Fitness effects and resistance mechanisms of beneficial mutations in *Pseudomonas aeruginosa*

by

Barbara Kubica

**THESIS SUBMITTED TO THE
FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE**

**DEPARTMENT OF BIOLOGY
UNIVERSITY OF OTTAWA**

Acknowledgements

Great things are not done by impulse, but by a series of small things brought together.
–Vincent van Gogh

Over the past few years, there have been several individuals who have supported and assisted me during my graduate endeavours at the University of Ottawa. Most importantly, I owe my most sincere gratitude to my supervisor, Rees Kassen, for his mentorship, guidance, expertise, and patience. Rees – it has been both an honour and a privilege to be one of your students. Thank you for giving me so much of your valuable time, for continually challenging me with many aspects of evolutionary biology, and for allowing me the space and freedom to pursue my own interests at the science-policy interface. By doing so, you have allowed me to become a better researcher and science critic.

Thank you to my committee members, Scott Findlay and Thien-Fah Mah from the University of Ottawa and Mark Forbes from Carleton University, for your counsel, insightfulness, and feedback on my progress throughout my degree. Scott – a very special thank you for also providing guidance in my final years of my BSc and supporting my entrance into graduate studies.

To the current and past members of Rees Kassen's lab, notably Susan Bailey, Jeremy Dettman, Anita Melnyk, Mariam Taha, and Alex Wong, thank you for your encouragement, advice, and most importantly, camaraderie. Having spent so much time together, you have been my backbone – helping me fine-tune my research protocols, assist me with various experiments, and provide endless statistical help. Nonetheless, I am most indebted to the lasting friendships I have made with each of you which have been a very important part of my graduate life at the University of Ottawa.

Further thanks go to my many colleagues in Howard Rundle's lab, including Devin Arbuthnott, Matthieu Delcourt, Jacqueline Sztepanacz, and Alison White, for your optimism, pleasantries, and shared conversations during lunch, in the office, and beyond.

Financial support for my degree was provided by a scholarship from the University of Ottawa. In addition, much of the research expenses, were generously covered by a grant from the Natural Sciences and Engineering Research Council of Canada to Rees Kassen.

Finally, I am grateful for the unconditional support of my parents, to whom this is dedicated.

Barbara Kubica
Ottawa, July 2012

Abstract

Beneficial mutations are the fuel for adaptation but remain poorly studied. Extreme value theory has been used extensively as a model for understanding the distribution of fitness effects (DFE) among beneficial mutations prior to selection, however there is little empirical data available to test the most compelling predictions of the theory. This thesis aims to test the prediction that the DFE among beneficial mutations is a negative exponential, with many mutations of small effect and few of large effect. To do this, I collected a wide range of *Pseudomonas aeruginosa* mutants resistant to the antibiotic ciprofloxacin and measured their fitness in the absence of the drug. I reject the exponential as an adequate descriptor of the data and show that the DFE among beneficial mutations is better described as being located in the Weibull domain of attraction, which is characterised by having a right-hand bound on the magnitude of the largest fitness effects. I also sequenced a number of genes known to be targets of fluoroquinolone resistance to shed light on the range and number of genetic targets involved in resistance. I found a number of mutations, and often multiple mutations, in known genetic targets. Evidence suggests that mutations in regions of the genome not sequenced are also important in determining fitness under permissive conditions. Taken together, these results provide valuable insight into one of the most fundamental problems in evolution, the nature and fitness effects of beneficial mutations.

Résumé

Les mutations bénéfiques sont primordiales pour l'adaptation, pourtant elles demeurent encore peu étudiées. La théorie des valeurs extrêmes a été largement utilisée comme modèle pour comprendre la distribution de la fréquence génique (DFG) des mutations bénéfiques avant la sélection. Néanmoins, il existe peu de données empiriques testant les prédictions les plus convaincantes de cette théorie. Cette thèse vise à tester la prédiction que la DFG des mutations bénéfiques suit une loi exponentielle négative, avec de nombreuses mutations de faible effet et peu de mutations à grand effet. Pour ce faire, j'ai recueilli un large éventail de souches *Pseudomonas aeruginosa* résistantes à l'antibiotique ciprofloxacine et mesuré leurs valeurs sélectives en absence de l'antibiotique. Je rejette l'hypothèse de l'exponentielle comme descripteur adéquat des données et montre que la DFG des mutations bénéfiques est mieux décrite comme étant située dans le domaine d'attraction de Weibull, qui est caractérisé par une courbe dont la concavité est orientée en bas à droite sur l'ampleur des plus grandes valeurs sélectives. J'ai également séquencé un nombre de gènes identifiés comme étant des cibles potentiels de la résistance aux fluoroquinolones pour éclairer la portée et le nombre de déterminants génétiques impliqués dans la résistance antibiotique. J'ai trouvé plusieurs mutations ponctuelles, ainsi que des mutations successives, localisées dans des cibles génétiques connues. Ces expériences suggèrent que les mutations ayant lieu dans des régions du génome non-séquencées sont aussi importantes pour la détermination de la valeur sélective sous des conditions permissives. De manière générale, ces résultats fournissent des paramètres clés pour comprendre un des problèmes fondamentaux en évolution, la nature et les effets des mutations bénéfiques.

Contents

List of tables

List of figures

List of abbreviations and symbols

A	population carrying capacity
AIC	Akaike's information criterion
ANOVA	analysis of variance
CF	cystic fibrosis
CI	confidence interval
DFE	distribution of fitness effects
E	environment
EVT	extreme value theory
FGM	Fisher's geometric model
G	genotype
GPD	generalized Pareto distribution
κ ($\hat{\kappa}$)	shape parameter
LB	Luria-Bertrani medium
LogL	log-likelihood
LRT	likelihood ratio test
μ	maximum linear population growth rate
MEE	microbial experimental evolution
MIC	minimum inhibitory concentration
MLM	mutational landscape model
n	number of phenotypic trait values under selection
NCBI	National Institute for Biotechnology Information
O_S, O_P	optimum under selective and permissive conditions
OD_{630}	optical density at 630 nm, i.e. absorbance
PA14	*Pseudomonas aeruginosa*, strain UCBPP-PA14
QRDR	quinolone-resistance determining region
ρ	Spearman's rank correlation coefficient
rpm	revolutions per minute
s_m	selection coefficient of the mutant founder
s_{opt}	distance from fitness optimum
SCFM	synthetic cystic fibrosis medium
std	standard deviation
r	phenotypic effect of a mutation
τ ($\hat{\tau}$)	scale parameter
w	mutant fitness
w_m	fitness of the mutant founder

Aims and outline of the thesis

The fixation of beneficial mutations fuels adaptation while enabling organisms to adapt to stressful conditions, as is the case in nearly all opportunistic human pathogens surviving exposure to antibiotics. Understanding the phenotypic and genotypic changes occurring among beneficial mutations in novel conditions, i.e. antibiotic- supplemented and free environments, is critical in circumventing the growing concerns associated with the emergence and persistence of antibiotic resistance.

This thesis aims to contribute to the empirical work on the nature and fitness of beneficial mutation in the pathogenic *Pseudomonas* species as a model. Chapter 1 reviews theoretical models of adaptation, restricting much attention to the distribution of fitness effects (DFE) associated with beneficial mutations, and exemplifies microbial experimental evolution studies, which have focused on this more limited body of theory. While the rarity of beneficial mutations makes it difficult to empirically evaluate theories of the DFE among beneficial mutations in natural populations, Chapter 2 details a two-pronged experimental and genomic approach using *P. aeruginosa* that (a) applies a likelihood framework to evaluate the DFE among beneficial mutations, and (b) screens for putative genetic mechanisms conferring resistance to a quinolone antibiotic.

Chapter 2 was written as a stand-alone scientific manuscript in the style of an academic journal article. As such, there is some overlap with the background material presented in Chapter 1, although it was endeavoured to keep this to a minimum. The purpose of Chapter 1 was to provide a more in-depth treatment of the subject than was possible in the manuscript form.

CHAPTER 1

A review of the distribution of pleiotropic fitness effects among beneficial mutations

Introduction

The process of adaptive evolution relies on the origination of mutations through natural selection, generating advantageous new traits that survive and multiply and neutral or disadvantageous traits that tend to die out over time. Modelling evolutionary dynamics of adaptation requires that we be able to predict and estimate the distribution of fitness effects (DFE) among mutations exposed to selection—including those which are beneficial, neutral, and deleterious. Nonetheless, the rarity of beneficial mutations in natural populations has imposed a challenge for characterizing the complete DFE.

Two theoretical approaches, one based on phenotypes and the other based on DNA sequences, have been used to try to provide some first insight into the DFE among beneficial mutations (Orr 2005). While the frequency of mutations and genomic regions affected differ between species, theorists have used both models to conceptualize some general principles, and reasonable predictions, about the expected DFE. Even so, both models have failed to consider the other explicitly. Adaptive theories based on the evolution of phenotypes are not embedded in a clear genotypic context; concurrently, those based on DNA sequence differences are not embedded in a clear phenotypic context (Chevin et al. 2010).

The DFE imposed by the accumulation of new mutations is important far beyond our intrinsic interest of understanding biological adaptation. Insofar as the fixation (and functional form) of new mutations shapes and increases diversity, the DFE among beneficial mutations has implications at the fundamental level on the rate and extent of adaptation (Orr 1998). Further, it applies in a range of problems associated with the quantitative genetics of mutations, including patterns of inheritance (linkage),

inbreeding, clonal interference, species hybridization, and adaptation of DNA sequences (Rokyta et al. 2008). Even more applied contexts relate to medical settings where the emergence and persistence of antibiotic resistance has increasingly hampered therapeutic options against pathogenic infections (MacLean et al. 2010a; MacLean et al. 2010b). In fact, the optimization of strategies to control or eliminate antibiotic resistance depends of our understanding of the extent and fitness effects of beneficial mutations available to a human pathogen.

Here, we review theoretical models of adaptation—pointedly restricting our attention to the DFE associated with beneficial mutations—and highlight some recent empirical studies, which have looked at the (pleiotropic) fitness effects among adaptive mutations in pathogenic microbial populations.

Theoretical models of adaptation

The DFE among new mutations has been studied along two directions, one focusing on phenotype and the other on DNA sequences (Orr 2005). Fisher (1930) developed one of the first mathematical frameworks to contribute to our understanding of phenotypic adaptive evolution, which has come to be known as Fisher's geometric model (FGM). In this model, organisms are located in an n-dimensional 'phenotypic space', where n represents the number of orthogonal phenotypic trait values under selection. The 'fittest' phenotype is located at the intersection of a combination of locally optimal n values relevant to organismal fitness, also known as the phenotypic optimum. The assumption is that the wild type is already fairly well adapted, such that its distance from the local phenotypic optimum is small. Mutations that are beneficial bring a phenotype closer to the optimum, while those that are deleterious move the phenotype further away. The probability p that a mutation is favourable, thus moving an organism towards the optimum, depends on the phenotypic effect of the mutation r and the distance to the optimum from the starting genotype s_{opt} (Hamilton 2009). This can be described by a probability density function:

$$p = \frac{1}{2}\left(1 - \frac{r}{s_{opt}}\right)$$

where FGM assumes a continuous probability distribution (e.g. Gaussian) where organismal fitness declines with distance to the optimum. The model predicts that mutations of infinitesimally small r have a probability of being beneficial that is approximately ½. Larger effect mutations (those with higher r) have a lower probability of being beneficial. FGM also illustrates a net phenotypic effect where mutations affect many or all of the traits related to organismal fitness, a phenomenon known as pleiotropy. The general consensus is that it is harder for a mutation to be beneficial when it affects more traits (i.e. it is more pleiotropic).

While the Fisherian model demonstrates how evolution proceeds favouring beneficial mutations of small effect along a continuously-valued random phenotypic trait distribution, it fails to capture two factors affecting the quantitative genetics of mutations: (i) the finite number of mutational events available for selection, and (ii) the shape of the DFE among mutations. Sequence-space models, most notably Gillespie's mutational landscape model (MLM), do however address these points. The MLM reflects the genetic basis of adaptation (i.e. transitions between nucleotide substitutions) and assumes that in a beneficial sequence, genotypes can mutate by a *single* mutational step only (Betancourt & Bollback 2006). For a hypothetical DNA sequence L, where three possible nucleotide (mutation) sites exist, distance to the fitness optimum is reduced when mutations are fixed at any of the $3L$ sites conferring greater fitness. As in FGM, the frequency of fitness effects among new mutations can thus be approximated by a continuous probability distribution. The most predominant limitation of Gillespie's model is however its unaccountability of the fraction of mutations with infinitesimally small effects (Betancourt & Bollback 2006).

Gillespie (1984) proposed that while the overall shape of the underlying fitness distribution of new mutations remains unknown, the combined facts that beneficial mutations are rare and that the wild type allele is usually of high fitness, might allow predictions about the DFE among beneficial mutations. As a population evolves from a given starting genotype, mutations having higher fitness relative to the ancestor are

considered beneficial and drawn from the extreme right tail of an underlying fitness distribution of mutations (Eyre-Walker & Keightley 2007). Gillespie (1983, 1984, 1991) used a branch of statistics dealing with the extreme draws (or 'extreme spacings') of probability distributions—extreme value theory (EVT)—to characterize the statistical properties of the DFE among beneficial mutations. Applications of EVT in genetics assume, *a priori*, that the wild type is already fairly well adapted to a given environment (i.e. close to the optimum), and that single-step beneficial mutations are even fitter than the wild type and so drawn from the rightmost tail. As such, the distribution of beneficial mutations should exhibit one of three tail behaviours (Orr 2005), illustrated in Figure 1.1.

