organism with different genetic backgrounds to identical microcosms [119]. Parallel lineages are allowed to propagate independently in batch culture under the same chosen experimental conditions, which allows the selection of improved phenotypic traits such as fitness [120]. By manipulating natural selection in a controlled way under laboratory conditions, selection experiments allow us to gain a better understanding of fundamental evolutionary processes that underlie adaptation [121]. Whole-genome sequencing of the endpoint isolates of these resistant mutants should allow us to gain insight into the genetic mechanisms that contribute to their adaptive evolution and phenotypic alterations.

1.6 Research objectives and hypotheses

1.6.1 System-wide molecular mechanisms that contribute to the fitness cost of rifampicin resistance

In Chapter 2, I aim to dissect the molecular mechanisms that contribute to the variations in the fitness cost of rifampicin resistance in *P. aeruginosa*. I hypothesize that most of the variation in the fitness cost of rifampicin resistance can be attributed to the variation in transcriptional efficiencydue to compromised RNAP activity, which can be considered the direct effect of *rpoB* mutations. The remaining fitness cost, which cannot be accounted for by the variations in transcriptional efficiency, is expected to stem from the indirect, pleiotropic effects of *rpoB* mutations.

To quantify the extent to which the direct effects of *rpoB* mutations on RNAP activity can account for the variations in fitness cost, I shall perform phenotypic measurements of transcriptional efficiency and competitive fitness in different *rpoB* mutants and their common rifampicin sensitive ancestral strain. The linear correlation between the two measured quantities should provide an indication of the degree to which the variations in mutant fitness can be explained by the variations in transcriptional

efficiency. In this study, I shall focus on the effects of *rpoB* mutations on the expression levels of essential genes using gene expression analyses using quantitative real-time PCR (qPCR) and RNA-Seq. Genes are classified as essential or nonessential based on their functional significance with respect to the effects of their gene products on the organism's fitness, and by definition, a gene is considered to be essential if a knock-out of this gene results in a lethal phenotype in a given environment [122]. Given the fitness effects of essential genes, an overall reduction in the expression levels of essentials genes is expected to generate a fitness cost in resistant mutants.

In addition to essential gene expression, I aim to expand my investigation of the pleiotropic effects of rifampicin resistance mutations by examining if the activation of global stress responses can be observed, an example of which is the ppGpp-mediated stringent response mentioned earlier. Preliminary evidence from literature suggests that the stringent response is likely to be activated by *rpoB* mutations, because the ppGpp binding site in the β-subunit of RNAP is in very close proximity (27Å) to the rifampicin-binding domain [123, 124]. The activation of the stringent response directly alters the promoter selectivity of RNAP by inhibiting the transcription of ribosomal genes, but at the same time, stimulating the expression of proteolytic and amino acid biosynthesis enzymes, which translates into retarded growth rate at a phenotypic level [38]. In *E. coli*, four different rifampicin resistance mutations result in the constitutive activation of the stringent response in *rpoB* mutants even in the absence of any physiological stress and nutritional starvation [125], presumably due to structural alterations in mutant RNAP enzyme. However, it is unclear if this phenomenon is specific to these four *rpoB* mutations. It is hoped that RNA-Seq transcriptome profiling of *rpoB* mutants will reveal any preferential transcription of stringent response regulated genes in *rpoB* mutants of *P. aeruginosa.* If *rpoB* mutants are adversely affected by the activation of global stress responses such as the stringent response, it is conceivable that adjunctive treatments could be designed with a view to reducing the selective advantage of resistant mutants during rifampicin treatment by exploiting the fitness cost that stems from the activation of these stress responses, thereby constraining the evolution of rifampicin resistance [3].

1.6.2 Understanding the role of fitness cost of resistance in influencing compensatory adaptation

The objective of Chapter 3 is to understand the role and importance of fitness cost of resistance in determining which resistant mutant populations are more likely to undergo compensatory adaptation. More specifically, I hypothesize that mutant populations carrying resistance mutations with higher fitness costs are more likely to undergo compensatory adaptation. In support of this view, Poon & Chao (2005) reported that DNA bacteriophage φX174 populations carrying deleterious mutations with more severe fitness effects were significantly more likely to fix compensatory mutations during a selection experiment [41], which highlights the important role of fitness cost in influencing the evolutionary dynamics of viral mutants carrying deleterious mutations.

To achieve this objective, I shall adopt an experimental evolution approach in conjunction with whole-genome sequencing to uncover genome-wide mutations that have accumulated in the endpoint populations of different *rpoB* mutants of *P. aeruginosa* that were serially passaged in the absence of rifampicin. Whole-genome sequencing can also shed light on the genetic basis of general adaptation, which is currently poorly understood given the preponderance of compensatory mutations that have been reported in literature. By including 3 independent lineages of each mutant carrying the same *rpoB* mutations in the selection experiment, I aim to quantify the degree to which parallel evolution can be observed at a genome-wide level [118].

1.6.3 Dissecting the fitness cost of pNUK73 small plasmid carriage in *P. aeruginosa*

In addition to chromosomal antibiotic resistance mutations, mobile genetic elements such as plasmids and lysogenic phages acquired through horizontal gene transfer can also confer resistance to various antibiotics [3, 15, 25, 126]. Mobile genetic elements are typically associated with fitness costs [3, 15, 126]. The synthesis of additional nucleic acids and proteins constitutes metabolic burdens on the bacterial host [3, 25]. In addition, intragenomic conflicts between the host chromosomal genes and mobile genetic elements may interfere with normal cellular physiology [3, 25, 126]. In Chapter 4, I aim to apply the integrative experimental approach that I have developed in other sections of this thesis in a different context to dissect the molecular underpinnings of the fitness cost associated with carriage of the pNUK73 small plasmid.

The pNUK73 plasmid was engineered from the cryptic natural pNI10 plasmid, which was originally isolated from *Pseudomonas fulva*, to function as a shuttle vector [127]. pNUK73 carries three genes; namely a replication protein gene (*rep*), a kanamycin/neomycin resistance gene and the *lacZ* reporter gene. A recent study demonstrated that pNUK73-bearing *P. aeruginosa* PAO1 host (PAO1/pNUK73) showed a significantly lower competitive fitness relative to wild-type PAO1 [128]. After being serially passaged for 300 generations, endpoint isolates of PAO1/pNUK73 showed wild-type competitive fitness level, suggesting that the fitness cost associated with pNUK73 carriage has been alleviated. Whole-genome sequencing of several endpoint clones of PAO1/pNUK73 revealed that the PAO1 host adapted to pNUK73 by fixing mutations in three chromosomal genes, which include a putative helicase (PA1372) and two putative serine/threonine protein kinases (PA4673.15 and PA4673.16).

In the work presented in Appendix A, we demonstrated that all three chromosomal genes that carry compensatory mutations are recently acquired genes in the PAO1 genome (bioinformatic analyses, Appendix A). Using RNA-Seq, the transcriptome profiles of parental PAO1/pNUK73 and endpoint isolates carrying compensatory mutations were obtained. Amongst approximately 700 genes that were differentially expressed in parental PAO1/pNUK73 relative to the PAO1 wild-type strain, key genes

associated with the SOS response pathway exhibited significant upregulation. Interestingly, compensatory mutations in the three recently acquired chromosomal genes silenced both *rep* expression and the SOS response. Since the expression levels of the other two plasmid-encoded genes did not show any significant reduction in the compensated mutants relative to the PAO1/pNUK73 initial population, we hypothesized that most of the fitness cost associated with pNUK73 carriage should stem from *rep* expression alone.

Recently acquired genes that are carried by plasmids have the potential to function inefficiently within the new genomic backgrounds of their bacterial host and disrupt existing regulatory and physiology networks in the host cell [126]. It is likely that the fitness cost of pNUK73 carriage stems from conflicts between the gene products of the three recently acquired chromosomal genes (PA1372, PA4673.15 and PA4673.16) and the Rep plasmid replication protein. We aim to understand the molecular mechanism through which intragenomic conflict generates a fitness cost in PAO1/pNUK73.

Given the significant upregulation of the SOS regulon in PAO1/pNUK73, we hypothesized that the SOS response is a potential molecular mechanism that contributes to the fitness cost of pNUK73 carriage. The SOS response is an important global stress-response in response to DNA damage [129]. During normal growth, the auto-repressor protein LexA represses the transcription of the SOS regulon [130], whose gene products facilitate DNA repair or tolerance to DNA damage [131]. The SOS response can be induced by single-stranded DNA (ssDNA) from stalled replication forks or from direct DNA damage due to exposure to DNA-damaging agents such as UV irradiation and antibiotics (e.g. ciprofloxacin) [129, 132]. As illustrated in Figure 1.4 (left panel), the SOS response is activated by *recA* expression in response to the accumulation of significant DNA damage. RecA is an ATPase that mediates DNA homologous recombination by forming a helical nucleoprotein filament with ssDNA [133-135] and induces the autocatalytic cleavage of LexA [131, 136]. This inactivation of LexA repressor proteins results in the controlled derepression and expression of stress response genes in the SOS regulon [130]. The activation of the SOS response carries a significant fitness cost due to

the expression of the SfiA (SulA) protein, which strongly inhibits cell division [137] and promotes cell filamentation [138].

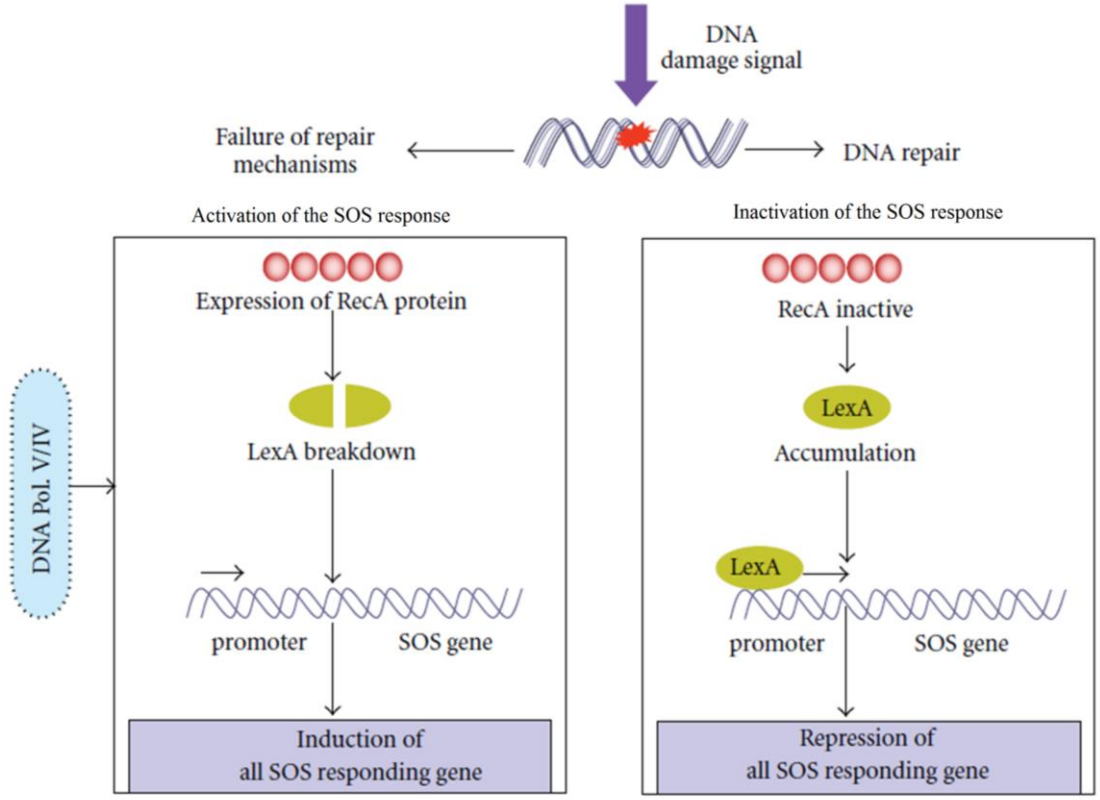


Figure 1.4: Schematic summarizing how the SOS response is activated or inactivated

In the presence of significant DNA damage (left panel), the expression of RecA results in the autocatalytic cleavage of LexA repressor proteins and the induction of the SOS regulon. The expression of the error-prone DNA polymerase V/IV is a part of the mutagenic SOS response. When the SOS response is not activated (right panel), LexA directly represses the expression of SOS genes at a transcriptional level [139].

Figure adapted from Rastogi *et al* (2010)

To test our hypotheses, I shall insert an IPTG-inducible *rep* construct into a neutral chromosomal site in the genome of PAO1 to verify if the induction of *rep* expression can generate a significant fitness cost. Since the upregulation of two key SOS regulator genes, *recA* and *lexA*, is a molecular signature for detecting the activation of the SOS response [140], I aim to quantify the gene expression levels of

these two genes by quantitative real-time PCR to ascertain if the activation of the SOS response can be observed in conjunction with *rep* expression. These gene expression profiles will be compared against that of PAO1/pNUK73 to assess the extent to which the two *rep*-expressing strains show phenotypic and genotypic similarities, which will provide useful insight into the molecular mechanisms through which pNUK73 carriage generates a fitness cost. My contributions to this work are outlined in the Contributions of Authors section on page 4.

2. Linking system-wide impacts of RNA polymerase mutations to the fitness cost of rifampicin resistance in *Pseudomonas aeruginosa*

Qin Qi [1], Gail M. Preston [2], R. Craig MacLean [1]

1. Department of Zoology, University of Oxford, United Kingdom;
2. Department of Plant Sciences, University of Oxford, United Kingdom

Abstract

Fitness costs play a key role in the evolutionary dynamics of antibiotic resistance in bacteria by generating selection against resistance in the absence of antibiotics. Although the genetic basis of antibiotic resistance is well understood, the precise molecular mechanisms linking the genetic basis of resistance to its fitness cost remain poorly characterized. Here, we examine how the system-wide impacts of mutations in the RNA polymerase gene *rpoB* shape the fitness cost of rifampicin resistance in *Pseudomonas aeruginosa*. Rifampicin resistance mutations reduce transcriptional efficiency, and this explains 76% of the variation in fitness among *rpoB* mutants. The pleiotropic consequence of *rpoB* mutations is that mutants show altered relative transcript levels of essential genes. We find no evidence that global transcriptional responses impact the fitness cost of rifampicin resistance as revealed by RNA-Seq. Global changes in gene expression levels associated with *rpoB* mutations are subtle, demonstrating that the transcriptional regulatory network of *P. aeruginosa* is robust to decreased transcriptional efficiency associated with *rpoB* mutations. On a smaller scale, we find that rifampicin resistance mutations increase the expression of RNA polymerase due to decreased termination at an attenuator upstream of *rpoB*, and we argue that this helps to minimize the cost of rifampicin resistance by buffering against reduced RNAP activity. In summary, our study shows that it is possible to dissect the molecular mechanisms underpinning variation in the cost of antibiotic resistance, and highlights the importance of genomic buffering in providing robustness against resistance mutation.

Introduction

Antibiotic resistance conferred by chromosomal mutations is typically accompanied by fitness costs, expressed in terms of reduced growth rate, competitive ability and virulence of resistant mutants compared to their antibiotic-sensitive ancestors in the absence of antibiotics [3, 15, 23]. Fitness costs play a key role in the evolutionary dynamics of antibiotic resistance by reducing the rate at which resistance mutations spread in bacterial populations that are exposed to antibiotics and accelerating the rate at which resistance mutations are lost once antibiotic use is stopped [3]. At least 80 studies have quantified the fitness costs associated with chromosomal antibiotic resistance mutations [141]. Intriguingly, some of these studies show that different mutations that confer resistance to the same antibiotic can be associated with very different fitness costs [29, 33-35, 65]. How can we explain this variation in the fitness effects of resistance mutations? Elegant experiments have shown that the variation in the fitness effects of mutations that confer resistance to fusidic acid and peptide deformylase inhibitors can be attributed to variation in the rate of protein synthesis [36] and translation initiation [37] respectively. Beyond these examples, the molecular basis of fitness in antibiotic resistant mutants remains unresolved. In this paper, we use a systems-level approach to dissect the mechanistic basis of fitness in rifampicin resistant mutants of the pathogenic bacterium *Pseudomonas aeruginosa*.

Rifampicin is a broad-spectrum antibiotic [142] that inhibits bacterial RNA synthesis by specifically targeting a small but highly conserved pocket in the β-subunit of RNA polymerase (RNAP) encoded by *rpoB* [55]. When rifampicin binds to the rifampicin-binding pocket within the DNA/RNA channel of a wild-type RNAP, the path of nascent RNA transcripts is directly blocked and elongation cannot proceed beyond the first three nucleotides [58, 59]. Mutations in *rpoB* can result in alterations to the structure of the rifampicin-binding pocket and confer rifampicin resistance by decreasing the binding affinity between rifampicin and RNAP [60]. Previous studies have demonstrated that rifampicin resistance mutations carry variable fitness costs in *Pseudomonas aeruginosa* [65], *Escherichia coli* [64], *Mycobacterium tuberculosis* [21] and *Staphylococcus aureus* [48]. First, by altering the structure

of a highly conserved domain of RNAP, *rpoB* mutations are thought to generate a direct fitness cost by compromising transcriptional efficiency. In support of this argument, Brandis *et al* established that one *rpoB* mutant of *Salmonella enterica* with a slower growth rate in a rich medium also showed a reduction in transcriptional efficiency relative to the wild-type [66]. Similarly, Reynolds showed that three costly rifampicin resistant mutants of *E. coli* had reduced transcriptional efficiency using a semi-quantitative β-galactosidase reporter gene assay [46].

In addition to the direct effect of rifampicin-resistance mutations on RNAP activity, it is likely that *rpoB* mutations generate an indirect, pleiotropic cost by altering the transcriptional profiles of mutants at a genome-wide scale. Since RNAP is required for the transcription of every gene, *rpoB* mutations are thought to be global regulatory mutations that alter the transcriptional profiles of mutants [69]. Previous work has shown that *rpoB* mutations have pleiotropic effects on carbon catabolism in *E. coli* [73], *B. subtilis* [32] and *P. aeruginosa* [65], as well as lipid metabolism in *M. tuberculosis* [74]. Pleiotropy has been demonstrated to generate a fitness cost in both *S. cerevisae* [68] and *P. fluorescens* [143], suggesting that pleiotropic effects of *rpoB* mutations on gene expression may also contribute to the cost of rifampicin resistance. However, this hypothesis has not been tested in previous work.

Although the genetic and biochemical basis of rifampicin resistance has been well characterized, an important question concerning the fitness costs still remains - why do mutations in the same gene that confer resistance to the same antibiotic carry different fitness costs? If a reduction in RNAP activity is the principal cause of decreased fitness, what are the contributions of the direct effect (i.e. compromised transcriptional efficiency) and pleiotropic effects of resistance to the fitness cost of resistance? In this study, we address this problem by attempting to link the direct and pleiotropic effects of rifampicin resistance mutations with fitness. First, we used a luciferase reporter gene to quantify the direct effects of *rpoB* mutations on transcriptional efficiency using luciferase reporter activity as a proxy measurement. Second, to determine the pleiotropic effects of *rpoB* mutations on gene expression, we quantified the average transcript levels of a random sample of essential genes in resistant mutants relative to the rifampicin sensitive ancestral strain. In addition, the whole-genome

transcriptional profiles of two common *rpoB* mutants and the ancestral strain were acquired through RNA-Seq to uncover the genome-wide effects of resistance mutations. Using this approach, we showed that the direct effect of *rpoB* mutations on transcriptional efficiency is responsible for most of the variations in the fitness cost of rifampicin resistance mutations. The pleiotropic effects of *rpoB* mutations on essential gene expression is likely to be a consequence of the reduction in transcriptional efficiency, and do not generate a fitness cost in their own right. We also demonstrated that the upregulation of the *rpoBC* operon is one of the most profound effects of *rpoB* mutations and elucidated a key molecular mechanism which is responsible for the increased expression of this operon despite the reduction in RNAP activity and transcriptional efficiency.

Materials and Methods

Generating the IPTG-inducible luciferase-tagged ancestral strain and isolating *rpoB* mutants

The DNA fragment between the SalI and BamHI restriction sites from the pUC18-mini-Tn7T-Gm-*LAC* mini-Tn7 vector (accession no. AY599234) [144], which contains the *lac* operon promoter and its repressor gene, was cloned upstream of the *luxCDABE* operon in the pUC18-mini-Tn7T-Gm-*lux* vector (accession no. AY962893) [144] at the same sites to generate the mini-Tn7 transposon delivery vector shown in Figure S2.6, Supplementary Materials. Sucrose-treated electrocompetent *P. aeruginosa* wild-type (PAO1) cells were co-transformed with this vector and a helper plasmid pUX-BF13 [145] that encodes the Tn7 site-specific transposition pathway to generate the PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain, which bears an IPTG-inducible *luxCDABE* reporter gene construct at the *attTn7* site.

Using a fluctuation test [146], *rpoB* mutants were evolved from the ancestral strain (Table 2.1A). The ancestral strain was grown overnight in LB medium at 37°C and diluted 10^6 fold. The diluted cultures were transferred to 96-well Nunc microplates (Fisher Scientific, USA) and incubated overnight at 37°C. Overnight cultures were pin-replicated onto M9KB agar containing 64 μg ml^{-1} of rifampicin and incubated at 37°C overnight. Rifampicin-resistant mutants were isolated the following day. Genomic DNA was extracted from each mutant strain using the Wizard Genomic DNA Extraction Kit (Promega, USA). Two regions of *rpoB* were sequenced using procedures described in a previous work [47].

Isolation of *rpoB* mutants with intragenic compensatory mutations

rpoB mutants with intragenic compensatory mutations were evolved from the rifampicin-resistant mutants isolated in this study through a selection experiment and were identified using Illumina whole-genome sequencing, which we have described in a separate study (refer to Annex A, Supplementary Materials, Chapter 2). Briefly, overnight cultures of the original *rpoB* mutant strains

were diluted 1000-fold in fresh M9KB culture medium [147] without rifampicin and incubated overnight with shaking at 30°C. The serial transfer was repeated for 30 consecutive days under the same experimental conditions. Genomic DNA was extracted from 3 colonies of the initial and final populations from the selection experiment using the DNeasy Blood & Tissue Kit (QIAGEN, Netherlands). Paired-end high-throughput sequencing was performed using the Illumina HiSeq™ System for Next Generation Sequencing (Illumina, USA) by the Wellcome Trust Centre for Human Genetics, University of Oxford. Raw whole-genome sequencing data were analyzed using an experimentally validated in-house computational pipeline [148]. The compensated mutants used in this study are listed in Table 2.1B.

Table 2.1: *rpoB* mutations identified in this study

Table 2.1A: *rpoB* mutations identified in rifampicin resistant mutants obtained through fluctuation test in this study

Point mutation	Amino acid substitution
A1553T	Q518L
A1553G	Q518R
A1562G	D521G
G1561A	D521N
A1592T	H531L
A1592G	H531R
C1591T	H531Y
C1607T	S536F
C1736T	S579F

Table 2.1B: *rpoB* mutations and intragenic compensatory mutations in compensated mutants evolved through the selection experiment

Point mutations	Amino acid substitutions
C1591T + A1592G	H531C
A1553G + G1584T	Q518R + E528D
C1591T + A1718G	H531Y + N573S

Competitive fitness assay

The fitness cost of *rpoB* mutations was determined using standard competitive fitness assays. The PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain, *rpoB* mutant strains (Table 2.2) and a GFP-tagged derivative strain of PAO1, which had been generated by inserting a mini-Tn7 construct carrying the *gfpmut3* gene [accession no. DQ493877] into the chromosomal mini-Tn7 site of PAO1 [144], were pre-cultured in M9KB medium overnight at 37°C with continuous shaking. On the next day, all pre-cultures were diluted in M9KB and re-grown to early exponential phase at 30°C. The non-fluorescent strains (the ancestral strain and *rpoB* mutants) were individually mixed with the GFP-tagged control strain, diluted 1:200 in fresh M9KB medium and incubated in Nunc 96-well microplates (Thermo Scientific, USA) at 30°C overnight for approximately 24 hours with continuous shaking.

The proportion of fluorescent and non-fluorescent cells in each competition was determined using an Accuri C6 flow cytometer (BD Biosciences, USA) according to the manufacturer's instructions before and after overnight incubation. The competitive fitness of each non-fluorescent strain was calculated as the ratio of population doublings of each mutant relative to the GFP-tagged control strain it was competing against. The relative competitive fitness of the *rpoB* mutants was obtained by standardizing their competitive fitness to that of the ancestral strain within each round of competition. Three biological and three technical replicates of each strain were assayed.

Table 2.2: The number of independent *rpoB* mutant isolates used in luciferase reporter gene assays, competitive fitness assays and qPCR experiments

Table 2.2A: The original rifampicin resistant *rpoB* mutant isolates isolated from the fluctuation test

***rpoB* mutation**	**Number of independent *rpoB* mutant isolates assayed ***		
	Luciferase reporter gene assays	**Competitive fitness assays**	**qPCR**
Q518L	2	2	1
Q518R	1	1	1
D521G	7	2	1
D521N	2	2	1
H531L	2	2	1
H531R	2	2	2
H531Y	3	2	1
S536F	1	1	1
S579F	2	2	2

Table 2.2B: Compensated mutants isolated from the selection experiment

***rpoB* mutation**	**Number of independent *rpoB* mutant isolates assayed ***		
	Luciferase reporter gene assays	**Competitive fitness assays**	**qPCR**
[#] H531C	4	4	1
Q518R + E528D	2	2	1
H531Y + N573S	2	2	1

* The number of *rpoB* mutant isolates used in competitive fitness assays form a subset of those tested in luciferase reporter gene assays. Likewise, the number of isolates assayed using qPCR is a subset of those tested in competitive fitness assays. Small variations were observed amongst individual *rpoB* mutant isolates with the same *rpoB* mutations in all experiments. Three biological replicates of each isolate were assayed (n = 3).

Two H531C isolates were evolved from the H531R *rpoB* mutant. The other two H531C isolates were evolved from the H531Y *rpoB* mutant.

Bacterial luciferase reporter gene assay

The PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain and the *rpoB* mutant strains (Table 2.2) were pre-cultured in M9KB medium overnight at 37°C with continuous shaking (225 rpm). The overnight pre-cultures were diluted 200-fold in fresh M9KB and transferred to black 96-well Costar microplates with clear bottoms (Corning, USA). Luciferase (*lux*) expression was induced using 1 mM IPTG. Bacterial cultures were incubated in a Synergy 2 microplate reader (Bio-Tek, USA) at 30°C, during which luminescence emission (RLU) per OD_{600} were measured every 20 minutes with shaking prior to each read. The maximal gradient of each RLU/ OD_{600} curve was calculated using 10 consecutive data points during the early exponential phase. Within each assay, the relative *lux* reporter activity in each *rpoB* mutant strain was obtained by normalizing each gradient to that of the ancestral strain control in the same plate. For each strain, three biological and three technical replicates were assayed.

RNA extraction, reverse transcription & quantitative real-time PCR

Total RNA samples were extracted from the PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain, *rpoB* mutant derivative strains (Table 2.2A) and three *rpoB* mutants with compensatory mutations (Table 2.2B). Three biological replicates of total RNA were isolated from each strain, with the inclusion of the ancestral strain control in every round of RNA extraction. The strains were inoculated in M9KB medium and pre-cultured overnight at 37°C with continuous shaking. The overnight cultures were diluted in M9KB with 1 mM IPTG to induce luciferase expression and incubated at 30°C with shaking. RNA was extracted after two hours during early exponential phase using the SV Total RNA Isolation System (Promega, USA) according to the manufacturer's instructions with minor modifications. To eliminate genomic DNA, DNase treatment was performed during and after RNA extraction. cDNA was synthesized using the GoScript Reverse Transcription System (Promega, USA). Quantitative real-time PCR (qPCR) assays were performed using the relative quantification method. Each gene was amplified using the Fast SYBR Green Master Mix (Applied Biosystems, USA) on the StepOnePlus Real-time PCR System (Applied Biosystems, USA).

Given the compromised RNAP activity and transcriptional efficiency in *rpoB* mutants, a key challenge of this study was to identify a suitable group of stably expressed internal reference genes for normalizing target gene expression levels. We used the *GeNorm* algorithm [149] to assess the expression stability of six candidate internal reference genes in all sample groups and shortlisted 16S rRNA, *atpA*, *acpP* and *rpoD* as our internal reference genes. These results were further verified using the *NormFinder* algorithm [150]. Normalization factors were calculated based on the geometric mean of the uncorrected transcript levels of the four internal reference genes. Next, the relative transcript levels of target genes in each cDNA sample were determined by normalizing the uncorrected transcript levels of every target genes to the normalization factors. The oligonucleotide sequences of all qPCR primers used in this study are shown in Table S2.2, Supplementary Materials. Finally, the relative transcript levels of target genes in *rpoB* mutants were normalized to that of the ancestral strain within the same batch of biological replicates. Further experimental details are described in Annex A, Supplementary Materials.

RNA-Seq

Two biological replicates of RNA samples extracted from the ancestral strain, D521G and H531R mutants were sequenced at the Wellcome Trust Centre for Human Genetics, University of Oxford. Ribosomal RNA was depleted using the Ribo-Zero Magnetic Kit for Gram-negative bacteria (Epicentre, USA). cDNA libraries were prepared using the NEBNext® mRNA Library Prep Master Mix Set for Illumina (New England BioLabs, USA) according to the manufacturer's instructions with modifications. Paired-end high-throughput sequencing of the cDNA libraries was performed using the Illumina MiSeq™ System for Next Generation Sequencing (Illumina, USA).

Raw sequence reads from FASTQ files were mapped to the reference genome of the PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain using TopHat Version 1.4 [151] to generate BAM files. Raw gene counts were generated from the BAM files using the *htseq-count* Python script with default parameters [152] and analyzed using BioConductor [153]. The gene counts data were filtered to

eliminate genes with an average read count of lower than 10 reads in all sequenced samples. Differential expression analysis of the remaining 4765 genes was performed using *DESeq2* [154]. The read counts of genes in the *rpoB* mutant strains were normalized to those in the ancestral strain from within the same batch of biological replicates. Logarithmic fold-changes in transcript levels were calculated relative to the ancestral strain and raw p-values for each gene were calculated using a generalized linear model based on the negative binomial distribution [155]. Raw p-values were corrected for multiple testing using the Benjamini-Hochberg procedure with a significance threshold of $p < 0.05$ [156].

Results and Discussion

Determining the fitness cost carried by rifampicin resistance mutations

rpoB mutants were evolved from a rifampicin sensitive ancestral strain PAO1::mini-Tn7-*p*LAC-*lux* carrying a chromosomally-integrated reporter construct designed to measure luciferase reporter activity as a proxy for transcriptional efficiency measurement. To determine the fitness cost associated with rifampicin resistance mutations, we directly competed each resistant mutant (Table 2.2A) and the ancestral strain against a GFP-tagged derivative of the rifampicin sensitive PAO1 control strain in rifampicin-free culture medium. The average competitive fitness of resistant mutants was lower than that of the ancestor as shown in Figure 2.1 (mean fitness = 0.842, s.e = 0.059) and varied significantly between *rpoB* mutations (one-way ANOVA: $p < 0.0005$, $F_{9, 20} = 5.82$). With the exceptions of S536F and S579F, all *rpoB* mutations carried significant fitness costs compared to the ancestral strain (Dunnett's test: $p < 0.05$).

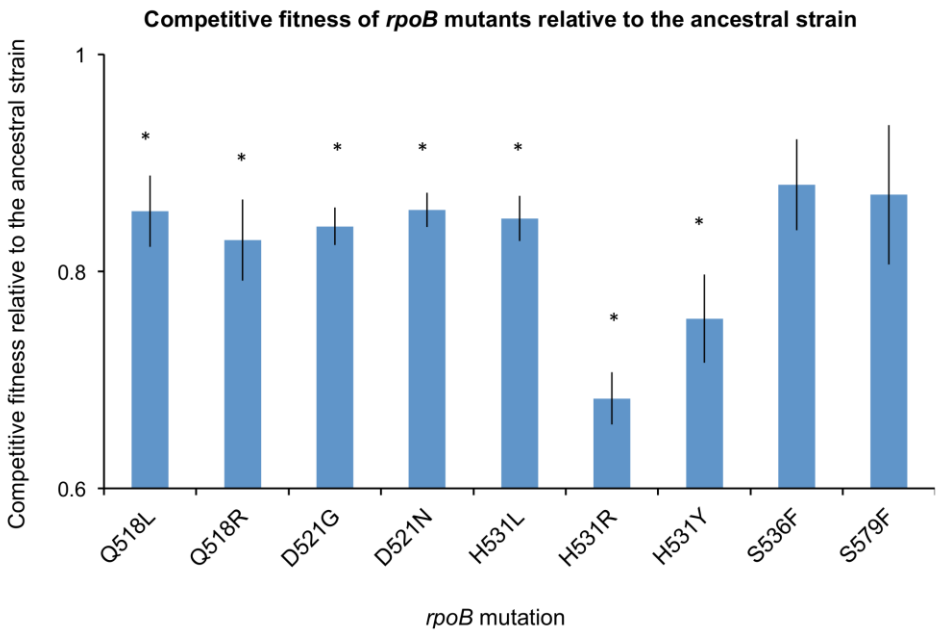


Figure 2.1: Rifampicin resistance carries a fitness cost.

Bars show the competitive fitness of rifampicin resistant mutants (mean ± s.e; n = 3) relative to the rifampicin sensitive ancestral strain, as determined by competition assays in rifampicin-free culture medium. On average, rifampicin resistance mutations decreased fitness by 15.8%, and individual mutations that carry significant fitness costs are shown with an asterisk, as determined by Dunnett's test (*$p < 0.05$).

Rifampicin resistance generates fitness costs by reducing transcriptional efficiency

Rifampicin resistance mutations alter a highly conserved domain on RNAP that is involved in transcription, suggesting that resistance mutations are likely to generate a fitness cost by compromising the enzymatic activity of RNA polymerase. To test this idea, we estimated luciferase reporter activity in *rpoB* mutants relative to their rifampicin sensitive ancestor using a chromosomally integrated luciferase reporter gene, which offers high sensitivity measurements [157] over a wide dynamic range [158] – see Figure 2.2A for an example. On average, luciferase reporter activity was lower in resistant mutants compared to the ancestor as shown in Figure 2.2B (mean *lux* reporter activity = 0.672, s.e = 0.040) and varied significantly between *rpoB* mutations (one-way ANOVA: $p < 0.0001$, $F_{9, 20} = 80.3$). Specifically, we found that all *rpoB* mutations except S536F showed significantly lower luciferase reporter activity compared to the ancestral strain (Dunnett's test: $*p < 0.05$).

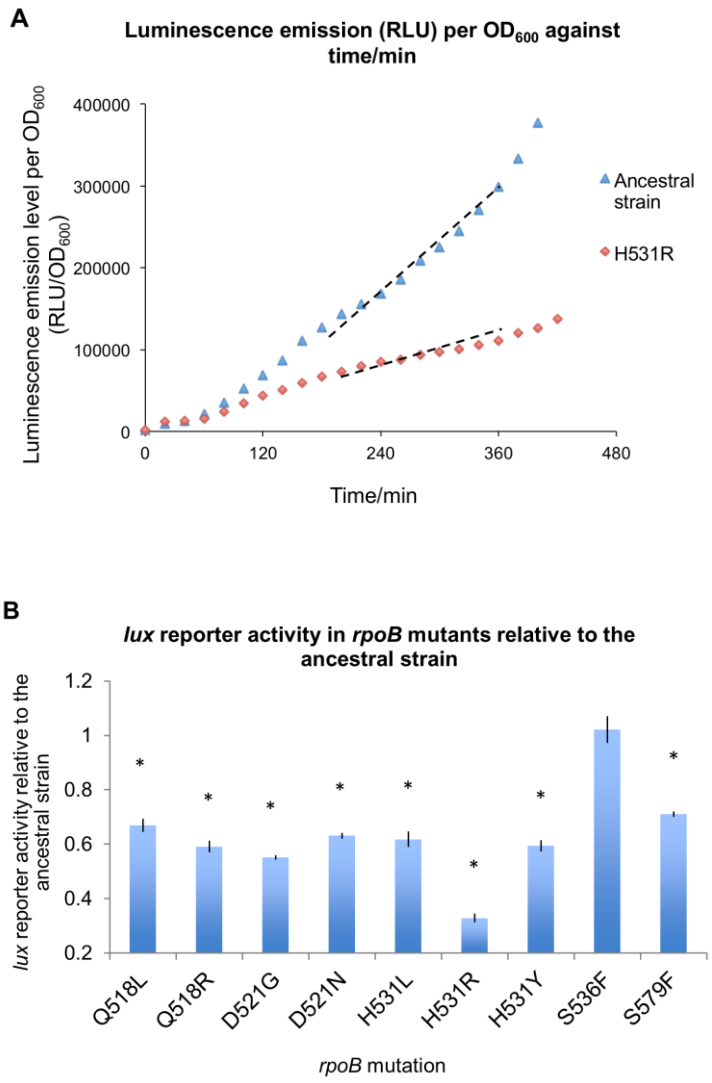


Figure 2.2: **Rifampicin resistance *rpoB* mutations reduce transcriptional efficiency.**

Panel **A**. This graph shows an example of how luciferase (*lux*) reporter activity was determined in an *rpoB* mutant strain (H531R) relative to the ancestral strain. IPTG-induced luminescence emission (RLU) per optical density (OD_{600}) was monitored continuously and the maximal gradient on the RLU/ OD_{600} curves during early exponential phase was calculated (dotted lines). Absolute *lux* reporter activity in the *rpoB* mutants were normalized to that of the ancestral strain control within each assay to obtain relative *lux* reporter activity, which is used as proxy for relative transcriptional efficiency measurements.

Panel **B** shows *lux* reporter activity in *rpoB* mutants relative to the rifampicin sensitive ancestral strain (mean ± s.e; n = 3), which is used as an estimate for relative transcriptional efficiency. All mutations except S536F resulted in significant reductions in the relative transcriptional efficiency of *rpoB* mutants, as determine by Dunnett's test (*$p < 0.05$).

Previous work has shown that bacteria can compensate for the cost of rifampicin resistance by fixing secondary mutations in RNA polymerase genes *rpoA*, *rpoB* or *rpoC* [66, 104, 159]. If the decreased transcriptional efficiency observed in rifampicin resistant mutants is a cause of decreased fitness, compensatory mutations should recover transcriptional efficiency. To test this idea, we adopted an experimental evolution approach and allowed *rpoB* mutants to evolve in rifampicin-free culture medium for a period of 30 days, or approximately 300 generations. Through whole-genome sequencing of evolved isolates, we identified strains that carried only a rifampicin resistance mutation and second site RNAP mutation (Table 2.1B). To establish the causal link between compromised RNAP function and fitness cost, we measured both the luciferase reporter activity and the competitive fitness in these compensated mutant isolates (Table 2.2B). Figure 2.3A shows that all compensated mutants showed a statistically significant increase in mean *lux* reporter activity compared to the original *rpoB* mutants they had been evolved from (two-tailed, two-sample *t*-test: *$p<0.05$). Likewise, all compensated mutants also showed an increase in their mean relative competitive fitness (Figure 2.3B). Except for Q518R + E528D, the difference between the mean relative competitive fitness of each pair of compensated mutant and the original *rpoB* mutant is statistically significant (two-tailed, two-sample *t*-test: *$p<0.05$).

The key assumption of our reporter-based assay is that luminescence emission levels reflects variation in expression levels of *lux* operon genes, so that measuring relative rate of luminescence production through time is equivalent to measuring the rate of *lux* mRNA transcript synthesis. To test this assumption, we quantified the relative mRNA transcript levels of the *luxC* gene in mutants relative to the ancestral strain using quantitative real-time PCR (qPCR). Relative *luxC* transcript levels was highly correlated with luminescence emission (RLU) normalized to optical density (OD_{600}) measured immediately prior to RNA extraction (Figure S2.1, Supplementary Materials; one-way ANOVA: $p = 0.0006$, $F_{1, 8} = 30.1$, $r^2 = 0.790$). In this analysis, we excluded a statistically significant outlier (Q518R) identified using the least trimmed squares method. Q518R showed a substantial decrease in its normalized luminescence emission despite a slight increase in *luxC* expression relative to the ancestral strain.

