Winter 2012

The distribution of direct and indirect fitness effects of beneficial mutations

Rachel Katherine Staples
University of New Hampshire, Durham

THE DISTRIBUTION OF DIRECT AND INDIRECT FITNESS EFFECTS OF BENEFICIAL MUTATIONS

BY

RACHEL KATHERINE STAPLES

B.S., University of New Hampshire, 2010

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

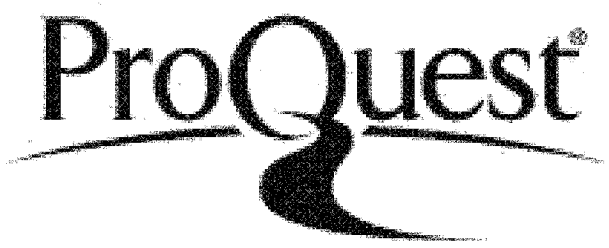
the Requirements of the Degree of

Master of Science

In

Microbiology

December, 2012



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

This thesis has been examined and approved.

Thesis Director, Vaughn S. Cooper,
Associate Professor, Microbiology

W. Kelley Thomas,
Hubbard Professor in Genomics and Director

Jessica A. Bolker,
Associate Professor, Zoology

12/11/12
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

THE DISTRIBUTION OF DIRECT AND INDIRECT FITNESS EFFECTS OF BENEFICIAL MUTATIONS

By

Rachel Katherine Staples

University of New Hampshire, December, 2012

Rare, beneficial mutations that increase an organism's fitness provide the basis by which adaptation proceeds. Current theory predicts that the individual fitness effects of these beneficial mutations are exponentially distributed, suggesting that mutations conferring a small fitness increase are more numerous than those of large benefit. However, there is little empirical evidence describing measurable fitness effects of individual mutations, nor their availability or effects across a range of environments. We experimentally evolved a single strain of the cystic fibrosis pathogen *Burkholderia cenocepacia* under both physically structured (biofilm) and unstructured (planktonic) conditions, collected a sample of mutants, and measured the fitness effect of each in direct competition with the ancestor. Fitness was also measured in a variety of alternative environments to quantify the pleiotropic, or indirect, effects of each mutation. We found that the distribution of direct mutational effects was better modeled by an extreme value distribution with a truncated, Weibull-like domain of attraction, rather than exponential. A clustering of high fitness values and parallel evolution at the nucleotide

level indicate that mutations greatly increasing fitness are more readily available to an adapting population than previously assumed. Pleiotropic effects were generally positive, although mutants did experienced a fitness trade-off under some alternative conditions, suggesting that highly beneficial mutations in a structured environment are likely specific to that biofilm environment and may ultimately narrow the organism's niche breadth. We also found that the magnitude of direct and pleiotropic fitness effects were strongly correlated, indicating that mutations of higher initial benefit in the selective environment also drastically influence fitness in alternative environments, the negative effects of which may bar their success under fluctuating conditions.

CHAPTER I

THE DISTRIBUTION OF FITNESS EFFECTS OF BENEFICIAL MUTATIONS

Introduction

All organisms adapt by acquiring traits that are beneficial in the selective environment and are generated by mutation. Theory predicts that the majority of all possible mutations will have neutral or detrimental effects on an organism, yet some small subset will benefit the organism, increasing its fitness. Thus, the fitness effects of these rare beneficial mutations would likely be located within the right tail of a probability distribution (Figure 1.1). The size and shape of that tail is dictated by the number of available beneficial mutations, as well as their individual effects (Gillespie 1984; Orr, 2003; Beisel et al., 2007). Accurately defining the spectrum of beneficial fitness effects available for natural selection to act upon remains crucial to understanding and modeling many aspects of adaptation, including the probability of beneficial mutations occurring and rising in frequency, the overall rate of fitness increase in a population, and the rate of adaptation to a new environment (Haldane, 1927; Fisher, 1930; Gillespie, 1983, 1984; Orr, 2003; Patwa and Wahl, 2008; Perfeito *et al.*, 2007; Rozen *et al.*, 2002).