The prediction under EVT is that the DFE among beneficial mutations will be exponentially distributed above a threshold (i.e. the wild type), when mutations fall in the so-called Gumbel domain of attraction (Orr 2002, 2003, 2005) and invariant—the DFE will be exponentially distributed regardless of the fitness rank of the wild type. Two other forms for the behaviour of right tails of distributions exist; the Fréchet domain characterized by a truncated 'heavier' upper tail, i.e. one falling off more slowly than the exponential, and the Weibull domain characterized by a 'lighter' upper tail than the exponential (Beisel et al. 2007). While these domains of attraction are possible descriptors of the DFE among beneficial mutations, the Weibull and Fréchet types, they have both been criticized as being inappropriate biologically (Orr 2006). Joyce et al. (2008) have nonetheless generalized the theory to include the possibility of the DFE to fall into any of the three domains of attraction using the generalized Pareto distribution (GPD). The GPD encompasses all three tail behaviours under EVT (Pickands 1975) and the cumulative distribution function is given by:

$$F(x \mid \kappa, \tau) = \begin{cases} 1-(1+\kappa x/\tau)^{-1/\kappa}, x \geq 0, \text{if } k > 0 & \text{Fréchet} \\ 1-(1+\kappa x/\tau)^{-1/\kappa}, 0 \leq x < -\dfrac{t}{k}, \text{if } k < 0 & \text{Weibull} \\ 1-e^{-\kappa/\tau}, x \geq 0, \text{if } k = 0 & \text{Gumbel} \end{cases}$$

where the form of the GPD is determined by two parameters: scale τ and shape κ. The weight of the tail, characterized by the shape parameter, corresponds to the Gumbel

domain when $\kappa = 0$ (Gumbel 1958), the Fréchet domain when $\kappa > 0$ (Leadbetter et al. 1983), and the Weibull domain when $\kappa < 0$ (Embrechets et al. 1997).

Building on the "gradualist" view that it is a small fraction of favourable mutations that make up the "stuff" of evolution, Orr suggests that beneficial mutations will be less frequent under more 'complex' conditions since they must also escape stochastic loss (or drift) (Orr 1998; Orr 2000). As a population evolves from a given starting genotype, its evolutionary trajectory will depend on the starting fitness rank of the wild type (i.e. its distance from a fitness optimum) and the probability of beneficial mutations to survive drift while at low frequencies (2002).

Martin and Lenormand (2006a,b) used a generalized version of FGM to extend Orr's work into a multivariate Gaussian framework. On the basis of arbitrary selective and mutational interactions, Martin and Lenormand contend that the shape of the DFE is dependant on (i) the distance to the optimum, (ii) the direction to the optimum, and (iii) the number of phenotypic traits under selection, and varies in a predictable way across species and environments. Notably, although the DFE among beneficial mutations is expected to be beta-distributed in general, when selection acts on a large number of traits simultaneously the DFE becomes exponential.

Most recently, considerable focus has been placed on the role of pleiotropy associated with beneficial mutations (e.g. do beneficial mutations that improve fitness in one environment also improve fitness in a novel environment?). Understanding the prevalence, magnitude, and form of pleiotropy among beneficial mutations is of particular importance when considering the evolution of niche specialization (resource use), limits to adaptation, and extent of organismal complexity (Ostrowski et al. 2005; Wang et al. 2010). In terms of applied considerations, this is again relevant to the growing concerns of antibiotic resistance where beneficial mutations conferring resistance may be paying little or no (fitness) costs in permissive (antibiotic-free) environments. In this instance, determining the fraction of resistance mutations that are also beneficial in the absence of antibiotic is critical in designing effective therapies to treat pathogenic infections. If in fact there is little or no cost associated with the fixation of beneficial mutations in permissive environments, implying positive pleiotropic fitness effects

among these environments, than the practice of discontinuing antibiotic treatment once resistance has (emerged and) evolved will be ineffective. In parallel with the predictions under EVT, the question then arises, whether pleiotropic fitness effects are also correlated in magnitude, i.e. jointly exponentially distributed?

Martin et al. (in prep) have posited that the correlated effects of beneficial mutations in novel environments depend in a predictable way on the evolutionary history of the wild type in the given environment. Martin et al.'s fitness landscape model is an extension of FGM to include multiple optima. In the simplest case the model considers two environments (selective *s* and permissive *p*) and assumes that each has a distinct phenotypic optimum (O_S and O_P in the selective and permissive environments, respectively). The pleiotropic fitness effect in the second environment of a mutation beneficial in a focal environment thus depends on the position of the starting genotype relative to each optima, as illustrated in Figure 1.2. Geometrically, the fitness correlation ρ between mutations arising in two environments is approximately equal to the cosine of the angle θ between the wild type and the selective gradients towards each environment. The correlated fitness effect will be beneficial when ρ is positive ($\theta < \pi/2$) and deleterious when ρ is negative ($\theta > \pi/2$).

The classic view is that a single mutation—whether it be beneficial, neutral, or deleterious—affects all phenotypic traits, also known as universal pleiotropy. Empirical studies suggest however that there may be a more succinct link between phenotypic traits and genotypic changes, a phenomenon known as (restricted) modular pleiotropy (Wagner & Zhang 2011). Modular pleiotropy focuses on 'modules', where specific sets of genes have pleiotropic effects on the same set of phenotypic traits (Wagner et al. 2007). In this regard, molecular genetics may influence the underlying fitness distribution such that identifying phenotype-genotype associations between both trait and gene modules may be equally critical in understanding the DFE among beneficial mutations.

Empirical studies of adaptation

Both the phenotypic and genotypic models discussed above have been used to frame questions and derive predictions regarding two key features of the DFE among beneficial

mutations: the number of mutations available for selection (mutational neighbourhood and mutation supply rate) and the functional form (or shape) of the distribution. Here, we necessarily focus on testing the latter facet empirically (e.g. is the DFE among beneficial mutations exponentially distributed?) as it is critical to a broad range of issues in adaptive evolution. Due to their rarity and difficult detectability in nature, properties of beneficial mutations have however remained poorly understood and the theoretical expectations of the DFE have been tested much less frequently.

An experimental approach known as microbial experimental evolution (MEE) has permitted directly observing the first step of adaptation. MEE studies measure quantitatively the extent of adaptation. Large population sizes and short generation times make bacteria ideal model systems for MEE studies. Additionally, bacteria may be cryogenically frozen in non-evolving states allowing for future (additional) replication of experiments.

One MEE approach used to infer the DFE among beneficial mutations involves collecting spontaneous mutations among bacterial populations through fluctuation assays (Luria & Delbrück 1943). Fluctuation-style experiments challenge an isogenic, sensitive bacterial base population to adapt to a strong selective pressure (e.g. antibiotics), allowing for the isolation of independent clones under controlled laboratory conditions. Initially, a large number of isogenic cells are spread onto agar plates containing a selective pressure such as an antibiotic (MacLean et al. 2010). Any viable colonies that grow represent individual resistant strains, which are, by definition, beneficial in the presence of the drug. In laboratory experiments, resistance evolution is usually acquired through the accumulation of chromosomal mutations provided there are no plasmids in the founding strain(s). The changes in DNA sequence induced by resistance mechanisms often, though not always, carry a physiological cost, generally reflected by reduced bacterial growth (fitness) (Andersson & Hughes 2010; Sousa et al. 2012; Trindade et al. 2010). Assaying the fitness of collected mutations in the selected, as well as novel environments, may thus provide insight into which mutations are likely to be of clinical importance, those displaying cost-free resistance and/or positive pleiotropic effects. Specifically, it is possible to address the following questions: (1) How does bacterial growth among beneficial

mutations differ across environments? In particular, what are the pleiotropic fitness (and mutational) effects in novel environments? (2) Is the fraction of beneficial mutations jointly exponentially distributed across environments? (3) Are mutational events associated with an imposed selective pressure similar among beneficial mutations?

Identifying the relative frequency of fitness effects among beneficial mutations allows us to test the most compelling predictions of the theory that the DFE among beneficial mutations is (i) invariant regardless of the fitness rank of the wild type, and (ii) exponentially distributed (across environments) according to EVT. To date, testing the accuracy of these predictions has yielded conflicting results.

In the following examples, experimentalists challenged sensitive ancestral bacterial or viral strains in conventional fluctuation-style assays to generate a large library of independently evolved antibiotic-resistant genotypes conferring increased fitness under permissive conditions. Kassen and Bataillon (2006) observed exponentially distributed fitness effects among mutations conferring resistance to the antibiotic nalidixic acid in *Pseudomonas fluorescens* in both permissive and selective (antibiotic-supplemented) environments. Using *Pseudomonas aeruginosa*, MacLean and Buckling (2009) observed exponential distributions among mutations conferring resistance to the antibiotic rifampicin only when fitness of the wild type was high. When fitness of the wild type was low, the DFE was no longer exponentially distributed, violating the assumption under EVT that the starting genotype is usually of high fitness. There is however no statistical theory that predicts the shape of the DFE among beneficial mutations when the wild type is of low fitness. The exponential has been further rejected by Roktya et al. (2008) and Bataillon et al. (2011) who showed that the Weibull domain of attraction best described the DFE among beneficial mutations in two viral data sets and *P. fluorescens* mutants, respectively. One caveat is that Roktya et al. (2008) used viral data sets rather than bacterial ones, as in afore-mentioned studies, and it is still unclear whether DFEs vary among organisms and/or environments. Taken together, these studies suggest that, in reality, the shape of the DFE among beneficial mutations remains difficult to predict and requires more empirical attention.

Further, it appears as though no published predictive theory exists to date on the pleiotropic effects of beneficial mutations in novel environments. What has been observed empirically, is that beneficial mutations that increase fitness in one environment will not necessarily decrease it in a novel environment (Ostrowski et al. 2005; Fitzsimmons et al. 2009; Kassen & Bataillon 2006). Research by Ostrowski et al. (2005) on a collection of thirty *Escherichia coli* genotypes evolved in a glucose limiting-environment showed that mutations that are beneficial in this environment were also beneficial in four of the five other (carbon source) environments. Likewise, positive pleiotropic fitness effects are often the case with antibiotic resistance where mutations that confer resistance to one antibiotic are beneficial in both the presence and absence of the drug. Kassen and Bataillon (2006) amassed a collection of independently derived strains of *P. fluorescens*, containing mutations conferring resistance to nalidixic acid. This collection of mutants contained a large amount of genetic variation reflected in a large range of fitness costs to resistance, including a substantial number of mutants that were fitter than the sensitive wild type under novel conditions. The results of their study suggest the presence of 'universally superior' genotypes that are both antibiotic resistant and fitter than the sensitive ancestor in the absence of antibiotic. These superior genotypes, arising in a single mutational step, appear to occur far more frequently than previously thought, imposing an obvious challenge for the future treatment of pathogenic infections. This positive pleiotropy has equally been shown by Fitzsimmons et al. (2009), where a collection of sixty-six independently derived *P. fluorescence* genotypes conferring resistance to nalidixic acid exhibited positively correlated parametric growth measures of fitness (growth rate and population carrying capacity) across four novel carbon source limiting-environments. While these studies present patterns of pleiotropy among beneficial mutations in the phenotypic 'space', the genetic targets of these mutants is often elusive.

Limitations and drawbacks of such experimental work do exist, including: the lack of large enough collections of favourable mutations ensuring powerful tests of theoretical predictions, the inability to detect mutations of small effect, and the potential for genetic pseudoreplication among mutations captured using fluctuation-style assays.

Pseudoreplication occurs when collected strains have acquired identical mutations and are then treated as replicate (rather than independent) measurements. Despite that, combining genetic tools (e.g. targeted or whole genome sequencing) with these experiments permits verifying the latter (genetic uniqueness of collected mutations, as required by theory), identifying mutational changes affecting fitness, and ultimately, validating and/or readjusting the DFEs estimated by means of paralleled MEE work.

The end result is that, despite the recent focus on the properties and functional form of new mutations, little is known about the exact DFE among those mutations that are beneficial. Empirically, the exponential and Weibull have been shown as adequate descriptors of the DFE among beneficial mutations while no studies have provided any evidence for a DFE located in the Fréchet domain of attraction. More data still needs to be collected to describe the DFE among beneficial mutations and MEE studies combined with genetic approaches may provide valuable insight into one of the most fundamental problems in evolution, the nature and fitness effects of beneficial mutations.

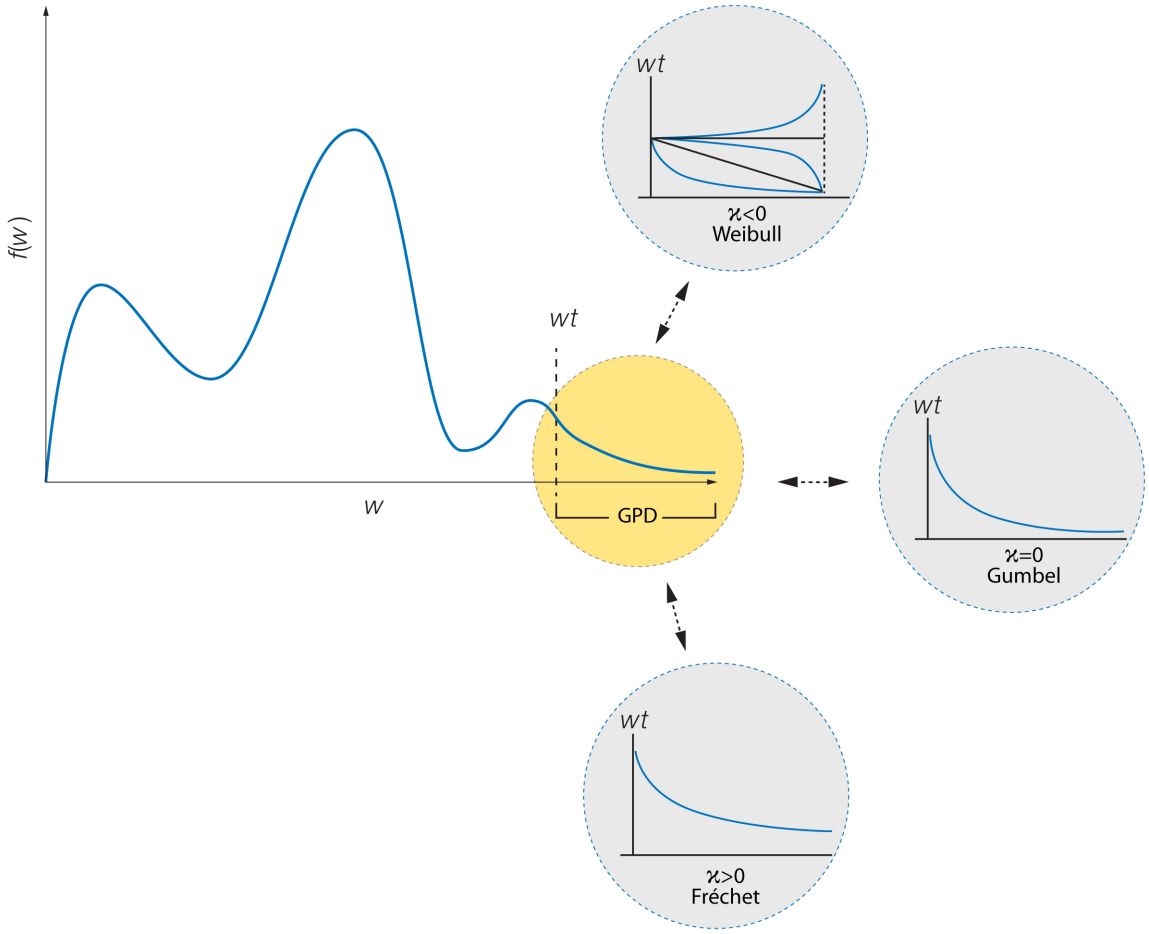

Figure 1.1 | **The distribution of fitness effects among beneficial mutations.** As a population evolves from a given starting genotype, mutations having higher fitness relative to the wild type (*wt*) are considered beneficial and drawn from the extreme rightmost tail of an arbitrary underlying fitness distribution (*w*, fitness value; *f(w)*, frequency of a certain fitness value). According to extreme value theory, a generalized Pareto distribution (GPD) encompasses three possible limiting right tail behaviours, or domains of attraction: the Weibull (truncated), Gumbel (exponential), and Fréchet (heavy tailed).