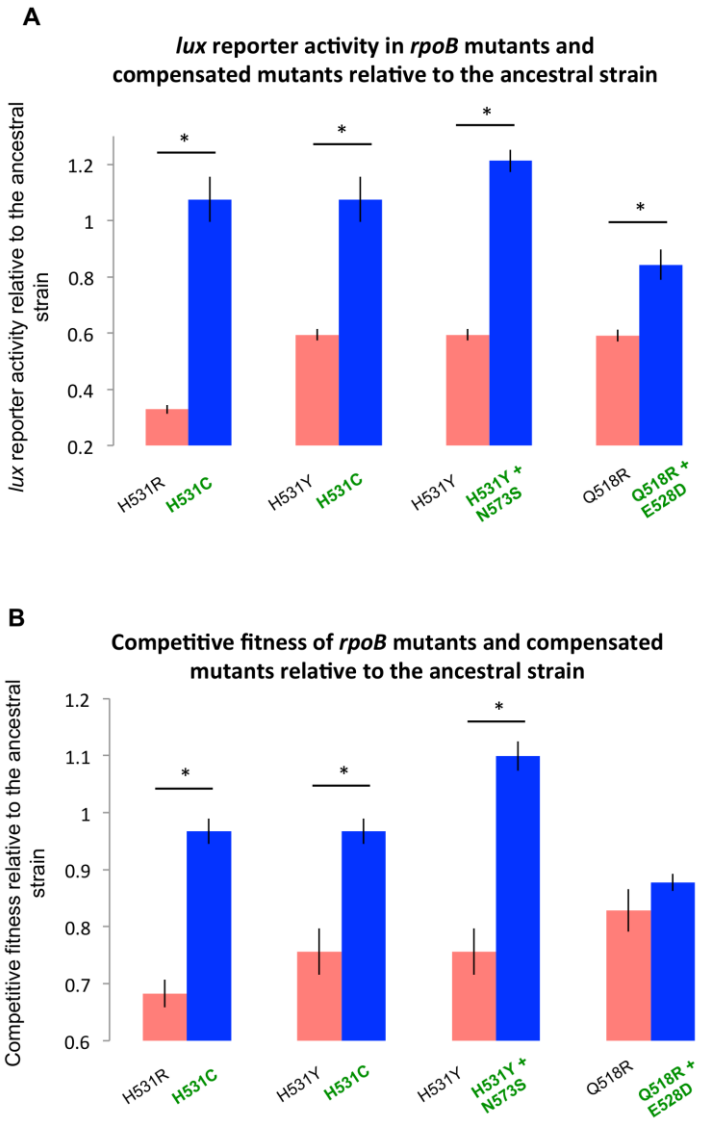


Figure 2.3: Compensatory mutations restore both the transcriptional efficiency and competitive fitness of resistant mutants.

This figure shows impact of compensatory mutations on **A**. transcriptional efficiency (mean ± s.e; n = 3) and **B**. competitive fitness (mean ± s.e; n = 3) relative to the rifampicin-sensitive ancestral strain. The data are displayed as pairs of the original rifampicin resistant mutants from the fluctuation test (shown in red) and compensated strains (shown in blue) carrying both an *rpoB* mutation and a second-site compensatory mutation within *rpoB*. Two-tailed, two-sample *t*-test: *p<0.05

To measure the contribution of reduced transcriptional efficiency to the cost of resistance, we tested for a correlation between luciferase reporter activity and competitive fitness (Figure 2.4). To ensure that each *rpoB* mutant strain represents an independent data point, compensated mutants Q518R + E528D and H531Y + N573S are excluded from all correlation analyses in this study. Consistent with our hypothesis, we found that decreasing transcriptional efficiency increased the fitness cost of resistance (one-way ANOVA: $p = 0.0005$, $F_{1, 9} = 28.8$, $r^2 = 0.762$). This direct effect of *rpoB* mutations on the transcriptional efficiency can account for 76.2% of the variance in fitness amongst mutants.

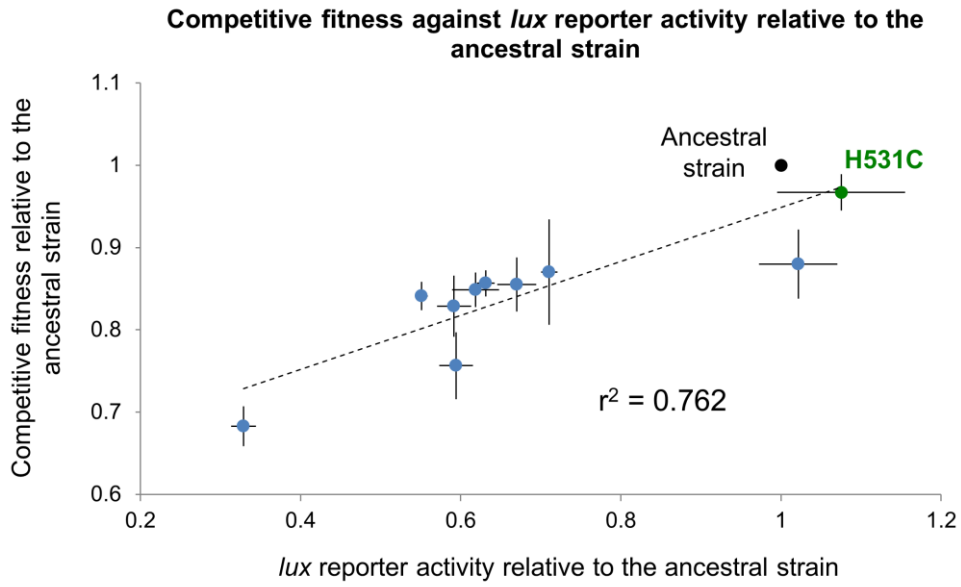


Figure 2.4: Variation in the fitness cost of rifampicin resistance can be explained by variation in transcriptional efficiency.

This graph shows the correlation between relative competitive fitness (mean ± s.e; n = 3) and relative transcriptional efficiency of resistant mutants (mean ± s.e; n = 3). 76.2% of the variation in competitive fitness can be explained by variation in transcriptional efficiency, as determined by linear regression (dashed line).

The ancestral strain is shown in black, and the compensated mutant H531C in green. To ensure that every *rpoB* mutant strain is an independent data point, compensated mutants Q518R + E528D and H531Y + N573S were excluded from the correlation analysis.

Rifampicin resistance mutations alter the transcriptional profile of essential genes

Since RNA polymerase is a master regulator of gene expression, changing the structure of RNA polymerase has the potential to alter gene expression at a genome-wide scale. To test the hypothesis that *rpoB* mutations generate a pleiotropic fitness cost by altering gene expression, we quantified the mRNA transcript levels of a random sample of 15 essential genes in *rpoB* mutants relative to the ancestral strain using qPCR. The distinction between measuring transcriptional efficiency and relative transcript levels is subtle, but important. Measuring luciferase reporter activity (Figure 2.2) provides a measure of the enzymatic efficiency of RNA polymerase. Relative transcript levels, on the other hand, provide an estimate of how resistance mutations alter the transcriptional profiles of the *rpoB* mutants. In other words, our luciferase reporter activity measurements provide a proxy for how resistance mutations alter the rate at which RNA polymerase can produce transcripts, whereas relative transcript level measurements reveal which genes are being differentially transcribed by RNA polymerase mutants and shed insight into the transcript-specific effects of the *rpoB* mutations. Crucially, we expected the impact of RNA polymerase mutations on relative transcript levels to vary between genes owing to the diversity of transcriptional regulatory mechanisms in bacterial genomes [72].

In general, we found that rifampicin resistance mutations had pervasive, but relatively subtle effects on the expression of essential genes (Figure 2.5). To ensure that every *rpoB* mutant strain is an independent sample group, compensated mutants Q518R + E528D and H531Y + N573S were excluded from our two-way ANOVA analysis. First, we found significant variation in the average change in relative transcript levels between essential genes (two-way ANOVA: $p<0.0001$, $F_{14, 420} = 5.29$). Second, we found significant variation in the average change in relative transcript levels between resistance mutations (two-way ANOVA: $p<0.0001$, $F_{10, 420} = 6.73$). Finally, we found a significant essential gene * *rpoB* mutation interaction (two-way ANOVA: $p<0.0001$, $F_{140, 420} = 2.26$), implying that different *rpoB* mutations have different impacts on the expression of different genes. This analysis suggests that rifampicin resistance mutations had variable pleiotropic effects on the relative transcript levels of essential genes, which is perhaps not surprising given the central role of

RNAP in gene expression and the diversity of transcriptional regulatory mechanisms for regulating gene expression.

Interestingly, the compensated mutants (H531C, H531Y + N573S and Q518R + E528D) clustered much more closely with the ancestral strain than the *rpoB* mutants they had been evolved from (H531R, H531Y and Q518R) as indicated by our two-way clustering analysis (Figure 2.5), implying that these intragenic compensatory mutations have the dual effect of restoring RNAP activity in terms of transcriptional efficiency, as well as restoring the expression levels of essential genes to ancestral strain levels.

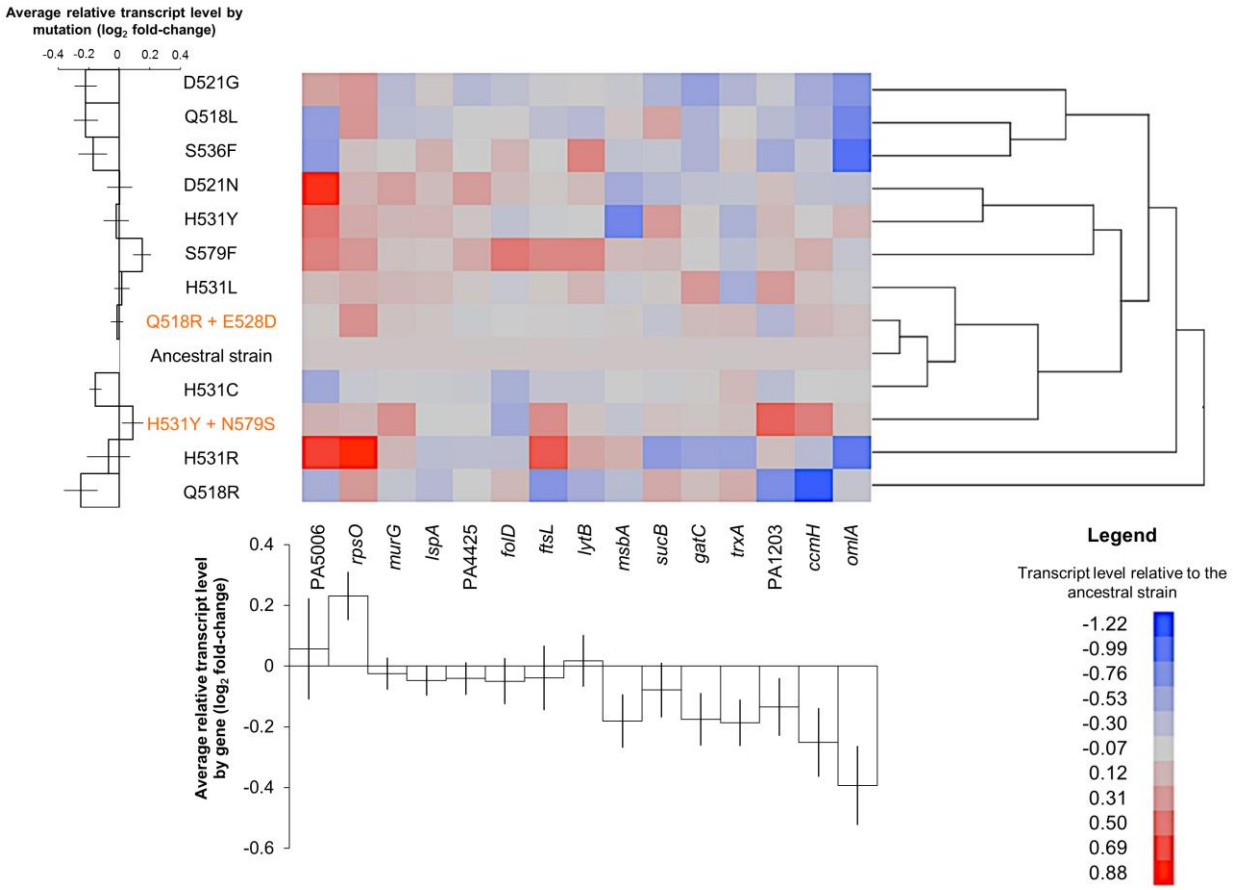


Figure 2.5: Impact of *rpoB* mutations on relative essential gene transcript levels

The heatmap shows the average transcript levels of 15 randomly selected essential genes in the *rpoB* mutants and three mutants with intragenic compensatory mutations relative to that in the ancestral strain, expressed in terms of $\log_2$ fold-change. Significant variation in relative transcript levels is observed between *rpoB* mutations (vertical axis: mean ± s.e; n = 11) and between essential genes (horizontal axis: mean ± s.e; n = 15). Interaction between mutations and genes also contributed significantly to the observed variation in the relative transcript levels of essential gene. Q518R + E528D and H531Y + N573S were excluded from the two-way ANOVA analysis (highlighted in orange font). The dendrogram from the two-way clustering analysis shows that the compensated mutants cluster more closely with the ancestral strain compared to the original *rpoB* mutants they had been evolved from.

Previous work in *E. coli* strongly supports the idea that the expression levels of essential genes are under stabilizing selection [160]. If this is the case, both increases and decreases in the expression of essential genes are expected to carry a fitness cost. Therefore, we quantified the pleiotropic effects of rifampicin resistance mutation by calculating the average magnitude of fold-change in expression for the fifteen essential genes for each *rpoB* mutant. We did not find any evidence of a correlation between pleiotropy and fitness after correcting for variation in transcriptional efficiency of *rpoB* mutants (Pearson partial correlation coefficient, $r = -0.541$, $p = 0.106$), which implies that there is no overall correlation between pleiotropic changes in essential gene expression levels and fitness.

We then hypothesized that the average magnitude of deviation in the expression levels of essential genes in *rpoB* mutants from ancestral strain levels depends on the extent to which transcriptional efficiency is altered. Amongst the *rpoB* mutants evolved from the fluctuation test (Table 2.1A), the average magnitude of change in essential gene expression showed significant variations between *rpoB* mutations (one-way ANOVA: $p < 0.0001$, $F_{9, 140} = 4.47$). In addition, we observed that decreasing transcriptional efficiency increases the average magnitude of deviation in essential gene expression levels (Figure 2.6; one-way ANOVA: $p = 0.0303$, $F_{1, 9} = 6.59$, $r^2 = 0.423$). Taken together, our results support the view that the altered transcriptional profile of essential genes in *rpoB* mutants is likely to be a consequence of altered transcription rate and does not generate a fitness cost in its own right.

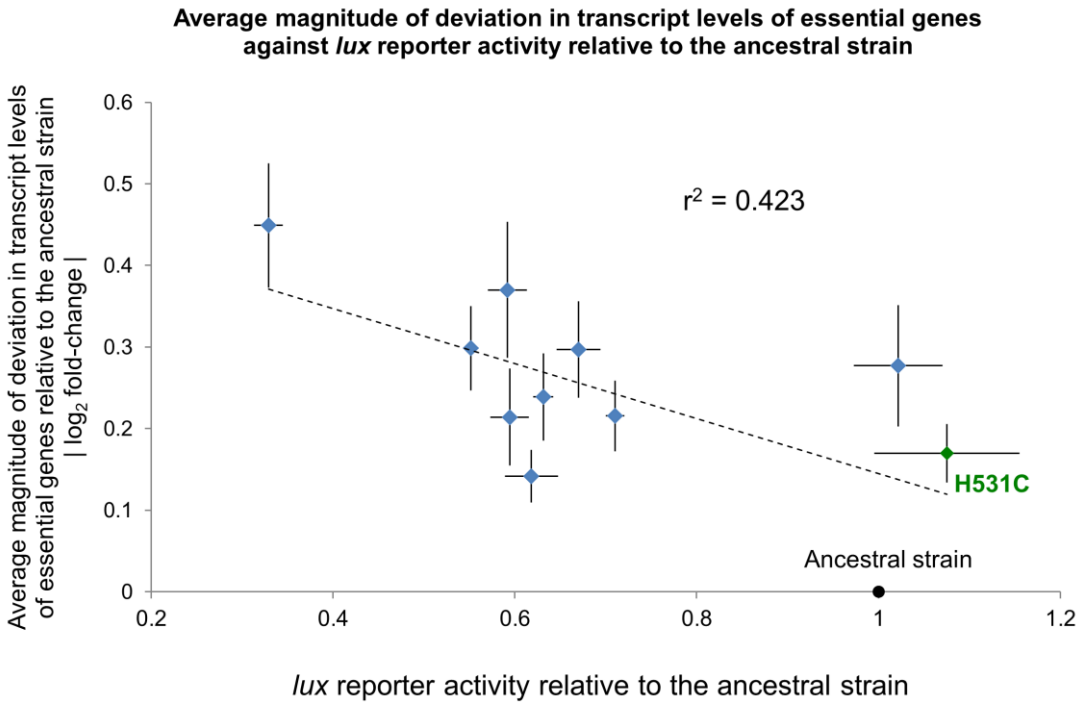


Figure 2.6: Decreasing transcriptional efficiency increases average magnitude of deviation in gene expression levels of essential genes in *rpoB* mutants relative to the ancestral strain.

This graph shows the correlation between relative transcriptional efficiency (mean ± s.e; n = 3) and the average magnitude of deviation in gene expression levels of essential genes in resistant mutants (mean ± s.e; n = 15). Decreased transcriptional efficiency is associated with elevated average deviation in gene expression levels in *rpoB* mutants, as determined by linear regression (dashed line). The ancestral strain is shown in black, and the compensated mutant H531C in green. (Q518R + E528D and H531Y + N573S were omitted from the correlation analysis.)

RNA-Seq transcriptome profiling of two *rpoB* mutants

To extend our qPCR measurements of gene expression to a genome-wide scale, we used RNA-Seq to obtain the transcriptome profiles of the ancestral strain and two *rpoB* mutants that carry *rpoB* mutations D521G and H531R, which are common mutations that carry significant fitness costs. At a genome-wide scale, both of the rifampicin resistant mutants displayed similar expression profiles to the rifampicin sensitive strain, as demonstrated by the highly correlated expression scores between the mutants and the sensitive strain (Panels A and B, Figure 2.7). In the D521G mutant, only 77 genes, or 1.36% of the total number of genes in the genome showed significant differential expression, while in H531R, 475 genes or 8.36% of all genes were differentially expressed ($p < 0.05$ after correcting for false discovery rate). In line with this observation, we found that genome-wide expression profiles of the two mutants were very similar to each other (Panel C, Figure 2.7). For example, 59 genes showed significant differential expression ($p < 0.05$) in both sequenced mutant strains (Figure S2.4, Supplementary Materials), which means most of the differentially expressed genes in the D521G mutant are similarly affected in the H531R mutant (Annex B, Supplementary Materials).

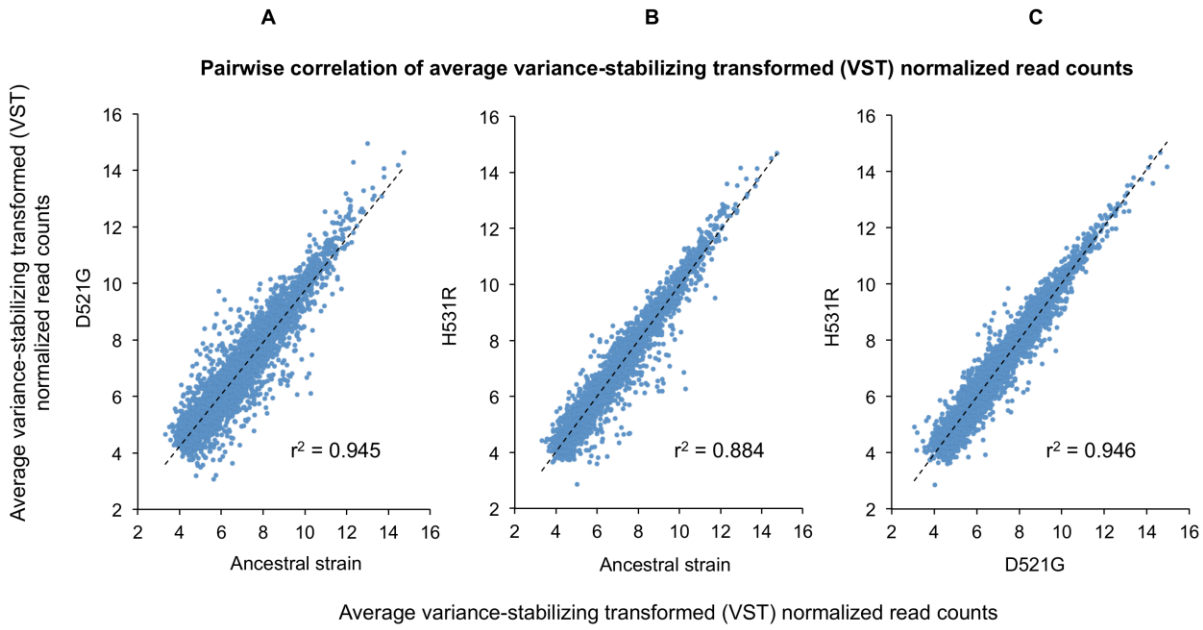


Figure 2.7: *rpoB* mutations D521G and H531R led to small, but pervasive changes in transcriptome profile relative to the ancestral strain at a genome-wide scale.

Panels **A**-**C** show pairwise comparisons between the transcriptome profiles of two rifampicin resistant mutants, D521G and H531R, and the rifampicin sensitive ancestral control strain, as determined by RNA-Seq. Each data point represents a single gene. The horizontal and vertical axes show the average variance-stabilization-transformed (VST) normalized read counts for each strain.

At a genome-wide level, we calculated the average magnitude of fold-change for the expression of three categories of genes in *rpoB* mutants D521G and H531R relative to that of the ancestral strain – 'essential' genes, highly conserved 'core' genes that play vital roles in bacterial physiology and lineage-specific 'accessory' genes that probably provide bacteria with niche-specific benefits. Based on the whole-genome sequencing data for 36 *Pseudomonas* species [101], 1435 core genes were identified [148], out of which 196 are also essential genes identified by Liberati *et al* [161]. Accessory genes include all remaining genes in the genome that are neither core nor essential. In agreement with the idea that the expression levels of essential and core gene are under stabilizing selection (48), we found that the average magnitude of deviation in gene expression levels are higher for accessory genes than for the combined group of essential and core genes in both the D521G and H531R mutants (two-tailed, two-sample *t*-test; D521G: $t = -4.10$, d.f = 3.64×10^3, $p < 0.0001$; H531R: $t = -5.68$, d.f = 3.86×10^3, $p < 0.0001$).

For all three categories of genes, the average magnitude of deviation in gene expression levels from ancestral strain levels is significantly greater in the H531R mutant compared to D521G (Figure S2.3, Supplementary Materials: two-tailed, two-sample *t*-test; essential genes: $t = -7.21$, d.f = 4.66×10^3, $p < 0.0001$; core genes: $t = -16.3$, d.f = 1.97×10^3, $p < 0.0001$; accessory genes: $t = -24.0$, d.f = 5.00×10^3, $p < 0.0001$). Since the H531R mutant has lower transcriptional efficiency than the D521G mutant, these results suggest that reduced transcriptional efficiency is linked to altered gene expression patterns at a genome-wide scale, which supports our qPCR-based measurements of gene expression levels shown in Figure 2.6.

Differential expression of the *rpoBC* operon in *rpoB* mutants

Previous work in *E. coli* demonstrated that an increase in the expression of the *rpoBC* operon was observed in the presence of a sub-lethal dose of rifampicin [162]. A similar upregulation of *rpoBC* was observed when a partial restriction of RNAP activity was induced in an *E. coli* mutant strain with conditionally defective RNAP [163]. Given the adverse effects of *rpoB* mutations have on RNAP

function, we hypothesized that the expression levels of *rpoBC* are higher in rifampicin-resistant mutants relative to the rifampicin-sensitive ancestral strain and that the extent of *rpoBC* upregulation should depend on the extent to which RNAP function is impaired. Consistent with this idea, we found that relative transcript levels of *rpoB* correlated significantly with transcriptional efficiency estimated using the luciferase reporter gene, and that *rpoB* mutants with the highest reduction in RNAP activity show the largest increase in *rpoB* expression (Figures 2.8A and 2.8B, one-way ANOVA: $p = 0.0091$, $F_{1, 9} = 10.9$, $r^2 = 0.549$). Interestingly, *rpoB* and *rpoC* were among the 59 genes that were significantly upregulated in both D521G and H531R from our RNA-Seq dataset, suggesting that the upregulation of *rpoBC* is one of the most profound effects of rifampicin resistance mutations, at least in strains carrying mutations that lead to substantial reductions in transcriptional efficiency.

We investigated the molecular mechanism underpinning the differential expression of *rpoBC* by quantifying the transcriptional read-through of the intergenic region between the *rplJL* and *rpoBC* operons. In *P. aeruginosa*, a partial terminator (attenuator) P26 lies immediately downstream of *rplL* (Figure 2.8C). In *E. coli*, non-bacteriostatic concentrations of rifampicin, which partially inhibited RNAP activity and decreased transcriptional efficiency, were shown to have a small stimulatory effect on the expression of *rplJL* [164], but a large stimulatory effect on the expression of *rpoBC* [165, 166]. Morgan & Hayward (1987) demonstrated that this was a direct effect of an increase in read-through of the t_{L7} partial Rho-independent terminator between *rplL* and *rpoB*, which is the *E. coli* homologue of P26 and has a termination efficiency of 80% in the wild-type strain [166].

Since many *rpoB* mutations are known to alter the termination phenotypes of RNAP in *E. coli* [63], we hypothesized that the observed increase in the expression of *rpoBC* in rifampicin resistant mutants with compromised RNAP activity can be attributed to an increase in read-through of the P26 partial terminator in *P. aeruginosa*. To test this idea, we quantified the transcript levels of two regions spanning the start and end of P26 (Figure 2.8C) using qPCR. One amplicon (*rplL*_P26) spans the junction between *rplL* and P26, while the other (P26_intergenic) spans the junction between P26 and the start of the remaining intergenic region. The extent of P26 read-through was characterized by the

transcript levels of the P26_intergenic amplicon relative to those of the *rplL*_P26 amplicon within each *rpoB* mutant strain.

Figure 2.8D shows that the percentage change in P26 read-through is significantly correlated with the relative expression of *rpoB* in rifampicin resistant mutants (one-way ANOVA: $p = 0.0013$, $F_{1, 9} = 21.3$, $r^2 = 0.702$), confirming that altered P26 read-through contributes to the differential expression of *rpoB* in rifampicin resistant mutants. To establish the causality of the correlation between *rpoB* expression and P26 read-through, we used qPCR to measure the relative expression levels of *rpoB* and alteration in P26 read-though in the three compensated mutants listed in Table 2.2B relative to the ancestral strain. Consistent with our hypothesis, the compensated mutants H531C and H531Y + N573S showed statistically significant reductions in relative *rpoB* expression levels compared to the original *rpoB* mutants they had been evolved from (Panel A in Figure S2.5, Supplementary Materials). The extent of P26 read-through also showed a statistically significant decrease in H531C compared to H531R (Panel B in Figure S2.5, Supplementary Materials). Taken together, these observations support the idea that alterations in P26 read-through, and consequently *rpoBC* expression, depend on the extent to which RNAP function is compromised in *rpoB* mutants. However, we emphasize that altered P26 read-through may not be the only mechanism that contributes to altered expression of *rpoBC*.

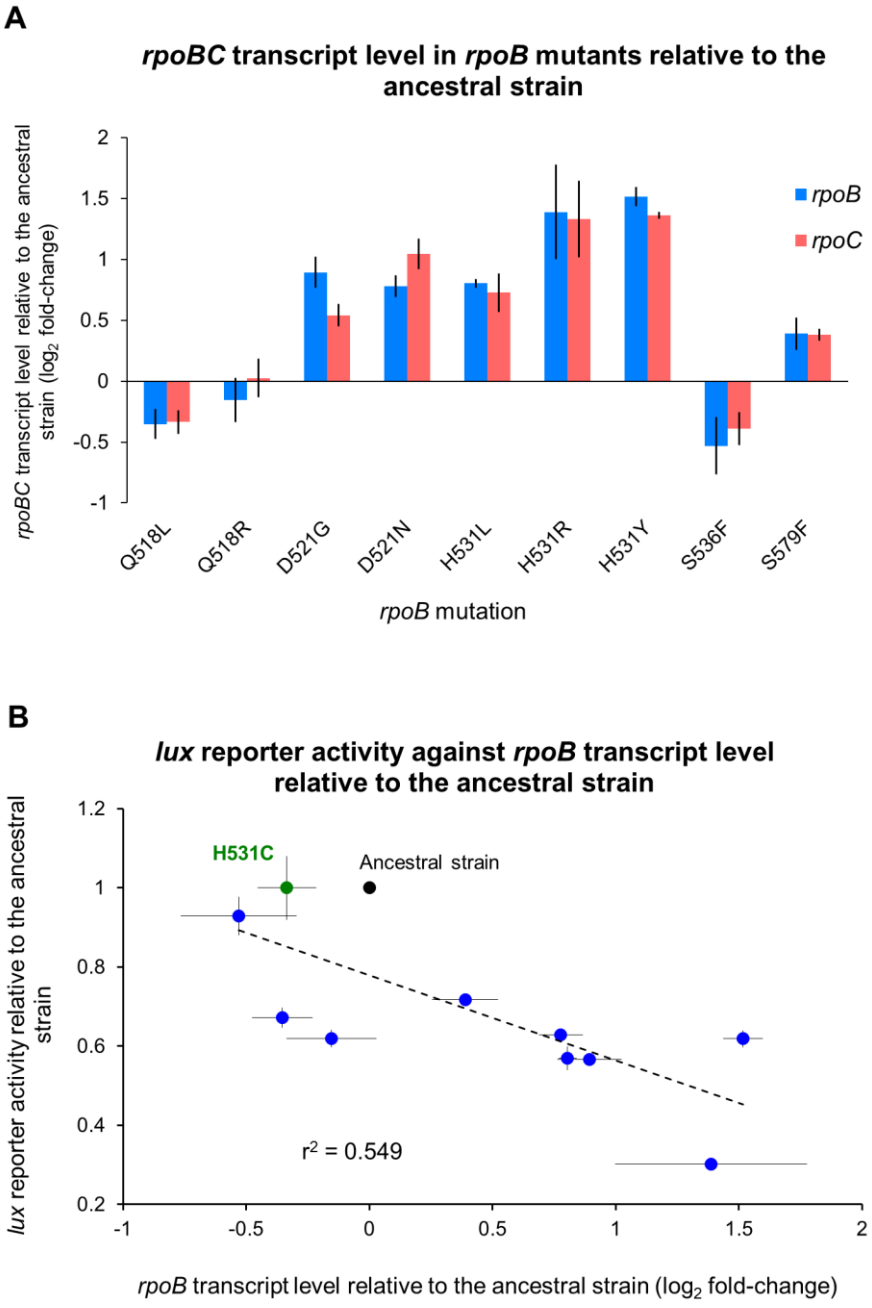


Figure 2.8: *rpoB* mutations increase the expression of the *rpoBC* operon via increased read-through at a partial terminator.

Panel **A** shows the transcript levels of *rpoB* and *rpoC* ($\log_2$ fold-change) in rifampicin resistant mutants relative to the ancestral strain (mean ± s.e; n = 3), as determined by qPCR.

Panel **B** shows the *rpoB* expression and transcriptional efficiency in resistant mutants relative to the ancestral strain (mean ± s.e; n = 3). Decreased transcriptional efficiency is associated with elevated *rpoB* expression level, as determined by linear regression (dashed line).

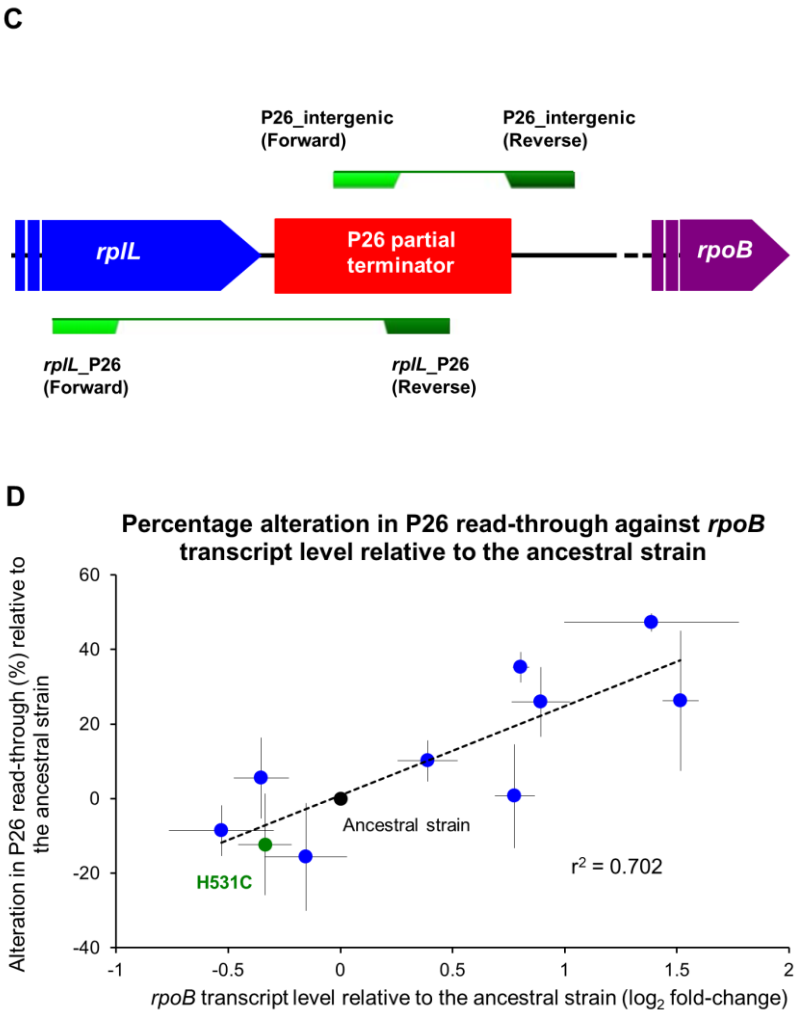


Figure 2.8 (continued): *rpoB* mutations increase the expression of the *rpoBC* operon via increased read-through at a partial terminator.

Panel **C** shows a schematic of the qPCR strategy used to quantify read-through at the P26 partial terminator, which is located immediately upstream of *rpoB*. One amplicon (*rplL*_P26) spans the junction between *rplL* and P26, while the other (P26_intergenic) spans the junction between P26 and the start of the remaining intergenic region. The transcript levels of the P26_intergenic amplicon were normalized to those of *rplL*_P26 to quantify the extent of the partial terminator read-through.

Panel **D** shows the correlation between changes in P26 read-through (mean ± s.e; n = 3) relative to the rifampicin sensitive ancestral strain and *rpoB* expression relative to the ancestral strain (mean ± s.e; n = 3). Decreased termination at P26 can explain 70.2% of the variation in relative *rpoB* transcript levels, as determined by linear regression (dashed line).

Conclusions

In summary, we found that rifampicin resistance resulted in significant reductions in transcriptional efficiency (Figure 2.2), and this direct effect of rifampicin resistance explains 76.2% of the variation in the fitness of rifampicin resistant mutants (Figure 2.4). The indirect, pleiotropic consequence of compromised RNAP activity was that rifampicin resistance mutations showed altered transcriptional profiles at a genome-wide scale. Pleiotropic changes in relative transcript levels were modest in comparison to the large changes in transcriptional efficiency associated with rifampicin resistance, and this helps to explain why we do not detect any evidence that global transcriptional responses contribute to the cost of rifampicin resistance. For example, the D521G mutation was associated with a 44.9% reduction in transcriptional efficiency, which is typical of rifampicin resistance mutations (Figure 2.2), but this mutation resulted in significant changes in the expression of only 71 genes (Figure S2.5, Supplementary Materials), which corresponds to 1.25% of the *P. aeruginosa* genome.

One of the most important conclusions of this study is that the transcriptional regulatory network of *P. aeruginosa* displays a remarkable level of robustness to *rpoB* mutations. First, we found that rifampicin resistance mutations were associated with in an increase in the expression of *rpoBC*, implying that the loss of RNAP activity in *rpoB* mutants was buffered by increased RNAP abundance via increased read-through at a partial terminator upstream of *rpoB*. Transcription is a complex, multi-step biochemical process that can be divided into the three distinct stages of initiation, elongation and termination. *rpoB* mutations which affect any stage of transcription, promoter selectivity of the RNAP core enzyme or interactions between RNAP and transcriptional regulatory elements such as transcriptional attenuators can potentially affect the gene expression profiles of mutants. It is unclear if reduced attenuation is specific to the P26 partial terminator, or if this phenomenon can be observed at a genome-wide scale. We plan to further investigate these questions using the RNA-Seq datasets we have obtained for the D521G and H531R mutants.

Second, we found that pleiotropic changes in the average magnitude of gene expression levels are much smaller for core and essential genes than for accessory genes. The molecular mechanisms underpinning this are not clear, but it has been shown that essential genes and genes with a high degree of evolutionary conservation show much less cell-to-cell variation in expression than accessory genes in *E. coli* [160]. This suggests that bacteria have evolved very general mechanisms to minimize variation in the expression of essential and core genes. If this is the case in rifampicin resistant mutants of *P. aeruginosa*, the small pleiotropic changes in the expression of essential genes in *rpoB* mutants might have very little to do with altered RNAP activity *per se*.

Chromosomal antibiotic resistance mutations often alter the structure of highly conserved enzymes with pivotal biochemical functions, such as ribosomes, RNAP and DNA gyrase. Intuition suggests that alterations to these cellular components are likely to have extensive pleiotropic effects at a genome-wide scale. In line with our findings, previous studies have found that the fitness effects of resistance mutations correlate strongly to the direct effects of resistance mutations [36, 37]. Our study raises the intriguing possibility that despite the strong correlation between the direct effects of rifampicin resistance mutations and fitness, the buffering capacity of *rpoB* mutants against pleiotropic changes in essential gene expression levels could diminish the fitness cost that are associated with these pleiotropic effects. One motivation for studying the costs of antibiotic resistance is that understanding the costs of resistance could theoretically allow for the development of new therapeutics aimed at exploiting weaknesses of resistant bacteria. If bacteria are well buffered against resistance mutations, then resistant mutants probably lack Achilles' heels.

Acknowledgements

This work was supported by funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant (StG-2011-281591) and from the Royal Society. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics funded by Wellcome Trust grant reference 090532/Z/09/Z and Medical Research Council Hub grant G0900747 91070 for generation of the high-throughput sequencing data. The authors would like to thank Dr Cedric Patthey for assessing the quality of RNA samples for high-throughput sequencing.

Author Contributions

Conceived and designed the experiments: QQ, GMP, RCM

Contributed reagents/materials/equipment/analysis tools: RCM, GMP

Performed the experiments: QQ

Analyzed the data: QQ, RCM

Wrote the paper: QQ, GMP, RCM

The authors declare that there are no conflicts of interest.

Supplementary Materials for Chapter 2

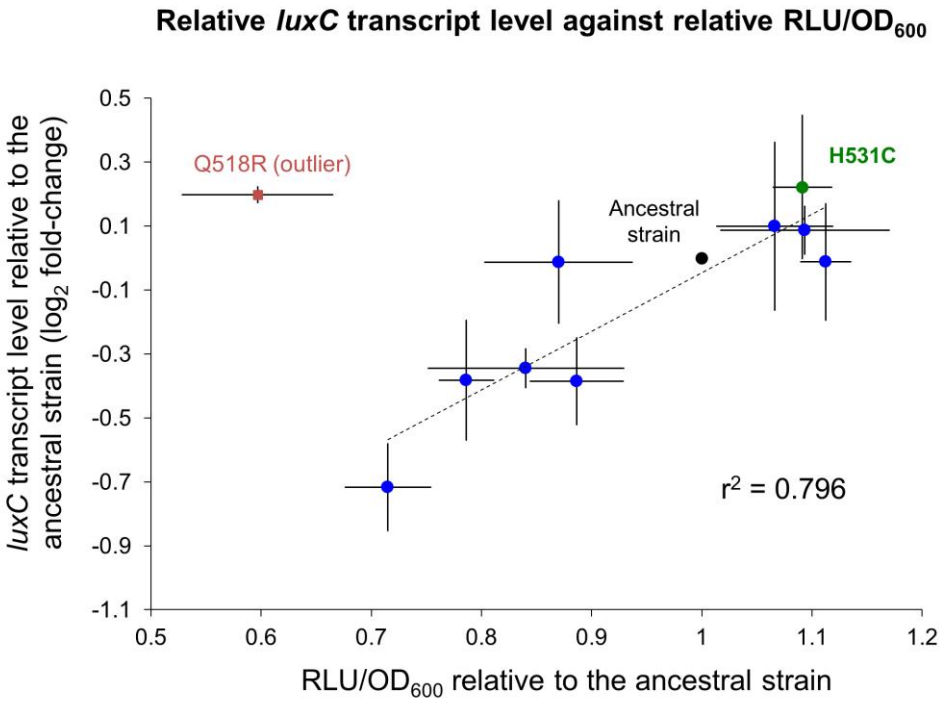


Figure S2.1: Experimental validation of the luciferase reporter gene.