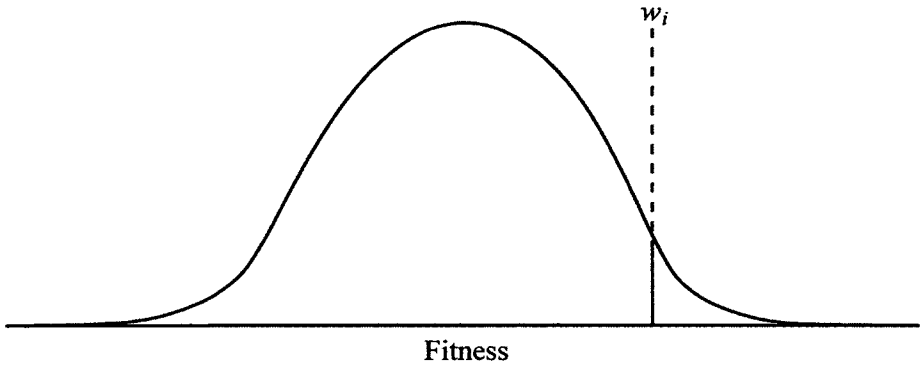


Figure 1.1: A theoretical probability distribution of fitness effects associated with all possible mutations occurring on a given genotype. The majority of mutations are predicted to have fitness effects worse than or equal to the ancestral genotype (w_i). Therefore, beneficial mutations are expected to be rare and their associated fitness gains located in the right tail.

Little is known about the individual fitness effects of beneficial mutations, nor the frequency at which they occur, because they are difficult to detect and measure. For this reason, most models of adaptation assume that the majority of beneficial mutations will slightly increase fitness, while an increasingly smaller percentage will have much larger fitness effects, fitting an exponential distribution. This exponential distribution has also been derived, using extreme value theory (EVT), and rationalized as a means of predicting the effects of beneficial mutations (Gillespie, 1984; Orr, 2003). Select studies measuring the fitness effects of mutations collected prior to selection have supported this theory (Kassen, 2006; MacLean, 2009). An exponential-like distribution of effects was also reported for a set of beneficial mutations rising to high frequency (Rozen *et al.*, 2002). However, the exponential distribution was derived from EVT under the assumption of relatively high ancestral fitness in that environment (Gillespie 1984), which may not be entirely realistic for all models of adaptation, particularly those under strong selection (Barrett *et al.*, 2006). If ancestral fitness in the selective environment is

low, a larger portion of mutations are expected to increase fitness and their effects may no longer be confined to the far right tail. In this case, extreme value theory would no longer apply and the fitness effects may be better described by a non-exponential distribution, like those reported by a number of studies focusing on beneficial mutations and viral evolution in different hosts or environments (Rokyta et al. 2008; Sanjuán *et al.*, 2004; Vale *et al.*, 2012).

Empirically testing the distribution of effects remains difficult because new beneficial alleles may easily be lost from a population due to random sampling or competitive interactions with other mutants (Haldane, 1927; Wilke, 2004). The latter problem is particularly prevalent in large asexual populations in which co-occurring beneficial mutations cannot assemble by recombination and instead compete with one another. Known generally as the Hill-Robertson effect (Hill and Robertson, 1966), in asexual populations this process is known as clonal interference, and it is expected to favor mutations conferring a larger fitness increase, while those of smaller benefit are likely to be lost (Gerrish and Lenski, 1998). These phenomena make it difficult to identify and sample a broad range of beneficial mutations, but may be overcome by evolving large populations that are easily manipulated.

Microbes provide an excellent opportunity to study the effects of beneficial mutations for many reasons, one of which is an inherently large population size (Elena and Lenski, 2003). Small populations produce less genetic variation, most of which will be lost during a bottleneck, therefore the eventual success of a beneficial allele in a small population is more the result of it randomly surviving drift than its individual fitness benefit. Genetic variation will also be lost in larger populations, but a greater amount of

initial variation and a wider bottleneck allow more beneficial mutants to survive and ultimately be collected (Kimura & Ohta, 1969; Nei *et al.,* 1975).