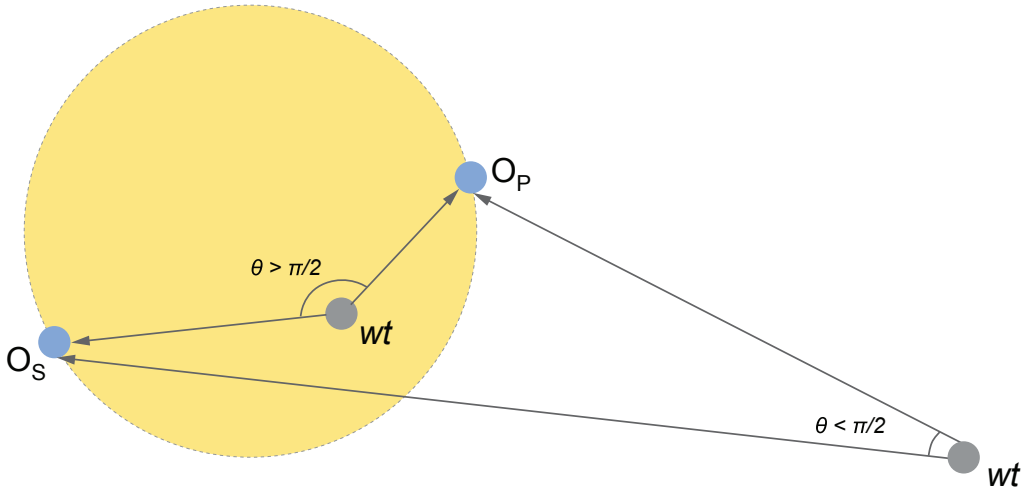

Figure 1.2 | **Geometric interpretation of correlated fitness effects between two environments.** Martin et al.'s fitness landscape model is an extension of Fisher's geometric model to include multiple optima (in prep). In the simplest case the model considers two environments. Fitness is determined by *n*-traits (here, *n*=2; the two dimensions of the plane) and each environment corresponds to a distinct phenotypic optimum (O_S and O_P in the selective and permissive environments, respectively). According to Martin et al. the fitness correlation ρ between mutations arising in both environments is conditioned on the position of the wild type (*wt*) relative to each phenotypic optimum in the fitness landscape. ρ is approximately equal to the cosine of the angle θ between the wild type and the selective gradients towards each environment. The yellow shaded area delimits where ρ is negative ($\theta > \pi/2$) and where there is a trade-off between environments. ρ is positive when $\theta < \pi/2$.

CHAPTER **2**

Fitness effects and resistance mechanisms of beneficial mutations in *Pseudomonas aeruginosa*

Introduction

Beneficial mutations are the fuel for adaptation, but remain poorly understood. Adaptation occurs when increases in fitness among organisms, by which natural selection acts on, lead to the fixation of beneficial mutations over time. Beneficial mutations are thought to be rare with small effects (Fisher 1930), making them difficult to detect in natural populations. Partly because of this, they have been studied much less extensively than deleterious and neutral mutations (Orr 2005), which have a negative or no effect on the fitness of an organism, respectively. Understanding the distribution of fitness effects (DFE), particularly the DFE among beneficial mutations, may serve important in positing a general theory of adaptive evolution (Bataillon et al. 2011).

DFE among beneficial mutations

As a population evolves from a given starting genotype, mutations having higher fitness relative to the ancestor are considered beneficial and drawn from the extreme right tail of an underlying fitness distribution of mutations (Eyre-Walker & Keightley 2007), illustrated in Figure 1.1. While the overall shape of the DFE remains unknown, extreme value theory (EVT) has been used extensively as a model for understanding draws from the rightmost tail of this distribution, that is, the DFE among beneficial mutations, prior to selection. Gillespie (1983, 1984, 1991) posited that the combined facts that beneficial mutations are rare and that the starting genotype is usually of high fitness, allows us to make predictions about the relative frequency of beneficial mutations. According to EVT, three possible limiting tail behaviours, or domains of attraction, exist. The prediction is

that the distribution of fitness values higher than the starting genotype are characterized by a negative exponential tail, with many mutations having a small effect on fitness and few having a large effect (Orr 2003). Nonetheless, two other forms for right tails of distributions exist; the Fréchet domain characterized by a truncated 'heavier' upper tail, i.e. one falling off more slowly than the exponential, and the Weibull domain characterized by a 'lighter' upper tail than the exponential (Beisel et al. 2007).

A generalized Pareto distribution (GPD) encompasses all three possible domains of attraction under EVT (Pickands 1975). The cumulative distribution function is given by:

$$F(x \mid \kappa, \tau) = \begin{cases} 1-(1+\kappa x/\tau)^{-1/\kappa}, x \geq 0, \text{if } k > 0 & \text{Fréchet} \\ 1-(1+\kappa x/\tau)^{-1/\kappa}, 0 \leq x < -\frac{t}{k}, \text{if } k < 0 & \text{Weibull} \\ 1-e^{-\kappa/\tau}, x \geq 0, \text{if } k = 0 & \text{Gumbel} \end{cases}$$

where the form of the GPD is determined by two parameters: scale τ and shape κ. The weight of the tail, characterized by the shape parameter, corresponds to the domain when $\kappa = 0$ (Gumbel 1958), the Fréchet domain when $\kappa > 0$ (Leadbetter et al. 1983), and the Weibull domain when $\kappa < 0$ (Embrechets et al. 1997).

Nonetheless, the rarity of beneficial mutations makes it difficult to evaluate theories for the DFE among these types of mutations because it necessarily leads to reduced statistical power for deciding among competing hypotheses. To date, there is few empirical data available to test the most compelling prediction of the theory that the DFE among beneficial mutations is exponentially distributed, and that which exists to date has yielded conflicting results (e.g. Kassen & Bataillon 2006; McDonald et al. 2010; Rokyta et al. 2008; Bataillon et al. 2011).

Resistance mechanisms

Irrespective of the underlying fitness distribution, changes at the level of DNA sequences affect the fixation of beneficial mutations. Beneficial mutations enable organisms to adapt to stressful conditions, as is the case in nearly all opportunistic human pathogens surviving exposure to antibiotics. In the past, the most common practice has

been to remove an antibiotic from general use once resistance evolved. Nonetheless, while antibiotic resistance is commonly associated with fitness costs (MacLean et al. 2010a), there is a growing recognition that this is not always the case (Andersson & Hughes 2010). Most beneficial mutations are thought to pay little or no cost of resistance or have positive pleiotropic effects (Kugelberg et al. 2005) leading to higher fitness in both antibiotic- supplemented and free environments. The persistence of cost-free resistance mutations has been associated with parallel evolution (repeated and independent evolution of the same mutational change) (Wong & Kassen 2011), and the fixation of compensatory mutations (second-site mutations, which decrease the cost incurred by the initial resistance mutation) (Björkman 2000; Andersson & Hughes 2010).

The implications of these resistance mechanisms are especially critical in clinical settings, where the emergence and prevalence of antibiotic resistance is high, making it difficult to control and eliminate once it has evolved. This is evident with the bacterium *Pseudomonas aeruginosa*, the most common respiratory pathogen colonizing the lung of cystic fibrosis (CF) patients. Once quinolones commonly used to treat *P. aeruginosa* infections among CF patients are removed from general use, resistance persists and impedes treatment and management of infection. Inhibitory activities of quinolones in *P. aeruginosa* have been attributed to the presence of mutations in the quinolone resistance-determining regions (QRDR) of type-II topoisomerases and putative efflux regulatory operons (Wong & Kassen 2011), as summarized in Table 2.1. Mutated type-II topoisomerases, DNA gyrase (*gyrA* and *gyrB*) and topoisomerase IV (*parC* and *parE*), have been associated with negative supercoiling (or relaxing positive supercoils) and the decatenation of covalently closed circular double-stranded DNA, respectively. Mutated efflux regulatory genes *mexR* and *nfxB* have been associated with the suppression of structural genes of the *mexA-mexB-oprM* and *mexC-mexD-oprJ* operons, leading to increased pump activity and extrusion of antibiotics outside the cell. Targeted sequencing methods allow for surveying the frequency and differences in mutational processes affecting bacterial fitnesses of known resistance mutations (Wong & Kassen 2011). Shedding light on the range and number of genetic targets involved in resistance is critical in designing strategies to minimize the emergence and persistence of resistance.

Here we address both the DFE and resistance mechanisms using a large collection of independently derived single-step *P. aeruginosa* mutants conferring resistance to the quinolone antibiotic ciprofloxacin. Together with the antibiotic sensitive ancestor, we assay the fitness effects of these strains in a commonly used laboratory medium and one of clinical relevance designed to nutritionally mimic the CF lung and screen for putative genetic targets associated with quinolone resistance. Since the mutations assayed confer resistance they are, by definition, beneficial in the presence of the drug. We first estimate the DFE among novel beneficial genotypes and then using a subset of all strains determine the type and frequency of point mutations occurring in the aforementioned six known genetic targets of quinolone resistance. The purpose of examining the mutations themselves is two-fold. Firstly, it allows us to confirm that the mutants in this collection are genetically unique as required by theory, that is, that the same resistance mutation hasn't arisen independently more than once. Secondly, it provides insight into which mutations are likely to be of clinical importance, those displaying cost-free resistance and/or positive pleiotropic effects.

Methods

Bacterial strains, media, and chemicals

All experiments were performed with *P. aeruginosa*, strain PA14, a highly virulent clinical isolate. Bacterial cultures were grown at 37° C in Luria-Bertrani (LB) medium and a synthetic cystic fibrosis medium (SCFM) designed to nutritionally mimic the CF lung prepared according to Palmer et al. (2007) (*see Appendix A*). Supplements of 2 μg ml^{-1} of ciprofloxacin, purchased from Sigma-Aldrich, were used in antibiotic-supplemented LB and SCFM media.

Susceptibility testing

Minimum inhibitory concentrations (MIC) of the wild type *P. aeruginosa* were carried out in a 96-well flat-bottom polystyrene plate (Costar, Corning Incorporated, Corning, NY) using a broth microdilution method (Qaiyumi, 2007) in LB and SCFM

media. MICs were performed in triplicate for each fixed concentration of antibiotic × medium combination. Suspensions of the wild type were prepared by re-suspending one colony of an overnight culture from LB agar in LB and SCFM broths. The initial bacterial density of ~1 x 10^9 bacteria ml^{-1} was diluted with LB or SCFM to obtain an inoculum of ~1 x 10^5 bacteria ml^{-1} for each medium. Each microwell of the 96-well plate containing 100 µl of ciprofloxacin at different fixed concentrations was inoculated with 100 µl of the bacterial suspension. The concentrations of ciprofloxacin ranged from 0.087 µg ml^{-1} to 5 µg ml^{-1}. A positive control for growth containing no ciprofloxacin was also included for each replicate. The microplate was incubated for 24 h in a BioTek PowerWave XS2 Microplate Reader (Bio-Tek Instruments Inc., Winooski, Vermont, USA), and absorbance at 630 nM (OD_{630}) was determined using Gen5 Data Analysis software (Bio-Tek Instruments Inc.). The microplate was shaken for 5 min every 30 min and readings were recorded every 30 min following inoculation for 24 h. The MIC was defined as the lowest concentration that did not show growth. The mean MIC value for all three replicates was 0.65 µg ml^{-1} in LB and 1.48 µg ml^{-1} in SCFM.

Mutant isolation

Approximately 500 cells of the wild type *P. aeruginosa* were inoculated into each microwell of a 96-well flat-bottom polystyrene plate (Costar, Corning Incorporated, Corning, NY) containing 195 µl of LB medium, which was then incubated at 37° C in an orbital shaker set to 150 revolutions per minute (rpm) to maintain culture homogeneity and adequate aeration. Following 24 h incubation, the cultures were spread onto LB agar plates containing 2 µg ml^{-1} ciprofloxacin, a concentration which is just above the MIC for the PA14 strain (LB, 0.65 µg ml^{-1}; SCFM, 1.48 µg ml^{-1}), and were incubated for 48 h. This concentration of antibiotic ensured that mutants collected weren't limited to those conferring either very high or low levels of resistance, or even false-positives. From those plates containing viable colonies, a single colony was used to inoculate 6 mL of LB medium supplemented with 2 µg ml^{-1} ciprofloxacin as to ensure only one genotype was used. Following 24 h incubation, this culture was cryogenically frozen in a non-evolving state and stored at -80° C for further analysis in 50 per cent culture : 50 per cent glycerol

by volume. A total of 614 independently-derived, single-step strains resistant to ciprofloxacin were collected using this conventional fluctuation-style assay (Luria & Delbrück 1943).