A significant correlation between *luxC* transcript levels in $\log_2$ fold-change (mean ± s.e; n = 3) and luminescence emission (RLU) per OD_{600} (mean ± s.e; n = 3) in *rpoB* mutants relative to the ancestral strain. A statistically significant outlier (Q518R) was excluded from this analysis.

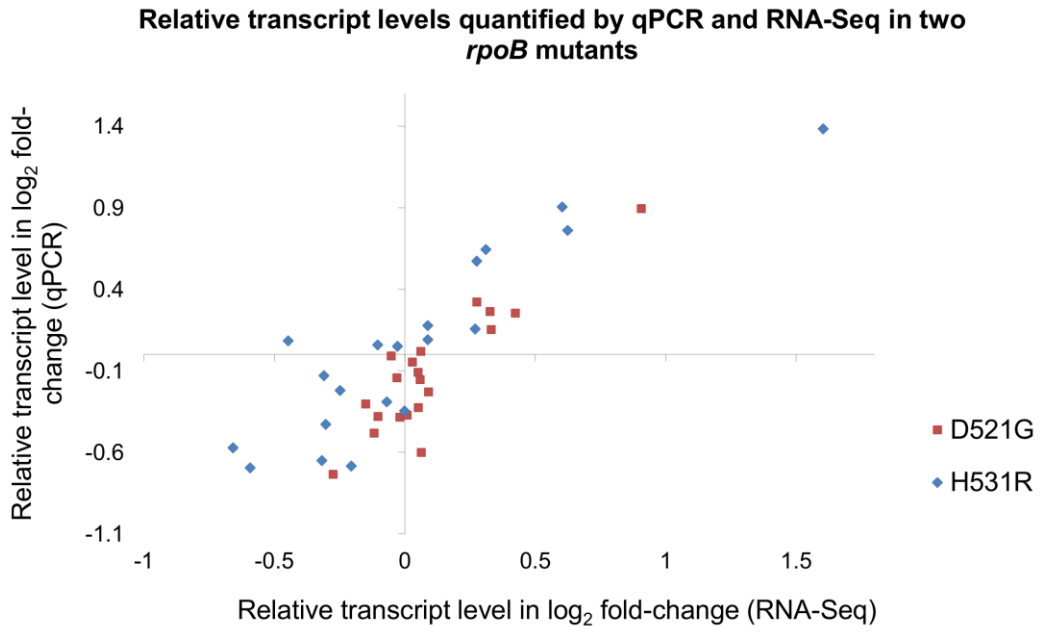


Figure S2.2: Linear correlation between gene expression measurements from qPCR and RNA-Seq.

The relative transcript levels of *rpoB*, the 15 randomly selected essential genes and internal reference genes (excluding 16S rRNA) in the two *rpoB* mutants quantified by qPCR were correlated against those quantified by RNA-Seq. *rpoC*, which is co-expressed and co-regulated with *rpoB*, was excluded from this analysis. One-way ANOVA: D521G mutant: $p < 0.0001$, $F_{1, 18} = 71.7$, $r^2 = 0.799$; H531R mutant: $p < 0.0001$, $F_{1, 18} = 71.1$, $r^2 = 0.798$

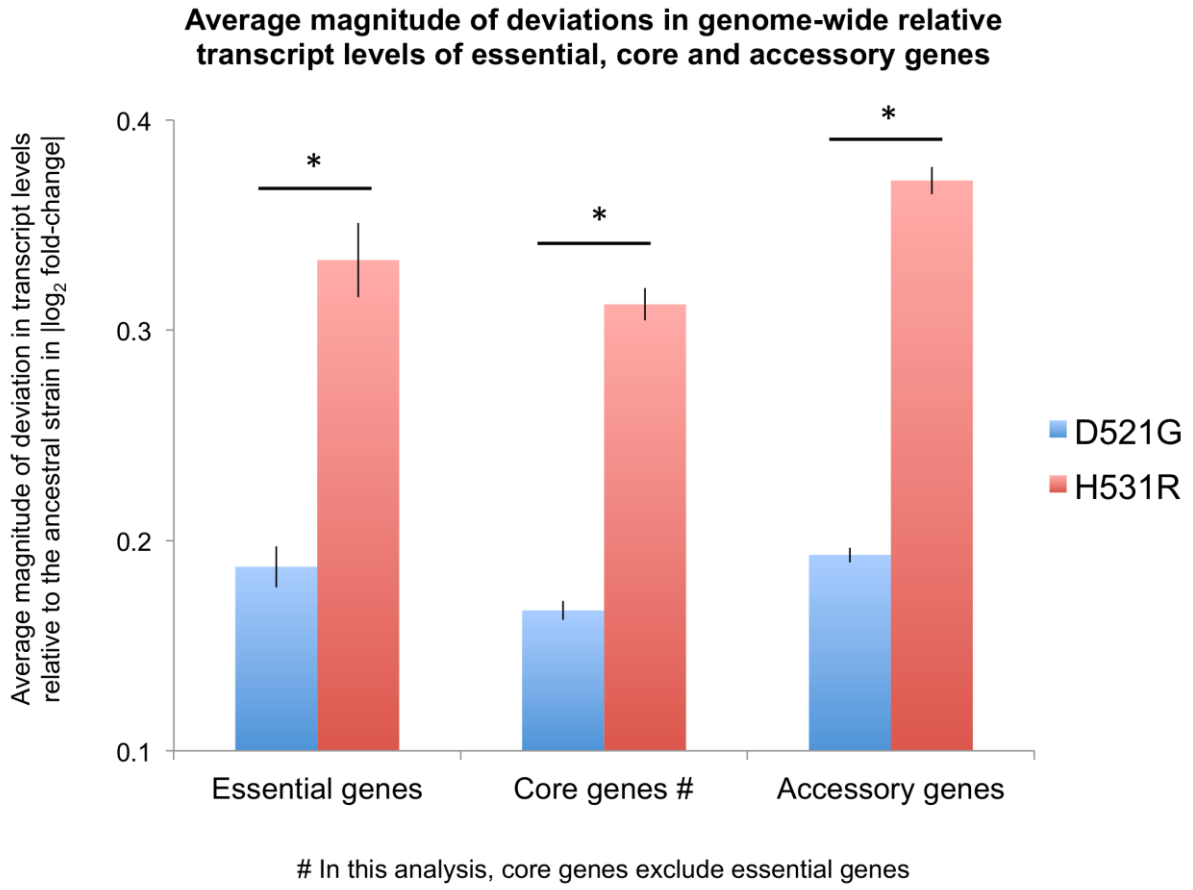


Figure S2.3: Genome-wide average expression levels of essential, core and accessory genes in resistant mutants D521G and H531R

Regardless of the classification of genes, the average magnitude of deviations in gene expression levels relative to ancestral levels is significantly greater in the H531R mutant compared to the D521G mutant (two-tailed, two-sample *t*-test, *p<0.0001*).

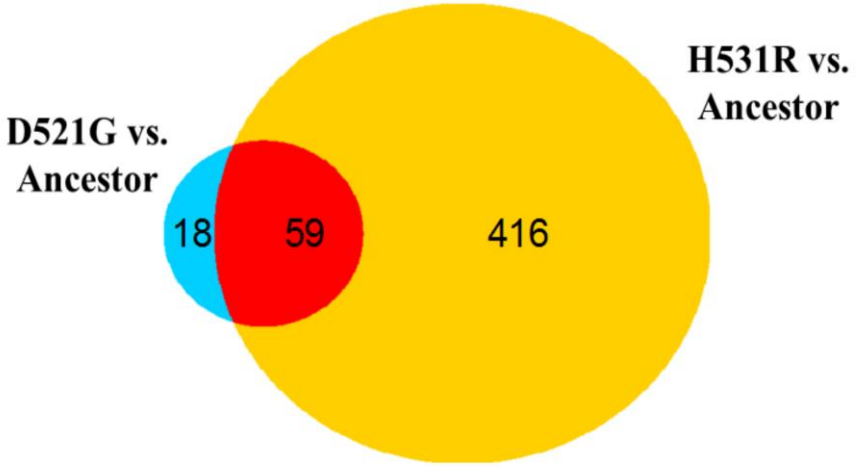


Figure S2.4: Number of differentially expressed genes in the D521G and H531R mutants relative to the ancestral strain.

This Venn diagram shows the number of differentially expressed genes in the two *rpoB* mutants relative to the ancestral strain ($p < 0.05$) after Benjamini-Hochberg correction.

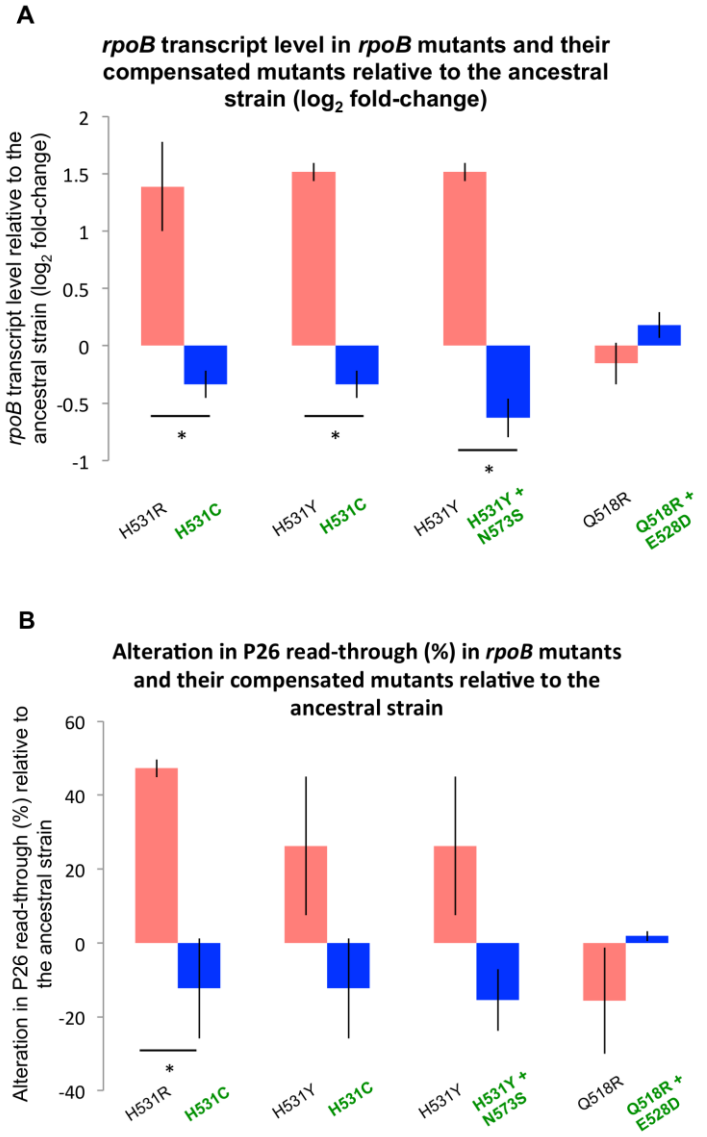
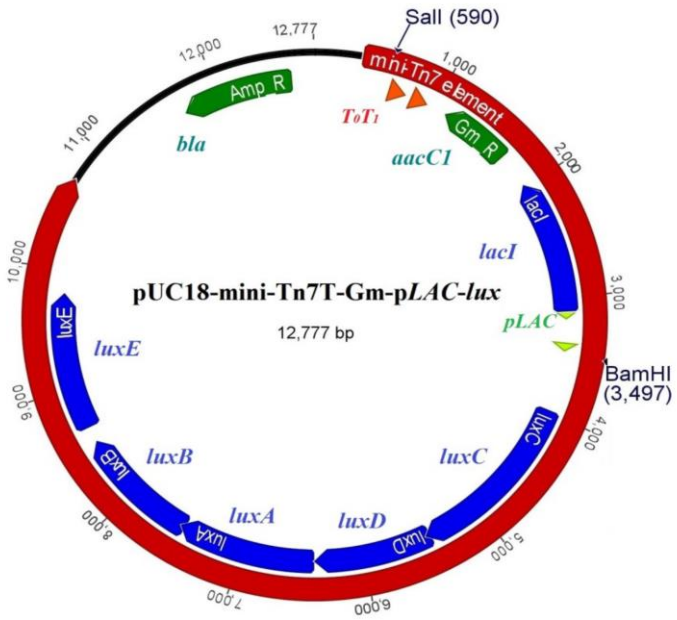


Figure S2.5: Compensatory mutations reduce mean relative *rpoB* expression levels and extent of P26 read-through.

Panel **A**. Except for Q518R + E528D, all compensated mutants showed a statistically significant reduction in mean relative *rpoB* expression compared to the original *rpoB* mutants they had been evolved from (two-tailed, two-sample *t*-test: $p<0.05$). The error bars indicate standard error of the mean.

Panel **B**. All compensated mutants showed a decrease in mean percentage alteration in P26 read-through compared to the *rpoB* mutants without compensatory mutations. The difference between the mean percentage alteration in P26 read-through in H531C and H531R is statistically significant (two-tailed, two-sample *t*-test: $*p<0.05$). The error bars indicate standard error of the mean.



Key:

- T_0T_1: transcriptional terminators
  - T_0 from bacteriophage λ
  - T_1 from the *rrnB* operon in *E. coli*
- mini-Tn7 element: mini-Tn7 transposon for chromosomal insertion
- *aacC1*: gentamycin resistance gene
- *lacI*: *lac* repressor gene
- *pLAC*: *lac* operon promoter
- *bla*: ampicillin resistance gene

Figure S2.6: Vector map for the pUC18-mini-Tn7-*p*LAC-*lux* mini-Tn7 transposon delivery plasmid.

We constructed a mini-Tn7 delivery plasmid carrying an IPTG-inducible luciferase reporter gene (*luxCDABE*) to generate the chromosomally tagged PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain.

Table S2.1: A list of essential genes for gene expression analysis by quantitative real-time PCR.

Names of fifteen randomly selected essential genes in *P. aeruginosa* PAO1 strain whose expression levels in *rpoB* mutants were determined relative to the ancestral strain.

Essential gene	Gene product
ccmH	Cytochrome C-type biogenesis protein CcmH
folD	Folate biosynthesis enzyme 5,10-methylenetetrahydrofolate dehydrogenase
ftsL	Cell division protein FtsL
gatC	Glutamyl-tRNA amidotransferase subunit C
lspA	Prolipoprotein signal peptidase LspA
lytB	LytB protein
msbA	Transport protein MsbA
murG	UPD-N-acetylglucosamine transferase required for the membrane steps of peptidoglycan synthesis
omlA	Outer membrane lipoprotein OmlA precursor
rpsO	30S ribosomal protein S15
sucB	E2 subunit of dihydrolipoamide succinyltransferase
trxA	Thioredoxin
PA1203	Hypothetical protein
PA4425	Sedoheptulose 7-phosphate isomerase GmhA
PA5006	Hypothetical protein

Table S2.2: Oligonucleotide sequences of quantitative real-time PCR primers.

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')	Reference
rpoD	CAGGTGGCGTAGGTGGAGAA	GGGCGAAGAAGGAAATGGTC	[167]
acpP	ACTCGGCGTGAAGGAAGAAG	CGACGGTGTCAAGGGAGT	[168]
16S rRNA	CGGYCCAGACTCCTACGGG	TTACCGCGGCTGCTGGCAC	[169]
gyrB	CCTGACCATCCGTCGCCACAAC	CGCAGCAGGATGCCGACGCC	
luxC	ATACGGTAGGGCAAAGATGG	AGGCCGCCTTTAGAACATAA	[170]
atpA	TACGCCAGATCACGCCCGGA	GGCGGCGTCTACGGTATGGC	This study
groEL	ATCCGCGAAATGCTGCCGGT	ACAGCCGCGACCTTGACGAT	
rpoB	CAGGCCTTCGCTTTCGGCCA	ACCGTCGTGTCCGTTGCGTC	
rpoC	GCGCTGCTGTTCCAGGTCGT	GATCCGCGCCTTCTCGTCCG	
ccmH	CAGGAACTGCGCTGCCCGAA	CTCGTTGACCGGCGGCTTGT	
folD	AAGCGCAAGGACTGCGAGGA	TCGTCGTTCAGGCGGTCGAT	
ftsL	GAGCAGAGCACCTGGACCGC	CGCCACCATCCTGACCTCGG	
gatC	CACCACCGAGACCCTGAACA	ACAGACCTTCTTCCACGGCG	
lspA	AACACCGGTGCCGCCTTCAG	ACGTGGCCAAGGACCATGCG	
lytB	CCAGGATGCCGTGAAGGAAC	GCAGGCGGTTGGAGTTGGAA	
msbA	TACGCGGTGCCGTTGCTGAT	TGGCGAGGAAGAAGTTGCCGA	
murG	ATTGCCCGCCTTCCTGGTGC	AGCGCAGGGTTTCGGGATGC	
omlA	ATGGGCAACCCGCTCATCGT	ACGGTGGTGCTGCCTTCCTT	
rpsO	CAGGTAGCCCTGCTGTCCGC	GACCACGGCGGGAATGGTGA	
sucB	AAGCCGGGCGAAGCAGTCAA	TGCTCAGAACGGTGTCGCCT	
trxA	GCCAGGGACTATCAGGGCAA	TTGAACAGCATCAGGGTAGGGA	
PA1203	GGAGTGGCCGCTGGAGGGTA	GATCGGGCAGGCATTGGCGA	
PA4425	CAGCGGCAATTCGGCCAACG	CAGCAGCAGCGAAGCCATGC	
PA5006	GCGTGCGTGCGGAAACTCCC	CGGCAACTGGGCGAAGAACAC	
*rplL*_P26	GAAGCCGAAGCCGCCAAGAA	AGCCACCAGCCATCAGCCTT	
P26_ intergenic	CGAGCGTCAAGCGAAAGGCT	GCGAGTCCACCCAGAACGGAA	

Annex A: Additional experimental details

1. Evolving and identifying *rpoB* mutants with intragenic compensatory mutations

Rifampicin resistant mutants with compensatory mutations within the *rpoB* gene were evolved during a selection experiment and identified using whole-genome sequencing in a separate study by Q Qi, M Toll-Riera, GM Preston and RC MacLean (Chapter 3).

2. RNA extraction, reverse transcription & quantitative real-time PCR

An ancestral strain control sample was included in every RNA extraction experiment to allow the gene expression levels of target genes in all *rpoB* mutant sample groups to be normalized relative to the ancestral strain. Each strain was inoculated in M9KB medium and grown overnight at 37°C with continuous shaking (225 rpm). The overnight cultures were diluted in M9KB with 1 mM IPTG to induce luciferase expression and incubated at 30°C for two hours with shaking. Prior to RNA extraction, single time point measurements of luminescence emission levels (RLU) per OD_{600} were taken using the Bio-Tek Synergy 2 microplate reader.

Bacterial cultures were mixed with RNAprotect Bacteria Reagent (QIAGEN, Netherlands). Total RNA was extracted using the SV Total RNA Isolation System (Promega, USA) during the early exponential phase. To eliminate genomic DNA, an on-column DNase I digest (Promega, USA) and an additional DNase treatment using the Ambion Turbo DNA-free kit (Life Technologies, USA) were performed during and after the RNA isolation procedure respectively according to the manufacturers' instructions. First-strand cDNA samples were synthesized from 72.5 ng of total RNA templates using the GoScript Reverse Transcription System with random hexamer primers (Promega, USA). Negative control reactions, which contained all components for reverse transcription except reverse transcriptase, were performed in parallel to verify the absence of genomic DNA in cDNA samples.

Quantitative real-time PCR assays were performed using the StepOnePlus Real-time PCR System (Applied Biosystems, USA). Each target gene was amplified using the Fast SYBR Green Master Mix (Applied Biosystems, USA), 50 nM forward and reverse oligonucleotide primers (synthesized by Eurofins MWG Operon, Germany) in three biological and two technical replicates. The oligonucleotide sequences of all the qPCR primers used in this study are shown in Table S2.2. The amplification efficiency, linearity and specificity of each qPCR primer pair were determined three times independently using a serially diluted mixture of experimental cDNA samples.

The *GeNorm* algorithm was applied to assess the average expression stability of six candidate genes to select the most stably expressed group of internal reference genes, which indicated that the most stably expressed genes were 16S rRNA, *acpP*, *atpA* and *rpoD*. The results were further verified using the *NormFinder* algorithm [150], which selects the best internal reference genes by analyzing variations in uncorrected transcript levels at both intra-group (variations between the three biological replicates within each *rpoB* mutant sample group) and inter-group (variations across different *rpoB* mutant sample groups) levels. *NormFinder* showed that *acpP* was the most stably expressed gene, while *acpP* and *atpA* formed the most stable pair of internal reference genes.

To experimentally validate the luciferase reporter gene and demonstrate that variations in luminescence emission (RLU) per OD_{600} reflects variations in mRNA transcript levels of the luciferase reporter gene in *rpoB* mutants relative to the ancestral strain, we correlated relative RLU/OD_{600} measurements against relative *luxC* transcript levels and observed a significant correlation after excluding a statistically significant outlier (Figure S2.1, Supplementary Materials).

Annex B: List of genes that show common differential expression in two *rpoB* mutants sequenced by RNA-Seq

Significance threshold: $p < 0.05$ after Benjamini-Hochberg correction

Key:
Downregulation
Upregulation

Gene name	Locus tag in PAO1	Classification of gene	Expression level relative to the ancestral strain (log_2 fold-change)	
			H531R	D521G
PA0050	PA0050	Essential	-2.18684	-0.90771
PA0122	PA0122	Accessory	-1.28088	-0.87225
opdC	PA0162	Accessory	1.20633	0.933208
aer2	PA0176	Accessory	-1.76184	-0.91709
PA0178	PA0178	Accessory	-1.21783	-0.91378
PA0179	PA0179	Accessory	-1.25571	-0.79626
PA0457	PA0457	Accessory	1.335954	1.065658
nirN	PA0509	Accessory	-1.39106	-1.29512
PA0510	PA0510	Accessory	-1.04659	-1.04948
nirJ	PA0511	Accessory	-1.24411	-1.13651
PA0512	PA0512	Accessory	-1.00894	-1.08486
PA0513	PA0513	Accessory	-1.26739	-1.1847
nirL	PA0514	Accessory	-1.1392	-1.14486
PA0515	PA0515	Accessory	-0.98015	-1.03708
nirF	PA0516	Accessory	-1.04053	-1.23624
nirM	PA0518	Accessory	-1.7293	-1.71523
nirS	PA0519	Accessory	-1.52142	-1.66339
nirQ	PA0520	Accessory	-1.62827	-1.55765
norC	PA0523	Accessory	-1.52219	-1.70251
PA0525	PA0525	Accessory	-1.17286	-1.79333
PA0588	PA0588	Core	-1.11941	-0.80209
ampDh3	PA0807	Accessory	3.304342	1.169189
PA0858	PA0858	Core	0.868666	0.7633
PA1041	PA1041	Accessory	-1.55489	-0.89288
PA1123	PA1123	Accessory	-1.91551	-0.89629
PA1196	PA1196	Accessory	1.886591	0.880185
aprA	PA1249	Accessory	-2.94284	-1.25458
PA1663	PA1663	Accessory	-1.64142	-0.74532

Gene name	Locus tag in PAO1	Classification of gene	Expression level relative to the ancestral strain (log_2 fold-change)	
			H531R	D521G
lasA	PA1871	Accessory	-1.29397	-0.85118
hmgA	PA2009	Core	2.445865	1.598093
liuA	PA2015	Accessory	1.78669	0.919628
liuR	PA2016	Core	2.047642	1.288935
PA2068	PA2068	Accessory	-1.16016	-1.08457
PA2069	PA2070	Accessory	-1.85655	-1.29446
chiC	PA2300	Accessory	-1.72479	-1.26315
gcvP2	PA2445	Accessory	1.771366	0.762797
PA2630	PA2630	Core	1.189041	0.798811
aroP1	PA3000	Accessory	0.91321	0.886733
rmf	PA3049	Accessory	-1.30333	-0.91531
PA3181	PA3181	Accessory	-1.01001	-0.947
gapA	PA3195	Core	-0.93693	-0.74197
PA3284	PA3284	Accessory	-0.73807	-1.11523
PA3310	PA3310	Accessory	1.728817	0.840713
nosZ	PA3392	Accessory	-2.9544	-2.49696
nosD	PA3393	Accessory	-2.12592	-1.87034
rhlA	PA3479	Accessory	-1.00903	-1.00535
PA3839	PA3839	Accessory	1.239984	0.714966
exoS	PA3841	Accessory	-1.38238	-0.90203
narK2	PA3876	Accessory	1.74644	0.854221
PA3941	PA3941	Core	1.389278	0.981531
PA4141	PA4141	Accessory	-1.33585	-0.97524
rpoC	PA4269	Accessory	1.612703	0.843496
rpoB	PA4270	Essential	1.604633	0.907186
PA4358	PA4358	Accessory	-0.87766	-0.72621
desB	PA4888	Accessory	1.288385	1.237716
PA4889	PA4889	Accessory	1.869719	1.004198
PA4916	PA4916	Accessory	-1.10015	-0.91477
PA4939	PA4939	Core	1.062812	0.701101
PA5266	PA5266	Accessory	-2.28754	-0.83102

3. Limits to compensatory adaptation in rifampicin resistant populations of *Pseudomonas aeruginosa*

Qin Qi *[1], Macarena Toll-Riera[1, 2], Gail M. Preston[3], R. Craig MacLean[1]

1. Department of Zoology, University of Oxford, United Kingdom
2. Current affiliation: Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland
3. Department of Plant Sciences, University of Oxford, United Kingdom

Abstract

Compensatory adaptation is thought to play a key role in the long-term maintenance of antibiotic resistance by eliminating the fitness cost associated with resistance acquisition. Although compensatory mutations are beneficial whenever resistance carries a cost, population genetic theory predicts that compensatory adaptation should occur most easily when resistance carries a large cost. To test this prediction, we performed experimental evolution on rifampicin resistant mutants of *Pseudomonas aeruginosa* in conjunction with extensive whole-genome sequencing. Our results revealed a striking systematic difference in the genetic basis of adaptation. Evolved populations founded by costly rifampicin resistant mutants adapted via compensatory mutations in RNA polymerase genes, whereas populations founded by low cost mutations showed an increase in fitness as a result of selection for generally beneficial mutations, predominantly in the *lasR* transcriptional regulator. Moreover, we found evidence that clonal interference between lineages carrying generally beneficial mutations and compensatory mutations constrained compensatory adaptation in populations founded by low cost resistance mutations. The idea that compensatory adaptation maintains resistance has become engrained in evolutionary models of antibiotic resistance, but our study suggests that compensation is likely to be the exception, rather than the rule, except when resistance carries very large costs.

Introduction

Antibiotic resistance is generally accompanied by a fitness cost, expressed in terms of reduced growth rate, competitive fitness and virulence in the absence of antibiotics [3, 39]. Unfortunately, bacteria can adapt to the fitness cost of antibiotic resistance by second-site compensatory mutations that offset the cost of resistance without compromising antibiotic tolerance [3, 24, 39]. Compensatory mutations restore mutant fitness by fully or partially reversing the deleterious effects associated with the original resistance mutations through epistatic interactions [40-44]. Intragenic compensatory mutations restore the structure and function of mutant enzymes or altered RNA, whereas extragenic compensatory mutations restore the structure and function of mutant multi-subunit enzymes or organelles [42]. In a well-characterized example of compensatory adaptation, various populations of a streptomycin resistant mutant of *Salmonella typhimurium* carrying a point mutation in the ribosomal gene *rpsL* were found to have acquired intragenic, as well as extragenic mutations in 3 other ribosomal genes which offset the fitness cost of the original mutation [42, 109, 111].

Compensatory adaptation is thought to play a key role in the evolutionary dynamics of resistance by eliminating selection against resistance in the absence of antibiotics, for example when patients stop antibiotic use, or when pathogens transmit between hosts via antibiotic-free environments. For example, compensatory adaptation has been shown to increase the transmission rate of rifampicin resistant *Mycobacterium tuberculosis* [104, 159, 171], and this may explain why mutations that compensate for the cost of rifampicin resistance have evolved on multiple occasions in clinical isolates of this important human pathogen [159, 172, 173]. More generally, the idea that compensatory adaptation offsets the costs of deleterious mutations has become engrained in bacterial evolution. For example, compensatory adaptation has been shown to offset cost associated with gene loss [174], spontaneous deleterious mutation [175], plasmid carriage [27, 28, 176, 177] and ecological specialization [143]. In this paper, we extend this body of work by asking the question of when antibiotic resistant bacterial populations will evolve compensatory adaptations that offset the fitness cost of antibiotic resistance.

Population genetic theory predicts that fitness costs should play a key role in the evolutionary dynamics of compensatory adaptation. It has been estimated that there are on average 12 possible compensatory mutations per deleterious mutation in bacteria [106]. Therefore, compensatory mutations are likely to arise at a rate of approximately 1×10^{-9} per genome per generation, given an estimated genomic mutation rate of 8.9×10^{-11} per base pair per generation [178]. The rarity of compensatory mutations is expected to impose a limit on the rate of compensatory adaptation, especially because most compensatory mutations are likely to be lost due to stochastic drift when they are initially at low frequency [179]. If compensatory mutations can offset the cost of resistance, as has been observed in many experimental evolution studies to date [46, 108, 109], then compensatory mutations will most likely escape drift in populations where the costs of resistance are large and compensatory mutations carry a large fitness benefit. Therefore, we expect that low fitness costs will constrain the rate of compensatory adaptation by limiting the effective rate at which compensatory mutations are introduced into bacterial populations.

Furthermore, escaping genetic drift does not guarantee that beneficial mutations will ultimately be fixed by selection in bacterial populations. In asexual populations, different beneficial mutations arising independently in different individuals cannot be combined into a single, fitter genotype and must overcome competition with alternative beneficial mutations through clonal interference [180-184]. In contrast with compensatory mutations, generally beneficial mutations contribute to an increase in fitness regardless of the genetic background in which they arise [43, 105] and do not specifically ameliorate the deleterious effects of resistance mutations. Generally beneficial mutations are likely to be an important source of competition for compensatory mutations. Different studies have reported widely different estimates for the overall rate of occurrence of generally beneficial mutations in bacteria, but published estimates fall between approximately 1×10^{-8} [182, 185] and 1×10^{-5} [186, 187], implying that generally beneficial mutations are likely to be 10 to 10000 times more common than compensatory mutations. Crucially, the impact of clonal interference from generally beneficial mutations on compensatory adaptation is predicted to depend on the cost of resistance. When resistance mutations are very costly, compensatory mutations will have large effects on fitness.

Therefore, competition with generally beneficial mutations that are common, but confer a small fitness advantage, is unlikely to impede the fixation of compensatory mutations. In contrast, when resistance carries a small cost, compensatory mutations and generally beneficial mutations will have similar effects on fitness, and clonal interference will slow the rate of fixation of compensatory mutations in bacterial populations (Figure 3.1). In summary, population genetic theory predicts that the rate of compensatory adaptation should increase with the cost of resistance as a result of (i) decreased stochastic loss of rare compensatory mutations and (ii) decreased clonal interference from generally beneficial mutations.

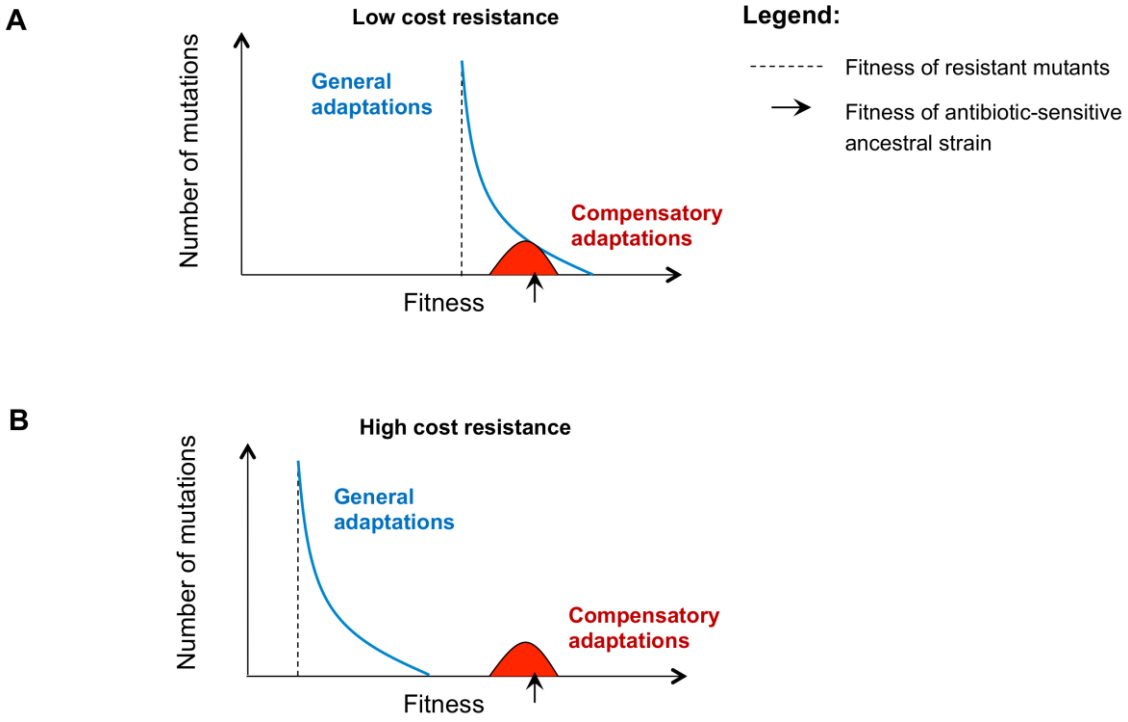


Figure 3.1. Schematic illustrating constrains on compensatory adaptation

This figure shows a conceptual schematic of distribution of fitness effects of beneficial mutations that can be produced by an antibiotic resistant mutant. We assume that generally beneficial mutations are relatively common, and that the fitness effects of beneficial mutations follow an approximately exponential distribution [188]. We assume that compensatory mutations tend to offset most of the cost of resistance, and are relatively rare [178]. Under these assumptions, compensatory adaptations will occur more rapidly when resistance carries a large cost (Panel **B**) than when resistance carries a low cost (Panel **A**).

To test these predictions, we allowed populations founded by isogenic strains of *P. aeruginosa* carrying 8 different rifampicin resistance mutations that are associated with different fitness costs to evolve for 300 generations in the absence of antibiotics. Rifampicin resistance evolves by substitutions in *rpoB* that prevent rifampicin from binding to RNA polymerase (RNAP) [51]. Previous work has shown that bacteria can adapt to the cost of rifampicin resistance by fixing second site mutations in *rpoA*, *rpoB* and *rpoC*, making rifampicin resistance a tractable and well-characterized model system for studying compensatory adaptation [3, 112]. Using experimental evolution, we show that populations founded by different rifampicin resistant mutants adapt and converge in fitness. Whole-genome sequencing of evolved populations shows a systematic difference in the genetic basis of adaptation. Specifically, we find that evolved populations founded by costly rifampicin resistant mutants adapt via compensatory mutations in RNAP genes, whereas populations founded by low cost mutations increase fitness as a result of selection for generally beneficial mutations.

Results & Discussion

Rifampicin resistant mutants recover the fitness cost of resistance

To study the evolution of compensatory adaptation, we allowed three independent populations of eight different *rpoB* mutants to evolve in rifampicin-free culture medium for approximately 300 generations. The increase in average fitness demonstrates that selection was effective over the time of our experiment (Figure 3.2: two-tailed paired *t*-test, $t = 5.47$, d.f = 7, $p = 0.0009$). Overall, the average competitive fitness in endpoint populations converged to that of the rifampicin-sensitive ancestral strain regardless of the initial fitness, demonstrating that selection is able to eliminate the initial fitness cost associated with rifampicin resistance.

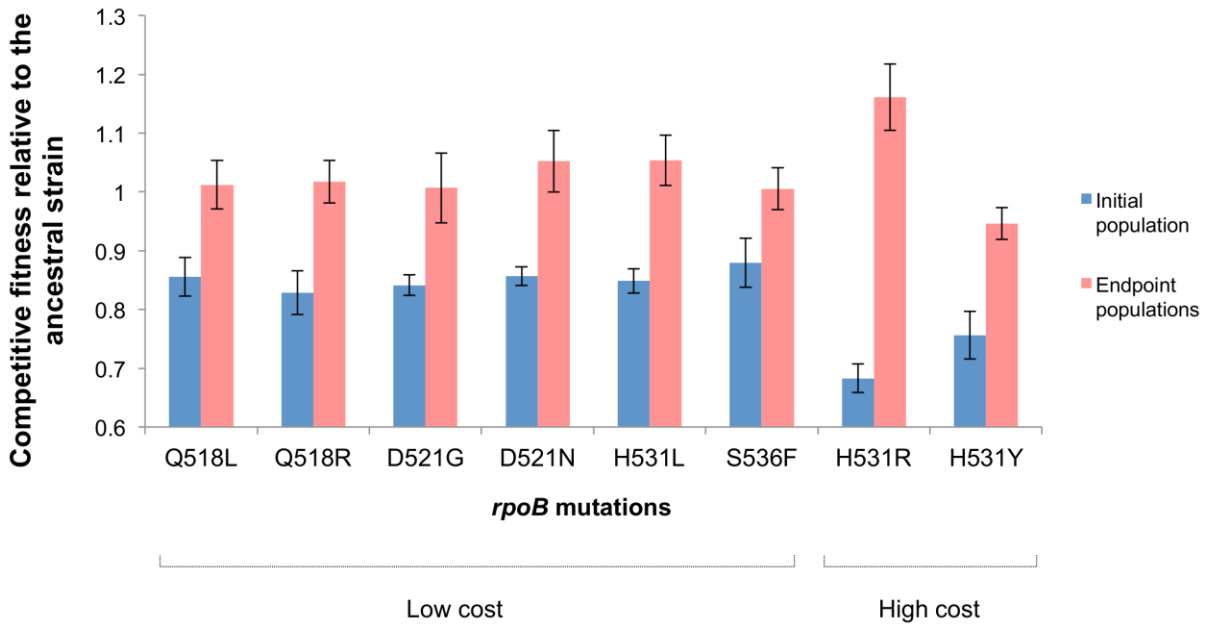


Figure 3.2. Adaptation in rifampicin resistant populations

This figure shows competitive fitness (mean ± s.e; n = 3) of the eight *rpoB* mutants used to found the selection experiments and the average fitness (mean ± s.e; n = 3) of three endpoint populations from the selection experiment. Fitness increased in populations founded by each *rpoB* mutant (two-tailed, two-sample *t*-test, $p < 0.05$), and the fitness of evolved lines converged on that of the rifampicin sensitive ancestral strain.

Testing for compensatory adaptation using whole-genome sequencing

To determine the mechanistic basis of adaptation, we performed whole-genome sequencing on 3 colonies isolated from each evolved endpoint population. Across the 72 clones of the 24 endpoint populations we sequenced (Table 3.1), a total of 63 independent mutations were identified. Single base-pair substitutions (SNPs) were the most common form of mutations (n = 37), but we also identified a large number of short insertion and deletion (INDEL) polymorphisms (< 1 kb; n = 20) and a small number of large deletions and duplications (> 1 kb; n = 7).