Bacterial populations, and their environment, can also be easily manipulated in a favorable manner (Elena and Lenski, 2003; Cooper, 2002). In this way, environmental conditions may be altered to restrict certain interactions, such as clonal interference. Physical structure in the environment may reduce these negative interactions, and preserve a larger number of beneficial mutants. Structure has been shown to help maintain genetic diversity within a microbial population by localizing interactions, reducing overall competition, and decreasing the effects of selection (Korona *et al.*, 1994; MacLean *et al.*, 2004; Perfeito *et al.*, 2006). We predicted that mutants of lower benefit would be more likely to survive clonal interference in a structured environment, allowing us to collect a sample of mutants that more accurately represents all beneficial mutations available to an adapting population. This broad collection could then be used to measure individual fitness values in direct competition with the ancestor, and ultimately describe the distribution of beneficial effects.

Methods

Strains and Culture Conditions

The *Burkholderia cepacia* complex (BCC) is comprised of seventeen closely related species commonly isolated from the environment and capable of chronically infecting persons with cystic fibrosis (Isles *et al.,* 1984; Mahenthiralingam *et al.*, 2000; Vanlaere *et al.*, 2008; Vanlaere *et al.*, 2009). One of the most common species associated with increased patient morbitiy and mortality, *Burkholderia cenocepacia,* is also known for its ability to form robust biofilms on biotic and abiotic surfaces (Reik *et al.*, 2005;

Speert *et al.*, 2002; Springman *et al.*, 2009). *B. cenocepacia* HI2424, the strain used in this study, was isolated from an onion field and frozen at -80°C, and remains naive to a laboratory environment (LiPuma *et al.*, 2002).

A Tn7 vector was used to introduce the gene *lacZ* to HI2424, conferring β-galactosidase activity and allowing colonies to be distinguished by their blue color when plated on 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal) (Ellis, 2008).

All evolution experiments and fitness assays were carried out in 18 x 150mm test tubes and incubated at 37°C while rotating on a roller drum, unless otherwise stated.

Experimental Evolution and Collection of Mutants

Biofilm evolved populations were created by reviving Lac+ marked and Lac- unmarked *B. cenocepacia* HI2424 in 5mL of Tryptic Soy Broth from frozen stocks. All cultures were grown overnight, then diluted 1:100 into 5mL of M9 minimal media supplemented with 3% galactose (3% GMM). A 7mm polystyrene bead was also added to the progenitor cultures for all biofilm evolution cultures, and incubated for 24 hours. Beads were then removed and all attached cells were vortexed off in Phosphate Buffered Saline. Seven individual populations were then seeded using a 1:1 mixture of oppositely marked ancestor in 3% GMM containing a white bead, and incubated for 24 hours. The bead was then moved to a new tube of fresh media containing a black bead, and again incubated. Experimental evolution continued in this fashion, transferring the 24 hour bead to fresh media containing an oppositely marked bead, selecting for daily biofilm formation and dispersal. All populations were sampled every other day by removing all cellular content from the 48 hour bead, diluting 1:100,000 in PBS, and plating on ½ strength Tryptic soy (Tsoy) agar plates containing X-gal (Figure 1.2).