Fitness assays

Fitness of each resistant mutant and the sensitive ancestor was assayed in the selection medium (LB with ciprofloxacin) and three alternative media (LB, SCFM, and SCFM with ciprofloxacin), as the change in cell density of each genotype grown in pure culture following 24 h incubation. Genotypes were first grown overnight from frozen stocks, and 10 µl of the overnight culture was inoculated into each microwell of a 96-well flat-bottom polystyrene plate (Costar, Corning Incorporated, Corning, NY) containing 190 µl of the selection medium. Following 48 h incubation in an orbital shaker (150 rpm), to ensure that resistant mutants were comparably acclimated to the assay environment, 10 µ l of the culture was inoculated into each microwell of a sterile 96-well plate containing 190 µl of one of four assay media. Cell density was measured as the OD_{630} on an ELX800 Microplate Reader (Bio-Tek Instruments Inc., Winooski, Vermont, USA) using Gen5 Data Analysis software (Bio-Tek Instruments Inc.). Readings were recorded every 12 h following inoculation for 48 h. Fitness was estimated as the difference between final and initial OD_{630} readings. Enough plates were used to ensure two replicate wells of each mutant genotype × environment combination (2 replicates × 614 strains × 4 media environments = 4912 wells). Each plate included one well with the wild type.

PCR amplification

Sixty-six strains from the collection of 614 were chosen for genetic analysis. Twenty-five of these strains were selected and classified as 'beneficial' because they produced population densities above the 95% CI of the sensitive ancestor grown in permissive LB medium after 48 h growth. The remaining forty-one strains were chosen from the larger collection, where twenty-four strains were classified as 'neutral' in regards to the ancestor

and seventeen strains were classified as 'costly' because they produced lower population densities than the ancestor in permissive LB medium.

Genomic DNA was extracted from 500 µl of the selected sixty-six *P. aeruginosa* isolates, using the Wizard® SV Genomic DNA Purification System (Promega Corporation, Madison, Wisconsin, USA) as per the manufacturer's instructions. PCR amplification of the QRDRs of *gyrA*, *gyrB*, *parC*, and *parE*, and the multidrug efflux operon sites of *mexR* and *nfxB* was carried out with primers designed and kindly provided by Alex Wong. The oligonucleotide primers were used to amplify fragments of *gyrA* positions 2014754 to 2017519, *gyrB* positions 4275 to 6695, *parC* positions 5843612 to 5845876, and *parE* positions 5847418 to 5849307 containing the QRDR region, and fragments of *mexR* positions 486057 to 486500 and *nfxB* positions 5421533 to 5422051 for the putative efflux pump regulators. Details of primer sets used for PCR amplification are reported in Table 2.2. The target gene loci were amplified using a thermal cycler PCR system (Mastercycler ep, Eppendorf, Hamburg, Germany). The 25 µl reaction mixture contained 12.5 µl Go*Taq*® Colorless Master Mix (Promega Corporation, Madison, Wisconsin, USA), 1.25 µl each antisense and sense primers, 9 µl distilled water, and 1 µl DNA template. The PCR amplification protocol comprised of denaturation for 2 minutes at 95°C, then 35 cycles of denaturation for 30 seconds at 95°C, 30 seconds for 57°C, elongation for 1 minute at 72°C, followed by a final incubation step for 5 minutes at 72°C.

Sequence determination

The PCR-amplified DNA was sent to the Genome Québec Innovation Centre at McGill University for sequencing on the forward strand. Nucleotide sequences were aligned and converted to their corresponding protein sequences using *MEGA* version 5 (Tamura et al. 2011) and Sequencher® (Gene Codes Corporation, Ann Arbor, MI USA). The sequences were accepted once compared and matching closely with the *P. aeruginosa* UCBPP-PA14 genome, as reported in the National Institute for Biotechnology Information (NCBI) databank (Geer et al. 2010). Previously published sequences of target genes used for aligning were obtained from the NCBI databank under accession number

ABJ12393.1 for *gyrA*; ABJ14960.1 for *gyrB*; ABJ15660.1 for *parC*; ABJ14351.1 for *parE*; ABJ15391.1 for *mexR*; and YP_793059 for *nfxB*.

Data analysis

All analyses were conducted using *R* v. 2.9.2 (R Development Core Team 2009). Logistic growth models were fit to the optical density readings for each replicate of genotype × environment combination to determine two biological growth measures, growth rate μ and carrying capacity A, using the `gcFitModel` function in the `grofit` package in *R* (Kahm 2010). Mutant fitness w was calculated from the change in OD_{630} over the course of 24 h growth. Selection coefficients of each mutant were calculated as the relative difference in fitness; $s_{opt} = (w_{opt} - w_m)/w_m$ where w_m is the fitness of the mutant founder. Excluded from data analyses were mutants that had no or very little growth. All analyses were thus performed on 326 mutant strains that grew in both replicates in all four environments.

Mean parametric growth measures and standard errors for differences of means were calculated in all four environments. Inasmuch as antibiotic-containing environments were supplemented with 2 μg ml^{-1} of ciprofloxacin, a concentration above the MIC for the PA14 strain, growth of the wild type was only observed and measured in permissive LB and SCFM environments. Spearman's rank correlations were calculated to evaluate the relationship between μ, A, and w among mutant strains in pairwise permissive and antibiotic-supplemented environments. To further determine whether strains with one high parametric growth measure in one environment also had high values in the other environments, Spearman's rank correlations were also calculated between pairs of parametric growth measures in all four environments. Average values of replicates were used.

A standard two-factor analysis of variance (ANOVA) was performed for each parametric growth measure to detect variability among strains across all four environments, with genotype and environment as random main effects. Focusing explicitly on variance contributed by mutation, the unmutated sensitive ancestor was excluded from these analyses. The variance components of genotype, environment, their

interaction, and residual error were estimated upon μ, A, and *w* (one model for each of μ, A, and *w*). The restricted maximum likelihood method (`lmer`) in the `lme4` package in *R* (Bates & Sarkar 2007) was then used to calculate the variance components associated with each main effect attributable for each term in the model.

To test the robustness of the prediction under EVT that the data are from an exponential distribution in the absence of antibiotic, likelihood-ratio methods `evalrtmc` developed by Beisel et al. (2007) were used. One approach of EVT is to consider the distribution of fitness values above a threshold, i.e. the wild type. Mutant strains were classified as 'beneficial' when they produced higher absolute fitness values than the common ancestor grown in permissive environment. Since evolution favours the fixation of mutations with large effects, it is possible that those of small effect are missed over the course of the experiment. Thus, the empirical sample of fitness measures obtained in both LB and SCFM environments were scaled to the smallest observed selection coefficient, that being the smallest beneficial mutation, instead of the common ancestor. Shifting relative to the smallest observation decreased the sample size (and degrees of freedom) by 1 in each environment. Within each permissive environment, the likelihood of the rescaled data for μ, A, and *w* was obtained under the GPD, which encompasses all three possible domains of attraction under EVT. Using a likelihood ratio test, the likelihood of the rescaled data was optimized under the null model where κ was restricted ($\kappa = 0$) depicting an exponential distribution, and then under the alternative model where κ was free to vary ($\kappa \neq 0$). *P*-values were obtained from 10,000 bootstrap samples with the fitted scale. Akaike's information criterion (AIC, Akaike 1974) was then used to compare the log likelihood of both models. The model that is most consistent with the observations is the one with the lowest AIC. Given the model that maximizes the likelihood of observing the empirical data, expected distributions of fitness effects for fixed beneficial mutations, as characterized by one of the three possible domains of attraction, were fit using the `fitdistr` function in the `MASS` package in *R* (Venables & Ripley 2002).

Results

Biological growth parameters

Biological growth parameters of a collection of 614 single-step antibiotic-resistant *P. aeruginosa* strains were assayed in the presence and absence of ciprofloxacin in LB and SCFM environments (Table 2.3). The mean absolute fitness of the wild type was 0.945 ± 0.097 (mean ± std) in LB (95% CI 0.893 – 1.007) and 0.406 ± 0.036 (mean ± std) in SCFM (95% CI 0.379 – 0.432) environments. Mutants grew better in permissive LB and SCFM environments rather than their corresponding ciprofloxacin-supplemented environments (Figure 2.1). A substantial number of mutants were found having higher parametric growth measures than the wild type in permissive environments (LB, $\mu = 105$, $A = 14$, $w = 89$; SCFM, $\mu = 300$, $A = 167$, $w = 319$).

Relationship between μ, A and w

Mutant strains with high fitness w in a permissive environment also had significantly high fitness in an antibiotic-supplemented environment (LB × LB with ciprofloxacin, $\rho = 0.409$, $P < 0.0001$; SCFM × SCFM with ciprofloxacin, $\rho = 0.430$, $P < 0.0001$; Table 2.4; Figure 2.2). Mutant strains with one high parametric measure in one environment also had significantly high values in the other environments (Table 2.5). The strongest correlation in each environment was between w and A among all strains (LB, $\rho = 0.828$, $P < 0.0001$; LB with ciprofloxacin, $\rho = 0.873$, $P < 0.0001$; SCFM, $\rho = 0.949$, $P < 0.0001$; SCFM with ciprofloxacin, $\rho = 0.954$, $P < 0.0001$, Table 2.5).

Genotype × environment interaction

Environment explained significant variation in both μ and A in ANOVA tests. The main effects of environment accounted for more than 70% of the variance in μ and only 1% in A, with the genotype × environment interaction accounting for only 11 and 0.07% of the variance in μ and A, respectively (Table 2.6). Strains' values for $\rho_{\mu,A}$ were positively correlated between environments (ρ always > 0.282, $P < 0.0001$; Table 2.5)

The genotype × environment interaction term, along with each term independently, explained significant variation in w in ANOVA tests. The main effects of genotype and

environment together accounted for 30% of the variance in *w*, with their interaction accounting for only 2% (Table 2.6). Strains' values for $\rho_{w,\mu}$ and $\rho_{w,A}$ were positively correlated between environments (ρ always > 0.309, $P < 0.0001$; Table 2.5).

Analysis of the DFE among beneficial mutations

Eighty-nine mutant strains were classified as 'beneficial' in LB and 317 in SCFM, because they produced higher fitness values than the wild type grown in LB and SCFM, respectively. The remaining strains were classified as 'costly' because they produced lower fitness values than the wild type (Figure 2.3). The mean selection coefficient s_j associated with beneficial mutants was 0.086 ± 0.0702 (mean ± std) in LB and 0.3293 ± 0.1793 (mean ± std) in SCFM.

Using the maximum likelihood estimator of Beisel *et al.* (2007), the exponential was not an adequate estimator of the empirical data. In both permissive LB and SCFM environments, estimates for the shape parameter of the GPD were negative (*P* always < 0.0001; Table 2.7). The AIC values for each parametric growth measure and environment were lowest in the full model, further suggesting that the Weibull domain of attraction when $\kappa < 0$, provides a better fit of the data than the exponential (Table 2.7). For both parameters of the Weibull distribution, the expected distribution of fitness effects remained bell-shaped in both permissive environments (LB, $\tau = 1.069$, $\kappa = 13.501$; SCFM, $\tau = 0.806$, $\kappa = 4.489$; Figure 2.4).

DNA sequence variation

Sequencing was used to identify genetic targets known or thought to be involved in quinolone resistance in *P. aeruginosa*. Results are reported in Table 2.8. Among the sixty-six sequenced strains, having varying levels of fitness compared to the common ancestor, mutations were found in the QRDRs of *gyrB* and *parE*, and the putative regulators of efflux pumps *mexR* and *nfxB* (Figure 2.5). The presence of mutations varied greatly among all genotypes (Figure 2.6). Four distinct mutations occurring at four amino acid sites (109, 404, 439, and 671) were identified within *gyrB*, two mutations at two sites (45

and 98) within *parE*, one mutation at one site (126) within *mexR*, and sixteen mutations at six sites (14, 21, 39, 64, 78, and 175) within *nfxB*. Only one double mutant was identified having a mutation in both *gyrB* and *nfxB*. No further mutations were found in the QRDRs of either *gyrA* or *parC* genes.

Discussion

Adaptation depends on mutations wherein natural selection acts to generate new traits that increase organismal fitness and purge deleterious traits that decrease fitness. The ability to predict the overall process of adaptive evolution is often thought to depend on the DFE among beneficial mutations (Orr 1998). Inasmuch as the first step of adaptation relies on the fixation of beneficial mutations, understanding the DFE among these mutations has become the focus of much theoretical work (Eyre-Walker & Keightley 2007; Orr 2003; Orr 2006). Even so, empirical studies looking at the phenotypic and genotypic properties of beneficial mutations have lagged as a sheer consequence of their rarity and difficult detectability in natural populations (Fisher 1930). Accordingly, the challenge of being able to predict realistic parameter values of the DFE among beneficial mutations in novel environments remains.

To help circumvent this lack of empirical data, we used bacterial populations to examine the DFE of beneficial mutations among isogenic pseudomonad quinolone-resistant mutants. The fitness of a large collection of resistant *P. aeruginosa* mutants was assayed under selective (ciprofloxacin-supplemented) and permissive (ciprofloxacin-free) conditions. By design, this procedure allows us to survey the pleiotropic costs of quinolone resistance in *P. aeruginosa* and test whether the DFE among beneficial mutations is a negative exponential. The latter is a prediction stemming initially from statistical properties characterizing Gillespie's mutational landscape models (1984) and then further justified according to EVT (Beisel et al. 2007).