Table 3.1: Genome-wide mutations in endpoint populations of rifampicin resistant mutants

***rpoB* mutation**	**Gene**	**Gene locus**	**Effects of mutation**	**Endpoint clones**								
				Population A			**Population B**			**Population C**		
				1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Q518L	Intergenic, upstream of PA0040		Insertion: 4 bp								■	
	PA0829		SNP: I172T				■					
	lasR	PA1430	Deletion: 66 bp				■		■			
			Deletion: 1 bp							■	■	
			Stop-gain: W88*					■				
			Stop-gain: Q98*	■	■	■						
			SNP: S129I									■
	bkdA1	PA2247	Deletion: 213 bp				■					
	rpoA	PA4238	SNP: S49F								■	
Q518R	*lasR*	PA1430	Stop-gain: K16*				■					
			SNP: R180W							■	■	
			SNP: V208E						■			
			SNP: A231V				■	■				■
	wzz	PA3160	SNP: T50A							■	■	
	arnT	PA3556	SNP: synonymous				■	■	■			
	rpoA	PA4238	SNP: D198V								■	
	rpoC	PA4269	SNP: D410G						■			
			SNP: G367S									■
	rpoB	PA4270	SNP: E528D	■	■	■	■	■				
D521G	*lasR*	PA1430	Deletion: 56 bp				■	■				
			SNP: T74P							■		■
			SNP: L84M								■	
			Stop-gain: S106*	■	■	■						
	tufB	PA4277	SNP: F6I				■	■	■			
	cupE6	PA4653	SNP: F124L						■			
D521N	*lasR*	PA1430	Deletion: 358 bp								■	■
			Deletion: 147 bp							■		
			Insertion: 2 bp	■								
			SNP: A227V						■			
	Several genes including *lasR*	PA1425 to PA1435	Deletion: 12457 bp					■				
	Several genes including *lasR*	PA1429 to PA1432	Deletion: 3873 bp				■					
	Several genes related to biofilm formation	PA2228 to PA2251	Deletion: 29106 bp		■							

Table 3.1: Genome-wide mutations in endpoint populations of rifampicin resistant mutants (Continued)

***rpoB* mutation**	**Gene**	**Gene locus**	**Effects of mutation**	**Endpoint clones**								
				Population A			**Population B**			**Population C**		
				1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
H531L	*fha1*	PA0081	Insertion: 6 bp						■			
	Several genes including *lasR*	PA1429 to PA1432	Deletion: 3139 bp	■								
	lasR	PA1430	Deletion: 217 bp		■							
			Deletion: 13 bp									■
			Insertion: 1 bp			■						
	Intergenic, upstream of PA1519 and PA1520		SNP	■	■	■						
	cupE6	PA4653	SNP: F124L				■	■	■			
			SNP: S185R							■	■	
H531R	*lasR*	PA1430	Deletion: 217 bp	■	■	■						
	Several genes	PA2225 to PA2228	Duplication: 2683 bp				■	■	■			
	Several genes	PA2339 to PA2386	Deletion: 55682 bp				■		■			
	rpoB	PA4270	Reversion							■	■	■
			SNP: H531C	■	■	■	■	■	■			
H531Y	PA4673.16 #		SNP: R311H									■
	PA1829		SNP: I295T			■						
	PA1944		SNP: S336A						■			
	Intergenic downstream of *pvcD* (PA2257) and *ptxR* (PA2258)		Deletion: 130 bp									■
	gacA	PA2586	Stop-gain: Q168*	■	■							
	Intergenic, upstream of *fptA* (PA4221)		SNP					■				
	rpoB	PA4270	SNP: N573S			■	■	■	■			
			SNP: H531C							■	■	■
S536F	*gacS*	PA0928	Stop-gain: Q402*			■						
	2 genes including *lasR*	PA1429 to PA1430	Deletion: 2058 bp							■	■	■
	lasR	PA1430	Deletion: 11 bp	■								
			Deletion: 1 bp			■						
			Insertion: 1 bp					■				
			Deletion: 20 bp		■							
			Deletion: 1 bp						■			
	PA2067		SNP: synonymous					■				
	Intergenic, upstream of *pvdG* (PA2425) and *pvdS* (PA2426)		SNP					■				
	selA	PA4808	SNP: synonymous					■				

A putative protein kinase gene from the phage-like insertion RGP42 reference genome. The site of insertion in the *P. aeruginosa* PAO1 reference genome is the tRNA-Met PA4673.1 phage integration attachment site.

The most striking feature of our results is the very high frequency of parallel mutations in the quorum-sensing transcriptional regulator gene *lasR*. 75% of all sequenced clones carried a mutation in *lasR*, with 26 independently evolved SNPs and short INDELs in *lasR*, as well as 3 independent large deletions (> 3 kb) spanning *lasR*. These missense, nonsense or frameshift mutations are expected to disrupt LasR function and result in a quorum-sensing-deficient phenotype in mutants [189, 190]. LasR is a key transcriptional regulator of quorum-sensing regulated genes in *P. aeruginosa*. Each LasR monomer consists of an N-terminal ligand-binding domain and a C-terminal DNA-binding domain [191]. LasR is stabilized through binding with the homoserine lactone autoinducer (AHL) ligands, which promotes dimerization of LasR [192]. The resultant LasR-AHL dimer complex binds to target DNA and activates the transcription of LasR-regulated quorum-sensing genes [193].

Previous work has shown that *lasR* mutations carry a general fitness benefit in laboratory culture medium by decreasing the frequency of cell lysis and cell death [194]. In addition, *lasR* mutants are known to be social cheats that enjoy a fitness advantage by avoiding the high metabolic cost associated with producing quorum-sensing signals and performing quorum-sensing regulated cooperative behavior [195]. To confirm that *lasR* mutations are beneficial, we measured the competitive fitness of 3 evolved clones that carry single mutations in *lasR* relative to their respective *rpoB* mutant ancestors. All evolved endpoint clones showed an increase in average competitive fitness compared to the initial *rpoB* mutants (Figure 3.3), with two showing a statistically significant increase in fitness (two-tailed, two-sample *t*-test, $*p < 0.05$). However, the observation that *lasR* mutations increase fitness does not provide any direct insight into whether these mutations are generally beneficial mutations or compensatory mutations since both types of mutations increase fitness. Our previous work has shown that the cost of rifampicin resistance arises from the deleterious effects of *rpoB* mutations on transcriptional efficiency, which was estimated using an IPTG-inducible luciferase reporter gene (refer to the attached manuscript for experimental method and results). Using the same experimental method, we found that mutations in *lasR* in the endpoint clones did not recover the compromised transcriptional efficiency associated with the original *rpoB* mutations (Figure S3, Supplementary Material), providing further evidence that *lasR* mutations are not compensatory.

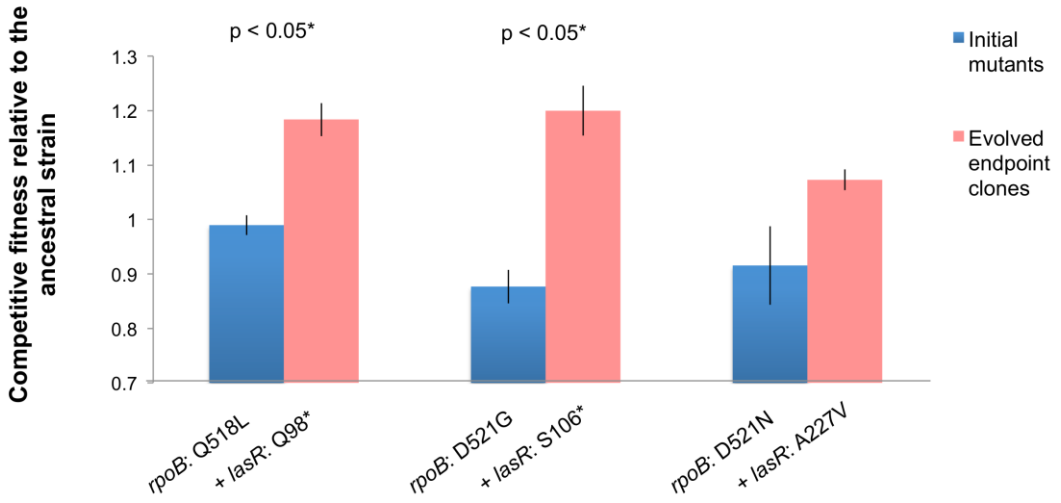


Figure 3.3: *lasR* mutations are beneficial.

The competitive fitness (mean ± s.e; n=3) of three initial *rpoB* mutants (blue bars) and three evolved endpoint clones (red bars) with single point mutations in the *lasR* gene were determined relative to the rifampicin-sensitive ancestral strain. Stop-gain mutations are denoted by the * symbol in *lasR*. All three evolved endpoint clones showed an increase in average competitive fitness compared to their ancestral *rpoB* mutants, two of which represent a statistically significant increase (two-tailed, two-sample *t*-test, *p < 0.05).

A small number of lineages have also acquired putative generally beneficial mutations in genes that are related to biofilm formation. For example, non-synonymous mutations in the *cupE6* gene were found in one endpoint population of the D521G *rpoB* mutant and two endpoint populations of H531L. CupE6 is a fimbrial adhesin-like protein in the chaperone-usher pathways (Cup) [196, 197]. Biofilm formation in *P. aeruginosa* involves fimbrial structures that are components of and assembled on the outer surface of bacterial cells by the gene products of Cup [198]. In one endpoint population of D521N and one endpopulation of H531L, deletion mutations were found in the *bkdA1* gene (PA2247), whose expression has been shown to be required for biofilm formation, but the functional roles of this gene in biofilm formation have yet to be characterized at a molecular level [199].

A number of lines of evidence demonstrate that the mutations that we identified outside of *lasR* were also predominantly beneficial mutations that spread as a result of positive selection. We found that the number of independently evolved non-synonymous mutations (n = 21) relative to the number of independently evolved synonymous mutations (n = 2) was much greater than expected under neutrality (K_a/K_s = 4.17, χ_1^2 = 4.39, p = 0.036). The small number of intergenic mutations that we observed (n = 4) made it difficult to test for selection on non-coding regions, which only comprise approximately 10% of the *P. aeruginosa* genome. However, we detected a selective sweep on an intergenic mutations immediately upstream (-17 bp) of the start site of a putative transcription factor (PA1520), demonstrating that at least ¼ of the observed intergenic mutations were beneficial. In these estimates of selection on coding and intergenic mutations, we excluded a presumably beneficial synonymous mutation in *arnT* that swept to fixation in one population.

We found 9 independently evolved point mutations in RNAP, including *rpoA* (n = 2), *rpoC* (n = 2), and *rpoB* (n = 5). Interestingly, one of the *rpoB* mutations represents a reversion mutation that swept to fixation in a lineage that was initiated with the H531R mutation. Several lines of evidence suggest that the remaining mutations in RNAP genes are compensatory mutations. First, we have demonstrated in a previous study that second-site intragenic mutations in *rpoB* can recover the reduction in transcriptional efficiency and fitness associated with rifampicin resistance mutation (refer to the accompanying manuscript). Second, we found evidence for parallel mutations at an amino acid level. Mutations leading the substitution of cysteine at position 531 of *rpoB* swept to fixation in two populations founded by H531R and one population founded by H531Y. Similarly, mutations leading to the E528D substitution in *rpoB* were also found in two independent populations founded by Q518R. Finally, many of the sites where we identified mutations in RNAP have been previously shown to be involved in compensation for the cost of rifampicin resistance in *Salmonella enterica* [66, 104] and *Mycobacterium tuberculosis* [21, 159, 171].

Constraints on compensatory adaptation

To test the hypothesis that low fitness costs act as constraint on compensatory adaptation, we tested for a negative correlation between the frequency of RNAP mutations that evolved during the selection experiment and the initial cost associated with each resistant mutant (Figures 3.2 and 3.4). We found a significant negative correlation between the number of compensatory mutations per genome and the initial cost associated with resistance (Spearman's rank correlation: r = 0.659, $F_{1, 6}$ = 11.6, p = 0.0145). This correlation is driven by the fact that almost all clones from populations that were founded by very costly resistance mutations, such as H531R and H531Y, carry compensatory mutations, whereas no compensatory mutations were found in populations that were founded by low cost resistance mutations, such as S536F. We included the reversion mutation in this analysis, because it is an adaptation to the cost of rifampicin resistance.

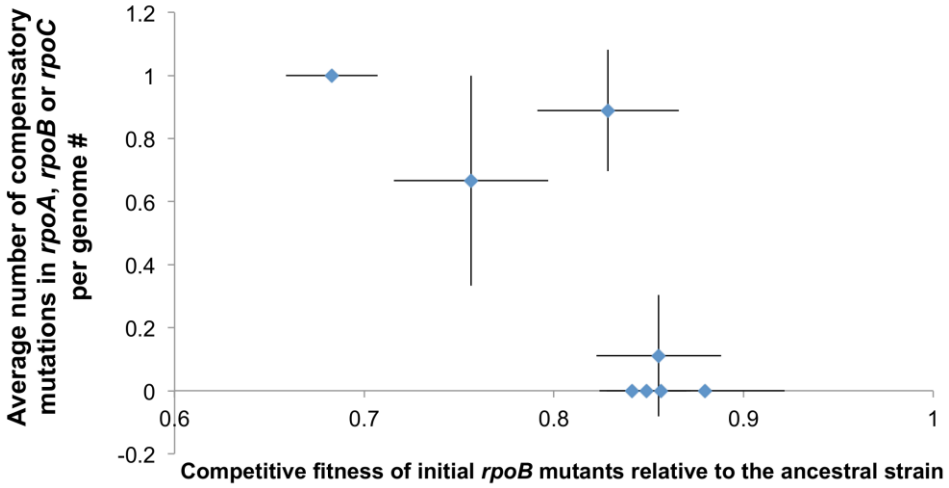


Figure 3.4. High fitness costs drive compensatory adaptation in rifampicin resistant populations of *P. aeruginosa*.

This figure shows the mean (mean ± s.e; n=3) number of compensatory mutations (# including the reversion mutation in *rpoB*) in RNA polymerase per evolved clone as a function of fitness cost associated with the 8 rifampicin resistance mutations (mean ± s.e; n=3). The number of compensatory mutations per clone increases with the cost of resistance, as judged by a rank-order correlation ($p < 0.05$).

According to the model that we have proposed, clonal interference between compensatory mutations and generally beneficial mutations limits the rate of compensatory adaptation in populations where the cost of resistance is low (Figure 3.1). To test for clonal interference, we re-constructed the phylogeny of the 3 clones sampled from endpoint populations. If clones carrying compensatory mutations co-occur with clones that have independently evolved generally beneficial mutations on a different branch of the phylogeny, then it can be inferred that clonal interference constrains the spread of compensatory mutations. Only 3 endpoint populations showed polymorphism for compensatory mutations in *rpoA, rpoB* or *rpoC*, which is a pre-requisite for detecting clonal interference, and in all three cases we found direct evidence for competition between lineages carrying generally beneficial mutations and compensatory mutations (Figure 3.5). More generally, we found evidence of co-existing clonal lineages in 16 of 24 populations, suggesting that clonal interference was widespread in our experiment.

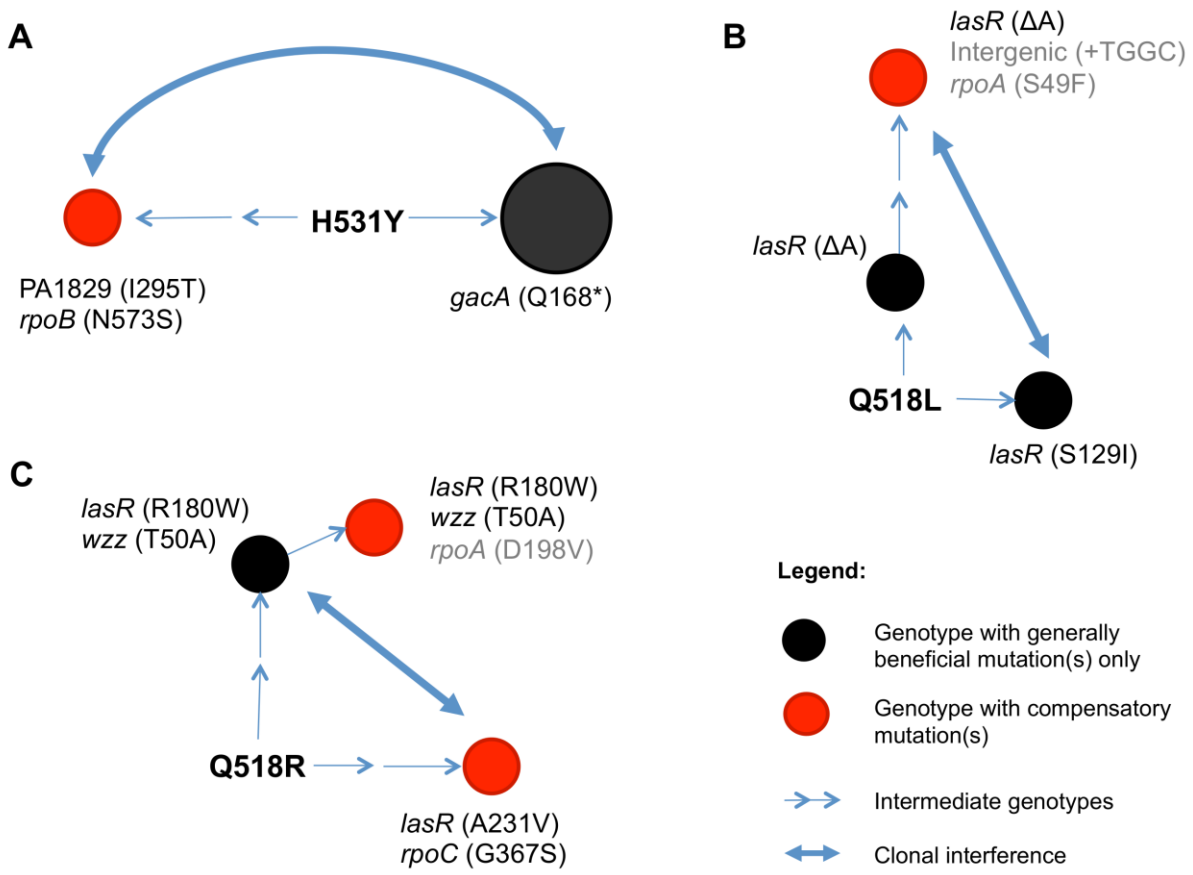


Figure 3.5: Clonal interference constrains compensatory adaptation.

This figure shows a schematic representing the phylogeny that we inferred for the three clones that we sampled from the 3 populations that were polymorphic for compensatory mutations. Circles show observed genotypes, and the diameter of the circle indicates the frequency of the clone in the population (n=1 or n=2). Red circles denote clones carrying compensatory mutations and black circles denote clones carrying generally beneficial mutations, but not compensatory mutations. Each thin arrow represents a mutation, and mutations carried by individual mutations are shown with text, with newly acquired mutations in grey and previously acquired mutations in black. Thick arrows show clonal interference between independently evolved lineages carrying generally beneficial and compensatory mutations.

Finding a large number of targets of positive selection combined with evidence of widespread clonal interference suggests that *P. aeruginosa* has access to a large number of beneficial mutations with relatively small effects on fitness, which fits with theoretical expectations. If this argument is correct, then we would expect that clonal interference should be more common in high fitness populations that have access to a large number of weakly beneficial mutations, than in low fitness populations, which also have access to rare and highly beneficial mutations. Consistent with this idea, we found that the frequency of clonal interference, inferred from the average number of independently evolved lineages per initial *rpoB* mutant population, was greatest in populations founded by low cost resistance mutations (Spearman's rank correlation: $r = 0.503$, $F_{1, 6} = 6.07$, $p = 0.0488$).

Conclusion

Compensatory adaptation is crucial to the long-term maintenance of antibiotic resistance alleles in bacterial populations. Here, we show that populations of *P. aeruginosa* are only capable of evolving compensatory adaptation over a timeframe of 300 generations when costs of resistance are very large (>15%), and this is at least partially attributable to the fact that selection for generally beneficial mutations impedes the evolution of compensatory adaptation by clonal interference. One important implication of this result is that selection for compensatory adaptation will be relatively ineffective under conditions where selection at other sites in the genome is strong. A number of important selective forces are likely to exert strong and recurrent selective pressure on populations of bacterial pathogens, such as antibiotic use [200, 201], bacteriophage [202] and the immune system [203], implying that clonal interference may impose an important limit to compensatory adaptation in clinical settings. Clonal interference is likely to play an especially important role in opportunistic pathogens due to the fact that opportunistic pathogens are essentially invading a novel niche to which they are initially poorly adapted. For example, adaptive evolution occurs when opportunistic [204, 205] and commensal [206] bacteria make the transition to a pathogenic lifestyle. In this respect, it is interesting to note that the best-characterized example of compensatory adaptation in a clinical

population, rifampicin resistance in *M. tuberculosis* [104, 159, 171], comes from an obligate pathogen that is presumably well adapted to life in mammalian hosts [207].

The idea that compensatory adaptation eliminates the cost of resistance and allows resistance genes to effectively persist in bacterial populations has become engrained in evolutionary models of antibiotic resistance [3, 15]. For example, it is often speculated that mutations that are found using whole-genome sequencing of antibiotic resistant clinical isolates are compensatory mutations [159, 172, 201, 208, 209]. Although there are some excellent examples of compensatory adaptation from clinical pathogens, a number of lines of evidence suggest that compensatory adaptation is the exception, rather than the rule, in clinical pathogen populations. The high frequency and variety of compensatory mutations that have been reported in many experimental evolution studies to date could be attributed to the low fitness of ancestral resistant mutants that were used to initiate these selection experiments, which favored compensatory adaptation [108, 109, 113-116]. Our study provides a simple explanation for why it can be difficult for bacterial populations to evolve compensatory adaptations, and we hope that this will pave the way for future work aimed at testing for limits to compensatory adaptation in clinical pathogens.

Materials & Methods

Selection Experiment

From our earlier study, rifampicin-resistant *rpoB* mutants were evolved from the rifampicin-sensitive PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain using a fluctuation test. Eight initial *rpoB* mutant strains were streaked on M9KB agar plates [147] and incubated overnight at 30°C. On the next day, 3 independent colonies per *rpoB* mutant strain were randomly selected for propagating three independent lineages (A, B and C) during the selection experiment. 1 ml of M9KB medium was inoculated with a colony in the absence of rifampicin. The bacterial cultures were incubated overnight with shaking (225 rpm) at 30°C. On the next day, 1 µl of each overnight culture was diluted 1000-fold in fresh M9KB medium and incubated overnight under the same experimental conditions. The serial transfer was repeated for 30 consecutive days, or approximately 300 generations. Glycerol stocks of populations from the selection experiment were prepared every 6 days.

Competitive fitness assay

In our previous study, we quantified the competitive fitness of the eight initial *rpoB* mutant strains relative to their rifampicin-sensitive PAO1::mini-Tn7-*p*LAC-*lux* ancestral strain using competitive fitness assays. We used the same method in this study to measure the competitive fitness of the endpoint populations of each of the propagated lineages in the selection experiment relative to the same ancestral strain. Briefly, the endpoint populations, the ancestral strain and a GFP-tagged derivative strain of PAO1 were pre-cultured in M9KB medium overnight at 37°C with continuous shaking. Overnight pre-cultures were diluted in M9KB and re-grown to early exponential phase at 30°C. Each non-fluorescent strain was mixed with the GFP-tagged strain, diluted 200-fold in M9KB medium and incubated in Nunc 96-well microplates (Thermo Scientific, USA) at 30°C overnight for approximately 24 hours with continuous shaking. The proportion of fluorescent and non-fluorescent cells in each co-culture was determined using an Accuri C6 flow cytometer (BD Biosciences, USA)

before and after the overnight incubation. The competitive fitness of each non-fluorescent strain was calculated as the ratio of population doublings of each endpoint population relative to the GFP-tagged control strain it was competing against [210]. The relative competitive fitness of the endpoint populations was obtained by standardizing their competitive fitness to that of the ancestral strain within each round of competition. Three biological and three technical replicates of each strain were assayed for each endpoint population.

To demonstrate that *lasR* mutations are beneficial under the conditions of the selection experiment, the competitive fitness of 3 evolved clones that carried different *lasR* mutations and the initial *rpoB* mutant strains were determined relative to the rifampicin-sensitive ancestral strain using a modified version of the competition assays, in which all strains were re-grown to exponential phase and co-cultured overnight with the GFP-tagged control strain in 5 ml Falcon tubes (BD Biosciences, USA) instead of Nunc 96-well microplates.

Whole-genome sequencing

The initial *rpoB* mutants and the endpoint populations from every lineage of the selection experiment were streaked on M9KB agar plates and incubated overnight at 30°C. On the next day, 3 independent colonies from each population were randomly selected to inoculate M9KB medium incubated at 30°C overnight. Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (QIAGEN, Netherlands) according to the manufacturer's protocol for Gram-negative bacteria.

Paired-end whole-genome sequencing with read length of 100 bp was performed using the HiSeq 2000 Sequencing System (Illumina, USA) at the Welcome Trust Centre for Human Genetics, University of Oxford. The average coverage was 50.3x (median coverage: 48.2x). Variants were identified using an experimentally validated in-house pipeline, previously described by San Millan *et al* (2014). The initial *rpoB* mutant isolates were found to have the same genetic background, with the exception of the mutations in *rpoB*. We discarded mutations that were already present in all the initial

rpoB mutant clones with respect to the *P. aeruginosa* PAO1 reference genome, analyzing only those mutations that have accumulated throughout the selection experiment.

Testing for positive selection

To test for selection on base substitutions in protein coding genes, we estimated the expected number of protein altering mutations, under the assumption that synonymous mutations are effectively neutral. First, we estimated the relative rate of all possible single base pair substitutions using data from 289 independently evolved SNPs obtained from whole-genome sequencing of *P. aeruginosa* selection experiments. Given these mutation rates, we used the nucleotide composition and codon usage of *P. aeruginosa* proteins to estimate the rates of non-synonymous and synonymous mutations expected under neutrality. Our estimated neutral rate was 2.52, owing to the biased GC content and codon usage in the *P. aeruginosa* PAO1 genome.

Acknowledgements

This work was supported by funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant (StG-2011-281591) and from the Royal Society.

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics funded by Wellcome Trust grant reference 090532/Z/09/Z and Medical Research Council Hub grant G0900747 91070 for generation of the high-throughput sequencing data.

Author Contributions

Conceived and designed the experiments: QQ, MTR, GMP, RCM

Performed the experiments: QQ

Analyzed the whole-genome sequencing data: MTR

Wrote the paper: QQ, MTR, GMP, RCM

Competing Financial Interests

The authors declare that there are no competing financial interests.

Supplementary Material for Chapter 3

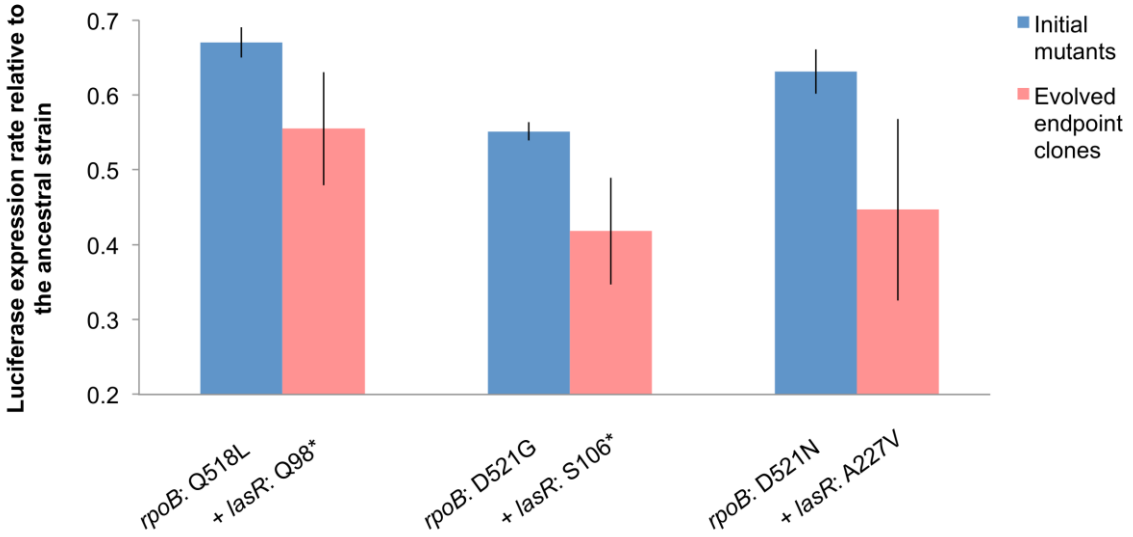


Figure S3: *lasR* mutations are not compensatory mutations.

Transcriptional efficiency (mean ± s.e; n=3) in three initial *rpoB* mutants (blue bars) and three evolved endpoint clones (red bars) with single point mutations in the *lasR* gene were estimated relative to that of the rifampicin-sensitive ancestral strain by measuring luciferase expression rate (Materials & Methods, Chapter 2). Stop-gain mutations are denoted by the * symbol in *lasR*. None of the three evolved endpoint clones showed an increase in average transcriptional efficiency compared to their ancestral *rpoB* mutants

4. Molecular mechanisms underpinning the fitness cost of pNUK73 small plasmid carriage in *Pseudomonas aeruginosa*

The study in Appendix A of this thesis was performed under the joint first authorship of Dr A San Millan and Dr M Toll-Riera. The emphasis of this chapter is on the experimental work I carried out within the context of this study. A brief summary of the results of the study shown in Appendix A is provided in the Introduction section of this chapter. The details of my contributions to this work are summarized on page 4 of this thesis.

Introduction

Bacteria can evolve resistance to antibiotics through acquisition of resistance genes carried on mobile genetic elements such as plasmids and lysogenic phages during horizontal gene transfer [3, 15]. The acquisition of plasmids is typically associated with fitness costs on the bacterial hosts [3, 15, 25, 126]. The molecular mechanisms that contribute to the fitness cost associated with plasmid carriage are complex, but many lines of evidence from literature support the idea that the expression of plasmid-encoded genes is the source of such fitness costs [15, 126]. Expressing plasmid-encoded genes imposes metabolic burdens on host cells and sequesters cellular machinery from performing critical housekeeping processes [211, 212]. In addition, newly acquired plasmid-encoded genes have not co-evolved with the existing genetic pathways and networks in the recipient genome and have the potential to function suboptimally within the genetic background of their new bacterial hosts [213].

pNUK73 is a medium host range shuttle vector that was engineered from the pNI10 plasmid, a cryptic natural plasmid originally isolated from *Pseudomonas fulva* [127]. Both pNUK73 and pNI10 are classified as small plasmids [214]. In contrast with large plasmids, small plasmids lack conjugative machinery, carry few potentially beneficial genes and rely on high copy numbers (usually >10) to avoid plasmid loss during segregation of host cells [214, 215]. pNUK73 has a plasmid size of 5128 bp and encodes 3 genes; namely a plasmid replication protein gene *rep*, an aminoglycoside 3'-

phosphotransferase gene *aphA1* (conferring resistance to both kanamycin and neomycin) and a β-galactosidase reporter gene *lacZ*.

A recent study demonstrated that the competitive fitness of the pNUK73-bearing *P. aeruginosa* PAO1 strain (PAO1/pNUK73) was significantly lower than that of the PAO1 wild-type strain (mean fitness of PAO1/pNUK73 relative to PAO1 = 0.786, s.e = 0.008, n = 18), suggesting that pNUK73 carriage was associated with a large fitness burden on the PAO1 host [128]. Using an experimental evolution approach, PAO1/pNUK73 was serially passaged in antibiotic-free LB medium for 300 generations. The competitive fitness of the endpoint population of PAO1/pNUK73 returned to the wild-type level (mean relative fitness = 1.018, s.e = 0.002, n = 3), suggesting that the fitness burden associated with pNUK73 carriage has been completely offset at a population level. However, the question remained as to whether it was the PAO1 host that adapted to the pNUK73 plasmid or *vice versa* during their co-existence. Laboratory experimental evolution studies examining other models of plasmid-carriage in bacteria have reported that compensatory adaptation can occur in the host, in the plasmid or both [27, 28, 216]. In the PAO1/pNUK73 model, the answer was revealed by whole-genome sequencing of endpoint clones of PAO1/pNUK73, which demonstrated that mutations in any of three specific genes in the PAO1 genome were responsible for the compensatory adaptation of the host [128]. All endpoint clones with compensatory mutations showed wild-type fitness levels. The three chromosomal genes that acquired compensatory mutations include a putative helicase (PA1372) and two contiguous putative serine/threonine protein kinases (PA4673.15 and PA4673.16).

In the manuscript included in Appendix A, we combined a transcriptomic profiling approach with bioinformatics to dissect the molecular mechanisms that contribute to the fitness cost of pNUK73 carriage in the *P. aeruginosa* PAO1 host. Bioinformatic analyses revealed that all the three genes in which compensatory mutations were found are recently acquired genes in the PAO1 genome. These compensatory mutations include various frameshift and non-synonymous mutations, all of which are thought to result in the inactivation of the three gene products. In addition, RNA-Seq transcriptome profiling revealed that 749 genes showed significant differential expression in parental PAO1/pNUK73 relative to the PAO1 wild-type strain ($p < 0.05$ after Benjamini-Hochberg

correction). In particular, key SOS response regulator genes such as *lexA* and *recA* showed significant upregulation in parental PAO1/pNUK73. In addition, ten out of the fifteen known LexA-regulated *P. aeruginosa* genes showed significant upregulation in PAO1/pNUK73 relative to the PAO1 wild-type strain. The remaining five Lex-A regulated genes also showed increased average expression, albeit at statistically insignificant levels. This suggests that the activation of the SOS response is an important transcriptional response associated with pNUK73 carriage. After 300 generations of co-existence, the expression levels of almost all the differentially expression genes, including the previously upregulated SOS response genes, returned to wild-type levels in the sequenced endpoint clones of PAO1/ pNUK73 carrying compensatory mutations. Interestingly, the expression levels of the replication protein gene (*rep*) per plasmid copy showed drastic and significant reductions in the compensated mutants relative to that in parental PAO1/pNUK73. This suggests that the compensatory mutations have the dual effect of decreasing *rep* expression and silencing the expression of genes associated with the SOS response in PAO1/pNUK73.

By combining all experimental and bioinformatic evidence, we hypothesized that the fitness cost associated with pNUK73 carriage in parental PAO1/pNUK73 stems from intragenomic conflicts between the pNUK73-encoded *rep* gene and the three recently acquired chromosomal genes in the PAO1 genome (PA1372, PA4673.15 and PA4673.16). By extrapolating from existing evidence in literature, the fitness cost of pNUK73 carriage is expected to stem from the expression of the plasmid-encoded genes [126]. However, it is evident that the pNUK73-encoded *lacZ* gene and the kanamycin/neomycin resistance gene *aphA1* are unlikely to have contributed to the fitness cost in PAO1/pNUK73, because *rep* is the only pNUK73-encoded gene whose expression levels showed significant reductions in the sequenced endpoint clones of PAO1/pNUK73 relative to that in the initial population. While the expression levels of *lacZ* increased slightly in the endpoint clones with compensatory mutations, those of *aphA1* remained invariant. Therefore, we hypothesized that *rep* expression alone is responsible for driving the fitness cost of pNUK73 carriage.

To test this idea, I cloned an IPTG-inducible copy of the *rep* gene into a neutral chromosomal site (mini-Tn7) in the plasmid-free PAO1 wild-type strain to generate the PAO1::*p*LAC-*rep* strain. The

objective is to characterize the genotypic and phenotypic effects of *rep* expression on the PAO1 host and exclude any possible confounding effects from the other two pNUK73-encoded genes. By manipulating the concentration of IPTG, we hope to control and modulate the expression levels of the *rep* gene downstream of the *lac* operon promoter (*p*LAC). The fitness cost on the PAO1 host is expected to increase with *rep* expression levels.

A specific molecular mechanism through which host fitness can be compromised due to small plasmid carriage is the SOS response. The ability of a plasmid replication protein to activate the SOS response in a bacterial host has been demonstrated in a similar bacterial host/plasmid model. Similar to pNUK73, pSC101 is a small plasmid which also depends on its plasmid-encoded replication protein (RepA) for its own replication [217]. Using the *E. coli*/pSC101 system as a model, Ingmer *et al* showed that an increase in RepA expression can delay host cell replication by activating the SOS response [218]. In addition to its primary function of initiating plasmid replication by binding to pSC101, RepA can also recruit host replication proteins such as DnaG and influence the rate at which the host genome is replicated [218]. When RepA is in excess, the interaction between RepA and DnaG has deleterious effects on host genome replication and triggers the activation of the SOS response. By retarding the rate of host cell division via the SOS response, RepA plays an important role in preventing the loss of pSC101 during segregation of host cells [218], but can also be viewed as the molecular origin of the fitness cost of pSC101 on the *E. coli* host.

In our PAO1/pNUK73 model, we hypothesized that the activation of the SOS response is an important consequence of *rep* expression that has the potential to contribute to the fitness of pNUK73 carriage. To test this idea, I quantified the expression levels of two key SOS response regulator genes *lexA* and *recA* in the PAO1::*p*LAC-*rep* strain. The expression levels of *lexA* and *recA* are expected to increase with *rep* expression levels in this strain.

Results and Discussions

Expression of *rep* generates a fitness cost

To test the hypothesis that *rep* expression is directly responsible for the fitness cost associated with pNUK73 carriage, we cloned the *rep* gene under the control of an IPTG-inducible *lac* operon promoter (*p*LAC) and integrated this inducible *rep* construct into the PAO1 genome using the mini-Tn7 system [144] to generate the PAO1::*p*LAC-*rep* with a putative ribosomal binding site (RBS) 5 bp upstream of the start codon of *rep* and the PAO1::*p*LAC-*rep1* strain without the putative RBS. Using the same system, we cloned *p*LAC into the PAO1 genome to generate the PAO1::*p*LAC negative control strain.The competitive fitness of the two PAO1 mini-Tn7 insertion strains bearing the p*LAC-rep* construct and that of the PAO1::*p*LAC negative control strain were determined relative to the PAO1 wild-type strain in LB medium containing 0 µM, 10 µM or 50 µM IPTG. The competitive fitness of the negative control strain and the PAO1::*p*LAC-*rep1* strain (without the putative ribosomal binding site) showed no statistical difference from that of PAO1 regardless of the presence and absence of IPTG (one-way ANOVA: $F_{6, 14} = 2.08$, $p = 0.121$).

The PAO1::*p*LAC-*rep* strain with the putative ribosomal binding site was the only strain that showed significant reductions in its competitive fitness relative to PAO1 across difference the concentrations of IPTG (one-way ANOVA: $F_{3, 8} = 3085$, $p < 0.0001$). The fitness of PAO1::*p*LAC-*rep* decreases with increasing IPTG concentration. The reduced fitness of PAO1::*p*LAC-*rep* relative to PAO1 even in the absence of IPTG induction (mean relative fitness = 0.846, s.e = 0.005, n = 3) is probably a consequence of leaky expression of *rep* from *p*LAC. At a phenotypic level, these relative fitness measurements provide the first line of evidence that *rep* expression is associated with a fitness cost in the absence of the other two pNUK73-encoded genes.

Interestingly, our data highlight the significance of the putative ribosomal binding site (5'-CGGAGG-3') upstream of *rep* in determining the fitness cost of *rep* expression. The significantly higher fitness cost of *rep* expression in PAO1::*p*LAC-*rep*, but not in PAO1::*p*LAC-*rep1*, is presumably due to the

essential role of this native ribosomal binding site in ensuring that *rep* mRNA transcripts are translated in the PAO1 host. Prior to this experiment, it was thought that the putative ribosomal binding site (5'-AGGAAA-3') [219], which is located between *p*LAC and the multiple cloning site in the original pUC18-mini-Tn7T-Gm-*LAC* delivery plasmid, should allow genes that are cloned downstream of *p*LAC to be translated. However, our results suggested otherwise, and for this reason, only the PAO1::*p*LAC-*rep* strain was used in all subsequent experiments.

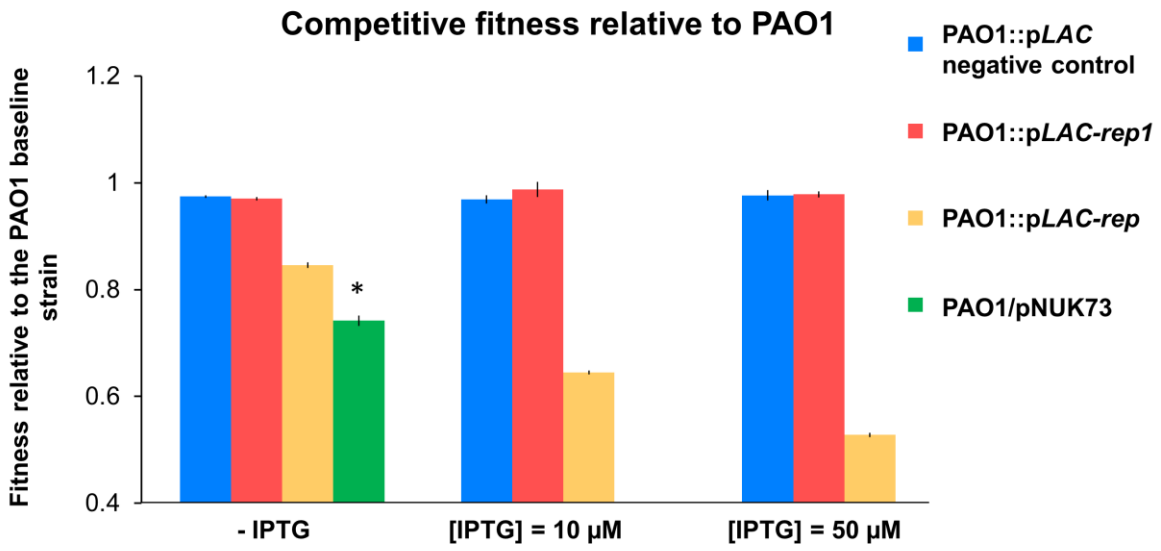


Figure 4.1: The competitive fitness of the PAO1::p*LAC* negative control strain (blue) and two mini-Tn7 insertion strains bearing a chromosomal copy of the *rep* gene (red and yellow) relative to the PAO1 baseline strain (mean ± s.e; n = 3). Error bars indicate standard error.