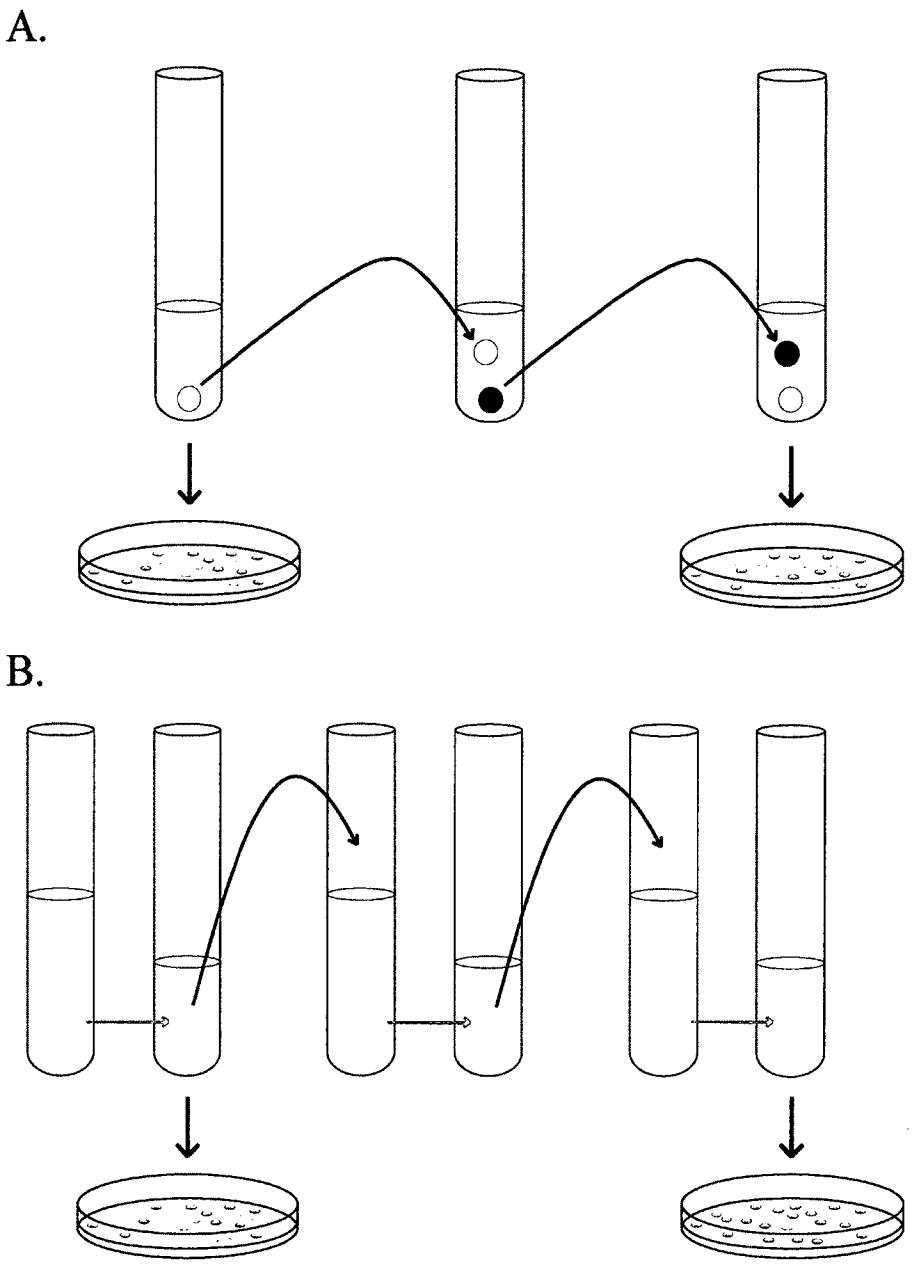


Figure 1.2: Experimental evolution model. A) Biofilm evolved populations: oppositely colored polystyrene beads were transferred every 24 hours to fresh media and plated every 48 hours. B) Planktonic evolved populations were transferred every 24 hours with a 1:100 dilution into phosphate buffered saline, then a 1:100 dilution into fresh media (1:10,000 dilution total) and plated every 48 hours.

Planktonically evolved populations were similarly founded by inoculation of a clone in Tryptic Soy Broth and then subcultured and preconditioned in 1% Galactose Minimal Media (1% GMM). Experimental replicates were then founded by adding equal amounts of oppositely marked ancestor to fresh 1% GMM, with an overall dilution of 1:100. After 24 hours, populations were diluted 1:100 into PBS, from which they were then diluted 1:100 into fresh media and again incubated. Transfers continued with daily 1:10,00 dilutions and populations were sampled every 4 days by plating on ½ Tsoy agar with X-gal (Figure 1.2).