The most predominant form of pleiotropy thought to be associated with antibiotic resistance is that mutations that increase fitness in an antibiotic-containing environment, for example by conferring resistance, will decrease fitness in an antibiotic-free

environment. This is because the acquisition of resistance mutations is expected to carry a fitness cost, often reflected by reduced bacterial growth under permissive conditions (Sousa et al. 2012; Trindade et al. 2010). Despite this, evidence suggests that this is not always the case, and that resistant mutations may in fact have positive pleiotropic effects (Ostrowski et al. 2005), and pay little or no cost of resistance in permissive environments (Andersson & Hughes 2010; Kugelberg et al. 2005). We found a substantial number of mutants that displayed a growth rate, carrying capacity, and fitness higher than that of the antibiotic-sensitive wild type in permissive conditions—and so show no evidence of bearing a cost of resistance. Moreover, those mutants that were fittest under selective conditions also displayed high fitness under permissive conditions. This result suggests that there is little or no cost associated with the persistence of antibiotic resistance in permissive environments among these mutants. While no-cost mutations may be responsible for continued resistance in permissive environments, this may be further complicated by the fixation of additional parallel and fitness-compensatory mutations, which have also been suggested as mechanisms, which reduce fitness costs associated with resistance (Andersson & Hughes 2010). Either way, the fact that cost-free mutants appear to evolve readily puts into question the utility of removing ciprofloxacin from general use once resistance has evolved. Instead, empirical studies are now focusing on the effects of combination therapies to treat pseudomonad infections (Chamot et al. 2003; Traugott et al. 2011). This has however resulted in much controversy. While the use of multiple (and extended-spectrum) antibiotics may allow for *in vitro* synergy between treatments resulting in improved outcomes, there is increased risk of the emergence of a 'superinfection'—infection emerging during (or as result of) antimicrobial treatment for the initial infection.

The higher frequency of beneficial mutations present in SCFM than LB, together with environment explaining significant variation in all parametric growth measures, suggests that the wild type is more maladapted to the SCFM environment. According to Fisher's geometric model (1930), a population located away from a fitness 'optimum' moves closer to it by the fixation of beneficial mutations. The large number of beneficial mutations available in SCFM suggests that the sensitive wild type is further from the

optimum. This is an important implication when considering the emergence and persistence of antibiotic resistance in CF patients, as it suggests that quinolone resistance should be highly prevalent in this clinical setting (i.e. the CF lung). Future work should thus test predictions about cost-free mutations and pleiotropic fitness effects in real-life settings, i.e. using clinical *P. aeruginosa* isolates from CF patients. Nonetheless, settings of clinical relevance (i.e. the CF lung) also introduce the added complexity of population dynamics and communal interactions. While the fitness of independently-derived isogenic resistant mutants was assayed here, it has been suggested that at the population level bacterial altruism may be involved in population-wide resistance (H. H. Lee et al. 2010). To increase survival capacity of the overall population, H.H. Lee at al. (2010) observed that *Escherichia coli* mutants highly resistant to norfloxacin (also a quinolone antibiotic) produced an indole, inducing various antibiotic-tolerance mechanisms giving protection to less resistant mutants and effectively, the population as a whole.

Despite the possible types of mutational and population-based strategies, which may reduce fitness costs of resistance, EVT has been used to provide insight into the DFE among beneficial mutations. It assumes, *a priori*, that the wild type is already fairly well adapted to a given environment (i.e. close to the optimum), and that single-step beneficial mutations are even fitter than the wild type and so drawn from the extreme right tail of an underlying fitness distribution exhibiting one of three tail behaviours (Orr 2005). The prediction is that the frequencies of fitness values among beneficial mutations will be exponentially distributed above a threshold (i.e. the wild type), when mutations fall in the so-called Gumbel domain of attraction. While two other domains of attraction exist, the Weibull and Fréchet types, they have both been criticized as being inappropriate biologically (Orr 2006).

The classic evolutionary view that the DFE relies on a large fraction of mutations of small effect and few of large effect appears to be true only under LB conditions, with many mutations of large effect in SCFM. Nonetheless, we reject the exponential as an adequate descriptor of the data in both environments, and show that the DFE among beneficial mutations is better described as being located in the Weibull domain of

attraction, confirming recent work in bacteria (Bataillon et al. 2011; McDonald et al. 2010; MacLean & Buckling 2009; Barrett et al. 2006).

The expected shape parameter describing our collection of beneficial mutants was estimated using Beisel's *et al.* (2007) maximum likelihood methods. Under both permissive conditions, the Weibull distribution, introduced by Weibull (1951), is characterized as bell-shaped with a right-truncated tail bound on the magnitude of the largest fitness effects, where the shape parameter κ affects the slope of the density function. The frequency decrease of mutations above the threshold, observed when $\kappa < 1$, may sometimes be mistaken for a stretched exponential distribution, with a finite right endpoint.

While the prediction of an exponentially distributed DFE under EVT assumes that the wild type is initially fairly well adapted, there is no statistical theory that predicts the shape of the DFE among beneficial mutations when the wild type is mal-adapted. Barrett et al. (2006) used a gamma distribution to observe what happens to the DFE among fixed and newly arising beneficial mutations when the data is no longer in the classic 'weak selective pressure' territory. Strongly favoured alleles (relative to the wild type) were less affected by stochastic loss than weakly favoured alleles. Under strong selective pressure, the DFE among fixed beneficial mutations closely reflected the DFE among newly arising beneficial mutations.

Theoretical predictions of the EVT framework depend on a number of assumptions, which may not hold in this study. In SCFM, the fitness of wild type is low, giving rise to a fraction of beneficial mutations having large effects on fitness such that the data may no longer fit the initial prediction that the DFE should be exponentially distributed (Orr 2003). While the inference that the domain of attraction among beneficial mutations is Weibull may consequently not be entirely appropriate, these results do provide an empirical glimpse of the shape of the overall DFE.

The selective pressure of antibiotics imposes not only changes at the phenotypic (or fitness) level, but also at the genetic, i.e. variations in the DNA sequence. Quinolone resistance has been attributed to the presence of mutations in the QRDRs of *gyrA*, *gyrB*, *parC*, and *parE*, and the putative efflux regulatory operons *mexR* and *nfxB*. Based on our

sequence assays, we found a number of mutations, and often multiple mutations, in these known genetic targets of resistance. The presence of mutations varied greatly among all genotypes insofar that beneficial mutants did not exhibit any uniquely different mutational changes otherwise absent among neutral or costly mutants. These results, combined with the fitness assays, confirm that mutants in this collection were not genetically pseudoreplicated. We did not, however, uncover any mutations at sites often associated with quinolone resistance in *gyrA* or *parC* genes. A single double mutant was also found, violating the assumption under EVT that the collection represents genotypes that are only a single beneficial allele from the sensitive ancestor (Gillespie 1984). This is also likely to be something of an underestimate because, by design, we only focused our attention on six known genetic targets of quinolone resistance. Using simulations of the DFE among beneficial mutations, Bataillon et al. (2011) suggest that this violation, however, should not drastically affect our estimation of the DFE among beneficial mutants. These results indicate that mutations in regions of the genome not sequenced and/or epistatic interactions between second-site mutations may help explain the fitness advantage observed in beneficial mutants under permissive conditions (Manna et al. 2011). Fortunately, recent advances in genomic technologies (e.g. rapid whole genome sequencing) may be able to reveal such mechanisms more readily.

All in all, future work on the adaptive trajectories of beneficial mutations should continue to try to bridge the gap between purely phenotypic and genotypic data. The aim of this would be to eventually be able to attribute a predictive value (i.e. fitness effect) to known genes, or even specific loci, responsible for conferring antibiotic resistance. A strategy might be to import statistical approaches (e.g. Bayesian estimates) to detect fitness differences among mutations and determine whether beneficial mutations are getting fixed at specific locations of a gene or DNA sequence. Notwithstanding, these results provide valuable insight into one of the most fundamental problems in evolution, the nature and fitness effects of beneficial mutations.

Target gene	Complete name	Description	Reference
gyrA	DNA gyrase subunit A	supercoiling of covalently closed circular double-stranded DNA	(Hooper 1999; Kugelberg et al. 2005; Wong & Kassen 2011; J. K. Lee et al. 2005)
gyrB	DNA gyrase subunit B	supercoiling of covalently closed circular double-stranded DNA	(Hooper 1999; Kugelberg et al. 2005; Wong & Kassen 2011; J. K. Lee et al. 2005)
parC	DNA topoisomerase IV subunit A	decatenation of circular double-stranded DNA	(Hooper 1999; Kugelberg et al. 2005; Wong & Kassen 2011; J. K. Lee et al. 2005)
parE	DNA topoisomerase IV subunit B	decatenation of circular double-stranded DNA	(Hooper 1999; Kugelberg et al. 2005; Wong & Kassen 2011; J. K. Lee et al. 2005)
mexR	transcriptional regulator (*MarR* family)	suppression of the *mexA-mexB-oprM* operon	(Aeschlimann 2003; Evans et al. 2001; Gotoh et al. 1998; Poole et al. 1996)
nfxB	transcriptional regulator (*LacR-GalR* family)	suppression of the *mexC-mexD-oprJ* operon	(Aeschlimann 2003; Gotoh et al. 1998)

Table 2.1 | **Inhibitory activities of fluoroquinolones against target genes of resistance in *P. aeruginosa*.**

Target gene	Amino acid positions[a]	Primer names	Primer sequences (5' to 3')	Notes[b]
DNA gyrase subunit A (*gyrA*)	2014754 to 2017519	GyrA–F	GTAAAACGACGGCCAGTGatgggcgaactggccaaag	M13(–20) sequencing
		GyrA–R	cagcaggttgggaatctt	
DNA gyrase subunit B *(gyrB*)	4275 to 6695	GyrB–F	GTAAAACGACGGCCAGTGatgagcagcatcgaatggc	M13(–20) sequencing; includes some upstream (recF)
		GyrB–R	gttcaggttacgcgtcagc	
		GyrB–2F	GTAAAACGACGGCCAGTGgccgagaagtctggcctgac	M13(–20) sequencing
		GyrB–2R	aaggccggttggcaacgatcg	
		GyrB–3F	GTAAAACGACGGCCAGTGaggcctggcgaagaaattc	M13(–20) sequencing
		GyrB–3R	cggagcgtgctcgttgacgtg	
DNA topoisomerase IV subunit A (*parC*)	5843612 to 5845876	ParC–F	GTAAAACGACGGCCAGTGatgagcgaatccctcgatc	M13(–20) sequencing
		ParC–R	cagcaggttgggcaggcg	
DNA topoisomerase IV subunit B (*parE*)	5847418 to 5849307	ParE–F	GTAAAACGACGGCCAGTGcgacgaaaccctcgactatcg	M13(–20) sequencing; includes some upstream
		ParE–R	caacaggttgcggaactcg	
Efflux pump regulator (*mexR*)	486057 to 486500	mexR–F	GTAAAACGACGGCCAGTGtacttacattcataggtg	M13(–20) sequencing; includes some upstream, downstream
		mexR–R	aagacttcggcatcaagatg	
Transcriptional regulator (*nfxB*)	5421533 to 5422051	nfxB–F	GTAAAACGACGGCCAGTGgcccacaacctgccaacgcg	M13(–20) sequencing; includes some upstream, downstream
		nfxB–R	ctgatcttcccgtgtgtc	

[a] Amino acid positions correspond to those of *P. aeruginosa* UCBPP–PA14 (GenBank accession no. CP000438).
[b] M13(–20) sequence, expressed in capital letters, added to forward primer.

Table 2.2 | **Details of primer sets used for PCR amplification.**

Environment	μ (OD/hr)	A (OD)	*w* (OD)
LB	0.402 (± 0.017)	1.654 (± 0.287)	0.733 (± 0.015)
LB with ciprofloxacin	0.105 (± 0.005)	0.267 (± 0.032)	0.165 (± 0.006)
wild type in LB	0.728 (± 0.114)	1.083 (± 0.091)	0.945 (± 0.097)
SCFM	0.454 (± 0.010)	0.827 (± 0.016)	0.728 (± 0.010)
SCFM with ciprofloxacin	0.193 (± 0.013)	0.426 (± 0.085)	0.243 (± 0.006)
wild type in SCFM	0.477 (± 0.137)	0.488 (± 0. 030)	0.406 (± 0.036)

Table 2.3 | **Mean parametric growth measures and standard errors for differences of means among *P. aeruginosa* mutants in four environments.** μ indicates growth rate, A carrying capacity, and *w* fitness in optical density (OD) units. Growth of the wild type was only measured in permissive LB and SCFM environments.

	μ (OD/hr)		A (OD)		*w* (OD)	
Environment	ρ	*P*	ρ	*P*	ρ	*P*
LB × LB with ciprofloxacin	0.041	0.505	0.207	<0.0001	0.409	<0.0001
SCFM × SCFM with ciprofloxacin	0.068	0.290	0.342	<0.0001	0.430	<0.0001
LB × SCFM	0.523	<0.0001	0.358	<0.0001	0.593	<0.0001
LB with ciprofloxacin × SCFM with ciprofloxacin	0.322	<0.0001	0.353	<0.0001	0.456	<0.0001

Table 2.4 | **Spearman's rank correlations (ρ) for parametric growth measures among *P. aeruginosa* mutants between pairwise permissive and selective environments.** μ indicates growth rate, A carrying capacity, and *w* fitness in optical density (OD) units. The wild type was excluded from the analyses.

Environment	μ (OD/hr) × A (OD)	μ (OD/hr) × *w* (OD)	A (OD) × *w* (OD)
LB	0.458	0.609	0.828
LB with ciprofloxacin	0.282	0.406	0.873
SCFM	0.341	0.309	0.949
SCFM with ciprofloxacin	0.610	0.494	0.954

Table 2.5 | **Spearman's rank correlations (ρ) for interactions between pairs of parametric growth measures among *P. aeruginosa* mutants in four environments.** μ indicates growth rate, A carrying capacity, and *w* fitness in optical density (OD) units. All values are significant ($P < 0.0001$). The wild type was excluded from the analyses.

		μ (OD/hr)				A (OD)				*w* (OD)			
Term	df	Sum Sq	*F* ratio	*P*	Total variance (%)	Sum Sq	*F* ratio	*P*	Total variance (%)	Sum Sq	*F* ratio	*P*	Total variance (%)
Environment	3	34.657	135.648	<0.0001	71.361	646.7	4.542	0.004	1.298	144.136	4389.151	<0.0001	25.713
Genotype	386	40.803	1.241	0.012	8.901	10443.1	0.570	1.000	9.487E-10	35.542	9.780	<0.0001	7.386
G × E	837	44.307	0.622	1.000	11.796	13789.2	0.347	1.000	0.073	32.668	3.055	<0.0001	2.738
Residual error	499	42.497			7.942	23680.2			98.629	7.115			64.163
Total	1725				100				100				100

Table 2.6 | **Results of two–factor ANOVAs and variance components analyses for parametric growth measures among *P. aeruginosa* mutants**. μ indicates growth rate, A carrying capacity, and *w* fitness in optical density (OD) units. The wild type was excluded from the analyses.