*The relative competitive fitness of the PAO1/pNUK73 strain (green) was quantified in the absence of IPTG only.

Competitive fitness data courtesy of Dr A San-Millan

rep* expression in PAO1/pNUK73 and IPTG-induced PAO1::*p*LAC-*rep

Using quantitative real-time PCR (qPCR), the relative expression levels ($\log_2$ fold-change) of the *rep* gene was determined in the PAO1/pNUK73 (initial population) and IPTG-induced PAO1::*p*LAC-*rep* strains relative to the leaky expression level of *rep* from *p*LAC in the uninduced PAO1::*p*LAC-*rep* reference strain. Figure 4.2 shows that the relative transcript levels of *rep* was higher in PAO1/pNUK73 than in IPTG-induced PAO1::*p*LAC-*rep* strains. When IPTG concentration was increased from 10 µM to 50 µM, the transcript levels of *rep* in the PAO1::*p*LAC-*rep* strain showed a statistically significant increase relative to the uninduced reference strain (two-tailed, two-sample *t*-test: $t = 3.32$, d.f = 3.72, p = 0.0327), which is consistent with our expectation that the relative transcript levels of *rep* should increase with the strength of *lac* promoter induction.

The relative expression level of *rep* in PAO1/pNUK73 is significantly higher compared to those in the IPTG-induced PAO1::*p*LAC-*rep* strains due to the difference in *rep* copy numbers. While there is a single copy of *rep* in the PAO1::*p*LAC-*rep* strain, the average copy number of pNUK73 in parental PAO1/pNUK73 is 11.0 (s.e = 1.89, n = 5) (San Millan *et al*, 2014). The average relative expression level of *rep* in PAO1/pNUK73 is 3.18 per plasmid copy (average $\log_2$ fold-change, propagated s.e = 0.55, n = 3), which is not statistically different from the average *rep* expression level in PAO1::*p*LAC-*rep* strains when induced with 50 µM IPTG (two-tailed, two-sample *t*-test: $t = 0.0837$, d.f = 2.17, p = 0.940).

However, the average fitness cost of pNUK73 carriage per plasmid copy (2.35%) is significantly smaller compared to that of PAO1::*p*LAC-*rep* strains induced using 50 µM IPTG (47.2%). We speculate that this may be explained by a subtle, but very important difference between these two strains. Direct binding between pNUK73 and its replication protein Rep is expected to buffer part of the fitness cost that is associated with *rep* expression. While the pNUK73 plasmid has binding sites for Rep, the PAO1::*p*LAC-*rep* strain has none. Therefore, the effective concentration of Rep is expected to be much higher in PAO1::*p*LAC-*rep* compared to that in PAO1/pNUK73 after the differences in *rep* gene copy number and transcriptional strength have been taken into account.

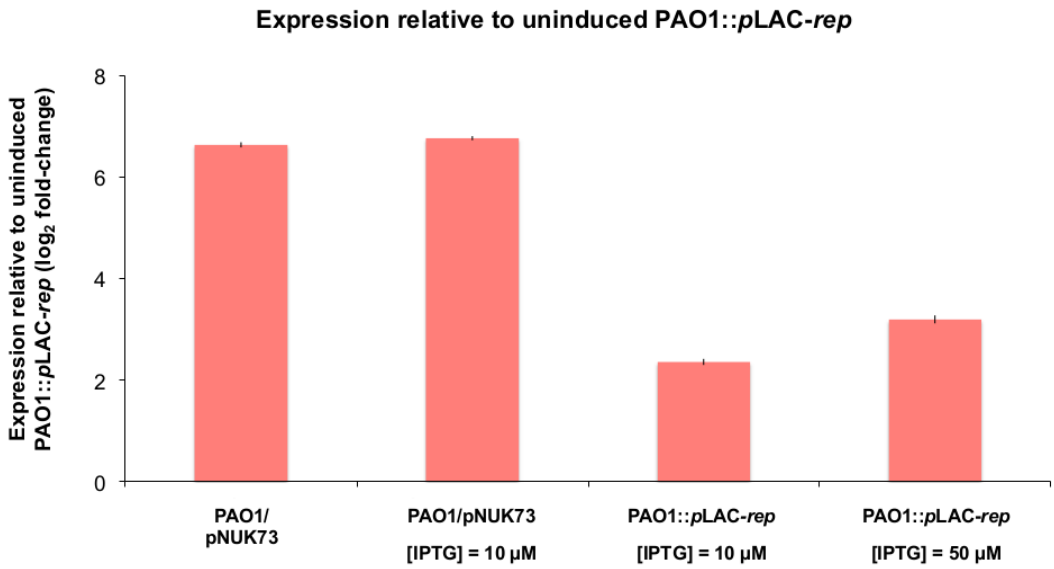


Figure 4.2: Transcript levels of *rep* (mean ± s.e; n = 3) expressed in $\log_2$ fold-change relative to the uninduced PAO1::*p*LAC-*rep* reference strain. Error bars indicate standard error.

Figure 4.1 shows that the competitive fitness of the uninduced PAO1::*p*LAC-*rep* strain was significantly lower compared to the PAO1::*p*LAC negative control strain. We hypothesized that this can be attributed to the leakiness of *p*LAC [220]. To demonstrate that leaky expression of *rep* can be observed in the uninduced PAO1::*p*LAC-*rep* strain, I performed absolute qPCR quantification to determine the copy numbers of *rep* and an internal reference gene *rpoD* in cDNA samples derived from this strain. *rpoD* is a constitutively essential gene which encodes the primary sigma factor of the exponential phase [221]. The copy number of *rep* transcripts due to the basal *rep* expression was (7.35 ± 0.86)% that of *rpoD* per unit volume of cDNA samples (n = 3). We can infer that this level of *rep* leaky expression was sufficient to elicit a fitness cost of 15.4% in the uninduced PAO1::*p*LAC-*rep* strain (Figure 4.1).

Expression of SOS response genes can be observed in conjunction with *rep* expression

Using qPCR, the relative expression levels ($\log_2$ fold-change) of two SOS response genes, *lexA* and *recA*, were determined in all the strains shown in Table 4.3 relative to the PAO1::*p*LAC negative control strain. Consistent with RNA-Seq results that were obtained in the first part of this work (Appendix A), the results in Figure 4.3 demonstrate that *lexA* and *recA* were upregulated in PAO1/pNUK73 relative to the *rep*-free negative control strain. Similarly, *lexA* and *recA* were upregulated in the PAO1::*p*LAC-*rep* strains relative to the negative control strain. Inducing the PAO1::*p*LAC-*rep* strain with 10 µM IPTG led to a significant increase in the relative expression level of *lexA* compared to the same strain without IPTG induction (two-tailed, two-sample *t*-test: $t = 9.05$, d.f = 3.78, $p = 0.001$) and a non-significant increase in the average relative expression level of *recA*. However, increasing IPTG concentration from 10 to 50 µM did not result in a further increase to *lexA* and *recA* expression despite the increase in *rep* expression (Figure 4.2). This suggests that the SOS response may not be the only molecular mechanism through which *rep* expression generates its fitness cost.

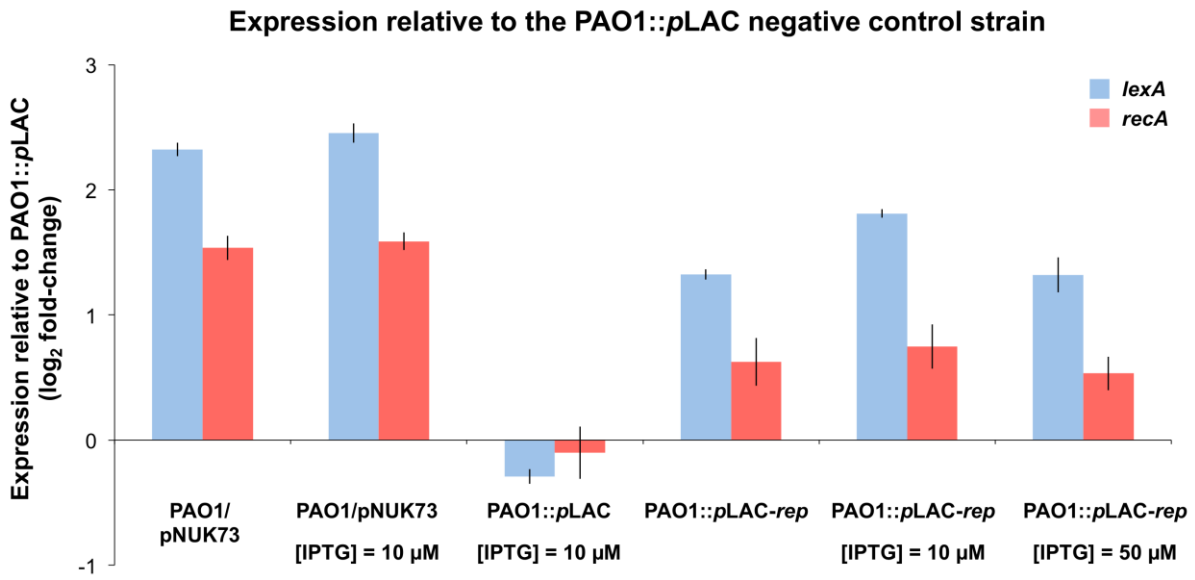


Figure 4.3: Transcript levels of *lexA* and *recA* (mean $\log_2$ fold-change ± s.e; n = 3) relative to the PAO1::*p*LAC negative control strain. Error bars indicate standard error.

Overall, the expression levels of *lexA* correlated significantly with that of *recA* in the seven strains (one-way ANOVA: p = 0.0007, $F_{1, 5}$ = 54.3, r^2 = 0.916), which reflect a similar pattern of gene expression for the two key SOS response regulator genes. At a phenotypic level, extensive cell filamentation was observed in strains that showed significant levels of reduction in fitness relative to the PAO1 baseline strain (refer to light microscopy images in Appendix A), which is a key phenotypic trait of bacterial cells whose growth rates have been adversely affected by the activation of the SOS response [138].

Conclusions

By combining these qPCR gene expression data with competitive fitness measurements and the RNA-Seq analysis, our results strongly support our hypothesis that *rep* expression drives the fitness cost in the PAO1 host strain. The causality of this relationship was established by the observation that in the presence of compensatory mutations in any one of the three recently acquired chromosomal genes, the relative expression level of *rep* decreased dramatically. In the presence of IPTG, the PAO1::*p*LAC-

rep strain showed highly significant upregulation of *rep* relative to its leaky expression level without IPTG. Amongst the many differentially expressed genes in PAO1/pNUK73 relative to PAO1, the activation of the SOS response is an important consequence of *rep* expression that has the potential to generate a direct fitness cost through phenotypic changes such as promoting cell filamentation. This inference is supported by the observation that both the IPTG-induced PAO1::*p*LAC-*rep* and PAO1/pNUK73 strains showed reduced competitive fitness (Figure 4.1) and increased expression of *lexA/recA* relative to the *rep*-free negative control strain (Figure 4.2). Moreover, our RNA-Seq analyses revealed that the expression levels of *lexA/recA* and all known LexA-regulated genes returned to wild-type levels in the presence of compensatory mutations.

In order to attribute the fitness cost directly to the activation of SOS response, future work should investigate the fitness effects of *rep* expression in an SOS-deficient mutant of the PAO1 wild-type strain. If *rep* expression in this SOS-deficient strain elicits no fitness cost, it can then be inferred that the activation of the SOS response is the direct mechanism through which *rep* expression generates the fitness cost within the PAO1 host. It should also be emphasized that at a mechanistic level, the molecular nature of the putative interactions between the three recently acquired chromosomal genes and the *rep* gene that gave rise to the extensively altered transcriptional response in PAO1/pNUK73, including the activation of the SOS response, remain unclear. Future research should focus on understanding the biochemical mechanisms that directly contribute to the observed intragenomic conflicts between these horizontally acquired genes, and whether our conclusions from this study can be further extrapolated to other models of interactions between bacterial hosts and plasmids.

Materials and Methods

Molecular cloning & transposon mutagenesis

The *rep* gene was PCR-amplified from pNUK73 plasmid DNA using primers shown in Table 4.1, with and without a putative ribosomal binding site (5'-CGGAGG-3') 5 bp upstream of the start codon

GTG for the *rep* gene. Each PCR product was cloned downstream of *p*LAC between the BamHI and HindIII restriction sites in the multiple cloning site in the pUC18-mini-Tn7T-Gm-*LAC* mini-Tn7 transposon delivery plasmid (accession no. AY599234) [144].

Primer	Oligonucleotide sequence (5'->3')	Restriction site	Description
*rep*_F	5'-TAAGGATCCTGG**CGGAGG**CGGCT-3'	PstI	With putative RBS
*rep*_R	5'- GCCAAGCTTTAACGTTCCCCTAACTT-3'	HindIII	
*rep1*_F	5'-TAAGGATCCGTGAAGGCACCGGATT-3'	PstI	Without putative RBS
*rep1*_R	5'- GACAAGCTTTAACGTTCCCCTAACTT-3'	HindIII	

Table 4.1: Oligonucleotide sequences of PCR primers used for molecular cloning, with restriction site sequences underlined. The putative ribosomal binding site (RBS) upstream of *rep* is shown in red.

To integrate the IPTG-inducible *rep* constructs into the chromosome of the plasmid-free PAO1 parental strain, sucrose-treated PAO1 was co-transformed with one of the two mini-Tn7 transposon delivery vectors that were constructed in this study and a Tn7-encoding transposase helper plasmid pUX-BF13 [145] via electroporation to generate the PAO1::*p*LAC-*rep* and PAO1::*p*LAC-*rep1* strains. Using the same procedure, a negative control strain (PAO1::*p*LAC) was generated, which is isogenic to the two mini-Tn7 transposon insertion strains bearing the *p*LAC-*rep* construct except for the absence of the *rep* gene downstream of *p*LAC.

Competitive fitness assays

The competitive fitness of the bacterial strains shown in Table 4.2 were determined relative to the PAO1 wild-type strain using competitive fitness assays by flow cytometry described in a previous work [215]. All strains were cultured in LB medium (BD Biosciences, USA) and all incubation steps were performed at 37°C with continuous shaking (225 rpm). Prior to each competitive fitness assay, IPTG was added to the LB growth medium at the concentrations indicated in Table 4.2 to allow for the induction of *p*LAC during the pre-culture stage. Three biological replicates of each strain were assayed.

Strain	Description	[IPTG]/μM
PAO1::*p*LAC	Negative control strain. Isogenic to the two mini-Tn7 *p*LAC-*rep* insertion strains, but lacks *rep* downstream of the *lac* operon promoter	0
		10
		50
PAO1::*p*LAC-*rep1*	PAO1 with an IPTG-inducible *rep* construct (without the putative RBS) downstream of *p*LAC at the mini-Tn7 chromosomal site	0
		10
		50
PAO1::*p*LAC-*rep*	PAO1 with an IPTG-inducible *rep* construct (with the putative RBS) downstream of *p*LAC at the mini-Tn7 chromosomal site	0
		10
		50
PAO1/pNUK73	Initial population of pNUK73-bearing PAO1	0

Table 4.2: Bacterial strains and concentration of IPTG for competitive fitness assays

Relative quantitative real-time PCR

Three biological replicates of total RNA samples were extracted from the bacterial strains shown in Table 4.3. All strains were inoculated in LB medium and pre-cultured overnight at 37°C with continuous shaking (225 rpm). IPTG concentrations of 10 μM and 50 μM were added to LB medium to induce the expression of *rep* in the PAO1::*p*LAC-*rep* strain. Detailed procedures for the RNA extraction can be found in Appendix A.

Strain	Description	[IPTG]/μM
PAO1::*p*LAC	Plasmid-free negative control strain. Isogenic to PAO1::*p*LAC-*rep*, but lacks *rep* downstream of the *lac* operon promoter	0
		10
PAO1/pNUK73	Initial population of the pNUK73-bearing PAO1 strain	0
		10
PAO1::*p*LAC-*rep*	Plasmid-free PAO1 strain with an IPTG-inducible *rep* construct downstream of the *lac* operon promoter inserted into the mini-Tn7 chromosomal site	0
		10
		50

Table 4.3: Bacterial strains and concentration of IPTG for quantitative real-time PCR (qPCR)

cDNA was synthesized using the GoScript Reverse Transcription System (Promega, USA). Quantitative real-time PCR (qPCR) was performed using the relative quantification method. Each gene shown in Table 4.4 was amplified using the Fast SYBR Green Master Mix (Applied Biosystems,

USA) and 100 nM oligonucleotide primers on the StepOnePlus Real-time PCR platform (Applied Biosystems, USA). Normalization factors were calculated from the geometric mean of the uncorrected transcript levels of three internal reference genes *acpP*, *atpA* and *rpoD*. The transcript levels of the target genes (*lexA*, *recA* and *rep*) was then determined by normalizing the transcript levels of every target genes to the normalization factors. Finally, the transcript levels of the target genes was normalized to that of the control group within the same batch of biological replicates. The transcript levels ($\log_2$ fold-change) of *lexA* and *recA* were determined relative to the PAO1::*p*LAC control strain, while the transcript levels of *rep* was normalized to that in the uninduced PAO1::*p*LAC-*rep* reference strain.

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')	Reference
acpP	ACTCGGCGTGAAGGAAGAAG	CGACGGTGTCAAGGGAGT	[168]
rpoD	CAGGTGGCGTAGGTGGAGAA	GGGCGAAGAAGGAAATGGTC	[167]
atpA	TACGCCAGATCACGCCCGGA	GGCGGCGTCTACGGTATGGC	This study
lexA	ATCGGAGCGAACTCAGGGT	CCGTATCAATCCCGCCTTCT	
recA	TTCGCTACCCACCACCTCGT	CAAGAACGCCAACTGCCTGG	
rep	AGCAGAGCGACGAGGAAGTG	ATAGCCACCCAGCCGCCATT	

Table 4.4: Oligonucleotide sequences of qPCR primers

Absolute quantitative real-time PCR

To quantify the extent of leaky expression of *rep* from *p*LAC in the absence of IPTG, we measured the copy number of *rep* and *rpoD* transcripts in cDNA samples derived from the uninduced PAO1::*p*LAC-*rep* reference strain using absolute qPCR quantification. Calibration standard curves for both genes were generated using genomic DNA extracted from the PAO1::*p*LAC-*rep* strain, whose DNA concentration was determined using the QuantiFluor dsDNA System (Promega, USA). The copy number of *rep* and *rpoD* were calculated based on the gDNA standard curves and the mass of gDNA per genome.

5. General Discussions

Thesis summary

The main objective of my doctoral research is to bridge part of the existing gap in our understanding of the molecular mechanisms that contribute to variations in the fitness cost associated with chromosomal antibiotic resistance mutations. Using rifampicin resistance in *P. aeruginosa* as a model, I applied a systems-level approach to dissect the molecular mechanisms that contribute to the variations in the cost of rifampicin resistance. A secondary research theme centers on exploring how the fitness cost influences the evolutionary dynamics of rifampicin resistant mutants in an antibiotic-free environment *in vitro*.

In Chapter 2, I performed phenotypic assays of *rpoB* mutants carrying different fitness costs and demonstrated that most of the variation (76.2%) in the fitness cost can be attributed to the direct effects of *rpoB* mutations on transcriptional efficiency. Gene expression analyses showed that even though rifampicin resistance mutations have variable impacts on the expression levels of essential genes, these variations do not generate a significant fitness cost in their own right. In addition, I demonstrated that at a genome-wide level, changes in the transcriptional profiles of the *rpoB* mutants are very subtle relative to the rifampicin-sensitive ancestral strain, suggesting that resistant mutants show robustness against altered transcriptional efficiency. In support of this argument, the upregulation of the *rpoBC* operon was found to be one of the most profound impacts of *rpoB* mutations, which may have contributed to the observed robustness in the transcriptome profiles of resistant mutants by increasing the abundance of RNAP core enzyme.

In Chapter 3 of this thesis, I showed that *rpoB* mutants of *P. aeruginosa* adapted to the cost of resistance during an experimental evolution study during which they were serially passaged in an antibiotic-free environment. Whole-genome sequencing of endpoint populations revealed an

interesting dichotomy in the mechanisms of adaptive evolution that was strongly influenced by the fitness cost of resistance. Resistant mutants with low fitness costs tend to undergo general adaptation. A high degree of parallel mutation was observed, particularly with regard to non-synonymous mutations in the transcriptional regulator gene *lasR* that are thought to give rise to a quorum-sensing deficient phenotype in mutants. Resistant mutants with high fitness costs are more likely to fix compensatory mutations in genes that encode the three subunits of the RNAP core enzyme (*rpoA, rpoB* and *rpoC).* I presented phylogenetic evidence to support the idea that general adaptation limits compensatory adaptation in resistant mutants as a consequence of clonal interference.

In Chapter 4, I worked in collaboration with three co-authors to dissect the molecular underpinnings of the fitness cost associated with pNUK73 small plasmid carriage in *P. aeruginosa*. We demonstrated the fitness cost associated with pNUK73-carriage is primarily due to the expression of *rep*, which code for the pNUK73 plasmid replication protein Rep, and is not due to the expression of the other two pNUK73-encoded genes. The activation of the SOS response can be observed in conjunction with *rep* expression, and is likely be an important molecular mechanism through which *rep* expression generates a fitness cost. Compensatory mutations in three recently acquired chromosomal genes in the PAO1 genome silenced both *rep* and SOS response genes, which demonstrated the causality of the relationship between *rep* expression level and competitive fitness. On a broader scale, our work presented in Appendix A showed that intragenomic conflict between recently acquired chromosomal genes and the *rep* gene is the source of the fitness cost in our PAO1/pNUK73 model.

Challenges of linking alterations in transcriptome profiles of *rpoB* mutants to fitness cost

Since the variations in transcriptional efficiency cannot account for all the variations in the competitive fitness of *rpoB* mutants, at least a part of the residual fitness cost of rifampicin resistance is expected to stem from the pleiotropic effects of *rpoB* mutations. The objective of performing genome-wide gene expression analysis in *rpoB* mutants was to identify specific metabolic, regulatory

or stress response networks that are highly affected by the pleiotropic effects of *rpoB* mutations. It was anticipated that at least some of these gene and cellular pathways could have the potential to influence mutant fitness when their expression levels are significantly perturbed relative to wild-type levels. However, linking such alterations in the transcriptional response to the fitness cost of resistance proved highly challenging in the context of this study.

When gene expression analysis reveals that a particular gene is differentially expressed in rifampicin resistant mutants, it can be difficult to decipher if the altered gene expression pattern is a cause or a consequence of the fitness cost of resistance. In our RNA-Seq dataset for *rpoB* mutants D521G and H531R (Annex B, Supplementary Materials, Chapter 2), it was found that the downregulation of the denitrification operon is one of the most striking pleiotropic effects on the transcriptome profiles of *rpoB* mutants relative to their rifampicin-sensitive ancestral strain. In Appendix B, I hypothesized that the reduced expression of the denitrification genes contributed to the fitness cost of resistance. To test this hypothesis, I cloned an additional copy of selected denitrification genes under the control of an IPTG-inducible promoter into a neutral chromosomal site in the genome of the two sequenced *rpoB* mutants and their rifampicin-sensitive ancestral strain. If the downregulation of these denitrification genes has a deleterious effect on the fitness of the *rpoB* mutants, increasing their expression levels is expected to offset a part of the fitness costs. However, my results indicated that the presence of an additional copy of these denitrification genes had no statistically significant effects on the fitness of the resistant mutants. Similar to the genome-wide changes that were observed in the expression levels of essential genes (Chapter 2), the altered expression of the denitrification operon could be a consequence of alterations in RNAP activity in the *rpoB* mutants and does not generate an additional fitness cost in its own right. Since the gene products of the denitrification pathway are required for energy production via anaerobic respiration, it is also possible that the differences in the expression levels of these denitrification genes merely reflect the differences in growth rate and energy requirement between the *rpoB* mutants and the ancestral strain. Likewise, the significant downregulation of virulence genes *lasA* and *aprA* observed in both *rpoB* mutants is likely to be a

consequence of reduced growth rate, which resulted in the reduced expression of these quorum sensing regulated genes.

Potential contributions of environmental factors to the fitness cost of rifampicin resistance

All experiments including competitive fitness assays, luciferase-based phenotypic measurements of transcriptional efficiency, as well as gene expression analyses by qPCR and RNA-Seq were carried out under aerobic conditions at 30˚C in the nutrient-rich M9KB liquid medium [65]. However, it is known that the fitness cost of resistance strongly depends on the environment in which it is measured [15, 24, 31, 32]. Therefore, the conclusions we have drawn with regard to the molecular mechanisms underpinning the fitness cost of rifampicin resistance are only valid under the specified experimental conditions.

Despite the deleterious effects of rifampicin resistance mutations on transcriptional efficiency and *in vitro* fitness in the context of my research, recent studies have shown that *rpoB* mutations can be beneficial under certain specific environmental conditions in the absence of rifampicin. In aging colonies of *S. enterica* and *E. coli* grown on solid medium, it was observed that the frequency of rifampicin resistant mutants increased relative to parental wild-type cells, which was a consequence of selection for *rpoB* mutations rather than stress-induced mutagenesis [222]. *rpoB* mutants enjoy a fitness advantage relative to the wild-type strain in these aging colonies and grow at a faster rate before the onset of growth limitation during the stationary phase [222]. In a laboratory experimental evolution experiment, Rodríguez-Verdugo *et al* propagated *E. coli* that had never been exposed to antibiotics for 2000 generations under thermal stress at 42.2˚C. Amongst the evolved clones, various parallel mutations in *rpoB* that conferred resistance to rifampicin were identified that are beneficial at the elevated temperature of 42.2˚C in glucose-limited medium, but had different fitness effects in other genetic backgrounds and experimental environments [70]. These results reinforce the idea that the fitness costs of rifampicin resistance mutations strongly depend on the environment under which the *in vitro* measurements of competitive fitness are taken. Mechanistic studies should be undertaken

to investigate why compromised RNAP activity due to *rpoB* mutations do not translate into fitness costs under these stressful growth conditions, but confer a selective advantage instead.

A recent meta-analysis suggests that *in vitro* estimates of fitness of chromosomal antibiotic resistant mutants show strong correlation with *in vivo* measurements of fitness in animal hosts [141]. However, there are also examples where discrepancies exist between fitness measurements obtained from *in vivo* and *in vitro* systems [15, 223]. There are many possible explanations for such discrepancies. In the context of this study for example, our RNA-Seq transcriptomic data for *rpoB* mutants grown *in vitro* cannot capture the transcriptional response of these mutants in an *in vivo* growth environment, where complex environmental factors, such as host immune response, are expected to exert influences on the gene expression profiles of bacterial pathogens [224]. In a clinical setting, administering multiple antibiotics during treatment can alter both the transcriptional response [224, 225] and cellular physiology [1, 226] amongst resistant bacterial populations. These confounding factors may further limit the applicability of our findings in a medical treatment context. To overcome these limitations, key measurements in Chapter 2 and the *in vitro* selection experiment from Chapter 3 could be performed in suitable *in vivo* models to ascertain if the main findings still hold true.

Potential avenues for future research

In Chapter 2, our gene expression analyses by qPCR and RNA-Seq were performed using RNA samples extracted during the exponential phase of planktonic growth. To extend these gene expression analyses, transcriptome profiling should be performed using RNA samples extracted from stationary phase cultures and static biofilm samples of *rpoB* mutants. During the exponential phase, the principal sigma factor RpoD is required for the transcription of a large number of housekeeping genes, whereas during the stationary phase, other sigma factors such as RpoS and RpoN play an important role in regulating numerous cellular processes such as quorum sensing, virulence and biofilm formation [227]. Given our experimental method for performing competitive fitness assays (refer to Materials & Methods, Chapter 2), all competitive fitness measurements obtained in this study

should reflect the combined fitness cost of rifampicin resistance during the exponential and stationary phases of planktonic growth. Analyzing the differences in the transcriptome profiles of *rpoB* mutants and the ancestral strain during the stationary phase can potentially help us unravel new candidate genes whose differential expression may have contributed to the cost of resistance. For instance, while there is no evidence from our current RNA-Seq transcriptome data to suggest that the ppGpp-mediated stringent response is activated in the two sequenced *rpoB* mutants during the exponential phase of growth, there remains the possibility that this global stress response might only be activated when *rpoB* mutants reach the stationary phase. The ultimate goal is to explore the possibility of whether these differences in transcriptome profiles can be exploited to maximize the cost of rifampicin resistance in a more clinically relevant context.

In a comparative transcriptome analysis, Dötsch *et al* (2012) demonstrated that the transcriptome profiles of static biofilm cultures of the *P. aeruginosa* PAO1 wild-type strain and planktonic cultures of the same strain obtained from the late stationary phase showed both similarities and differences, which suggests that biofilms are not just surfaced attached stationary phase cells [228]. Therefore, it is conceivable that the transcriptome profiles of *rpoB* mutants extracted from static biofilm samples will also show similarities and differences compared to those extracted during the exponential and stationary phases of planktonic growth. Characterizing the transcriptome profiles of *rpoB* mutants in biofilms may prove highly useful from a clinical perspective. *P. aeruginosa* biofilms are often found in the lungs of cystic fibrosis patients who suffer from acute and chronic infections [229]. Recent studies have demonstrated that administering a mixture of rifampicin and fusidic acid can successfully decolonize methicillin resistant *Staphylococcus aureus* (MRSA) from the lungs of cystic fibrosis patients [230, 231]. At the same time, it is conceivable that this treatment strategy can lead to positive selection for rifampicin resistance in *P. aeruginosa* biofilms. Analyzing the transcriptome profiles of *P. aeruginosa rpoB* mutants found in biofilms can potentially help clinicians devise adjunctive treatment strategies to minimize selection for rifampicin resistance.

A reduction in the transcriptional fidelity of RNAP that eventually result in erroneous protein synthesis has the potential to generate a fitness cost in prokaryotes and eukaryotes alike [232]. Future

computational analysis of our existing RNA-Seq dataset should be extended to examine if mutant RNAP in *rpoB* mutants show reduced transcriptional fidelity compared to wild-type RNAP. Conceptually, this can be accomplished by quantifying the transcriptional error rates in the two sequenced *rpoB* mutants relative to the rifampicin-sensitive ancestral strain at a genome-wide scale.

In this work, I demonstrated that an increase in the transcriptional read-through of P26, a Rho-independent partial terminator in the intergenic region upstream of *rpoB,* contributes to the upregulation of the *rpoBC* operon amongst *rpoB* resistant. However, the question remains as to whether such altered transcriptional read-through of partial terminators is a phenomenon that is specific to P26. It is possible that altered RNAP activity could have variable effects on the transcriptional read-through of other Rho-independent and Rho-dependent partial terminations in the genome that can affect the expression levels of downstream genes. Such global alterations in partial terminator read-through could also have contributed to the robustness that has been observed in the transcriptome profiles of *rpoB* mutants D521G and H531R. Future work should focus on the quantification of potential alterations in transcriptional read-though at a genome-wide scale using our existing RNA-Seq dataset from Chapter 2.

At a broader level, it would be interesting to investigate if the robustness mechanism observed in the transcriptome of rifampicin resistant mutants against compromised transcriptional efficiency is a phenomenon that is unique to the rifampicin resistance model. Understanding the molecular underpinnings of the fitness cost of resistance could potentially allow for the development of therapeutics with a view to exploiting the genetic weaknesses of antibiotic resistant mutants. However, if antibiotic resistant mutants show intrinsic robustness against the deleterious effects of resistance mutations on the gene products that are affected by these mutations, searching for suitable molecular targets to maximize the cost of resistance can become significantly more challenging. Therefore, this phenomenon should be investigated in other models of chromosomal antibiotic resistance such as mutations that are responsible for resistance to fusidic acid and peptide deformylase inhibitors (PDFIs). Since both types of mutations have deleterious effects on protein translation rate in

resistant mutants [36, 37], quantitative high-throughput proteomic profiling of resistant mutants is expected to unravel any potential robustness against compromised translation rate.

Technical considerations from a molecular biology perspective

In Chapter 2 of this thesis, our gene expression analyses by qPCR revealed that the expression levels of essential genes in the *rpoB* mutants showed small fold-changes with respect to the ancestral strain. This was further validated by our RNA-Seq data for two *rpoB* mutants (D521G and H531R), which demonstrated that the average relative expression levels for essential gene are more robust against compromised transcriptional efficiency compared to those for non-essential genes. Given the overall subtle changes in the relative expression of the essential genes in *rpoB* mutants, it is imperative that internal reference genes should be evaluated and selected using a rigorous methodology [233]. The objective is to ensure that no bias is introduced when the transcript levels of essential genes is normalized to that of internal reference genes within each *rpoB* mutant sample group, such that subtle changes in relative gene expression can be reliably quantified relative to the ancestral reference strain. To overcome this technical challenge, I verified the expression stability of six candidate internal reference genes using two published algorithms – GeNorm [149] and NormFinder [150]. This allowed for the selection of 3-4 most stably expressed candidate genes, which were collectively used as internal reference genes during the normalization procedure. The strong correlation between relative gene transcript levels measured using qPCR and RNA-Seq (Figure S2.2, Supplementary Materials, Chapter 2) demonstrates the success of this normalization strategy, which allowed me to apply the same experimental procedure to our qPCR-based gene expression analysis in Chapter 4.

In Chapter 4 and Appendix B of this thesis, I cloned several genes downstream of an IPTG-inducible *lac* promoter and inserted these IPTG-inducible gene constructs into the mini-Tn7 site of the PAO1 genome. The objective was to use different concentrations of IPTG to modulate the transcriptional levels of candidate genes found immediately downstream of the *lac* operon promoter. In Chapter 4, I demonstrated that the leaky expression of *rep* from the *lac* operon promoter (*p*LAC) in the absence of

IPTG was sufficient to generate a significant fitness cost in the PAO1::*p*LAC-*rep* strain relative to the PAO1::*p*LAC negative control strain. In view of the leakiness of *p*LAC, a more tightly regulated inducible promoter that results in a lower expression level of the downstream gene in the absence of the inducer compound would have been more desirable. An example of a tightly regulated inducible promoter that fulfills these criteria is the L-arabinose inducible arabinose operon promoter (*p*BAD), which has a significantly higher induction/repression ratio than *p*LAC [234]. In the context of our qPCR experiments in Chapter 4, replacing *p*LAC with *p*BAD would have allowed us to quantify the fitness effects stemming from a greater range of *rep* expression levels.

6. General Conclusions

During my doctoral research, I demonstrated that the variations in the *in vitro* fitness cost of rifampicin resistance mainly depend on the direct effects of *rpoB* mutants on transcriptional efficiency in the opportunistic pathogen *P. aeruginosa* in a nutrient-rich culture medium. *rpoB* mutants showed widespread, but subtle alterations in transcript levels of essential genes relative to ancestral strain levels. However, such alterations in gene expression levels do not contribute to the fitness cost of resistance. At a genome-wide level, there are no obvious metabolic, regulatory or stress response networks whose altered expression is likely to have given rise to the fitness cost, at least in the two sequenced *rpoB* mutants. Interestingly, genome-wide gene expression levels correlate highly with ancestral strain levels, which suggest that *rpoB* mutants are robust against the deleterious effects of *rpoB* mutations on transcriptional efficiency. From a clinical perspective, the absence of "smoking guns" in the transcriptome profiles of *rpoB* mutants and the robustness of *rpoB* mutants against compromised RNAP activity present challenges in our effort to exploit the genetic weaknesses of *rpoB* mutants and maximize the cost of resistance. However, it should be emphasized that the validity of these observations are limited given the specific conditions of the *in vitro* experiments of this study.

By adopting an experimental evolution approach, I demonstrated that general adaptation can limit the evolution of compensatory adaptation amongst *rpoB* mutants that carry low fitness costs due to clonal interference. My research challenges the conventional view that compensatory adaptation is a common mechanism of adaptation amongst antibiotic resistance populations. Rather than the norm, compensatory adaptation is more likely to be found in resistant mutant lineages that were initiated by high cost resistant isolates. These results have implications for our understanding of the evolutionary dynamics of antibiotic resistance and the interpretation of whole-genome sequence data from clinical bacterial pathogens.

Using the integrative experimental framework I developed in this thesis, I contributed to the team effort in dissecting the fitness cost associated with carriage of a small plasmid (pNUK73) in *P. aeruginosa.* Our results demonstrated that expression of the plasmid replication protein gene was

responsible for generating the fitness cost of resistance as a consequence of intragenomic conflicts with three recently acquired chromosomal genes. This demonstrates the usefulness of this integrative experimental framework for dissecting the fitness cost of antibiotic resistance in a different biological context from the rifampicin resistance model.

7. Appendices

Appendix A: Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*

Alvaro San Millan*†, Macarena Toll-Riera†, Qin Qi and R Craig MacLean
Department of Zoology, University of Oxford, United Kingdom

† These authors contributed equally to this study
*Email: alvaro.sanmillan@zoo.ox.ac.uk

Abstract

Horizontal gene transfer (HGT) plays a central role in the evolution of microbes by granting them access to an arsenal of adaptive traits. Despite the benefits associated with HGT, newly acquired genetic material can potentially represent a source of intragenomic conflict by interfering with the established genetic networks of the cell and producing deleterious fitness effects known as the "cost of HGT". The balance between the benefits and costs of HGT will determine the probability of fixation of the newly acquired DNA. In this work, we used *Pseudomonas aeruginosa* PAO1 and the small plasmid pNUK73 as a model system to investigate the molecular basis of the cost of HGT. We found that the high fitness cost produced by pNUK73 is linked to an important alteration of the transcriptional profile in PAO1, including the activation of the SOS response. In contrast, experimentally evolved strains carrying single compensatory mutations in the host chromosome recovered the original transcriptional profile. We showed that high-level expression of the pNUK73 replication protein gene *rep* was responsible for driving the cost of plasmid carriage and activating the SOS response. Crucially, we found that the cost of pNUK73 stems from interactions between the plasmid and three chromosomal genes with a strong signature of recent horizontal acquisition in *P. aeruginosa* PAO1. These results provide new insights into the evolutionary and molecular mechanisms responsible for the cost of HGT and suggest that interactions between mobile genetic elements may play an important role in the evolution of bacteria via HGT.

Introduction

Horizontal gene transfer (HGT) is a key source of genetic diversity in bacteria, contributing to the adaptation and diversification of prokaryotes [235, 236]. One recent and particularly concerning example of the ability of HGT to promote innovation in bacteria is the vast spread of antibiotic resistance among clinical pathogens [237, 238]. However, the introduction of novel genes in a pre-existing, well-tuned genetic background is also a source of conflict [126]. Recently acquired DNA usually produces a fitness cost in the host bacteria [141], but the origins and molecular mechanisms of these costs are poorly understood. The cost of HGT can arise from direct and indirect effects of the acquired DNA. Direct effects are, for example, the cost of the transfer process itself [239], the potential disruption of the bacterial genome by the integration of foreign DNA [240], and the metabolic costs associated with the replication and especially with the expression of the newly acquired genes [211, 212]. The indirect effects arise mainly from the interaction between the proteins encoded by the mobile genetic element (MGE) and the host [241, 242]. These interactions can lead to the disturbance of cellular networks [242, 243] and cytotoxic effects [244]. The costs of HGT will play a key role in the chances of establishment of newly acquired DNA in the host genome and in the bacterial population. Therefore, in order to understand the limits to evolution by HGT, it is crucial to understand the molecular basis of the cost of HGT.

In a recent work (San Millan *el al*, 2014), we investigated the compensatory adaptation between a small plasmid, pNUK73, and *Pseudomonas aeruginosa* PAO1. The newly acquired plasmid produced a large fitness cost in the naïve bacterial host. However, after 300 generations of coexistence, PAO1 completely compensated for the cost of pNUK73 carriage. Whole-genome sequencing showed that mutations in any of three specific genes were responsible for the observed compensation adaptation. The chromosomal genes that acquired compensatory mutations include a putative helicase carrying an UvrD-like helicase C-terminal domain (PA1372) and two contiguous putative serine/threonine protein kinases (PA4673.15 and PA4673.16) [101]. The modifications in these genes produced frameshifts and non-synonymous mutations, suggesting that the putative inactivation of any of these proteins was responsible for the compensatory adaptation to pNUK73 (Figure 1).