Mutants assumed to differ from the ancestral strain by a single mutation were identified by either a skew in the 1:1 ratio of oppositely marked ancestor, indicating a mutation had occurring on the majority background (Hegreness *et al.*, 2006), or the presence of an altered colony morphology. These altered morphologies, referred to as studded (ST) and wrinkly (W) based on their appearance, are known to be associated with biofilm adaptation (Poltak and Cooper, 2011). Individual evolution experiments were discontinued once the first mutant in that population was discovered and a single clone of each was then isolated and frozen.

Fitness Assays

Fitness effects of each mutant were measured by direct competition with the ancestor in three- or four-fold replication. All mutants and their oppositely marked ancestor were separately revived and preconditioned in their selective environment, then added 1:1 to fresh media in the method by which they were originally evolved. Biofilm competitions were seeded using half the contents of a single bead for both mutant and ancestor. Planktonic competitions were created by adding 50 µl each of mutant and

ancestor to 9.9 ml PBS, then transferring 50 µl to 5 ml 1% GMM for a 1:10,000 dilution overall. Cultures were immediately sampled by diluting biofilm competitions to 10^{-4} and planktonic competitions to 10^{-2}, then plating 100 µl on ½ Tsoy-Xgal and incubating. Competitions were again sampled after 24 hours. For biofilm competitions, the bead was removed, its contents vortexed into PBS, then diluted to 10^{-5} and plated on ½ Tsoy-Xgal. Planktonic competitions were diluted to 10^{-6} in PBS, then similarly plated. All plates were incubated for 24 hours at 37°C, then allowed to develop at room temperature for 24-48 hours before colonies were counted. The number of colony forming units (CFUs) at T = 0 and T = 24 hours were used to determine individual yield, accounting for dilutions. Yield was then used to calculate mutant and ancestor Malthusian parameters (m) as:

$$m = \ln\left(\frac{yield_{T=24}}{yield_{T=0}}\right)$$

The difference in Malthusian parameters is then defined as the selection rate (r) and used to calculate the difference in the rate of increase between ancestor and mutant over 24 hours (Lenski, 1991):

$$r = m_{evolved} - m_{ancestor}$$

Replicate selection rate values were then averaged, to calculate a mean selection rate constant for each mutant. The mean selection rate, reported as units/time, was then used to compare the measurable fitness effects of all mutants. An internal control was performed with each assay by competing the Lac + marked ancestor vs the Lac – marked ancestor. Control selection rate values significantly deviating from 0 indicated a bias favoring one of the marked ancestors, and any fitness values simultaneously obtained were discarded. The variance in replicate mutant and control selection rate values, defined as 95% confidence intervals, was used as a measurement of experimental error.

Statistical Analyses

According to extreme value theory, the right tail of a probability distribution of all fitness effects can be modeled using a generalized Pareto distribution (GPD). The GPD has the following cumulative distribution function, and in which the right tail defined by its shape (κ) and scale (τ) parameters as one of three functions:

$$F(x|\kappa,\tau) = \begin{cases} 1-\left(1+\frac{\kappa x}{\tau}\right)^{-1/\kappa}, & x \geq 0, \quad \text{if } \kappa > 0 \\ 1-\left(1+\frac{\kappa x}{\tau}\right)^{-1/\kappa}, & 0 \leq x < -\frac{\tau}{\kappa}, \quad \text{if } \kappa < 0 \\ 1-e^{-x/\tau}, & x \geq 0, \quad \text{if } \kappa = 0 \end{cases}$$

The shape parameter (κ) estimates the overall shape of the distribution as either an exponential distribution with many mutations of low benefit and few large increases (Gumbel; $\kappa = 0$); an exponential distribution in which the rate of decrease diminishes over time, resulting in a heavy right tail with many mutations of large benefit (Fréchet; $\kappa > 0$); or finally a non exponential distribution with a clustering mutations of higher benefit (Weibull; $\kappa < 0$). The scale parameter (τ) provides an estimate of the spread of the distribution, with large values indicating a broader range.