			Full model					Reduced model					
Parameter	Environment	n	$\hat{\tau}$	$\hat{\kappa}$	LogL	k	AIC	$\hat{\tau}$	$\hat{\kappa}$	LogL	k	AIC	LRT (P–value)
μ	LB	89	1.474	−0.449	−539.921	2	1083.842	1.091	0	−552.816	1	1107.632	25.790 $P < 0.0001$
	SCFM	319	2.073	−0.956	−2529.301	2	5062.602	1.115	0	−2624.559	1	5251.118	190.516 $P < 0.0001$
A	LB	89	1.744	−0.212	−334.238	2	672.476	1.454	0	−336.865	1	675.730	5.252 $P < 0.0001$
	SCFM	319	2.233	−0.996	−2789.757	2	5583.514	1.187	0	−2892.954	1	5787.908	206.393 $P < 0.0001$
w	LB	89	1.297	−0.999	−673.647	2	1351.294	0.741	0	−712.956	1	1427.912	78.617 $P < 0.0001$
	SCFM	319	1.906	−1	−3723.195	2	7450.390	1.018	0	−3841.881	1	7685.762	237.373 $P < 0.0001$

Table 2.7 | **Likelihood–ratio tests (LRT) for the Gumbel domain of attraction against a generalized Pareto distribution.** The form of this distribution is determined by two parameters: scale $\hat{\tau}$ and shape $\hat{\kappa}$. n is the number of beneficial *P. aeruginosa* mutants observed in permissive LB and SCFM environments. All values were scaled relative to the smallest observed beneficial mutant instead of the wild type, decreasing the sample size (and degrees of freedom) by 1 in each environment. Akaike's information criterion (AIC) was calculated as -2(LogL) + 2k where LogL is the maximized value of the likelihood function for the estimated model and k is the number of parameters in the statistical model. P–values are based on parametric bootstrapping (10,000 bootstrap samples from an exponential with the fitted scale).

			Resistance mutations					
Genotype	w_{LB}	w_{SCFM}	*gyrA*	*gyrB*	*parC*	*parE*	*mexR*	*nfxB*
PA14	0.945	0.406	–[a]	–	–	–	–	–
175	1.243	0.807	–	–	–	–	–	-
537	1.222	0.753	–	–	–	–	–	Leu14Pro (CTG41CCG)
512	1.174	0.734	–	–	–	–	–	-
113	1.160	0.891	–	–	–	–	–	-
222	1.159	0.971	–	–	–	–	–	His21Pro (CAC62CCC)
614	1.149	0.760	–	–	–	–	–	-
490	1.145	1.120	–	–	–	–	–	Gln64*stop* (CAG190TAG)
460	1.121	0.788	–	–	–	–	–	-
489	1.120	1.250	–	–	–	–	–	Thr39Pro (Acc115CCC)
488	1.113	0.789	–	–	–	Glu45Lys (GAA133AAA)	–	-
253	1.112	0.888	–	–	–	–	–	-
230	1.105	1.013	–	–	–	–	–	-
430	1.101	0.828	–	–	–	–	–	-
492	1.081	0.920	–	–	–	–	Glu126Glu (GAG378GAA)	-
247	1.064	0.739	–	–	–	–	–	-
172	1.061	0.818	–	–	–	–	–	-
511	1.060	0.599	–	–	–	–	–	-
57	1.041	0.837	–	Asn109Tyr (AAC325TAC)	–	–	–	Leu14Pro (CTG41CCG)
463	1.039	0.808	–	–	–	–	–	-
215	1.027	0.559	–	–	–	–	–	-
559	1.024	0.723	–	–	–	–	–	Thr39Pro (Acc115CCC)
437	1.015	1.192	–	–	–	–	–	NA[b]
254	1.014	0.867	–	–	–	–	–	Ala175Val (GCA523GTA)
299	1.008	0.737	–	–	–	–	–	-
246	1.005	0.859	–	–	–	–	–	-
433	0.992	0.796	–	–	–	–	–	-
199	0.990	0.628	–	–	–	–	–	-
160	0.989	0.951	–	–	–	–	–	-
504	0.988	1.071	–	–	–	–	–	-
602	0.987	0.769	–	–	–	–	–	-
323	0.982	0.900	–	–	–	–	–	-
477	0.979	0.456	–	–	–	–	–	-
207	0.978	0.846	–	–	–	–	–	-
445	0.972	0.835	–	–	–	–	–	-
162	0.970	0.673	–	–	–	–	–	His21Pro (CAC62CCC)
301	0.968	0.838	–	–	–	–	–	-
450	0.965	0.821	–	His671Tyr (CAC2011TAC)	–	–	–	-

351	0.958	1.110	–	–	–	–	–	His21Pro (CAC62CCC)
508	0.946	0.835	–	–	–	–	–	Leu14Gln (CTG41CAG)
439	0.932	0.927	–	–	–	–	–	-
455	0.929	0.767	–	–	–	–	–	-
467	0.925	0.623	–	–	–	–	–	Leu14Gln (CTG41CAG)
453	0.792	1.052	–	–	–	–	–	Thr39Pro (ACC115CCC)
102	0.762	1.005	–	–	–	–	–	His21Pro (CAC62CCC)
165	0.756	0.455	–	Gly404Arg (GGC1210CGC)	–	–	–	-
442	0.729	0.693	–	–	–	–	–	-
443	0.727	0.804	–	–	–	–	–	-
461	0.724	0.794	–	–	–	–	–	-
374	0.719	1.008	–	–	–	–	–	-
99	0.710	0.979	–	–	–	–	–	-
530	0.694	0.652	–	–	–	–	–	His21Pro (CAC62CCC)
574	0.668	0.494	–	–	–	–	–	-
217	0.633	0.841	–	–	–	–	–	Glu78*stop* (GAG232TAG)
557	0.621	0.491	–	–	–	–	–	-
580	0.610	0.679	–	–	–	Lys98Glu (AAG292GAG)	–	-
227	0.609	0.633	–	–	–	–	–	-
486	0.579	0.569	–	–	–	–	–	-
390	0.553	0.741	–	–	–	–	–	-
441	0.545	0.584	–	Asn439Ser (AAT1316AGT)	–	–	–	-
119	0.491	0.538	–	–	–	–	–	-
345	0.488	0.805	–	–	–	–	–	-
97	0.462	0.673	–	–	–	–	–	-
177	0.404	0.620	–	–	–	–	–	-
558	0.396	0.793	–	–	–	–	–	-
293	0.381	0.418	–	–	–	–	–	-
169	0.359	0.535	–	–	–	–	–	-

[a] –, no mutation detected.
[b] NA, data not available.

Table 2.8 | **Non-synonymous DNA sequence variations among *P. aeruginosa* mutants having varying degrees of fitness (*w*) in permissive LB environment.** Sequences assayed were the quinolone resistance-determining regions of *gyrA*, *gyrB*, *parC* and *parE*, as well as the DNA-binding sites of the putative efflux pump regulators *mexR* and *nfxB*. The deduced amino acid sequences were compared with that previously reported for *P. aeruginosa* UCBPP–PA14 (GenBank accession no. CP000438).

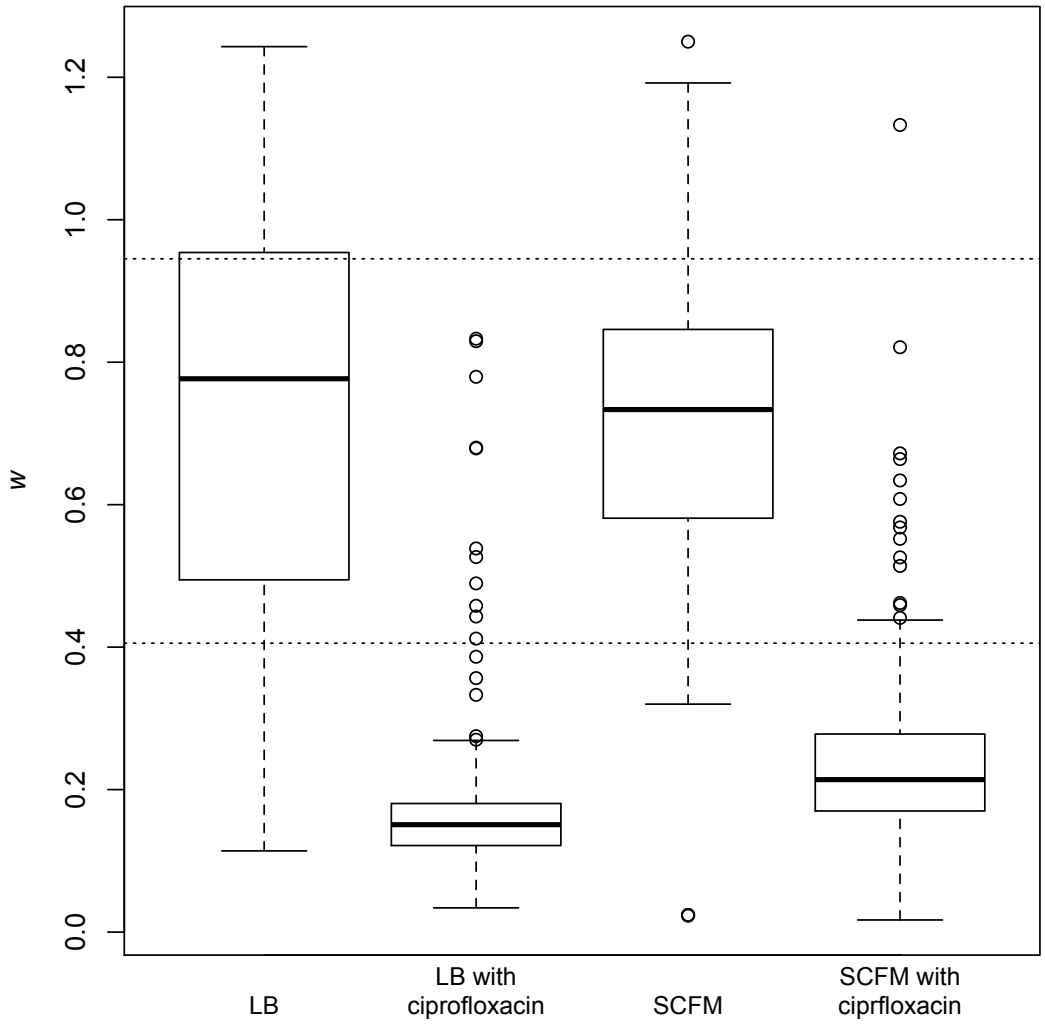

Figure 2.1 | **Fitness (*w*) among *P. aeruginosa* mutants in four environments.** Dotted horizontal lines mark wild type fitness in permissive environments (LB, upper line, $w = 0.945$; SCFM, lower line, $w = 0.406$). The wild type did not grow in antibiotic-supplemented environments.

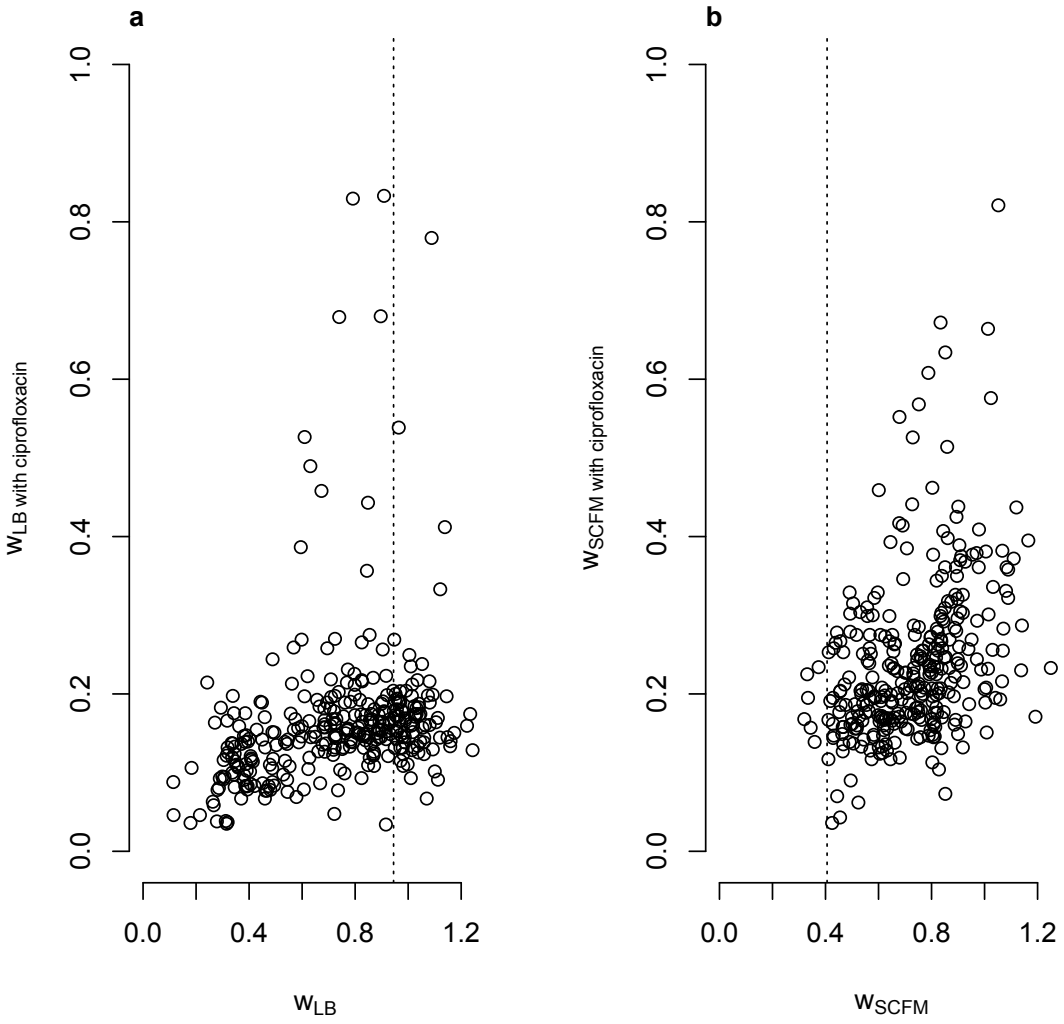

Figure 2.2 | **Relationship between fitness (*w*) among *P. aeruginosa* mutants assayed in permissive and antibiotic-supplemented LB (a) and SCFM (b) environments.** Dotted vertical lines mark wild type fitness in permissive environments (LB, $w = 0.945$; SCFM, $w = 0.406$). Strains located to the right of the dotted vertical lines represent 'beneficial' mutants, strains to the left represent 'costly' mutants (see *Methods* for definitions).