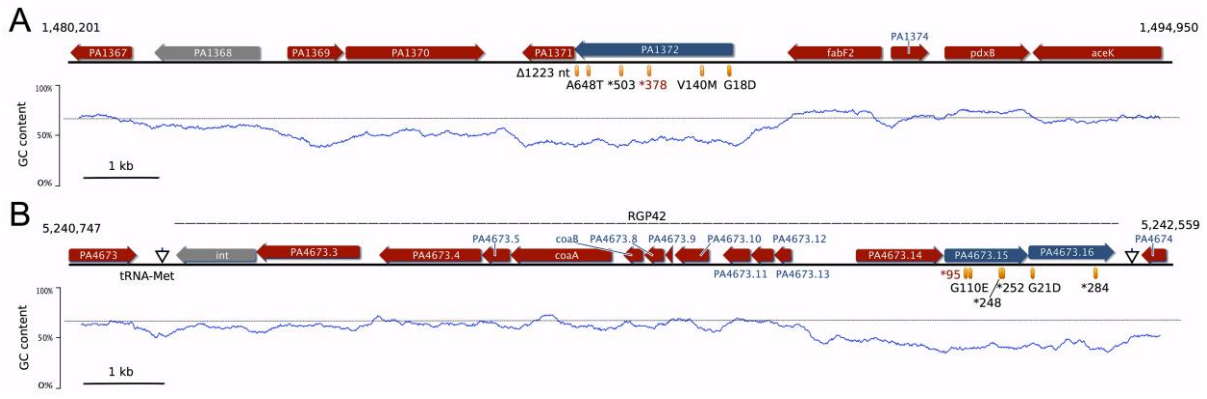


Figure 1: Genetic environment and mutations in PA1372 and PA4673.15-16

The schematic diagrams show the genetic environment of genes PA1372 (Panel A) and PA4673.15-16 (Panel B). Mutations in these genes are responsible for compensatory adaptation to plasmid pNUK73 in *P. aeruginosa* PAO1. The reading frames for genes are shown as arrows, with the direction of transcription indicated by the arrowhead. PA1372 and PA4673.15-16 are shown in blue. Genes involved in genetic transposition or integration are shown in grey. PA1368 is a transposase belonging to the IS4 family and PA4673.2 (*int*) is the RGP42 phage integrase. The rest of the genes are shown in red. The chart below the sequence represents GC content (blue line), with the black dotted line indicating the median GC content of the PAO1 genome (67%). The dashed line above the sequence in panel B indicates the position of the prophage RGP42 with the arrows pointing at the insertion site next to a tRNA-Met coding sequence. Yellow ellipses indicate the positions of the mutations responsible for compensatory adaptation to pNUK73 and the changes in the predicted protein sequence are described below them (an asterisk indicates a premature stop codon). The clones used in this study are those with the mutations highlighted in red. The numbers in the left and right above the sequences indicate the genomic location in the *P. aeruginosa* PAO1 genome (NCBI taxonomy ID: 208964).

In this study, we used the same tractable model system to investigate the molecular basis of the fitness costs associated with HGT and how these costs can be alleviated by compensatory adaptation. Using RNA-Seq, we examined changes in the global transcriptional profile of the PAO1 host as a consequence of carrying the pNUK73 plasmid. We found that this plasmid produces a highly

significant impact in the transcriptional profile of the host, including the induction of the SOS response via the expression of the plasmid replication protein Rep. The previously identified compensatory mutations in the putative helicase and kinases reduced the expression of *rep* and restored the original transcriptional profile. Interestingly, we found that the putative helicase and kinases responsible for the increased expression of *rep* are recently acquired genes in *P. aeruginosa*. Therefore, our results suggest that the interference between horizontally acquired elements drives most of the cost in the PAO1/pNUK73 system. We propose that this type of costly interactions between recently acquired MGE may be a general phenomenon in prokaryotes.

Results

Using RNA-Seq to dissect the cost of pNUK73

As a first approach to understanding the molecular basis of the cost produced by pNUK73 in PAO1, we used RNA-Seq to analyse the genome-wide transcriptional profiles of PAO1, PAO1/pNUK73 and two plasmid-bearing strains carrying compensatory mutations (Figure 1 & Supplementary Table 1). The mutant clones carried no other modifications compared to PAO1/pNUK73 apart from the compensatory mutation (San Millan *et al*, 2014). We found that pNUK73 significantly altered the expression of 749 genes in the host genome (Benjamini-Hochberg corrected p-value < 0.05), which represent 13% of the total number of genes in PAO1 (Figure 2A, Supplementary File 1). Gene set enrichment analyses (see Methods) showed that upregulated genes are mainly enriched in translation, ribosomal structure and biogenesis, while downregulated genes are enriched in metabolism (Supplementary File 2).

Interestingly, we found that plasmid-bearing strains carrying mutations in both the helicase (PA1372 with premature stop codon at position 378, Figure 1) or the kinase (PA4673.15 with premature stop codon at position 95, Figure 1) genes showed transcriptional profiles that were almost identical to those of plasmid-free PAO1 with a few very minor exceptions (Figure 2A). In the helicase mutant, PA2026 and PA3877, which encode a putative Na^+-dependent transporter and a nitrite extrusion protein respectively, were significantly downregulated (Supplementary File 1). Only one gene

(PA4673.16) was significantly downregulated in the plasmid-bearing kinase mutant clone. PA4673.16 encodes a second putative kinase that can also be mutated to compensate for the cost of pNUK73 carriage (Figure 1). PA4673.15 itself was also underexpressed, albeit non-significantly, suggesting that PA4673.15 regulates the expression of both kinases.

The pNUK73 small plasmid carries the backbone of the natural cryptic plasmid pNI10 isolated from *Pseudomonas fulva* [127] and codes for three genes, namely the Rep plasmid replication protein gene (*rep*), an aminoglycoside 3'-phosphotransferase gene (*aphA1*, which confers resistance to kanamycin and neomycin) and a *lacZ* reporter gene (Figure 2B). We compared the expression levels of these three genes between the parental PAO1/pNUK73 and the plasmid-bearing compensated mutants (Figure 2C). In both compensated mutants the expression levels of *aphA1* remained unchanged, while *lacZ* was slightly overexpressed (ca. 2.1-fold in both clones: helicase mutant, $p = 0.0002$; kinase mutant $p = 0.0004$). In contrast, compensatory mutations had dramatic effects on *rep* expression, which decreased 11.1-fold ($p < 10^{-28}$) and 7.4-fold ($p < 10^{-27}$) in the helicase and kinase mutant respectively relative to parental PAO1/pNUK73 (Figure 2B).

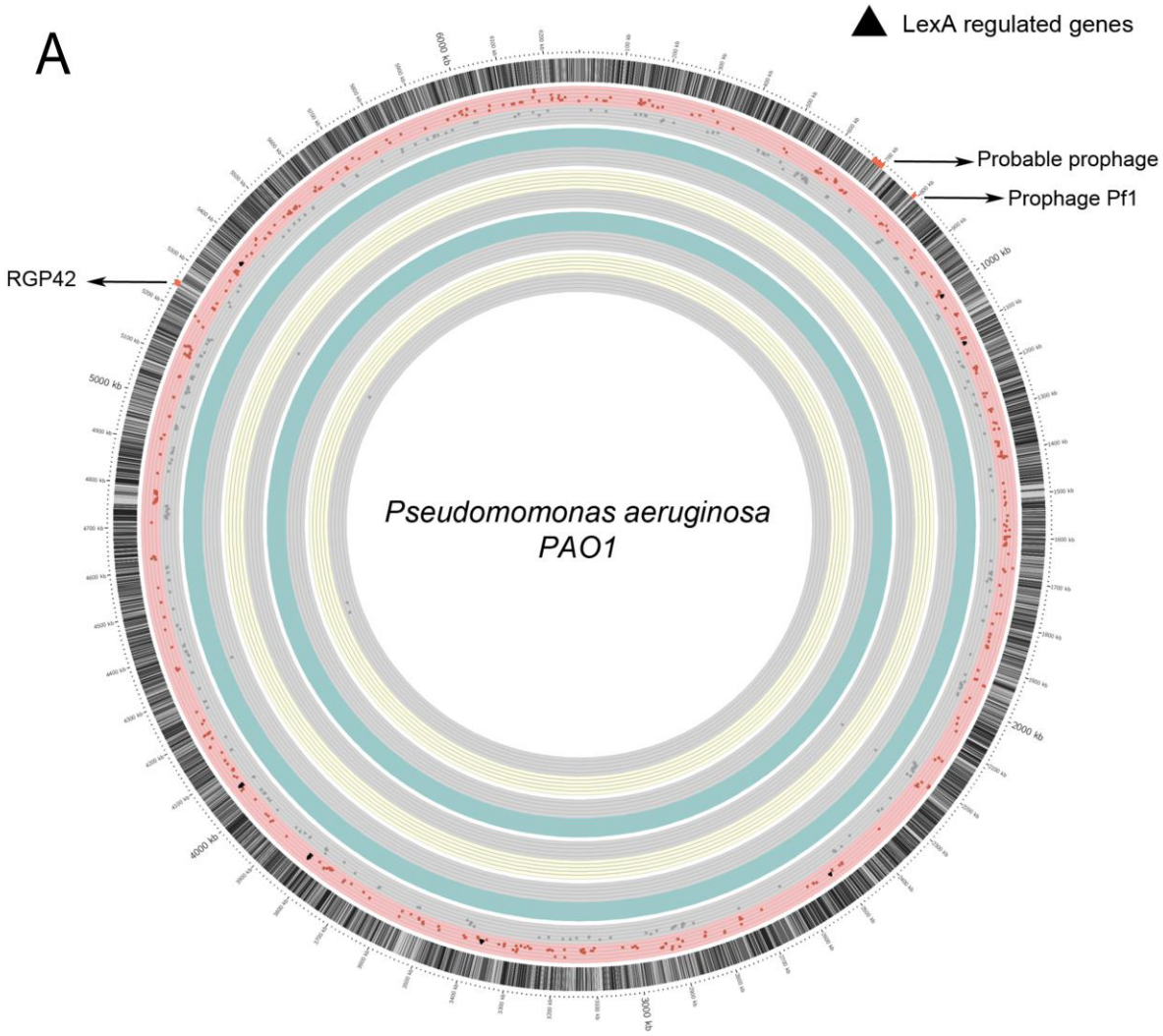


Figure 2: Gene expression map of the different strains analyzed using RNA-Seq

Panel A: Circos plot for the *P. aeruginosa* PAO1 genome. The outermost circle represents the GC content (light grey: GC content from 0% to 60%, grey: from 62% to 69%, dark grey: from 70% to 100%). From outside to inside, the 2nd to the 6th circles illustrate the changes in the transcriptional profiles in the following strains relative to wild-type PAO1. Within each circle, a scatterplot is drawn, with lines representing the degree of fold-change in $\log_2$. Each dot in the scatterplot represents a differentially expressed gene for each particular comparison. Triangles represent LexA-regulated genes.

- 2nd circle: parental PAO1/pNUK73 (red: up-, grey: downregulation)
- 3rd circle: helicase mutant/pNUK73 (green: up-, grey: downregulation)
- 4th circle: kinase mutant/pNUK73 (yellow: up-, grey: downregulation)
- 5th circle: plasmid-free helicase mutant (green: up-, grey: downregulation)
- 6th circle: plasmid-free kinase mutant (yellow: up-, grey: downregulation)

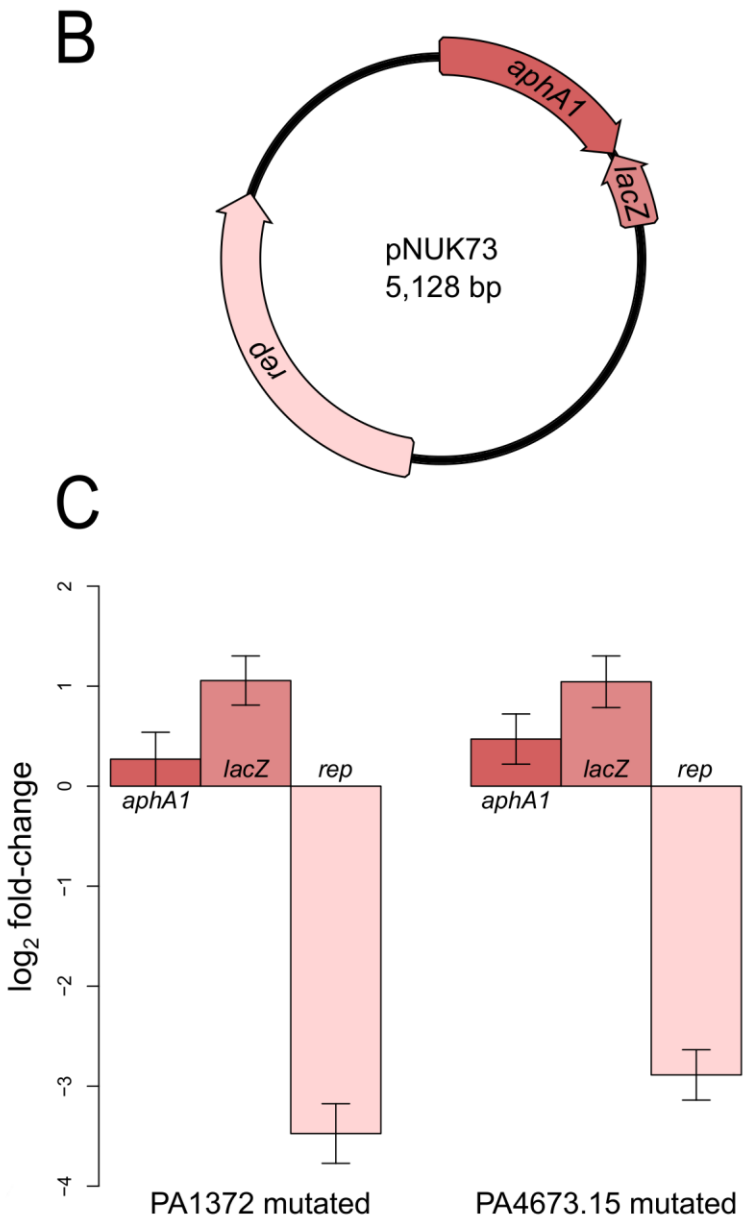


Figure 2: Gene expression map of the different strains analyzed using RNA-Seq (continued)

Panel B: Schematic representation of the pNUK73 plasmid. The reading frames for genes are shown as arrows, with the direction of transcription indicated by the arrowhead.

Panel C: Different in expression of pNUK73 genes in the helicase and kinase mutant strains compared to PAO1/pNUK73 (log_2 scale). The error bars represent the standard error of the mean.

The correlation between decreased *rep* expression levels and restored expression of chromosomal genes observed in the compensated mutants suggests that *rep* expression may be linked to both the cost of resistance and the genome-wide impact of pNUK73 on the expression of chromosomal genes (Figure 2). In an elegant study, Ingmer *et al* (2001) showed how an increase in the levels of the plasmid replication protein RepA encoded by the small plasmid pSC101 induces a delay in *E. coli* host cell division by activating the SOS response [218]. RepA binds primarily to its binding sites in pSC101 and recruits host replication proteins such as DnaA, DnaB and DnaG. However, when there is an excess of free RepA in the cell, the molecular interactions between RepA and DnaG stall chromosomal replication and inhibit cell division by inducing the SOS response. It is thought that this mechanism may help to limit the rate of segregational loss of the pSC101 plasmid by delaying cell division in host cells carrying low copy numbers of this plasmid, and therefore low number of Rep binding sites.

Consistent with the model proposed by Ingmer and colleagues, we found that pNUK73 induces the expression of the SOS response. Ten out the fifteen known LexA-regulated genes in *P. aeruginosa*, including *recA* and *lexA*, were significantly overexpressed in PAO1/pNUK73 compared to PAO1, and the remaining five also showed a non-significant increase in their expression levels (Supplementary Table 1) [129]. Furthermore, we found that plasmid carriage was associated with a delay in host cell division, measured in terms of a significant reduction in fitness (San Millan *et al*, 2014). As explained above, we found that compensatory mutations in the helicase and kinase both reduced *rep* expression and silenced the expression of the SOS response.

Linking *rep* expression to the cost of plasmid carriage

To test the hypothesis that *rep* expression is directly responsible for the fitness cost associated with pNUK73 carriage, we cloned the *rep* gene under the control of an IPTG-inducible *lac* operon promoter (*p*LAC) and integrated this inducible *rep* construct into the PAO1 genome using the mini-Tn7 system [144] to generate the PAO1::*p*LAC-*rep* strain. Using the same system, we cloned *p*LAC into the PAO1 genome to generate the PAO1::*p*LAC negative control strain. To determine the effect

of *rep* expression on fitness, we directly competed these two strains and parental PAO1/pNUK73 against a plasmid-free, GFP-tagged PAO1 strain across different IPTG concentrations (Figure 3A). The negative control strain did not show any reduction in fitness relative to the parental PAO1 under any induction regime, confirming that insertion of *p*LAC at the mini-Tn7 chromosomal site has no deleterious effects on the fitness of PAO1 (ANOVA: $p = 0.243$, $F = 1.427$, d.f = 1, 27). On the other hand, PAO1::*p*LAC-*rep* showed a significant reduction in relative fitness, even in the absence of IPTG induction (mean = 0.817, s.e = 0.005, n=3, two-sample *t*-test: $p < 0.001$, $t = 16.2$, d.f = 7.8). The cost of *rep* expression increased with IPTG concentration, and at the highest IPTG concentration (50 μM), the cost of *rep* construct expression (mean = 0.528, s.e = 0.004, n = 3) was even higher than that of pNUK73 carriage in the ancestral PAO1 strain (mean = 0.737, s.e = 0.008, n = 4, two-sample *t*-test: $p < 0.001$, $t = 22.2$, d.f = 4.3).

To link the fitness cost associated with *rep* expression to altered gene expression, we quantified the expression levels of *rep* and the SOS response genes *lexA* and *recA* in PAO1::*p*LAC-*rep* and PAO1/pNUK73 by qPCR. We found that the absolute transcript levels of *rep* stemming from leaky expression in the absence of IPTG was 7.35% (± 0.86) that of *rpoD* (n=3), a constitutively expressed essential gene whose expression is estimated to give rise to 1500 molecules of the RpoD protein per cell [221, 245]. This relatively high level of constitutive, leaky expression helps to explain why the fitness of the PAO1::*p*LAC-*rep* showed a reduction relative to the negative control strain even in the absence of IPTG induction. As expected from the fitness data, adding IPTG led to significant increases in *rep* expression (0 vs. 10 μM IPTG, two-sample *t*-test: $p = 0.01$, $t = 8.1$, d.f = 2.21, Figure 3B). The expression of SOS regulated genes in PAO1::*p*LAC-*rep* strain essentially mirrored the expression of *rep*. Constitutive leaky expression of *rep* induced the expression of *lexA* and *recA*, and IPTG induction resulted in a modest increase in the expression of SOS regulated genes (Figures 3C and 3D). Consistent with our RNA-Seq data, *rep*, *lexA* and *recA* were expressed at high levels in PAO1/pNUK73 (Figures 3B, 3C and 3D). These qPCR data broadly support our phenotypic measurements of competitive fitness and confirm that increasing *rep* expression is associated with decreased fitness and changes in *lexA/recA* expression that parallel those associated with pNUK73

carriage. The causality of this relationship was established by our RNA-Seq data, which revealed that compensatory mutations reduced the expression of *rep*. In addition, the upregulation of the SOS response genes shown by the gene expression analyses by qPCR and RNA-Seq support the idea that the SOS regulon is activated concomitantly with *rep* expression. In agreement with the gene expression analyses, we found that the presence of pNUK73 or the chromosomal copy of *rep* produced an increase in PAO1 cell size, which is associated with the delay in cell division induced by the SOS response [246] (Supplementary Figure 1).

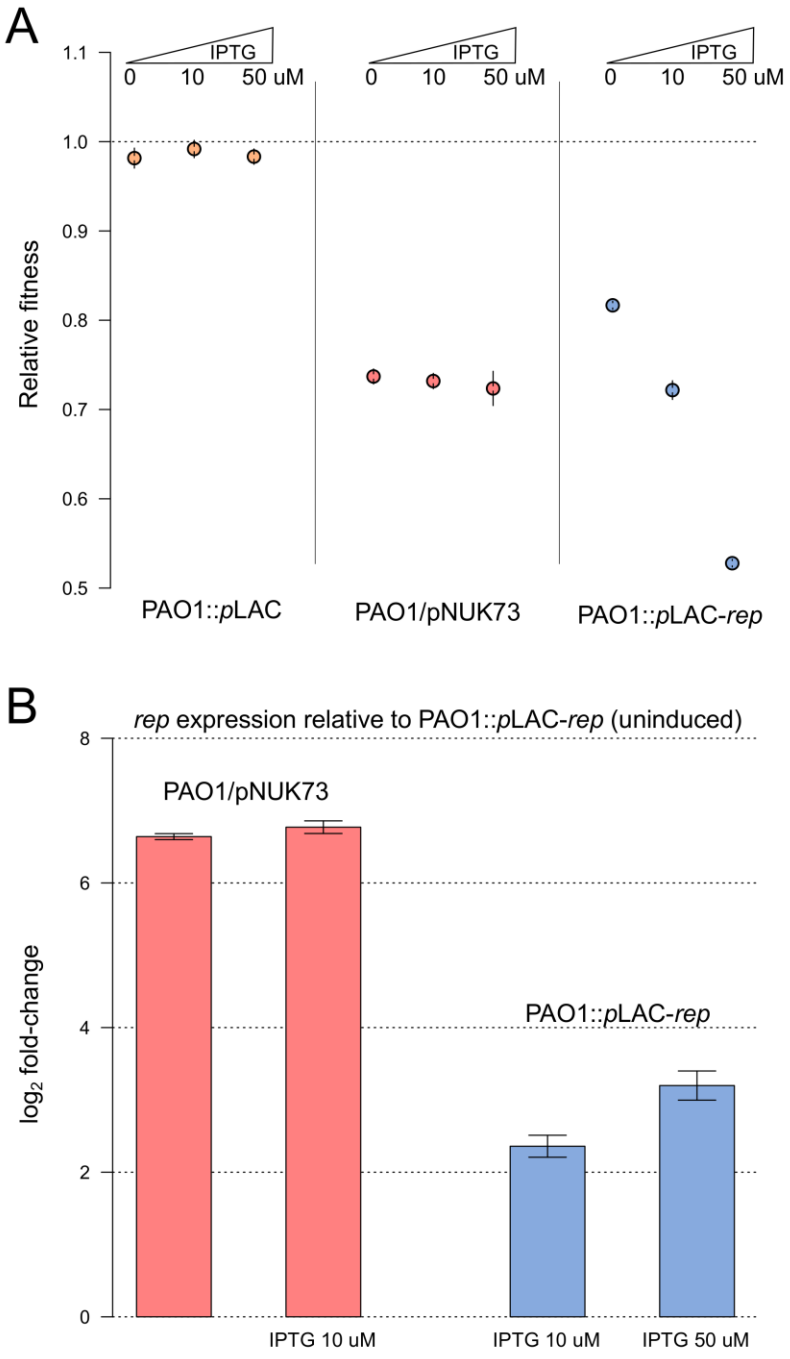


Figure 3: Expression of *rep* reduces fitness and activates the SOS response in PAO1

Panel A: Relative fitness compared to plasmid-free wild-type PAO1 for the PAO1::*p*LAC negative control strain (orange), parental PAO1/pNUK73 (red), PAO1::*p*LAC-*rep* (blue) under IPTG concentrations of 0 µM (first point), 10 µM (second point) and 50 µM (third point) .

Panel B: Relative expression (log_2 fold-change) of the *rep* gene strain in PAO1/pNUK73 (0 and 10 µM IPTG) and PAO1::*p*LAC-*rep* (10 and 50 µM IPTG) relative to the uninduced PAO1::*p*LAC-*rep* strain.

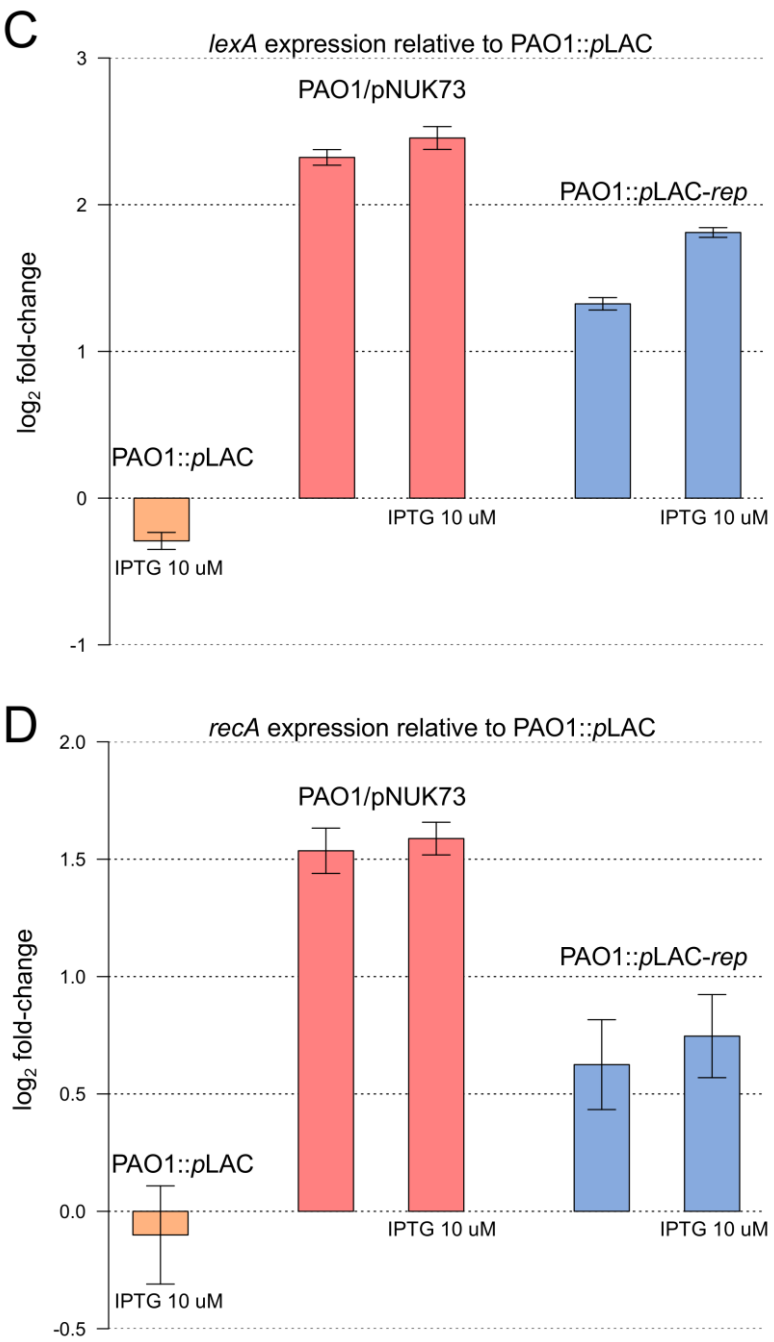


Figure 3: Expression of *rep* reduces fitness and activates the SOS response in PAO1 (continued)

Panels C and D: Expression levels of *lexA* and *recA* relative to the PAO1::*p*LAC negative control strain. Expression measurements were carried out in the presence of 10 μM IPTG and without IPTG induction. Each measurement was performed independently at least three times and the error bars represent the standard error of the mean.

At the same time, it is also evident that there are some quantitative differences between the PAO1/pNUK73 and PAO1::*p*LAC-*rep* strains with respect to how variations in *rep* expression levels can translate into differences in fitness costs. For example, *rep* expression in the presence of 10 μM IPTG in PAO1::*p*LAC-*rep* resulted in a fitness cost that is comparable to that of pNUK73 carriage (two-sample *t*-test: p = 0.317, *t* = 1.13, d.f = 4.2). However, the relative expression levels of *rep* and the two SOS genes (*lexA* and *recA*) were higher in PAO1/pNUK73 than in the IPTG-inducible strain (Figure 3). We can infer that pNUK73 effectively neutralizes some of the deleterious effects associated with *rep* expression, probably by sequestering rep protein on plasmid-carried Rep binding sites, which has been shown by Ingmer *et al* [218].

Helicase and kinases are recently acquired genes with no clear biological role in PAO1

Bioinformatic analyses show that the helicase and kinases genes that have acquired compensatory mutations exhibit a strong signature of recent horizontal acquisition in PAO1. First, while the guanine + cytosine (GC) content of the *P. aeruginosa* PAO1 genome is very high (median content of PAO1 genes = 67%), these three genes have particularly low GC contents (44% in PA1372; 40% in PA4673.15; 42% in PA4673.16), and fall in the bottom 1 percentile of GC content in the PAO1 genome (Supplementary Figure 1). Second, the three genes seem to be located in potentially mobile DNA regions. PA4673.15-16 genes are located inside the prophage RGP42 [78], and PA1372 is found inside a small low GC content island close to a gene coding for a putative transposase (PA1368) (Figure 1). Finally, these genes are very rare in all sequenced genomes of *P. aeruginosa*. Out of the 31 sequenced *P. aeruginosa* genomes available in GenBank, PA4673.15-16 is only present in one genome, while PA1372 is found in six. Taken together, these results indicate that the compensatory mutations responsible for the adaptation to pNUK73 occurred in genes that have been recently acquired by *P. aeruginosa* PAO1. Moreover, when performing similarity searches of these proteins against the non-redundant database of NCBI and against the European Nucleotide Archive (ENA), the hits were found to have very low percentage of identity (50% or lower). Interestingly, in those species where a hit was found, the genetic context was the same as in *P. aeruginosa* PAO1, i.e. PA4673.15-16 conserved their synteny and were preceded by PA4673.14, while PA1372 was found together with

PA1371. In addition, mobile DNA genes surrounded PA4673.14-16 and PA1731-2 in those genomes (i.e. recombinase, phage integrase and transposase).

Mobile genetic elements can spread through bacterial populations by acting as symbionts that increase the fitness of their bacterial hosts or by acting as parasites that selfishly exploit their hosts to maximize their own transmission [247]. Of course these two strategies are not mutually exclusive and many mobile elements can probably act as both symbionts and parasites, but it is important to draw this distinction because the evolutionary interpretation of the conflict between pNUK73 and the helicase or kinases depends crucially on whether these elements are symbionts or parasites. Previously, we have shown that mutations in the helicase and kinase do not alter bacterial fitness [128], suggesting that these genes are parasites, and not symbionts. To further test this idea, we measured the impact of helicase and kinase mutations on gene expression in the absence of pNUK73 carriage using RNA-Seq (Figure 2A). We found that the plasmid-free compensated mutants had transcriptional profiles that were almost identical to wild-type PAO1. We did not detect any differentially expressed genes in the plasmid-free helicase mutant, and only 4 genes were differentially expressed in the kinase mutant, including the two putative kinases (PA4673.15-16), a molybdenum cofactor biosynthetic protein A1 (PA3194) and a nitrite extrusion protein (PA3194). The last two genes were significantly underexpressed in both the plasmid-bearing helicase and the plasmid-free kinase mutants, suggesting that the kinases and the helicase may interact with each other.

Collectively, these transcriptomic results support the idea that helicase and kinase genes do not play any functional role in *P. aeruginosa*, perhaps because they are recently acquired genes that are not integrated in the transcriptional regulatory network of the bacterium. To further test this idea at a phenotypic level, we carried out two experiments designed to test for a functional role of the helicase and kinase genes. First, we tested the impact of helicase and kinase mutations on bacterial growth across a wider range of biological conditions using Biolog ECOLOG plates. We failed to find any disadvantage in average growth rate or stationary-phase density associated with the helicase or kinase inactivating mutations (Supplementary Figure 3). Second, to test if these mutations played a general role in adaptation to plasmids in PAO1, we analyzed the impact of helicase and kinase mutations on

the fitness burden associated with carrying 4 plasmids: Rms142 [248], pAKD1 [249], PAMBL-1 and PAMBL-2 [250]. We found no difference in the fitness burden associated with carrying these plasmids in the helicase and kinase mutants compared to the wild-type strain, indicating that these mutations are not general adaptations to plasmids (Supplementary Figure 4).

Conclusion

In this work we depicted the mechanisms implicated in the cost of and compensatory adaptation to the small plasmid pNUK73 in *P. aeruginosa* PAO1. We found that the detrimental effects produced by this plasmid arose from interactions between acquired genetic elements in PAO1. Even when the results presented here are constricted to this model system, we consider that there are several lines of evidence supporting the idea that the interactions between acquired genetic elements may play a central role in the cost of HGT. Future experimental and bioinformatic work will be necessary to test this hypothesis.

Materials & Methods

Bacterial strains, growth conditions and competitive fitness experiments

All bacterial strains were cultured in LB broth at 37°C with continuous shaking (225 rpm) and on LB agar plates at 37° C (Fisher Scientific, USA). Plasmids were electroporated into PAO1 and plasmid-bearing colonies were selected for as previously described [144, 215]. Competitive fitness experiments were performed as previously described [128]. We used Biolog EcoPlates (Biolog, USA) to assess bacterial growth in different environments according to the manufacturer's instructions. Flow cytometry was performed using an Accuri C6 Flow Cytometer Instrument (BD Accuri, USA) with the following parameters: flow rate: 66 µL/min, core size: 22 µm, events recorded per sample: 10,000 – 20,000.

RNA extractions and reverse transcription

Each bacterial strain (Supplementary Table 2) was inoculated in LB medium with IPTG induction (Sigma-Aldrich, USA) where required and grown overnight at 37°C with continuous shaking (225 rpm). The overnight cultures were diluted 1:100 in fresh LB with the same concentration of IPTG and incubated until they reached an OD_{600} of 0.5 under the same experimental conditions. The OD_{600} values of bacterial cultures were measured using the BioTek Synergy 2 microplate reader. Bacterial cultures were mixed with RNAprotect Bacteria Reagent (Qiagen, Netherlands) according to the manufacturer's instructions. Total RNA extraction was performed using the SV Total RNA Isolation System (Promega, USA). To remove genomic DNA, total RNA samples were DNase-treated during and after the RNA extraction procedure. For reverse transcription, first-strand cDNA samples were synthesized from 2.0 µg of total RNA templates using the GoScript Reverse Transcription System with random hexamer primers (Promega, USA). Negative control reactions, which contained all components for reverse transcription with the exception of the reverse transcriptase enzyme, were performed to verify the absence of genomic DNA in cDNA.

RNA-Seq analysis

RNA-Seq analysis was performed on the transcriptomic data six strains: PAO1 wild-type, parental PAO1/pNUK73, and two compensated mutants both carrying the plasmid and cured: a helicase mutant clone (PA1372, premature stop codon at position 378) and a kinase mutant clone (PA4673.15, premature stop codon at position 95). Two biological replicates of RNA samples obtained on different days were sequenced for each strain as described above. Library preparation (directional paired-end ribodepleted library) and sequencing was performed using Illumina MiSeq at the Welcome Trust Centre for Human Genetics, University of Oxford. Raw reads were initially filtered using the NGS QC Toolkit [251]. The ancestral strain used in this study has two additional features compared to the reference genome *P. aeruginosa* PAO1 (NC_002516.2), namely the insertion of the phage RGP42 (GQ141978.1) and the plasmid pNUK73. Filtered reads were mapped to the *P. aeruginosa* PAO1 reference genome and the reference sequences for the pNUK73 plasmid and the RGP42 phage using BWA [252]. Gene counts were obtained from BAM files using *HTSeq* [152]. Differential gene expression analysis with correction for batch effect was performed using *DESeq2* [154].

Quantitative real-time PCR

Quantitative real-time PCR (qPCR) assays were performed using the relative quantification method. Each gene shown in Supplementary Table 3 was amplified using the Fast SYBR Green Master Mix (Applied Biosystems, USA) and 100 nM oligonucleotide primers on the StepOnePlus Real-time PCR platform (Applied Biosystems, USA) in three biological replicates and two technical replicates. The amplification efficiency, linearity and specificity of each qPCR primer pair were verified three times using a serially diluted pool of experimental cDNA samples. The transcript levels ($\log_2$ fold-change) of the target genes *lexA, recA* and *rep* in all the test strains (Supplementary Table 2) were determined relative to the relevant control strains, i.e. the PAO1::*p*LAC control strain for *lexA* and *recA* expression; the PAO1::pLAC-*rep* strain without IPTG induction for *rep* expression. Normalization factors were calculated from the geometric mean of the transcript levels of three internal reference

genes (*acpP*, *atpA* and *rpoD*). The transcript levels of the target genes were determined by normalizing the transcript levels of every target genes to the normalization factors. Finally, the relative transcript levels for each target gene in each sample group were calculated after correcting for the batch effect. Further experimental details are described in Supplementary Materials.

To quantify the extent of leaky expression of *rep* from *p*LAC, we measured the copy number of *rep* and *rpoD* transcripts in cDNA samples derived from the un-induced PAO1::*p*LAC-*rep* strain using the absolute quantification method by qPCR. Calibrator standard curves for both genes were generated using genomic DNA extracted from the PAO1::*p*LAC-*rep* strain, whose concentration was determined using the QuantiFluor dsDNA System (Promega, USA). The copy number of *rep* and *rpoD* were calculated based on the gDNA calibrator standard curves and the mass of gDNA per genome.

Molecular Cloning & Transposon Mutagenesis

The *rep* gene with its presumed ribosomal binding site (5'-CGGAGG-3') 5 bp upstream of the start codon (GTG) was PCR-amplified from pNUK73 plasmid DNA using the forward primer 5'-TAAGGATCCTGGCGGAGGCGGCT-3' and reverse primer 5'-GCCAAGCTTTAACGTTCCCCTAACTT-3'. The PCR product was cloned downstream of *p*LAC between the BamHI and HindIII restriction sites in the multiple cloning site in the pUC18-mini-Tn7T-Gm-LAC (accession no. AY599234) delivery plasmid [144]. To integrate the IPTG-inducible *rep* construct into the mini-Tn7 site of the plasmid-free PAO1 parental strain used in this study, sucrose-treated PAO1 was co-transformed with the pUC18-mini-Tn7-*p*LAC-rep delivery vector and a Tn7 transposase-encoding helper plasmid pUX-BF13 [145] via electroporation to generate the PAO1::*p*LAC-*rep* strain. The negative control strain, PAO1::*p*LAC, which is isogenic to the PAO1::*p*LAC-*rep* strain except for the absence of *rep* downstream of *p*LAC, was created using the same procedure.

GO term and Pathway enrichment

GOEAST [253] was used to test for gene ontology enrichment among the group of differentially expressed genes. DAVID [254] was used to test for enrichment in several functional annotations.

Similarity searches

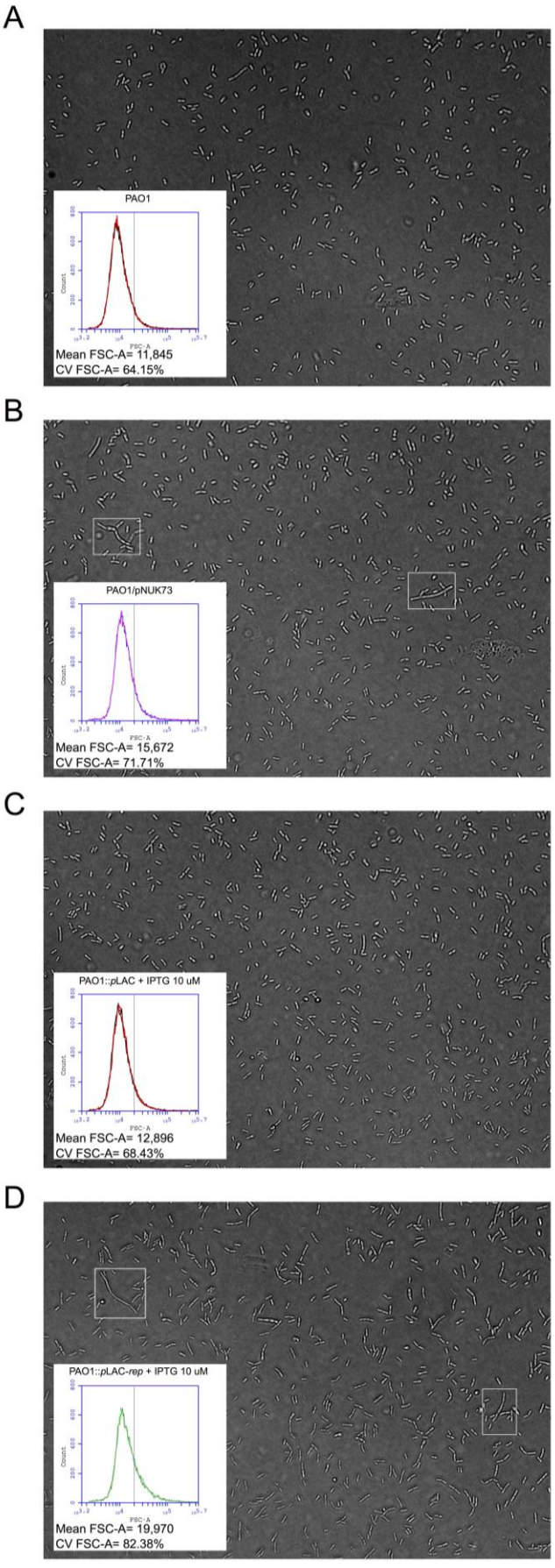
To study horizontal gene transfer, we used several bioinformatic approaches that rely on similarity searches. First, we downloaded 143 next-generation sequencing datasets deposited in the Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra), containing raw sequence reads from 48 *P. aeruginosa*, 63 *Pseudomonas* phages and 32 non-*aeruginosa Pseudomonas* datasets. We mapped the reads to our PAO1 reference genome using BWA [252] to check if there were reads mapping to our genes of interest. Secondly, we performed online BLASTP searches [255] against the non-redundant database in NCBI. Finally, we performed online searches using the European Nucleotide Archive (ENA) [256], which accesses all Ensembl Genomes.

Supplementary Materials

Supplementary Figure 1: Increase in cell size due to the activation of SOS response (next page)

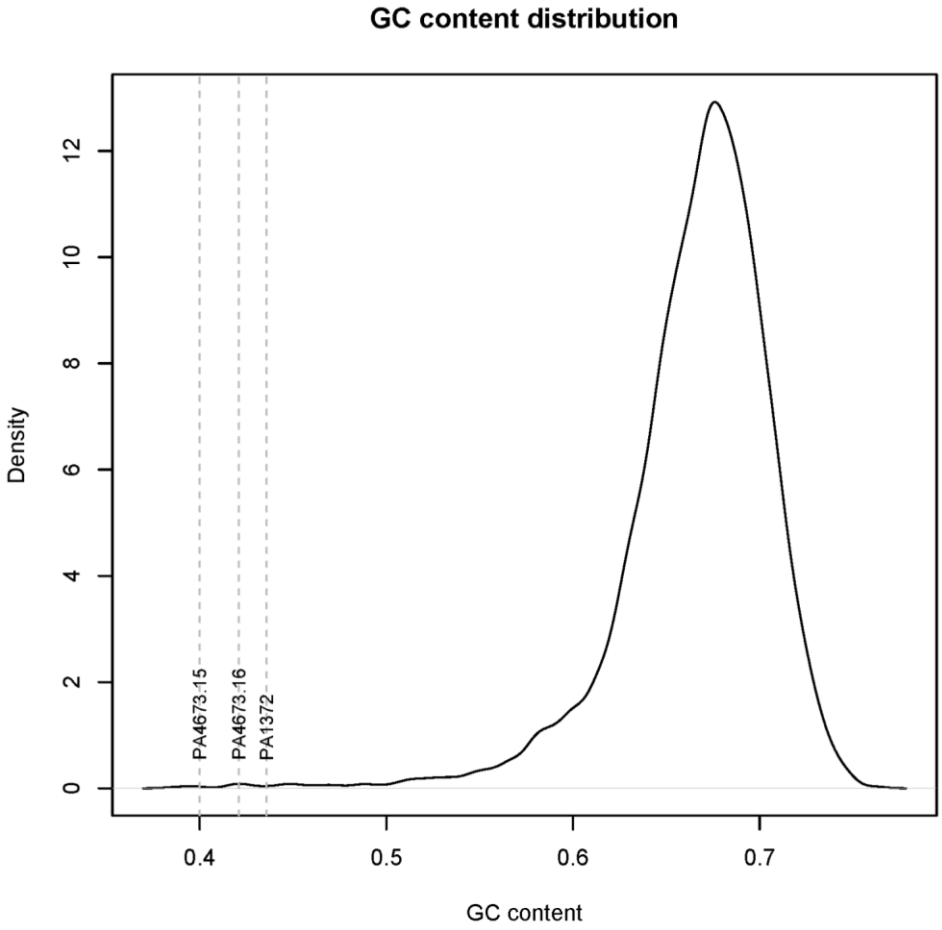
Each panel features an optical microscope image and a plot showing the distribution of Forward Scatter values (FSC-A) of three independent populations of PAO1 (Panel A), PAO1/pNUK73 (Panel B), PAO1::*p*LAC + IPTG 10 μM (Panel C) and PAO1::*p*LAC-*rep* + IPTG 10 μM (Panel D). We measured 20,000 events for each population using flow cytometry and used FSC-A as a proxy for cell size. The mean FSC-A of PAO1 increases significantly in response to pNUK73 (Panels A and B; two-sample *t*-test: $p = 0.002$, $t = 8.78$, d.f = 3.3) and induced *rep* expression (Panels C and D; two-sample *t*-test: $p < 0.001$, $t = 48.7$, d.f = 3.6), which provides phenotypic evidence that the SOS response is activated in conjunction with *rep* expression. White boxes show elongated cells. The dashed line in the FSC-A plot is intended to highlight the higher asymmetry (skewness) in the right tail of the distribution as a result of a higher proportion of elongated cells in B and D compared to A and C, respectively.

Supplementary Figure 1



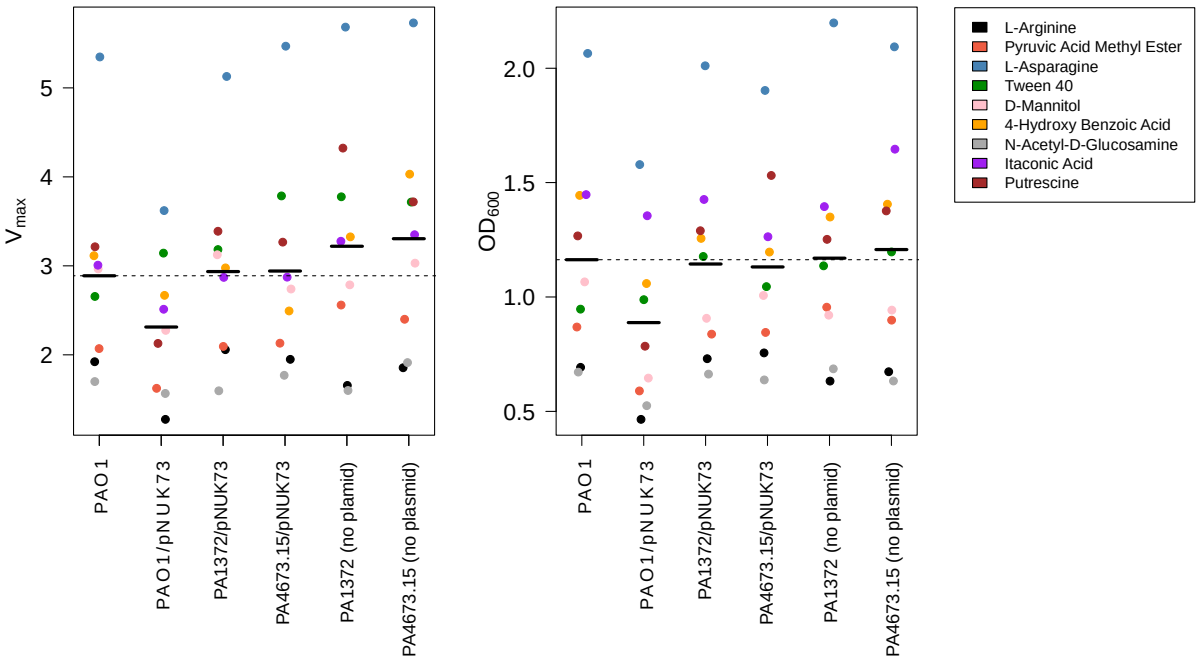
Supplementary Figure 2: Distribution of GC content in the *P. aeruginosa* PAO1 genome

The dotted lines indicate the GC content for the 3 genes carrying compensatory mutations responsible for the compensatory adaptation to carriage of the pNUK73 small plasmid.



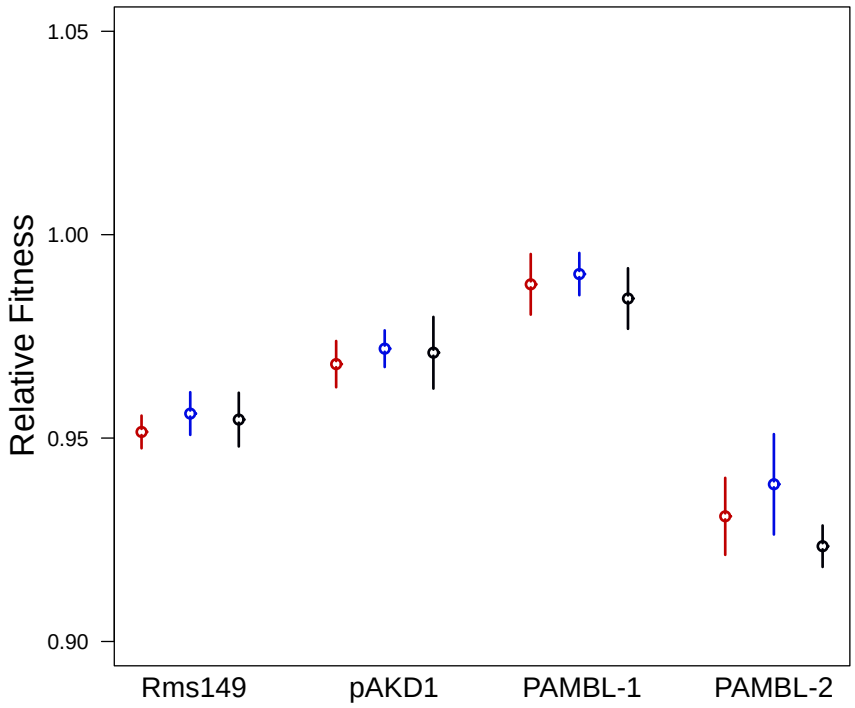
Supplementary Figure 3: Growth kinetics in Biolog EcoPlates

This graph shows the growth kinetics of PAO1, a helicase mutant clone (PA1372, premature stop codon at position 378) and a kinase mutant clone (PA4673.15, premature stop codon at position 95) with and without the plasmid pNUK73 in different carbon sources. We used the 31 different environments of the Biolog EcoPlates as the carbon sources. We observed measurable growth in 9 environments. The average of three independent measurements (five for the control PAO1) is represented by a dot in the chart. The black dashes represent the average value of V_{max} or OD_{max} for each clone and the dotted line represents the average for PAO1. Maximum velocity (V_{max}) and maximum OD_{600} (OD_{max}) of the clones in the different environments were compared to those of PAO1 as a proxy for bacterial fitness (growth rate and yield respectively). PAO1/pNUK73 showed a decrease in both V_{max} and OD_{max} (paired *t*-test, V_{max}: p = 0.021, *t*= 2.86; OD_{max}: p = 0.002, *t*= 4.49, d.f = 8) across the 9 environments. The helicase mutant (Paired *t*-test, V_{max}: p = 0.562, *t*= 0.60; OD_{max}: p = 0.667, *t* = 0.45, d.f = 8) and kinase mutant (paired *t*-test, V_{max}: p = 0.752, *t*= 0.33; OD_{max}: p = 0.584, *t*= 0.57, d.f = 8) compensated for the cost of pNUK73 throughout the different environments. Finally, the plasmid-free helicase mutant showed no differences in V_{max} or OD_{max} (paired *t*-test, V_{max}: p = 0.088, *t*= 1.94; OD_{max}: p = 0.841, *t*= 0.21, d.f = 8) compared to PAO1, while the plasmid-free kinase mutant presented no difference in OD_{max} and even a significant increase in V_{max} compared to PAO1 (paired *t*-test, V_{max}: p = 0.009, *t*= 3.40; OD_{max}: p = 0.305, *t* = 1.10, d.f = 8).



Supplementary Figure 4: Helicase and kinase mutations are not a general adaptation to plasmids

This graph shows the competitive fitness of parental PAO1/pNUK73 (red), helicase mutant of PAO1/pNUK73 with a premature stop codon at position 378 in PA1372 (blue) and kinase mutant of PAO1/pNUK73 with a premature stop codon at position 95 in PA4673.15 (black) relative to the plasmid-free PAO1 wild-type strain. Four different plasmids were used: Rms149, pAKD1, PAMBL-1 and PAMBL-2. We performed nine biological replicates of the competition experiment for each clone. There were no significant differences in the fitness of the three different clones carrying a particular plasmid (ANOVA: Rms149, p = 0.69, F= 0.163, d.f = 1, 25; pAKD1, p = 0.762, F = 0.093, d.f = 1, 25; PAMBL-1, p = 0.711, F = 0.141, d.f = 1, 25; PAMBL-2, p = 0.5881, F= 0.302, d.f = 1, 25).



Key:

- PAO1/pNUK73
- Helicase mutant of PAO1/pNUK73 with a premature stop codon at position 378 in PA1372
- Kinase mutant of PAO1/pNUK73 with a premature stop codon at position 95 in PA4673.15

Tables

Supplementary Table 1: pNUK73 activates the SOS response and increases the expression levels of LexA-regulated genes in parental PAO1/pNUK73 relative to the PAO1 wild-type strain

Gene locus in PAO1	Gene name	Fold-change # (log_2)	Benjamini-Hochberg corrected p-value
PA0069	*phl*	1.6	0.480
PA0669	*dnaE2*	1.3	0.673
PA0670	*imuB*	1.5	0.550
PA0671	*sulA2*	1.3	0.709
PA0922		5.8	3.02×10^{-7}*
PA1044		3.4	0.0165*
PA1045	*dinG*	2.2	0.0541
PA2288		7.7	3.93×10^{-9}*
PA3007	*lexA*	4.9	1.37×10^{-6}*
PA3413	*yebG*	3.7	8.07×10^{-4}*
PA3414		6.5	4.72×10^{-7}*
PA3616	*recX*	5.6	1.09×10^{-6}*
PA3617	*recA*	3.8	8.25×10^{-5}*
PA4763	*recN*	4.5	3.88×10^{-6}*
PA3008	*sulA*	5.3	1.25×10^{-6}*

Relative to the plasmid-free PAO1 wild-type strain.
* $p < 0.05$

Supplementary Table 2: Bacterial strains for RNA extraction

Strain name	Descriptions	Gene expression analyses
PAO1	Plasmid-free *P. aeruginosa* PAO1 wild-type strain.	RNA-Seq
PAO1/pNUK73	pNUK73 plasmid-bearing PAO1 strain.	RNA-Seq qPCR
PA1372 helicase mutant	PAO1 strain with a compensatory mutation in PA1372 (premature stop codon at position 378). With and without plasmid pNUK73.	RNA-Seq
PA4673.15 kinase mutant	PAO1 strain with the compensatory mutation in PA4673.15 (premature stop codon at position 95). With and without plasmid pNUK73.	RNA-Seq
PAO1::*p*LAC-*rep*	PAO1 strain with an IPTG-inducible *rep* gene downstream of *p*LAC inserted into the mini-Tn7 chromosomal site.	qPCR
PAO1::*p*LAC	Negative control strain that is isogenic to PAO1::mini-Tn7-Gm-*LAC-rep*, but lacks *rep* downstream of the *lac* operon promoter.	qPCR

Supplementary Table 3: Oligonucleotide sequences of quantitative PCR primers used in this study

Gene	Forward primer (5' → 3')	Reverse primer (5' → 3')	Reference
acpP	ACTCGGCGTGAAGGAAGAAG	CGACGGTGTCAAGGGAGT	[168]
atpA	TACGCCAGATCACGCCCGGA	GGCGGCGTCTACGGTATGGC	Chapter 2 of this thesis
rpoD	CAGGTGGCGTAGGTGGAGAA	GGGCGAAGAAGGAAATGGTC	[167]
lexA	ATCGGAGCGAACTCAGGGT	CCGTATCAATCCCGCCTTCT	This study
recA	TTCGCTACCCACCACCTCGT	CAAGAACGCCAACTGCCTGG	
rep	AGCAGAGCGACGAGGAAGTG	ATAGCCACCCAGCCGCCATT	

Supplementary File 1: Expression profile of the clones analyzed by RNA-Seq

Available on request

Supplementary File 2: Functional enrichment analysis of the genes differently expressed in PAO1/pNUK73 compared to PAO1 using DAVID and GOEAST software

Available on request

Acknowledgements

This work was supported by funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant (StG-2011-281591) and from the Royal Society. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The work of ASM was supported by a Marie Curie Intra-European Fellowship within the Seventh European Community Framework Programme (PIEF-GA-2011-298975). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics funded by Wellcome Trust grant reference 090532/Z/09/Z and Medical Research Council Hub grant G0900747 91070 for generation of the high-throughput sequencing data.

The authors would like to thank Professor N. Itoh for kindly providing the plasmid for this study. We are grateful to Dr C. Patthey for assessing the quality of RNA samples for RNA-Seq. We thank K. Heilbron for his assistance in analysing our Biolog EcoPlates data and designing Figure S3. We also thank J. A. Escudero, R. Niehus and Professor K.R. Foster for their comments on the manuscript.

The authors declare that there are no conflicts of interest.

Author Contributions

Conceived and designed the experiments: ASM, MTR, QQ, RCM

Performed the experiments: ASM, QQ

Analyzed the RNA-Seq data and performed bioinformatic analyses: MTR

Wrote the paper: ASM, MTR, QQ, RCM

Appendix B: Characterizing the fitness effects of increasing the expression of selected denitrification genes in two *rpoB* mutants

In Chapter 2, the transcriptome profiles of two *rpoB* mutants carrying mutations D521G and H531R were obtained using RNA-Seq. Amongst the 59 genes that showed significant differential expression in the two mutants relative to the rifampicin-sensitive ancestral strain, many genes from the denitrification gene cluster were found to be significantly downregulated (Annex B, Supplementary Materials, Chapter 2). *Pseudomonas aeruginosa* is a facultative anaerobe, which is capable of generating energy via aerobic respiration when oxygen availability is high, or via denitrification (nitrate respiration) under microaerobic or anaerobic conditions in the presence of nitrate [257, 258]. A cascade of denitrification enzymes in *P. aeruginosa* summarized in Figure 4 reduce nitrate (NO_3^-) to molecular nitrogen (N_2) via nitrite (NO_2^-), nitric oxide (NO) and nitrous oxide (N_2O) [259]. Interestingly, denitrification has also been observed under aerobic experimental conditions and is thought to be an electron-accepting mechanism that is supplementary to, or competitive with aerobic respiration [260]. Based on existing literature, I hypothesized that the downregulation of *nirCF* and *nirQOP* operons contributed to the fitness cost of resistance in the two sequenced *rpoB* mutants of *P. aeruginosa.*

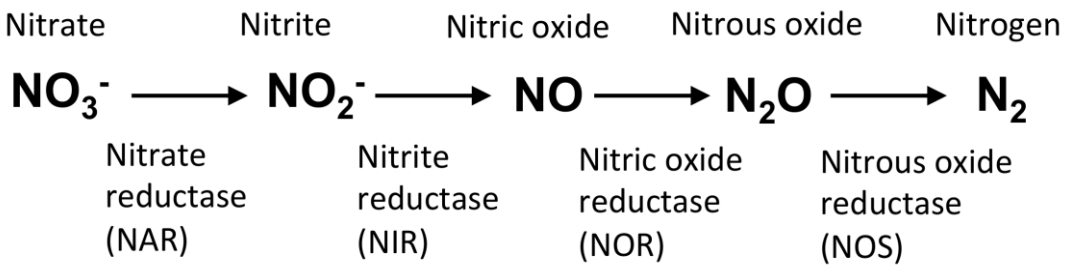


Figure 4: Schematic of the denitrification pathway in *P. aeruginosa*. The names of denitrification enzymes are shown below the arrows.

Figure adapted from van Alst *et al* [261]

The large gene cluster *nirSMCFDLGHJEN* encodes gene products that are required for the production of the active form of the periplasmic cytochrome cd_1 nitrite reductase (NIR) [262], which is a key denitrification enzyme that catalyzes the reduction of nitrite (NO_2^-) to nitric oxide (NO) [263]. NIR contains of two identical subunits, each of which has a heme *c* group and a heme d_1 cofactor [264, 265]. *nirF* plays an essential role in d_1 heme biogenesis [262], while *nirC* is an essential gene in *P. aeruginosa* PAO1 [161], which encodes a c-type cytochrome that mediates electron donation from the membrane electron pathway to NIR [266]. In an *in vivo* competitive fitness study, experimental mice were infected with a co-culture of the *P. aeruginosa* PA14 wild-type strain and a derivative mutant with the *nirF* gene disrupted by transposon insertion [267]. The authors demonstrated that the disruption of *nirF* carried a very high fitness cost. Given the profound fitness impacts of *nirCF* -disruption on the fitness of *P. aeruginosa*, I hypothesized that the significant downregulation of these two genes observed in the two sequenced *rpoB* mutants is associated with a fitness cost.

The *nirQ* gene in the *nirQOP* operon encodes an ATP-binding regulatory protein that is required for the activation of the nitric oxide reductase (NOR) enzyme, which converts nitric oxide (NO) to nitrous oxide (N_2O) [268, 269]. In *P. aeruginosa*, deletion of the *nirQOP* operon resulted in the abolition of anaerobic growth, presumably due to the significant reduction in NOR activity [270]. If denitrification can also be observed under aerobic conditions as suggested by Chen *et al* [260], a reduction in the expression of *nirQOP* is expected to reduce the amount of energy that can be derived from denitrification, which is expected to translate into a small fitness cost.

To test these ideas, I cloned an additional IPTG-inducible copy of *nirCF* or *nirQOP* into the mini-Tn7 chromosomal site of the PAO1 wild-type strain and the two rifampicin resistant mutants of PAO1 carrying the D521G and H531R mutations in *rpoB*. If the reduction in the expression levels of these two operons is a causal factor in generating a partial fitness cost, stimulating their expression levels is expected to partly increase the fitness of these *rpoB* mutants.

Materials and Methods

Molecular cloning & transposon mutagenesis

The *nirCF* and *nirQOP* operons with their presumed ribosomal binding sites (5'-AGGAGT-3' and 5'-ACGAGG-3' respectively) were PCR-amplified from the genomic DNA template extracted from the *P. aeruginosa* wild-type strain PAO1. The oligonucleotide sequences of PCR primers (synthesized by Sigma-Aldrich, United Kingdom) for molecular cloning are shown in Table 1. The PCR products were cloned downstream of the *lac* operon promoter between the PstI and HindIII restriction sites in the multiple cloning site in the pUC18-mini-Tn7T-Gm-*LAC* mini-Tn7 (accession no. AY599234) delivery plasmid [144].

Primer	Oligonucleotide sequence (5'->3')	Restriction site
*nirCF*_F	TAACTGCAGATGAAACCGTACGCACTG	PstI
*nirCF*_R	GACAAGCTTCTAGAGTCCGATGTGCTG	HindIII
*nirQOP*_F	TAACTGCAGCCTAGCGCAGGAGTAGGA	PstI
*nirQOP*_R	GACAAGCTTTCAGTTATTCATCCGCGC	HindIII

Table 1: Oligonucleotide sequences of PCR primers used in this chapter, with restriction site sequences underlined

The IPTG-inducible *nirCF* and *nirQOP* constructs were integrated into the mini-Tn7 chromosomal site of the PAO1 parental strain by co-transforming sucrose-treated PAO1 with one of the two mini-Tn7 delivery vector (pUC18-mini-Tn7-Gm-LAC-*nirCF* or pUC18-mini-Tn7-Gm-LAC-*nirQOP*) and a transposase-encoding helper plasmid pUX-BF13 [145] via electroporation. A negative control strain, PAO1::*p*LAC, which is isogenic to the IPTG-inducible *nir* operon insertion mutants of PAO1 except for the absence of the *nir* operon genes downstream of the *lac* promoter, was generated using the same procedure. Successful chromosomal insertion at the mini-Tn7 site downstream of the *glmS* gene was verified using procedures for colony PCR described by Choi & Schweizer [144].

Numerous attempts were made to insert the IPTG-inducible *nirCF* and *nirQOP* operons into the mini-Tn7 sites of rifampicin-resistant mutant derivatives of PAO1 via transposon mutagenesis. However, due to the lack of experimental success, *rpoB* mutants were evolved from the three rifampin-sensitive ancestral strains that were generated in this study (Table 2) using a fluctuation test described in Chapter 2 of this thesis.

Bacterial strain	**Descriptions**
PAO1::*p*LAC	Negative control strain with no *nir* operon genes downstream of the *lac* operon promoter
PAO1::*p*LAC-*nirCF*	IPTG-inducible *nirCF* operon at the mini-Tn7 chromosomal site
PAO1::*p*LAC-*nirQOP*	IPTG-inducible *nirQOP* operon at the mini-Tn7 site

Table 2: Bacterial strains generated using Tn7 transposon mutagenesis.

Competitive fitness assays by flow cytometry

The competitive fitness of the rifampicin-sensitive ancestral strains shown in Table 2 and their rifampicin-resistant derivatives carrying mutations D521G or H531R in *rpoB* were determined relative to the PAO1 wild-type strain by competitive fitness assays. Minor modifications were made to the experimental procedures for standard competitive fitness assays by flow cytometry described in the Materials & Methods section Chapter 2 to allow for IPTG induction of *nirCF* and *nirQOP* expression at the mini-Tn7 chromosomal site.

The rifampicin-sensitive ancestral strains (Table 2) and their *rpoB* mutant derivatives were pre-cultured overnight at 37˚C with shaking in M9KB culture medium containing IPTG at concentrations of 0 µM, 50 µM and 200 µM. On the next day, all strains were diluted in M9KB containing the same concentrations of IPTG used during pre-culture and re-grown to early exponential phase at 30˚C with shaking. Each non-fluorescent strains was individually mixed with a GFP-tagged control strain, diluted 1:200 in fresh M9KB medium containing the same concentration of IPTG added previously

and incubated at 30°C overnight for approximately 24 hours with continuous shaking. The competitive fitness of all strains was determined relative to the PAO1 wild-type strain.

Results and Discussion

Figures 5A and 5B show that overexpressing the five selected *nir* genes resulted in no statistically significant alteration in the competitive fitness of PAO1::*p*LAC-*nirCF* and PAO1::*p*LAC-*nirQOP* relative to that of PAO1::*p*LAC (one-way ANOVA at $p < 0.05$), regardless of the presence or absence of IPTG. Similarly, overexpressing *nirCF* and *nirQOP* in the two *rpoB* mutants (D521G or H531R) evolved from the rifampicin-sensitive PAO1::*p*LAC-*nirCF* and PAO1::*p*LAC-*nirQOP* ancestral strains resulted in no significant alteration in fitness compared to the rifampicin resistant derivatives of PAO1::*p*LAC carrying the same *rpoB* mutation, both in the presence and absence of IPTG (one-way ANOVA at $p < 0.05$). At a phenotypic level, these results provided preliminary evidence that inducing *nirCF* and *nirQOP* expression from an additional chromosomal copy of either *nir* operon cannot offset the fitness cost associated with these two *rpoB* mutations to a significant extent, which contradicts my hypothesis. It should be emphasized, however, that in the absence of any qPCR–based measurements of relative *nirCF* and *nirQOP* transcript levels, one cannot exclude the possibility that the overexpression of the five denitrification genes from the mini-Tn7 site may interfere with their expression from the native chromosomal site, which is a major drawback of this experimental approach.

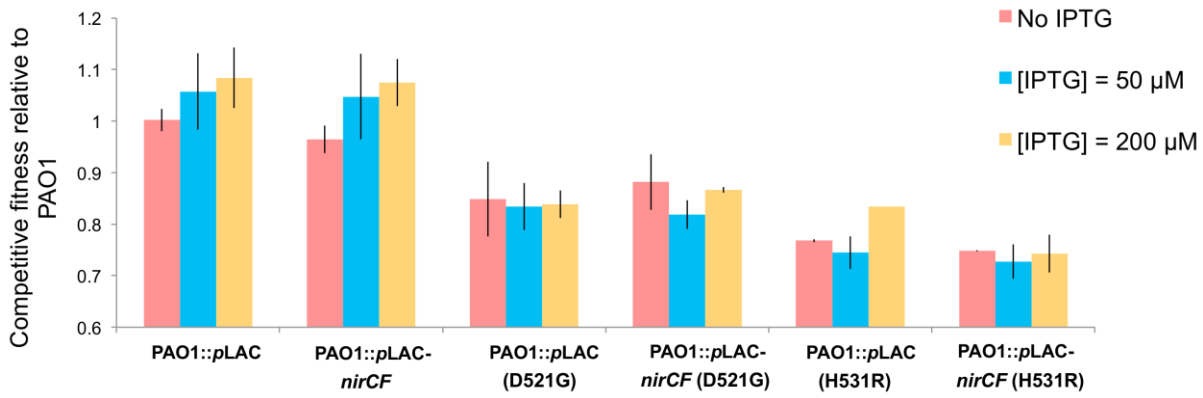


Figure 5A: Competitive fitness of PAO1::*p*LAC-*nirCF* and its *rpoB* mutant derivatives (D521G and H531R) relative to the PAO1 wild-type strain

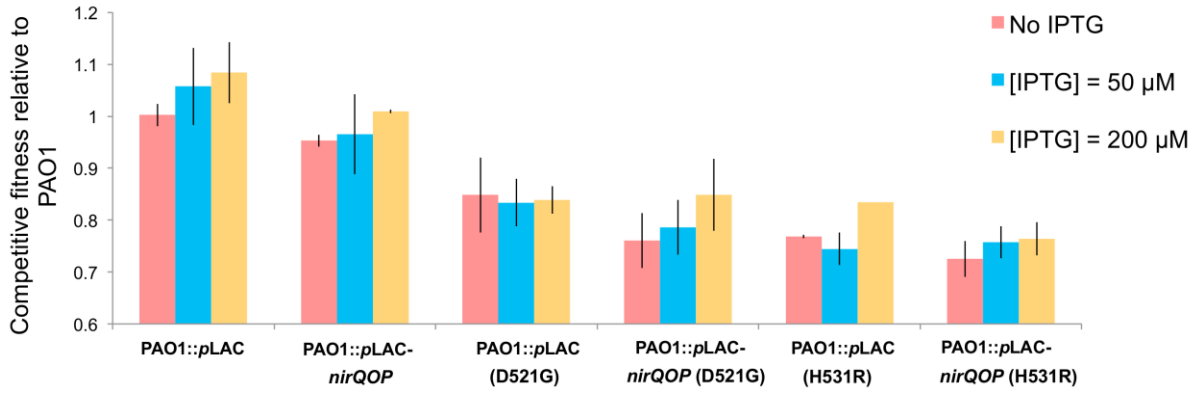


Figure 5B: Competitive fitness of PAO1::*p*LAC-*nirQOP* and its *rpoB* mutant derivatives (D521G and H531R) relative to the PAO1 wild-type strain

8. References

1. Kohanski, M., Dwyer, D., and Collins, J. (2010). How antibiotics kill bacteria: from targets to networks. Nat. Rev. Microbiol. *8*, 423-435.
2. Hancock, R.E. (2014). Collateral damage. Nat. Biotechnol. *32*, 66-68.
3. MacLean, R.C., Hall, A.R., Perron, G.G., and Buckling, A. (2010). The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. Nat. Rev. Genet. *11*, 405-414.
4. Arias, C.A., and Murray, B.E. (2009). Antibiotic-resistant bugs in the 21st century - a clinical super-challenge. N. Engl. J. Med. *360*, 439-443.
5. Levy, S.B. (2002). The antibiotic paradox : how the misuse of antibiotics destroys their curative power, 2nd Edition, Cambridge, MA; Great Britain: Perseus Pub.
6. Levy, S.B., and Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. Nat. Med. *10*, S122-129.
7. Okeke, I.N., Lamikanra, A., and Edelman, R. (1999). Socioeconomic and behavioral factors leading to acquired bacterial resistance to antibiotics in developing countries. Emerg. Infect. Dis. *5*, 18-27.
8. Alanis, A. (2005). Resistance to antibiotics: Are we in the post-antibiotic era? Arch. Med. Res. *36*, 697-705.
9. Boucher, H.W., Talbot, G.H., Bradley, J.S., Edwards, J.E., Gilbert, D., Rice, L.B., Scheld, M., Spellberg, B., and Bartlett, J. (2009). Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. Clin. Infect. Dis. *48*, 1-12.
10. Butler, M.S., Blaskovich, M.A., and Cooper, M.A. (2013). Antibiotics in the clinical pipeline in 2013. J. Antibiot. *66*, 571-591.
11. Imamovic, L., and Sommer, M.O. (2013). Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. Sci. Transl. Med. *5*, 204ra132.
12. Lewis, K. (2013). Platforms for antibiotic discovery. Nat. Rev. Drug Discov. *12*, 371-387.
13. Hughes, D. (2003). Exploiting genomics, genetics and chemistry to combat antibiotic resistance. Nat. Rev. Genet. *4*, 432-441.
14. Mitosch, K., and Bollenbach, T. (2014). Bacterial responses to antibiotics and their combinations. Env. Microbiol. Rep.
15. Andersson, D.I., and Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? Nat. Rev. Microbiol. *8*, 260-271.
16. Walsh, C. (2003). Antibiotics : Actions, Origins, Resistance. Washington, D.C.: ASM Press; [Oxford : Enfield, distributor].
17. Jatsenko, T., Tover, A., Tegova, R., and Kivisaar, M. (2010). Molecular characterization of Rif(r) mutations in *Pseudomonas aeruginosa* and *Pseudomonas putida*. Mutat. Res. *683*, 106-114.
18. Benveniste, R., and Davies, J. (1973). Mechanisms of antibiotic resistance in bacteria. Annu. Rev. Biochem. *42*, 471-506.
19. Yu, E.W., Zhang, Q., and Brown, M.H. (2013). Microbial efflux pumps : Current research. Caister Academic Press
20. Sandegren, L., and Andersson, D.I. (2009). Bacterial gene amplification: implications for the evolution of antibiotic resistance. Nat. Rev. Microbiol. *7*, 578-588.
21. Gagneux, S., Long, C.D., Small, P.M., Van, T., Schoolnik, G.K., and Bohannan, B.J. (2006). The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science *312*, 1944-1946.
22. Rozen, D.E., McGee, L., Levin, B.R., and Klugman, K.P. (2007). Fitness costs of fluoroquinolone resistance in *Streptococcus pneumoniae*. Antimicrob. Agents. Chemother. *51*, 412-416.
23. Andersson, D.I. (2006). The biological cost of mutational antibiotic resistance: any practical conclusions? Curr. Opin. Microbiol. *9*, 461-465.
24. Andersson, D.I., and Levin, B.R. (1999). The biological cost of antibiotic resistance. Curr. Opin. Microbiol. *2*, 489-493.

25. Lenski, R.E. (1998). Bacterial evolution and the cost of antibiotic resistance. Int. Microbiol. *1*, 265-270.
26. Dionisio, F., Conceicao, I.C., Marques, A.C., Fernandes, L., and Gordo, I. (2005). The evolution of a conjugative plasmid and its ability to increase bacterial fitness. Biol. Lett. *1*, 250-252.
27. Bouma, J.E., and Lenski, R.E. (1988). Evolution of a bacteria/plasmid association. Nature *335*, 351-352.
28. Dahlberg, C., and Chao, L. (2003). Amelioration of the cost of conjugative plasmid carriage in *Eschericha coli* K12. Genetics *165*, 1641-1649.
29. Paulander, W., Maisnier-Patin, S., and Andersson, D.I. (2009). The fitness cost of streptomycin resistance depends on *rpsL* mutation, carbon source and RpoS (σ^S). Genetics *183*, 539-546, 531SI-532SI.
30. Björkman, J., and Andersson, D.I. (2000). The cost of antibiotic resistance from a bacterial perspective. Drug Resist. Update *3*, 237-245.
31. Mariam, D.H., Mengistu, Y., Hoffner, S.E., and Andersson, D.I. (2004). Effect of *rpoB* mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. *48*, 1289-1294.
32. Perkins, A.E., and Nicholson, W.L. (2008). Uncovering new metabolic capabilities of *Bacillus subtilis* using phenotype profiling of rifampin-resistant *rpoB* mutants. J. Bacteriol. *190*, 807-814.
33. Sandegren, L., Lindqvist, A., Kahlmeter, G., and Andersson, D.I. (2008). Nitrofurantoin resistance mechanism and fitness cost in *Escherichia coli*. J. Antimicrob. Chemother. *62*, 495-503.
34. Binet, R., and Maurelli, A.T. (2007). Frequency of development and associated physiological cost of azithromycin resistance in *Chlamydia psittaci* 6BC and *C. trachomatis* L2. Antimicrob. Agents Chemother. *51*, 4267-4275.
35. Besier, S., Ludwig, A., Zander, J., Brade, V., and Wichelhaus, T.A. (2008). Linezolid resistance in *Staphylococcus aureus*: gene dosage effect, stability, fitness costs, and cross-resistances. Antimicrob. Agents Chemother. *52*, 1570-1572.
36. Macvanin, M., Johanson, U., Ehrenberg, M., and Hughes, D. (2000). Fusidic acid-resistant EF-G perturbs the accumulation of ppGpp. Mol. Microbiol. *37*, 98-107.
37. Zorzet, A., Pavlov, M.Y., Nilsson, A.I., Ehrenberg, M., and Andersson, D.I. (2010). Error-prone initiation factor 2 mutations reduce the fitness cost of antibiotic resistance. Mol. Microbiol. *75*, 1299-1313.
38. Potrykus, K., and Cashel, M. (2008). (p)ppGpp: still magical? Annu. Rev. Microbiol. *62*, 35-51.
39. Andersson, D.I. (2006). The biological cost of mutational antibiotic resistance: any practical conclusions? Curr. Opin. Microbiol. *9*, 461-465.
40. Hartman, P.E., and Roth, J.R. (1973). Mechanisms of suppression. Adv. Genet. *17*, 1-105.
41. Poon, A., and Chao, L. (2005). The rate of compensatory mutation in the DNA bacteriophage *phiX174*. Genetics *170*, 989-999.
42. Maisnier-Patin, S., and Andersson, D.I. (2004). Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. Res. Microbiol. *155*, 360-369.
43. Zhang, L., and Watson, L.T. (2009). Analysis of the fitness effect of compensatory mutations. HFSP J. *3*, 47-54.
44. Kimura, M. (1985). The role of compensatory neutral mutations in molecular evolution. J. Genet. *64*, 7-19.
45. Macvanin, M., Ballagi, A., and Hughes, D. (2004). Fusidic acid-resistant mutants of *Salmonella enterica* serovar typhimurium have low levels of heme and a reduced rate of respiration and are sensitive to oxidative stress. Antimicrob. Agents Chemother. *48*, 3877-3883.
46. Reynolds, M.G. (2000). Compensatory evolution in rifampin-resistant *Escherichia coli*. Genetics *156*, 1471-1481.

47. MacLean, R.C., and Buckling, A. (2009). The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. PLoS Genet. *5*, e1000406.
48. Wichelhaus, T.A., Böddinghaus, B., Besier, S., Schäfer, V., Brade, V., and Ludwig, A. (2002). Biological cost of rifampin resistance from the perspective of *Staphylococcus aureus*. Antimicrob. Agents Chemother. *46*, 3381-3385.
49. Lester, W. (1972). Rifampin - Semisynthetic derivative of rifamycin - Prototype for future. Annu. Rev. Microbiol. *26*, 85.
50. Maggi, N., Pasqualucci, C.R., Ballotta, R., and Sensi, P. (1966). Rifampicin: a new orally active rifamycin. Chemotherapy *11*, 285-292.
51. Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell *104*, 901-912.
52. Marianelli, C., Ciuchini, F., Tarantino, M., Pasquali, P., and Adone, R. (2004). Genetic bases of the rifampin resistance phenotype in Brucella spp. J. Clin. Microbiol. *42*, 5439-5443.
53. Kunin, C.M., Brandt, D., and Wood, H. (1969). Bacteriologic studies of rifampin, a new semisynthetic antibiotic. J. Infect. Dis. *119*, 132-137.
54. Verma, M., Kaur, J., Kumar, M., Kumari, K., Saxena, A., Anand, S., Nigam, A., Ravi, V., Raghuvanshi, S., Khurana, P., et al. (2011). Whole genome sequence of the rifamycin B-producing strain *Amycolatopsis mediterranei* S699. J. Bacteriol. *193*, 5562-5563.
55. Villain-Guillot, P., Bastide, L., Gualtieri, M., and Leonetti, J.P. (2007). Progress in targeting bacterial transcription. Drug Discov. Today *12*, 200-208.
56. Finberg, R.W., Moellering, R.C., Tally, F.P., Craig, W.A., Pankey, G.A., Dellinger, E.P., West, M.A., Joshi, M., Linden, P.K., Rolston, K.V., et al. (2004). The importance of bactericidal drugs: future directions in infectious disease. Clin. Infect. Dis. *39*, 1314-1320.
57. Deal, W.B., and Sanders, E. (1969). Efficacy of rifampin in treatment of meningococcal carriers. N. Engl. J. Med. *281*, 641-645.
58. Hartmann, G., Honikel, K.O., Knüsel, F., and Nüesch, J. (1967). The specific inhibition of the DNA-directed RNA synthesis by rifamycin. Biochim. Biophys. Acta *145*, 843-844.
59. Kessler, C., and Hartmann, G.R. (1977). The two effects of rifampicin on the RNA polymerase reaction. Biochem. Biophys. Res. Commun. *74*, 50-56.
60. Severinov, K., Soushko, M., Goldfarb, A., and Nikiforov, V. (1993). Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of *Escherichia coli* RNA polymerase. J. Biol. Chem. *268*, 14820-14825.
61. Jin, D.J., and Turnbough, C.L., Jr. (1994). An *Escherichia coli* RNA polymerase defective in transcription due to its overproduction of abortive initiation products. J. Mol. Biol. *236*, 72-80.
62. Jin, D.J., and Gross, C.A. (1991). RpoB8, a rifampicin-resistant termination-proficient RNA polymerase, has an increased Km for purine nucleotides during transcription elongation. J. Biol. Chem. *266*, 14478-14485.
63. Jin, D.J., Walter, W.A., and Gross, C.A. (1988). Characterization of the termination phenotypes of rifampicin-resistant mutants. J. Mol. Biol. *202*, 245-253.
64. Jin, D.J., and Gross, C.A. (1989). Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*. J. Bacteriol. *171*, 5229-5231.
65. Hall, A.R., Iles, J., and MacLean, R.C. (2011). The fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase. Genetics *187*, 817-822.
66. Brandis, G., Wrande, M., Liljas, L., and Hughes, D. (2012). Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. Mol. Microbiol. *85*, 142-151.
67. Stearns, F.W. (2010). One hundred years of pleiotropy: a retrospective. Genetics *186*, 767-773.
68. Cooper, T.F., Ostrowski, E.A., and Travisano, M. (2007). A negative relationship between mutation pleiotropy and fitness effect in yeast. Evolution *61*, 1495-1499.
69. Matsuo, M., Hishinuma, T., Katayama, Y., Cui, L., Kapi, M., and Hiramatsu, K. (2011). Mutation of RNA polymerase beta subunit (*rpoB*) promotes hVISA-to-VISA phenotypic conversion of strain Mu3. Antimicrob. Agents Chemother. *55*, 4188-4195.