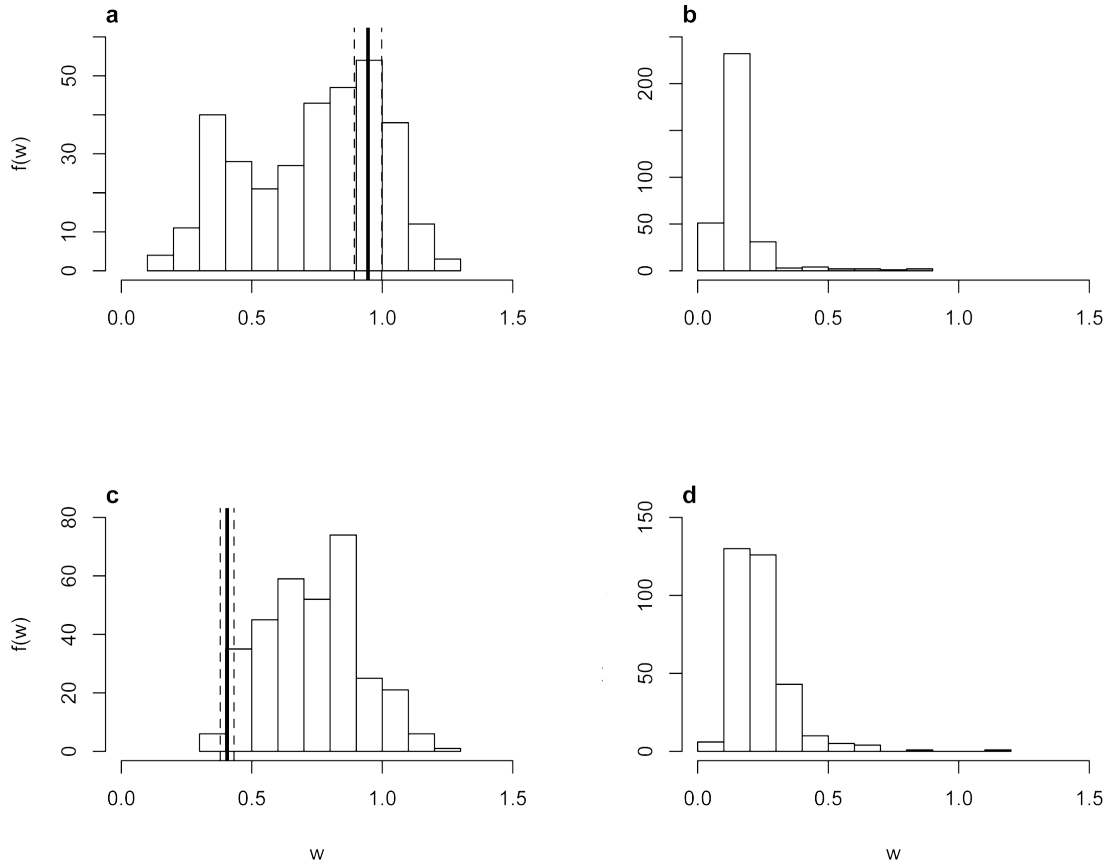

Figure 2.3 | **Frequency distributions of fitness (*w*) among *P. aeruginosa* mutants in four environments: LB (a), LB with ciprofloxacin (b), SCFM (c), and SCFM with ciprofloxacin (d).** The wild type did not grow in antibiotic-supplemented environments. Thick solid vertical lines mark wild type fitness in permissive environments (LB, $w = 0.945$; SCFM, $w = 0.406$); dashed lines denote 95% confidence limits. Strains located to the right of the solid vertical lines represent 'beneficial' mutants, strains to the left represent 'costly' mutants (see *Methods* for definitions).

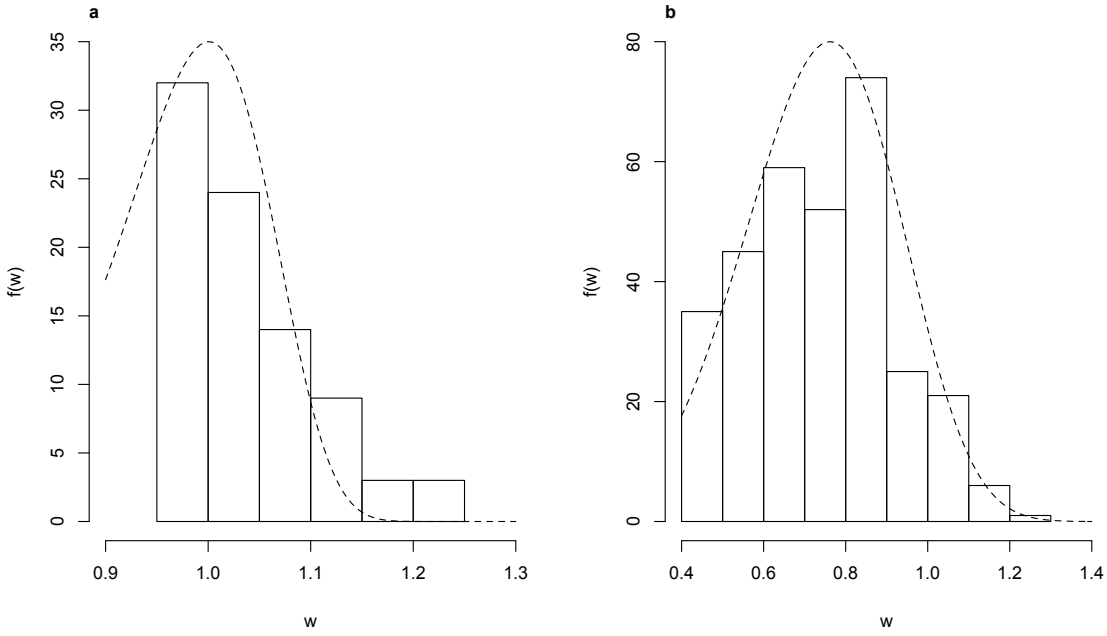

Figure 2.4 | **Observed (bars) and expected (curves) frequency distributions of fitness (*w*) among beneficial *P. aeruginosa* mutants assayed in permissive LB (a) and SCFM (b) environments.** Expected curves were fit for each environment using Weibull distributions with scale and shape parameters obtained from the likelihood analysis (LB, $\tau = 1.069$, $\kappa = 13.501$; SCFM, $\tau = 0.806$, $\kappa = 4.489$; *see Table 2.7*).

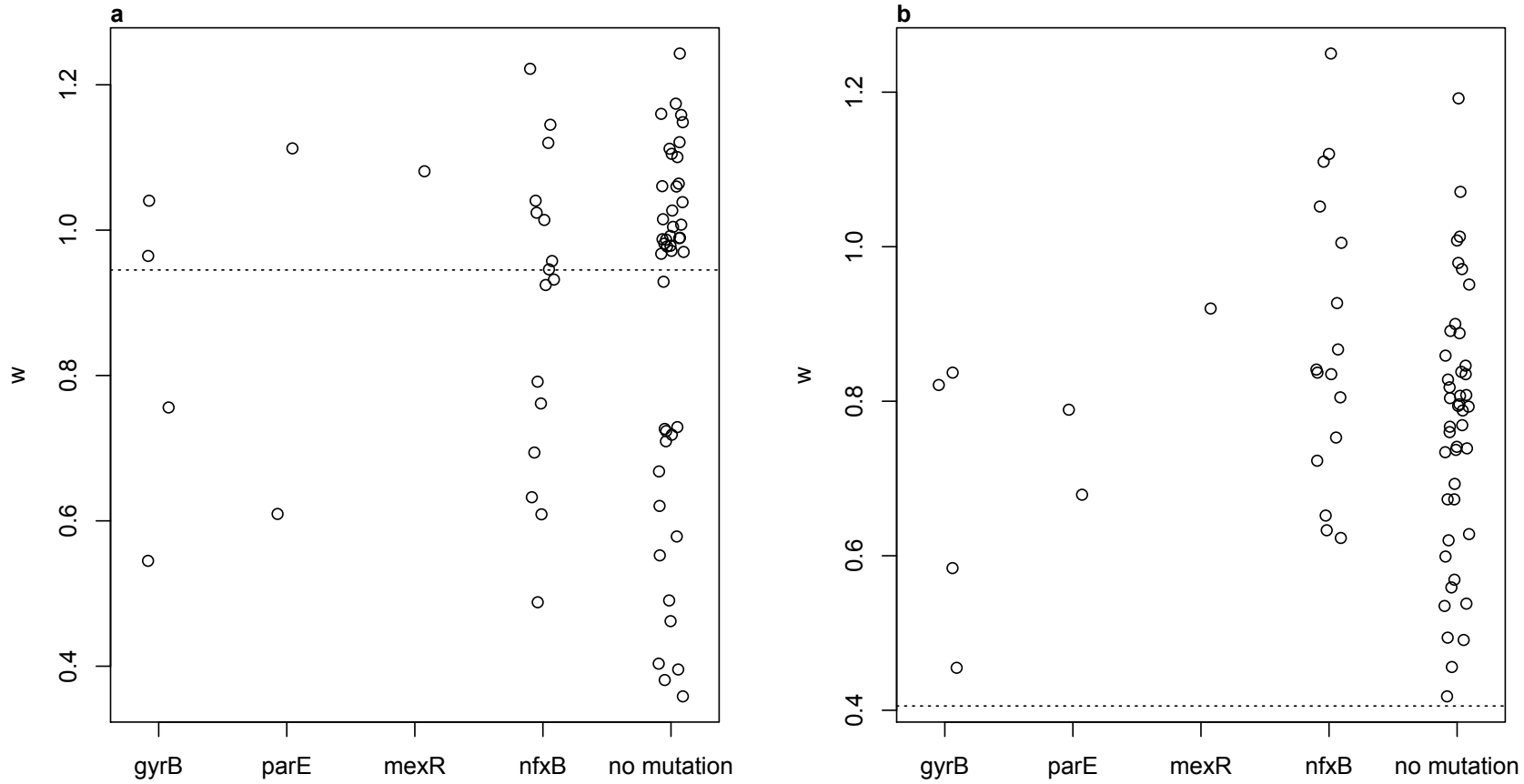

Figure 2.5 | **Distribution of non-synonymous mutations among *P. aeruginosa* mutants having varying degrees of fitness (*w*) in permissive LB (a) and SCFM (b) environments.** Sequences assayed were the quinolone resistance-determining regions (QRDR) of *gyrA*, *gyrB*, *parC* and *parE*, as well as the DNA-binding sites of the putative efflux pump regulators *mexR* and *nfxB*. No amino acid changes were found in the QRDRs of either *gyrA* or *parC* genes. Dotted vertical lines mark wild type fitness in permissive environments (LB, $w = 0.945$; SCFM, $w = 0.406$).

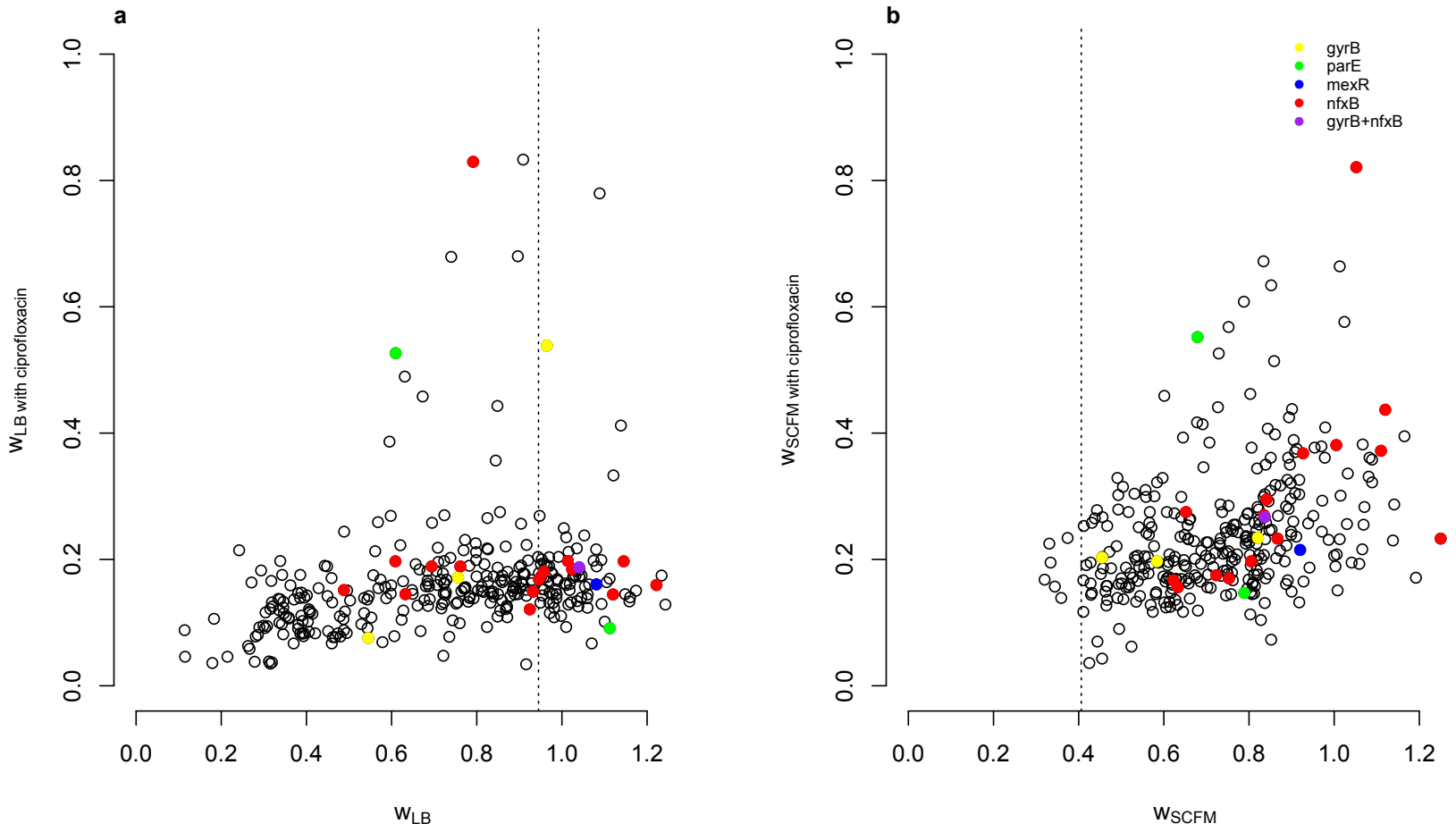

Figure 2.6 | **Distribution of genetic effects onto phenotypic (fitness, *w*) effects among *P. aeruginosa* mutants in permissive LB (a) and SCFM (b) environments.** Point mutations among sequenced isolates were located within the quinolone resistance-determining regions (QRDR) of *gyrB* and *parE*, as well as the DNA-binding sites of the putative efflux pump regulators *mexR* and *nfxB*. Dotted vertical lines mark wild type fitness in permissive environments (LB, $w = 0.945$; SCFM, $w = 0.406$).