70. Rodríguez-Verdugo, A., Gaut, B.S., and Tenaillon, O. (2013). Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. BMC Evol. Biol. *13*, 50.
71. Conrad, T.M., Frazier, M., Joyce, A.R., Cho, B.K., Knight, E.M., Lewis, N.E., Landick, R., and Palsson, B. (2010). RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. Proc. Natl. Acad. Sci. U.S.A. *107*, 20500-20505.
72. Krebs, J.E., Lewin, B., Kilpatrick, S.T., and Goldstein, E.S. (2013). Lewin's genes XI, 11th Edition, Burlington, Mass.: Jones & Bartlett Learning.
73. Hall, A.R. (2013). Genotype-by-environment interactions due to antibiotic resistance and adaptation in *Escherichia coli*. J. Evol. Biol. *26*, 1655-1664.
74. du Preez, I., and Loots, d.T. (2012). Altered fatty acid metabolism due to rifampicin-resistance conferring mutations in the *rpoB* gene of *Mycobacterium tuberculosis*: mapping the potential of pharmaco-metabolomics for global health and personalized medicine. OMICS *16*, 596-603.
75. Drake, D., and Montie, T.C. (1988). Flagella, motility and invasive virulence of *Pseudomonas aeruginosa*. J. Gen. Microbiol. *134*, 43-52.
76. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., et al. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature *406*, 959-964.
77. Lee, D.G., Urbach, J.M., Wu, G., Liberati, N.T., Feinbaum, R.L., and Ausubel, F.M. (2008). Combining genomic tools to dissect multifactorial virulence in *Pseudomonas aeruginosa*. 127-150.
78. Klockgether, J., Munder, A., Neugebauer, J., Davenport, C.F., Stanke, F., Larbig, K.D., Heeb, S., Schock, U., Pohl, T.M., Wiehlmann, L., et al. (2010). Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains. J. Bacteriol. *192*, 1113-1121.
79. Mikkelsen, H., McMullan, R., and Filloux, A. (2011). The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in *ladS*. PLoS ONE *6*, e29113.
80. Wiehlmann, L., Wagner, G., Cramer, N., Siebert, B., Gudowius, P., Morales, G., Kohler, T., van Delden, C., Weinel, C., Slickers, P., et al. (2007). Population structure of *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. U.S.A. *104*, 8101-8106.
81. Hentzer, M., Givskov, M., and Parsek, M.R. (2002). Targeting quorum sensing for treatment of chronic bacterial biofilm infections. Lab. Med. *33*, 295-306.
82. Eberhard, A. (1972). Inhibition and activation of bacterial luciferase synthesis. J. Bacteriol. *109*, 1101-1105.
83. Hentzer, M., Wu, H., Andersen, J.B., Riedel, K., Rasmussen, T.B., Bagge, N., Kumar, N., Schembri, M.A., Song, Z., Kristoffersen, P., et al. (2003). Attenuation of *Pseudomonas aeruginosa* virulence by quorum sensing inhibitors. EMBO J. *22*, 3803-3815.
84. Miller, M.B., and Bassler, B.L. (2001). Quorum sensing in bacteria. Annu. Rev. Microbiol. *55*, 165-199.
85. Fuqua, W.C., Winans, S.C., and Greenberg, E.P. (1994). Quorum sensing in bacteria - the LuxR-LuxI family of cell density-responsive transcriptional regulators. J. Bacteriol. *176*, 269-275.
86. Schuster, M., Lostroh, C.P., Ogi, T., and Greenberg, E.P. (2003). Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. J. Bacteriol. *185*, 2066-2079.
87. Bassler, B.L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. Curr. Opin. Microbiol. *2*, 582-587.
88. Pearson, J.P., Feldman, M., Iglewski, B.H., and Prince, A. (2000). *Pseudomonas aeruginosa* cell-to-cell signaling is required for virulence in a model of acute pulmonary infection. Infect. Immun. *68*, 4331-4334.

89. Davies, D.G., Parsek, M.R., Pearson, J.P., Iglewski, B.H., Costerton, J.W., and Greenberg, E.P. (1998). The involvement of cell-to-cell signals in the development of a bacterial biofilm. Science *280*, 295-298.
90. Parsek, M.R., and Greenberg, E.P. (2000). Acyl-homoserine lactone quorum sensing in Gram-negative bacteria: a signaling mechanism involved in associations with higher organisms. Proc. Natl. Acad. Sci. U.S.A. *97*, 8789-8793.
91. Costerton, J.W., Stewart, P.S., and Greenberg, E.P. (1999). Bacterial biofilms: a common cause of persistent infections. Science *284*, 1318-1322.
92. Koch, C., and Høiby, N. (1993). Pathogenesis of cystic fibrosis. Lancet *341*, 1065-1069.
93. Donlan, R.M. (2001). Biofilm formation: a clinically relevant microbiological process. Clin. Infect. Dis. *33*, 1387-1392.
94. Middleton, B., Rodgers, H.C., Cámara, M., Knox, A.J., Williams, P., and Hardman, A. (2002). Direct detection of N-acylhomoserine lactones in cystic fibrosis sputum. FEMS Microbiol. Lett. *207*, 1-7.
95. Hancock, R.E. (1998). Resistance mechanisms in *Pseudomonas aeruginosa* and other nonfermentative Gram-negative bacteria. Clin. Infect. Dis. *27 Suppl 1*, S93-99.
96. Hancock, R.E., and Speert, D.P. (2000). Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. Drug Resist. Update *3*, 247-255.
97. Rossolini, G.M., and Mantengoli, E. (2005). Treatment and control of severe infections caused by multiresistant *Pseudomonas aeruginosa*. Clin. Microbiol. Infec. *11 Suppl 4*, 17-32.
98. Morrison, A.J., and Wenzel, R.P. (1984). Epidemiology of infections due to *Pseudomonas aeruginosa*. Rev. Infect. Dis. *6 Suppl 3*, S627-642.
99. System, N.N.I.S. (2004). National Nosocomial Infections Surveillance (NNIS) System Report, Data Summary from January 1992 through June 2004, Issued October 2004. Am. J. Infect. Control *32*, 470-485
100. Spencer, D.H., Kas, A., Smith, E.E., Raymond, C.K., Sims, E.H., Hastings, M., Burns, J.L., Kaul, R., and Olson, M.V. (2003). Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. J. Bacteriol. *185*, 1316-1325
101. Winsor, G.L., Lam, D.K., Fleming, L., Lo, R., Whiteside, M.D., Yu, N.Y., Hancock, R.E., and Brinkman, F.S. (2011). *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. Nucleic Acids Res. *39*, D596-600.
102. Ryan, K.J., Ray, C.G. (2004). Sherris Medical Microbiology: An Introduction to Infectious Diseases, 4th Edition, (New York ; London: McGraw-Hill).
103. Brandis, G., Pietsch, F., Alemayehu, R., and Hughes, D. (2014). Comprehensive phenotypic characterization of rifampicin resistance mutations in *Salmonella* provides insight into the evolution of resistance in *Mycobacterium tuberculosis*. J. Antimicrob. Chemother.
104. Brandis, G., and Hughes, D. (2013). Genetic characterization of compensatory evolution in strains carrying *rpoB* Ser531Leu, the rifampicin resistance mutation most frequently found in clinical isolates. J. Antimicrob. Chemother. *68*, 2493-2497.
105. Estes, S., Phillips, P.C., and Denver, D.R. (2011). Fitness recovery and compensatory evolution in natural mutant lines of *C. elegans*. Evolution *65*, 2335-2344.
106. Poon, A., Davis, B.H., and Chao, L. (2005). The coupon collector and the suppressor mutation: estimating the number of compensatory mutations by maximum likelihood. Genetics *170*, 1323-1332.
107. Gifford, D.R., and MacLean, R.C. (2013). Evolutionary reversals of antibiotic resistance in experimental populations of *Pseudomonas aeruginosa*. Evolution *67*, 2973-2981.
108. Schrag, S.J., and Perrot, V. (1996). Reducing antibiotic resistance. Nature *381*, 120-121.
109. Maisnier-Patin, S., Berg, O.G., Liljas, L., and Andersson, D.I. (2002). Compensatory adaptation to the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. Mol. Microbiol. *46*, 355-366.
110. Albarracin Orio, A.G., Pinas, G.E., Cortes, P.R., Cian, M.B., and Echenique, J. (2011). Compensatory evolution of *pbp* mutations restores the fitness cost imposed by beta-lactam resistance in *Streptococcus pneumoniae*. PLoS Pathog. *7*, e1002000.

111. Maisnier-Patin, S., Paulander, W., Pennhag, A., and Andersson, D.I. (2007). Compensatory evolution reveals functional interactions between ribosomal proteins S12, L14 and L19. J. Mol. Biol. *366*, 207-215.
112. Barrick, J.E., and Lenski, R.E. (2013). Genome dynamics during experimental evolution. Nat. Rev. Genet. *14*, 827-839.
113. Nilsson, A.I., Zorzet, A., Kanth, A., Dahlstrom, S., Berg, O.G., and Andersson, D.I. (2006). Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes. Proc. Natl. Acad. Sci. U.S.A. *103*, 6976-6981.
114. Nagaev, I., Bjorkman, J., Andersson, D.I., and Hughes, D. (2001). Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. Mol. Microbiol. *40*, 433-439.
115. Kugelberg, E., Lofmark, S., Wretlind, B., and Andersson, D.I. (2005). Reduction of the fitness burden of quinolone resistance in *Pseudomonas aeruginosa*. J. Antimicrob. Chemother. *55*, 22-30.
116. Björkman, J., Nagaev, I., Berg, O.G., Hughes, D., and Andersson, D.I. (2000). Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. Science *287*, 1479-1482.
117. Sander, P., Springer, B., Prammananan, T., Sturmfels, A., Kappler, M., Pletschette, M., and Bottger, E.C. (2002). Fitness cost of chromosomal drug resistance-conferring mutations. Antimicrob. Agents Chemother. *46*, 1204-1211.
118. Dettman, J.R., Rodrigue, N., Melnyk, A.H., Wong, A., Bailey, S.F., and Kassen, R. (2012). Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. Mol. Ecol. *21*, 2058-2077.
119. Bell, G. (2008). Selection : the mechanism of evolution, 2nd Edition, (Oxford: Oxford University Press).
120. Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution - principles and applications for biotechnology. Microb. Cell Fact. *12*, 64.
121. Scheiner, S.M. (2002). Selection experiments and the study of phenotypic plasticity. J. Evol. Biol. *15*, 889-898.
122. Jordan, I.K., Rogozin, I.B., Wolf, Y.I., and Koonin, E.V. (2002). Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. Genome Res. *12*, 962-968.
123. Toulokhonov, I., Shulgina, I., and Hernandez, V. (2001). Binding of the transcription effector ppGpp to *Escherichia coli* RNA polymerase is allosteric, modular, and occurs near the N terminus of the β'-subunit. J. Biol. Chem. *276*, 1220-1225.
124. Reddy, P.S., Raghavan, A., and Chatterji, D. (1995). Evidence for a ppGpp-binding site on *Escherichia coli* RNA polymerase: proximity relationship with the rifampicin-binding domain. Mol. Microbiol. *15*, 255-265.
125. Zhou, Y.N., and Jin, D.J. (1998). The *rpoB* mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. *95*, 2908-2913.
126. Baltrus, D.A. (2013). Exploring the costs of horizontal gene transfer. Trends Ecol. Evol. *28*, 489-495.
127. Itoh, N., Kawanami, T., Nitta, C., Iwata, N., Usami, S., Abe, Y., and Koide, Y. (2003). Characterization of pNI10 plasmid in *Pseudomonas*, and the construction of an improved *Escherichia* and *Pseudomonas* shuttle vector, pNUK73. Appl. Microbiol. Biotechnol. *61*, 240-246.
128. San Millan, A., Peña-Miller, R., Toll-Riera, M., Halbert, Z., McLean, A.R., Cooper, B.S., and MacLean, R.C. (2014). Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. Nat. Commun. 5: 5208
129. Cirz, R.T., O'Neill, B.M., Hammond, J.A., Head, S.R., and Romesberg, F.E. (2006). Defining the *Pseudomonas aeruginosa* SOS response and its role in the global response to the antibiotic ciprofloxacin. J. Bacteriol. *188*, 7101-7110.
130. Rando, O.J., and Verstrepen, K.J. (2007). Timescales of genetic and epigenetic inheritance. Cell *128*, 655-668.

131. Friedberg, E.C., Walker, G.C., and Siede, W. (1995). DNA repair and mutagenesis, (Washington, D.C.: ASM).
132. Taddei, F., Matic, I., and Radman, M. (1995). cAMP-dependent SOS induction and mutagenesis in resting bacterial-populations. Proc. Natl. Acad. Sci. U.S.A. *92*, 11736-11740.
133. Bianco, P.R., Tracy, R.B., and Kowalczykowski, S.C. (1998). DNA strand exchange proteins: a biochemical and physical comparison. Front. Biosci. *3*, D570-603.
134. Bell, C.E. (2005). Structure and mechanism of *Escherichia coli* RecA ATPase. Mol. Microbiol. *58*, 358-366.
135. Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.
136. Little, J.W. (1984). Autodigestion of *lexA* and phage lambda repressors. Proc. Natl. Acad. Sci. U.S.A. *81*, 1375-1379.
137. Huisman, O., D'Ari, R., and Gottesman, S. (1984). Cell-division control in *Escherichia coli*: specific induction of the SOS function SfiA protein is sufficient to block septation. Proc. Natl. Acad. Sci. U.S.A. *81*, 4490-4494.
138. Justice, S.S., Hunstad, D.A., Cegelski, L., and Hultgren, S.J. (2008). Morphological plasticity as a bacterial survival strategy. Nat. Rev. Microbiol. *6*, 162-168.
139. Rastogi, R.P., Richa, Kumar, A., Tyagi, M.B., and Sinha, R.P. (2010). Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. J. Nucleic Acids *2010*, 592980.
140. Bergval, I.L., Klatser, P.R., Schuitema, A.R., Oskam, L., and Anthony, R.M. (2007). Specific mutations in the *Mycobacterium tuberculosis rpoB* gene are associated with increased *dnaE2* expression. FEMS Microbiol. Lett. *275*, 338-343.
141. Vogwill, T., and MacLean, R.C. (2014). The genetic basis of the fitness costs of antibicrobial resistance: a meta-analysis approach. Evol. Appl. (In press)
142. Wehrli, W. (1983). Rifampin: mechanisms of action and resistance. Rev. Infect. Dis. *5 Suppl 3*, S407-411.
143. MacLean, R.C., Bell, G., and Rainey, P.B. (2004). The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. Proc. Natl. Acad. Sci. U.S.A. *101*, 8072-8077.
144. Choi, K.H., and Schweizer, H.P. (2006). Mini-Tn7 insertion in bacteria with single *attTn7* sites: example *Pseudomonas aeruginosa*. Nat. Protoc. *1*, 153-161.
145. Bao, Y., Lies, D.P., Fu, H., and Roberts, G.P. (1991). An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of Gram-negative bacteria. Gene *109*, 167-168.
146. Luria, S.E., and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. Genetics *28*, 491-511.
147. MacLean, R.C., Perron, G.G., and Gardner, A. (2010). Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. Genetics *186*, 1345-1354.
148. Heilbron, K., Toll-Riera, M., Kojadinovic, M., and Maclean, R.C. (2014). Fitness is strongly influenced by rare mutations of large effect in a microbial mutation accumulation experiment. Genetics *197(3)*, 981-990.
149. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. *3*, RESEARCH0034.
150. Andersen, C.L., Jensen, J.L., and Ørntoft, T.F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. *64*, 5245-5250.
151. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.
152. Anders, S., Pyl, P.T., Huber, W. (2014). *HTSeq* — A Python framework to work with high-throughput sequencing data. Bioinformatics btu638v2-btu638.
153. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

154. Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.
155. Anders, S., McCarthy, D.J., Chen, Y., Okoniewski, M., Smyth, G.K., Huber, W., and Robinson, M.D. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nat. Protoc. *8*, 1765-1786.
156. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. *Series B (Methodological)*, 289-300.
157. Stewart, G.S., and Williams, P. (1992). *lux* genes and the applications of bacterial bioluminescence. J. Gen. Microbiol. *138*, 1289-1300.
158. Van Dyk, T.K., and Rosson, R.A. (1998). *Photorhabdus luminescens luxCDABE* promoter probe vectors. Methods Mol. Biol. *102*, 85-95.
159. Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., Kato-Maeda, M., Galagan, J., Niemann, S., and Gagneux, S. (2012). Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nat. Genet. *44*, 106-110.
160. Silander, O.K., Nikolic, N., Zaslaver, A., Bren, A., Kikoin, I., Alon, U., and Ackermann, M. (2012). A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*. PLoS Genet. *8*, e1002443.
161. Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., and Ausubel, F.M. (2006). An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. Proc. Natl. Acad. Sci. U.S.A. *103*, 2833-2838.
162. Bass, I.A., Danilevskaya, O.N., Mekhedov, S.L., Fedoseeva, V.B., and Gorlenko, Z.M. (1979). The effect of rifampicin upon the transcription of RNA polymerase β-gene in *Escherichia coli*. Mol. Gen. Genet. *173*, 101-107.
163. Little, R., and Dennis, P.P. (1979). Expression of RNA polymerase and ribosome component genes in *Escherichia coli* mutants having conditionally defective RNA polymerases. J. Bacteriol. *137*, 115-123.
164. Blumenthal, R.M., and Dennis, P.P. (1978). Gene expression in *Escherichia coli* B/r during partial rifampicin-mediated restrictions of transcription initiation. Mol. Gen. Genet. *165*, 79-86.
165. Hayward, R.S., and Fyfe, S.K. (1978). Non-coordinate expression of the neighbouring genes *rplL* and *rpoBC* of *Escherichia coli*. Mol. Gen. Genet. *160*, 77-80.
166. Morgan, B.A., and Hayward, R.S. (1987). Direct evidence for rifampicin-promoted readthrough of the partial terminator t_{L7} in the *rpoBC* operon of *Escherichia coli*. Mol. Gen. Genet. *210*, 358-363.
167. Savli, H., Karadenizli, A., Kolayli, F., Gundes, S., Ozbek, U., and Vahaboglu, H. (2003). Expression stability of six housekeeping genes: a proposal for resistance gene quantification studies of *Pseudomonas aeruginosa* by real-time quantitative RT-PCR. J. Med. Microbiol. *52*, 403-408.
168. Lenz, A.P., Williamson, K.S., Pitts, B., Stewart, P.S., and Franklin, M.J. (2008). Localized gene expression in *Pseudomonas aeruginosa* biofilms. Appl. Environ. Microbiol. *74*, 4463-4471.
169. Qin, X., Emerson, J., Stapp, J., Stapp, L., Abe, P., and Burns, J.L. (2003). Use of real-time PCR with multiple targets to identify *Pseudomonas aeruginosa* and other nonfermenting Gram-negative bacilli from patients with cystic fibrosis. J. Clin. Microbiol. *41*, 4312-4317.
170. Mesak, L.R., Yim, G., and Davies, J. (2009). Improved *lux* reporters for use in *Staphylococcus aureus*. Plasmid *61*, 182-187.
171. de Vos, M., Müller, B., Borrell, S., Black, P.A., van Helden, P.D., Warren, R.M., Gagneux, S., and Victor, T.C. (2013). Putative compensatory mutations in the *rpoC* gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. Antimicrob. Agents Chemother. *57*, 827-832.
172. Casali, N., Nikolayevskyy, V., Balabanova, Y., Ignatyeva, O., Kontsevaya, I., Harris, S.R., Bentley, S.D., Parkhill, J., Nejentsev, S., Hoffner, S.E., et al. (2012). Microevolution of extensively drug-resistant tuberculosis in Russia. Genome Res. *22*, 735-745.

173. Muller, B., Borrell, S., Rose, G., and Gagneux, S. (2013). The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. TIG *29*, 160-169.
174. Blank, D., Wolf, L., Ackermann, M., and Silander, O.K. (2014). The predictability of molecular evolution during functional innovation. Proc. Natl. Acad. Sci. U.S.A. *111*, 3044-3049.
175. Moore, F.B., Rozen, D.E., and Lenski, R.E. (2000). Pervasive compensatory adaptation in *Escherichia coli*. Proc. R. B Soc. B *267*, 515-522.
176. Lenski, R.E., Simpson, S.C., and Nguyen, T.T. (1994). Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. J. Bacteriol. *176*, 3140-3147.
177. Modi, R.I., Castilla, L.H., Puskas-Rozsa, S., Helling, R.B., and Adams, J. (1992). Genetic changes accompanying increased fitness in evolving populations of *Escherichia coli*. Genetics *130*, 241-249.
178. Wielgoss, S., Barrick, J.E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R.E., and Schneider, D. (2011). Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. G3 *1*, 183-186.
179. Haldane, J.B.S. (1932). The Causes of Evolution, Longmans, Green and Co.
180. Gerrish, P.J., and Lenski, R.E. (1998). The fate of competing beneficial mutations in an asexual population. Genetica *102-103*, 127-144.
181. Miralles, R., Gerrish, P.J., Moya, A., and Elena, S.F. (1999). Clonal interference and the evolution of RNA viruses. Science *285*, 1745-1747.
182. Imhof, M., and Schlotterer, C. (2001). Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. U.S.A. *98*, 1113-1117.
183. Lang, G.I., Rice, D.P., Hickman, M.J., Sodergren, E., Weinstock, G.M., Botstein, D., and Desai, M.M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature *500*, 571-574.
184. Lee, M.C., and Marx, C.J. (2013). Synchronous waves of failed soft sweeps in the laboratory: remarkably rampant clonal interference of alleles at a single locus. Genetics *193*, 943-952.
185. Barrett, R.D.H., MacLean, R.C., and Bell, G. (2006). Mutations of intermediate effect are responsible for adaptation in evolving *Pseudomonas fluorescens* populations. Biol. Lett. *2*, 236-238.
186. Perfeito, L., Fernandes, L., Mota, C., and Gordo, I. (2007). Adaptive mutations in bacteria: High rate and small effects. Science *317*, 813-815.
187. Stevens, K.E., and Sebert, M.E. (2011). Frequent beneficial mutations during single-colony serial transfer of *Streptococcus pneumoniae*. PLoS Genet. *7*.
188. Orr, H.A. (2002). The population genetics of adaptation: the adaptation of DNA sequences. Evolution *56*, 1317-1330.
189. Luján, A.M., Moyano, A.J., Segura, I., Argaraña, C.E., and Smania, A.M. (2007). Quorum-sensing-deficient (*lasR*) mutants emerge at high frequency from a *Pseudomonas aeruginosa mutS* strain. Microbiology *153*, 225-237.
190. Sandoz, K.M., Mitzimberg, S.M., and Schuster, M. (2007). Social cheating in *Pseudomonas aeruginosa* quorum sensing. Proc. Natl. Acad. Sci. U.S.A. *104*, 15876-15881.
191. Bottomley, M.J., Muraglia, E., Bazzo, R., and Carfì, A. (2007). Molecular insights into quorum sensing in the human pathogen *Pseudomonas aeruginosa* from the structure of the virulence regulator LasR bound to its autoinducer. J. Biol. Chem. *282*, 13592-13600.
192. Kiratisin, P., Tucker, K.D., and Passador, L. (2002). LasR, a transcriptional activator of *Pseudomonas aeruginosa* virulence genes, functions as a multimer. J. Bacteriol. *184*, 4912-4919.
193. Schuster, M., Urbanowski, M.L., and Greenberg, E.P. (2004). Promoter specificity in *Pseudomonas aeruginosa* quorum sensing revealed by DNA binding of purified LasR. Proc. Natl. Acad. Sci. U.S.A *101*, 15833-15839.
194. Heurlier, K., Dénervaud, V., Haenni, M., Guy, L., Krishnapillai, V., and Haas, D. (2005). Quorum-sensing-negative (*lasR*) mutants of *Pseudomonas aeruginosa* avoid cell lysis and death. J. Bacteriol. *187*, 4875-4883.

195. Diggle, S.P., Griffin, A.S., Campbell, G.S., and West, S.A. (2007). Cooperation and conflict in quorum-sensing bacterial populations. Nature *450*, 411-414.
196. Giraud, C., Bernard, C.S., Calderon, V., Yang, L., Filloux, A., Molin, S., Fichant, G., Bordi, C., and de Bentzmann, S. (2011). The PprA-PprB two-component system activates CupE, the first non-archetypal *Pseudomonas aeruginosa* chaperone-usher pathway system assembling fimbriae. Environ. Microbiol. *13*, 666-683.
197. Filloux, A. (2011). Protein secretion systems in *Pseudomonas aeruginosa*: an essay on diversity, evolution, and function. Front. Microbiol. *2*, 155.
198. Qaisar, U., Luo, L., Haley, C.L., Brady, S.F., Carty, N.L., Colmer-Hamood, J.A., and Hamood, A.N. (2013). The pvc operon regulates the expression of the *Pseudomonas aeruginosa* fimbrial chaperone/usher pathway (*cup*) genes. *PLoS* One *8*, e62735.
199. Finelli, A., Gallant, C.V., Jarvi, K., and Burrows, L.L. (2003). Use of in-biofilm expression technology to identify genes involved in *Pseudomonas aeruginosa* biofilm development. J. Bacteriol. *185*, 2700-2710.
200. Toprak, E., Veres, A., Michel, J.B., Chait, R., Hartl, D.L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. *44*, 101-105.
201. Mwangi, M.M., Wu, S.W., Zhou, Y., Sieradzki, K., de Lencastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E.D., et al. (2007). Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. Proc. Natl. Acad. Sci. U.S.A. *104*, 9451-9456.
202. Buckling, A., and Rainey, P.B. (2002). The role of parasites in sympatric and allopatric host diversification. Nature *420*, 496-499.
203. Suzuki, K., Meek, B., Doi, Y., Muramatsu, M., Chiba, T., Honjo, T., and Fagarasan, S. (2004). Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. Proc. Natl. Acad. Sci. U.S.A. *101*, 1981-1986.
204. Lieberman, T.D., Flett, K.B., Yelin, I., Martin, T.R., McAdam, A.J., Priebe, G.P., and Kishony, R. (2014). Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. Nat. Genet. *46*, 82-87.
205. Lieberman, T.D., Michel, J.B., Aingaran, M., Potter-Bynoe, G., Roux, D., Davis, M.R., Skurnik, D., Leiby, N., LiPuma, J.J., Goldberg, J.B., et al. (2011). Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nat. Genet. *43*, 1275-1280.
206. Young, B.C., Golubchik, T., Batty, E.M., Fung, R., Larner-Svensson, H., Votintseva, A.A., Miller, R.R., Godwin, H., Knox, K., Everitt, R.G., et al. (2012). Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. Proc. Natl. Acad. Sci. U.S.A. *109*, 4550-4555.
207. Sassetti, C.M., and Rubin, E.J. (2003). Genetic requirements for mycobacterial survival during infection. Proc. Natl. Acad. Sci. U.S.A. *100*, 12989-12994.
208. Billal, D.S., Feng, J., Leprohon, P., Legare, D., and Ouellette, M. (2011). Whole genome analysis of linezolid resistance in *Streptococcus pneumoniae* reveals resistance and compensatory mutations. BMC Genomics *12*, 512.
209. Niemann, S., Koser, C.U., Gagneux, S., Plinke, C., Homolka, S., Bignell, H., Carter, R.J., Cheetham, R.K., Cox, A., Gormley, N.A., et al. (2009). Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints. PLoS ONE *4*.
210. Lenski, R.E. (1991). Quantifying fitness and gene stability in microorganisms. Biotechnology *15*, 173-192.
211. Bragg, J.G., and Wagner, A. (2009). Protein material costs: single atoms can make an evolutionary difference. TIG *25*, 5-8.
212. Shachrai, I., Zaslaver, A., Alon, U., and Dekel, E. (2010). Cost of unneeded proteins in *E. coli* is reduced after several generations in exponential growth. Mol. Cell *38*, 758-767.
213. Wellner, A., and Gophna, U. (2008). Neutrality of foreign complex subunits in an experimental model of lateral gene transfer. Mol. Biol. Evol. *25*, 1835-1840.
214. Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010). Mobility of plasmids. Microbiol. Mol. Biol. Rev. *74*, 434-452.

215. San Millan, A., Heilbron, K., and MacLean, R.C. (2013). Positive epistasis between co-infecting plasmids promotes plasmid survival in bacterial populations. ISME J. *8*, 601-612.
216. De Gelder, L., Williams, J.J., Ponciano, J.M., Sota, M., and Top, E.M. (2008). Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetics *178*, 2179-2190.
217. Gamas, P., Burger, A.C., Churchward, G., Caro, L., Galas, D., and Chandler, M. (1986). Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. Mol. Gen. Genet. *204*, 85-89.
218. Ingmer, H., Miller, C., and Cohen, S.N. (2001). The RepA protein of plasmid pSC101 controls *Escherichia coli* cell division through the SOS response. Mol. Microbiol. *42*, 519-526.
219. Lahti, R., Perälä, M., Heikinheimo, P., Pitkäranta, T., Kukko-Kalske, E., and Heinonen, J. (1991). Characterization of the 5' flanking region of the *Escherichia coli ppa* gene encoding inorganic pyrophosphatase: mutations in the ribosome-binding site decrease the level of *ppa* mRNA. J. Gen. Microbiol. *137*, 2517-2523.
220. Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. *3*, 318-356.
221. Aramaki, H., and Fujita, M. (1999). *In vitro* transcription analysis of *rpoD* in *Pseudomonas aeruginosa* PAO1. FEMS Microbiol. Lett. *180*, 311-316.
222. Wrande, M., Roth, J., and Hughes, D. (2008). Accumulation of mutants in "aging" bacterial colonies is due to growth under selection, not stress-induced mutagenesis. Proc. Natl. Acad. Sci. U.S.A. *105*, 11863-11868.
223. Macvanin, M., Bjorkman, J., Eriksson, S., Rhen, M., Andersson, D.I., and Hughes, D. (2003). Fusidic acid-resistant mutants of *Salmonella enterica* serovar Typhimurium with low fitness *in vivo* are defective in RpoS induction. Antimicrob. Agents Chemother. *47*, 3743-3749.
224. Azhikina, T., Skvortsov, T., Radaeva, T., Mardanov, A., Ravin, N., Apt, A., and Sverdlov, E. (2010). A new technique for obtaining whole pathogen transcriptomes from infected host tissues. Biotechniques *48*, 139-144.
225. Davies, J., Spiegelman, G.B., and Yim, G. (2006). The world of subinhibitory antibiotic concentrations. Curr. Opin. Microbiol. *9*, 445-453.
226. Nonejuie, P., Burkart, M., Pogliano, K., and Pogliano, J. (2013). Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. Proc. Natl. Acad. Sci. U.S.A. *110*, 16169-16174.
227. Potvin, E., Sanschagrin, F., and Levesque, R.C. (2008). Sigma factors in *Pseudomonas aeruginosa*. FEMS Microbiol. Rev. *32*, 38-55.
228. Dötsch, A., Eckweiler, D., Schniederjans, M., Zimmermann, A., Jensen, V., Scharfe, M., Geffers, R., and Haussler, S. (2012). The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. PLoS ONE *7*, e31092.
229. Wagner, V.E., and Iglewski, B.H. (2008). *P. aeruginosa* biofilms in CF infection. Clin. Rev. Allergy Immunol. *35*, 124-134.
230. Goss, C.H., and Muhlebach, M.S. (2011). Review: *Staphylococcus aureus* and MRSA in cystic fibrosis. J. Cyst. Fibros. *10*, 298-306.
231. Garske, L.A., Kidd, T.J., Gan, R., Bunting, J.P., Franks, C.A., Coulter, C., Masel, P.J., and Bell, S.C. (2004). Rifampicin and sodium fusidate reduces the frequency of methicillin-resistant *Staphylococcus aureus* (MRSA) isolation in adults with cystic fibrosis and chronic MRSA infection. J. Hosp. Infect. *56*, 208-214.
232. Drummond, D.A., and Wilke, C.O. (2009). The evolutionary consequences of erroneous protein synthesis. Nat. Rev. Genet. *10*, 715-724.
233. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. *55*, 611-622.
234. Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. *177*, 4121-4130.

235. Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature *405*, 299-304.
236. Wiedenbeck, J., and Cohan, F.M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiol. Rev. *35*, 957-976.
237. Carattoli, A. (2013). Plasmids and the spread of resistance. Int. J. Med. Microbiol. *303*, 298-304.
238. Alekshun, M.N., and Levy, S.B. (2007). Molecular mechanisms of antibacterial multidrug resistance. Cell *128*, 1037-1050.
239. Koraimann, G., and Wagner, M.A. (2014). Social behavior and decision making in bacterial conjugation. Front. Cell. Infect. Microbiol. *4*, 54.
240. Canchaya, C., Fournous, G., and Brüssow, H. (2004). The impact of prophages on bacterial chromosomes. Mol. Microbiol. *53*, 9-18.
241. Pál, C., Papp, B., and Lercher, M.J. (2005). Horizontal gene transfer depends on gene content of the host. Bioinformatics *21 Suppl 2*, ii222-223.
242. Bagdasarian, M., Bailone, A., Bagdasarian, M.M., Manning, P.A., Lurz, R., Timmis, K.N., and Devoret, R. (1986). An inhibitor of SOS induction, specified by a plasmid locus in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. *83*, 5723-5726.
243. Gonçalves, G.A., Bower, D.M., Prazeres, D.M., Monteiro, G.A., and Prather, K.L. (2012). Rational engineering of *Escherichia coli* strains for plasmid biopharmaceutical manufacturing. Biotechnol. J. *7*, 251-261.
244. Geiler-Samerotte, K.A., Dion, M.F., Budnik, B.A., Wang, S.M., Hartl, D.L., and Drummond, D.A. (2011). Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. Proc. Natl. Acad. Sci. U.S.A. *108*, 680-685.
245. Tiburzi, F., Imperi, F., and Visca, P. (2008). Intracellular levels and activity of PvdS, the major iron starvation sigma factor of *Pseudomonas aeruginosa*. Mol. Microbiol. *67*, 213-227.
246. Bi, E., and Lutkenhaus, J. (1993). Cell division inhibitors SulA and MinCD prevent formation of the FtsZ ring. J. Bacteriol. *175*, 1118-1125.
247. Rankin, D.J., Rocha, E.P., and Brown, S.P. (2011). What traits are carried on mobile genetic elements, and why? Heredity *106*, 1-10.
248. Haines, A.S., Jones, K., Cheung, M., and Thomas, C.M. (2005). The IncP-6 plasmid Rms149 consists of a small mobilizable backbone with multiple large insertions. J. Bacteriol. *187*, 4728-4738.
249. Sen, D., Van der Auwera, G.A., Rogers, L.M., Thomas, C.M., Brown, C.J., and Top, E.M. (2011). Broad-host-range plasmids from agricultural soils have IncP-1 backbones with diverse accessory genes. Appl. Environ. Microbiol. *77*, 7975-7983.
250. Tato, M., Coque, T.M., Baquero, F., and Cantón, R. (2010). Dispersal of carbapenemase bla_{VIM-1} gene associated with different Tn*402* variants, mercury transposons, and conjugative plasmids in *Enterobacteriaceae* and *Pseudomonas aeruginosa*. Antimicrob. Agents Chemother. *54*, 320-327.
251. Patel, R.K., and Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE *7*, e30619.
252. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595.
253. Zheng, Q., and Wang, X.J. (2008). GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res. *36*, W358-363.
254. Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, P3.
255. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389-3402.
256. Leinonen, R., Sugawara, H., Shumway, M., and Collaboration, I.N.S.D. (2011). The sequence read archive. Nucleic Acids Res. *39*, D19-21.
257. Unden, G., Becker, S., Bongaerts, J., Schirawski, J., and Six, S. (1994). Oxygen regulated gene expression in facultatively anaerobic bacteria. Antonie van Leeuwenhoek *66*, 3-22.

258. Zumft, W.G. (1997). Cell biology and molecular basis of denitrification. Microbiol. Mol. Biol. Rev. *61*, 533-616.
259. Arai, H. (2011). Regulation and function of versatile aerobic and anaerobic respiratory metabolism in *Pseudomonas aeruginosa*. Front. Microbiol. *2*, 103.
260. Chen, F., Xia, Q., and Ju, L.K. (2003). Aerobic denitrification of *Pseudomonas aeruginosa* monitored by online NAD(P)H fluorescence. Appl. Environ. Microbiol. *69*, 6715-6722.
261. Van Alst, N.E., Wellington, M., Clark, V.L., Haidaris, C.G., and Iglewski, B.H. (2009). Nitrite reductase NirS is required for type III secretion system expression and virulence in the human monocyte cell line THP-1 by *Pseudomonas aeruginosa*. Infect. Immun. *77*, 4446-4454.
262. Kawasaki, S., Arai, H., Igarashi, Y., and Kodama, T. (1995). Sequencing and characterization of the downstream region of the genes encoding nitrite reductase and cytochrome c-551 (*nirSM*) from *Pseudomonas aeruginosa*: identification of the gene necessary for biosynthesis of heme d1. Gene *167*, 87-91.
263. Kawasaki, S., Arai, H., Kodama, T., and Igarashi, Y. (1997). Gene cluster for dissimilatory nitrite reductase (*nir*) from *Pseudomonas aeruginosa*: sequencing and identification of a locus for heme d1 biosynthesis. J. Bacteriol. *179*, 235-242.
264. Gudat, J.C., Singh, J., and Wharton, D.C. (1973). Cytochrome oxidase from *Pseudomonas aeruginosa*. I. Purification and some properties. Biochim. Biophys. Acta *292*, 376-390.
265. Kuronen, T., and Ellfolk, N. (1972). A new purification procedure and molecular properties of *Pseudomonas* cytochrome oxidase. Biochim. Biophys. Acta *275*, 308-318.
266. Hasegawa, N., Arai, H., and Igarashi, Y. (2001). Two c-type cytochromes, NirM and NirC, encoded in the nir gene cluster of *Pseudomonas aeruginosa* act as electron donors for nitrite reductase. Biochem. Biophys. Res. Commun. *288*, 1223-1230.
267. Skurnik, D., Roux, D., Aschard, H., Cattoir, V., Yoder-Himes, D., Lory, S., and Pier, G.B. (2013). A comprehensive analysis of *in vitro* and *in vivo* genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. PLoS Pathog. *9*, e1003582.
268. Jüngst, A., and Zumft, W.G. (1992). Interdependence of respiratory NO reduction and nitrite reduction revealed by mutagenesis of *nirQ*, a novel gene in the denitrification gene cluster of *Pseudomonas stutzeri*. FEBS Lett. *314*, 308-314.
269. Arai, H., Igarashi, Y., and Kodama, T. (1994). Structure and ANR-dependent transcription of the *nir* genes for denitrification from *Pseudomonas aeruginosa*. Biosci. Biotechnol. Biochem. *58*, 1286-1291.
270. Hayashi, N.R., Arai, H., Kodama, T., and Igarashi, Y. (1998). The *nirQ* gene, which is required for denitrification of *Pseudomonas aeruginosa*, can activate the RubisCO from *Pseudomonas hydrogenothermophila*. Biochim. Biophys. Acta *1381*, 347-350.