CHAPTER 3

Conclusions

Beneficial mutations are those which increase fitness relative to their ancestral genotype in a given environment. They tend to be rare in natural populations because, when they do occur, they are rapidly substituted. While they provide fuel for adaptation, little is known about the properties of beneficial mutations. This thesis explores the fitness effects and resistance mechanisms of beneficial mutations in the Gram-negative bacterium *Pseudomonas aeruginosa* and bears directly on the contemporary difficulties of understanding and eliminating antibiotic resistance once it has evolved in the aforementioned pathogenic population. Accomplished through experiments with a large library of independently derived quinolone resistant mutants in *P. aeruginosa*, the objectives of this thesis were two-fold; (a) to test the prediction under EVT that the DFE among beneficial mutations falls in the so-called Gumbel domain of attraction, and (b) to screen a selection of resistant strains for known genetic targets of quinolone resistance. Chapter 1 reviewed theoretical models of adaptation, restricting much attention to the DFE associated with beneficial mutations, and detailed examples of microbial experimental evolution studies focusing on this more limited body of theory. Using the EVT framework, Chapter 2 demonstrated that the DFE among beneficial mutations within a collection of quinolone resistant *P. aeruginosa* strains is not exponential, as characterized by the Gumbel domain of attraction. In lieu, the DFE among beneficial mutations, assayed in laboratory and clinically relevant media, fell within the Weibull domain of attraction. Furthermore, the presence of single and second-site mutations varied greatly among screened genotypes. The number and types of mutations present among beneficial genotypes was not enough to account for their increased fitness in permissive environments. These results suggest that mutations in regions of the genome not sequenced may be equally important in determining their pleiotropic fitness costs.

Among the gaps, future work should aim at incorporating data across non-replicate experiments to predict the distribution of joint fitness effects among beneficial mutations in novel environments as well as predicting the fitness effects among beneficial mutations containing known mutations conferring antibiotic resistance. Taken together, this thesis contributes to the growing body of empirical studies on the nature and fitness effects of beneficial mutations.

References

Aeschlimann, J.R., 2003. The role of multidrug efflux pumps in the antibiotic resistance of *Pseudomonas aeruginosa*. *Pharmacotherapy*, 23(7), p.916-924.

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on automatic control*, 19, p.716–723.

Andersson, D.I. & Hughes, D., 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature reviews. Microbiology*, 8(4), p.260-271.

Barrett, R.D.H., M'Gonigle, L.K. & Otto, S.P., 2006. The distribution of beneficial mutant effects under strong selection. *Genetics*, 174(4), p.2071-9.

Barrett, R.D.H., MacLean, R.C. & Bell, G., 2006. Mutations of intermediate effect are responsible for adaptation in evolving *Pseudomonas fluorescens* populations. *Biology letters*, 2(2), p.236-238.

Bataillon, T., Zhang, T. & Kassen, R., 2011. Cost of adaptation and fitness effects of beneficial mutations in *Pseudomonas fluorescens*. *Genetics*, 189(November), p.939-949.

Bates, D.M. & Sarkar, D., 2007. *lme4: Linear mixed-effects models using S4 classes*, R package version 0.99875-6.

Beisel, C.J. et al., 2007. Testing the extreme value domain of attraction for distributions of beneficial fitness effects. *Genetics*, 176(4), p.2441-2449.

Betancourt, A.J. & Bollback, J.P., 2006. Fitness effects of beneficial mutations: the mutational landscape model in experimental evolution. *Current opinion in genetics & development*, 16(6), p.618-623.

Björkman, J., 2000. Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science*, 287(5457), p.1479-1482.

Chamot, E. et al., 2003. Effectiveness of combination antimicrobial therapy for *Pseudomonas aeruginosa* bacteremia. *Antimicrobial agents and chemotherapy*, 47(9), p.2756-2764.

Chevin, L.-M., Martin, G. & Lenormand, T., 2010. Fisher's model and the genomics of adaptation: Restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution*, p.3213-3231.

Embrechets, P. et al., 1997. *Modelling extremal events for insurance and finance*. Springer, Berlin.

Evans, K. et al., 2001. MexR repressor of the mexAB-oprM multidrug efflux operon of Pseudomonas aeruginosa : Identification of MexR binding sites in the mexA-mexR intergenic region. *Journal of bacteriology*, 183(3), p.807-812.

Eyre-Walker, A. & Keightley, P.D., 2007. The distribution of fitness effects of new mutations. *Nature reviews. Genetics*, 8(8), p.610-618.

Fisher, R.A., 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford

Fitzsimmons, J.M. et al., 2009. Population consequences of mutational events: effects of antibiotic resistance on the r/K trade-off. *Evolutionary ecology*, 24(1), p.227-236.

Geer, L.Y. et al., 2010. The NCBI BioSystems database. *Nucleic acids research*, 38, p.D492-496.

Gillespie, J.H., 1983. A simple stochastic gene substitution process. *Theoretical population biology*, 23, p.202-215.

Gillespie, J.H., 1984. Molecular evolution over the mutational landscape. *Evolution*, 38(5), p.1116-1129.

Gillespie, J. H., 1991. *The causes of molecular evolution*. Oxford University Press, New York.

Gotoh, N. et al., 1998. Characterization of the MexC-MexD-OprJ multidrug efflux system in DeltamexA-mexB-oprM mutants of *Pseudomonas aeruginosa*. *Antimicrobial agents and chemotherapy*, 42(8), p.1938-1943.

Gumbel, E.J., 1958. *Statistics of extremes*. Columbia University Press, New York.

Hamilton, M.B., 2009. *Population genetics*. Wiley-Blackwell, West Sussex.

Hooper, D.C., 1999. Mechanisms of fluoroquinolone resistance. *Drug resistance updates*, 2(1), p.38-55.

Kahm, M., 2010. Grofit: Fitting biological growth curves. *Journal of statistical software*, 33(7), p.1-21.

Kassen, R. & Bataillon, T., 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature genetics*, 38(4), p.484-488.

Kugelberg, E. et al., 2005. Reduction of the fitness burden of quinolone resistance in *Pseudomonas aeruginosa*. *Journal of antimicrobial chemotherapy*, 55(1), p.22-30.

Leadbetter, M.R. et al., 1983. *Extremes and related properties of random sequences and processes*. Springer, New York.

Lee, H.H. et al., 2010. Bacterial charity work leads to population-wide resistance. *Nature*, 467(7311), p.82-85.

Lee, J.K. et al., 2005. Alterations in the GyrA and GyrB subunits of topoisomerase II and the ParC and ParE subunits of topoisomerase IV in ciprofloxacin-resistant clinical isolates of *Pseudomonas aeruginosa*. *International journal of antimicrobial agents*, 25(4), p.290-295.

Luria, S.E. & Delbrück, M., 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6), p.491-511.

MacLean, R.C. & Buckling, A., 2009. The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. *PLoS genetics*, 5(3), p.e1000406.

MacLean, R.C. et al., 2010a. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature reviews. Genetics*, 11(6), p.405-414.

Maclean, R.C. et al., 2010b. The evolution of antibiotic resistance: insight into the roles of molecular mechanisms of resistance and treatment context. *Discovery medicine*, 10(51), p.112-118.

Manna, F., Martin, G. & Lenormand, T., 2011. Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics*, 189(3), p.923-937.

Martin, G. & Lenormand, T., 2006a. General multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution*, 60(5), p.893-907.

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

Martin, G. et al., in prep. The correlated fitness effects of beneficial mutations are distributed according to fitness landscape theory in antibiotic resistant bacteria. p.1-49.

McDonald, M.J. et al., 2010. The distribution of fitness effects of new beneficial mutations in *Pseudomonas fluorescens*. *Biology letters*, p.1-4.

Orr, H.A., 1998. The population genetics of adaptation: The distribution of factors. *Evolution*, 52(4), p.935-949.

Orr, H.A., 2000. Adaptation and the cost of complexity. *Evolution*, 54(1), p.13-20.

Orr, H.A., 2002. The population genetics of adaptation: the adaptation of DNA sequences. *Evolution*, 56(7), p.1317-30.

Orr, H.A., 2003. The distribution of fitness effects among beneficial mutations. *Genetics*, 163, p.1519-1526.

Orr, H.A., 2005. The genetic theory of adaptation: a brief history. *Nature reviews. Genetics*, 6(2), p.119-127.

Orr, H.A., 2006. The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. *Journal of theoretical biology*, 238(2), p.279-285.

Ostrowski, E.A., Rozen, D.E. & Lenski, R.E., 2005. Pleiotropic effects of beneficial mutations in *Escherichia coli*. *Evolution*, 59(11), p.2343-2352.

Palmer, K.L., Aye, L.M, & Whiteley, M., 2007. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *Journal of bacteriology*, 189(22), p.8079-8087.

Pickands, III, J., 1975. Statistical inference using extreme order statistics. *Annals of statistics*. 3, p.119–131.

Poole, K. et al., 1996. Expression of the multidrug resistance operon mexA-mexB-oprM in *Pseudomonas aeruginosa*: mexR encodes a regulator of operon expression. *Antimicrobial agents and chemotherapy*, 40(9), p.2021-2028.

Qaiyumi, S., 2007. Chapter 4 Macro- and microdilution methods of antimicrobial susceptibility testing. *Antimicrobial susceptibility testing protocols*. CRC Press, p.75-79.

R Development Core Team., 2009. *R: A language and environment for statistical computing*. R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org.

Rokyta, D.R. et al., 2008. Beneficial fitness effects are not exponential for two viruses. *Journal of molecular evolution*, 67(4), p.368-376.

Sequencher® version 5.0 sequence analysis software, Gene Codes Corporation, Ann Arbor, MI USA. URL http://www.genecodes.com

Sousa, A., Magalhães, S. & Gordo, I., 2012. Cost of antibiotic resistance and the geometry of adaptation. *Molecular biology and evolution*, p.1-12.

Tamura, K. et al., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28(10), p.2731-2739.

Traugott, K. a et al., 2011. Monotherapy or combination therapy? The *Pseudomonas aeruginosa* conundrum. *Pharmacotherapy*, 31(6), p.598-608.

Trindade, S., Perfeito, L. & Gordo, I., 2010. Rate and effects of spontaneous mutations that affect fitness in mutator *Escherichia coli*. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1544), p.1177-1186.

Venables, W.N. & Ripley, B.D., 2002. *Modern applied statistics with S.* Fourth Edition. Springer, New York.

Wagner, G.P., Pavlicev, M. & Cheverud, J.M., 2007. The road to modularity. *Nature reviews. Genetics*, 8(12), p.921-931.

Wagner, G.P. & Zhang, J., 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature reviews. Genetics*, 12(3), p.204-213.

Wang, Z., Liao, B.-Y. & Zhang, J., 2010. Genomic patterns of pleiotropy and the evolution of complexity. *Proceedings of the National Academy of Sciences*, 107, p.18034-18039.

Weibull WA., 1951. Statistical distribution function of wide applicability. *ASME Journal of Applied Mechanics Transactions of the American Society of Mechanical Engineers*, p.293–297.

Wong, A. & Kassen, R., 2011. Parallel evolution and local differentiation in quinolone resistance in *Pseudomonas aeruginosa*. *Microbiology*, 157, p.937-944.

Appendix A: Synthetic cystic fibrosis medium recipe

Prepare buffered base.

6.5 ml 0.2 M NaH_2PO_4
6.25 ml 0.2 M Na_2HPO_4
0.348 ml 1 M KNO_3
1.084 ml 0.25 M K_2SO_4
0.122 g NH_4Cl
1.114 g KCl
3.03 g NaCl
2.3 g 10 mM MOPS
779.6 ml deionized water

Add amino acids from 100-mM stocks, prepared in dH_2O, to buffered base.

L-aspartate	8.27 ml	prepare this stock in 0.5M NaOH
L-tryptophan	0.13 ml	prepare this stock in 0.2M NaOH
L-tyrosine	8.02 ml	prepare this stock in 1.0M NaOH
L-threonine	10.72 ml	
L-serine	14.46 ml	
L-glutamate	15.49 ml	
L-proline	16.61 ml	
L-glycine	12.03 ml	
L-alanine	17.8 ml	
L-cysteine	1.60 ml	
L-valine	11.17 ml	
L-methionine	6.33 ml	
L-isoleucine	11.21 ml	
L-leucine	16.09 ml	
L-phenylalanine	5.3 ml	
L-ornithine	6.76 ml	
L-lysine	21.28 ml	
L-histidine	5.19 ml	
L-arginine	3.06 ml	

Adjust to pH 6.8 and filter sterilize.

After sterilization, add sterile components.

1 ml 3.6 mM $FeSO_4 \cdot 7H_2O$	prepare in 1 mg/ml stock
0. 606 ml 1 M $MgCl_2$	
1.754 ml 1 M $CaCl_2$	