To account for the predicted loss of smaller benefit mutations, mean fitness measurements were normalized by subtracting the lowest fitness value from all others. These values were then used with a statistical program designed by Beisel *et al.* (2007) that performed a likelihood-ratio test to determine whether the data fit the null, exponential distribution. The R program compared the likelihood that fitness effects were best described by a Gumbel, exponential distribution ($\kappa = 0$) or an alternative model ($\kappa \neq 0$) by

calculating negative twice the difference in log-likelihood (-2logΛ) (R Development Core Team, 2011). The test also estimates the value of both κ and τ, which are used to describe the actual distribution. P-values are calculated based upon 10,000 parametric bootstrap replicates, and significant values ($p<0.05$) rejected the null hypothesis in favor of an alternative distribution (Beisel *et al.*, 2007).

All other statistical tests and measurements were performed using measured fitness effects in JMP 9.

Genome resequencing

Full genome resequencing was used to identify mutations in each of the bead-evolved mutants. Genomic DNA was individually isolated from each mutant using the DNeasy Blood & Tissue Kit (Qiagen) protocol for gram-negative bacteria and prepared for Illumina sequencing with direct read barcodes (Nugen). Reads were mapped (Table S1) to the previously sequenced *B. cenocepacia* H12424 reference genome (DOE-Joint Genome Institute) and mutations were identified using the breseq pipeline (Barrick and Knoester, 2010).

Results and Discussion

Seven replicate populations were evolved under conditions selecting for a cycle of biofilm formation, dispersal, and reattachment, which produces a structured environment. From these populations, 18 mutants (BM) were selected based on their altered colony phenotype (Table 1.1). Because these same morphologies have been recovered from similarly evolved populations in the past (Poltak and Cooper, 2011), and are known to have a heritable genetic basis, they were immediately classified as mutants. All biofilm populations produced mutants on both the Lac+ and Lac- genetic background, which

likely explains why a skew in the 1:1 marker ratio was never seen. Of the mutants identified, 13 had a wrinkly phenotype; three were studded; and two mutants were found with a new colony phenotype referred to as "tiny mucoid" (TM). Seven mutants were isolated from two of the populations after just four days of serial transfer, while the other eleven mutants were isolated from the remaining five populations after eight days. Growth and generation time estimates for *B. cenocepacia* HI2424 under evolutionary conditions approximate ~8 generations every 24 hours (Traverse, 2012). Therefore, all mutants were isolated after only 32 or 64 generations of selection.

Nineteen populations were also evolved under planktonic conditions without a plastic bead, from which 19 mutants (PM) were isolated (Table 1.1). One mutant was isolated from each evolved population after either 8 or 12 days of evolution. Mutants were identified by a skew in the 1:1 marker ratio, after which a single colony of the majority background was chosen. Generation times under planktonic evolutionary conditions are calculated as $\log_2$(dilution factor). Therefore, the 1:10,000 daily dilution factor was used to estimate approximately 13 generations in 24 hours, thus mutants were collected after approximately 104 or 156 generations. Although mutations increasing fitness are theorized to be rare events, their quick appearance in both biofilm and planktonic environments may suggest otherwise. The fact that a single mutant was identified in each planktonic population, as opposed to the numerous mutants that appeared in biofilm populations, may indicate that fewer mutants were generally arising. Regardless, adaptive mutants were rapidly isolated from all populations, suggesting that the amount of overall beneficial variation available to these adapting populations was relatively abundant.

Biofilm-Adaptive Mutants				Planktonic-Adaptive Mutants			
Mutant	Population	Morphology	Marker	Mutant	Population	Morphology	Marker
BM1	BB1	Wrinkly	Lac +	PM1	PA1	Mucoid	Lac –
BM2	BB1	Wrinkly	Lac –	PM2	PA2	Mucoid	Lac +
BM3	BB2	Wrinkly	Lac +	PM3	PA3	Mucoid	Lac +
BM4	BB2	Wrinkly	Lac –	PM4	PA5	Mucoid	Lac +
BM5	BB2	Studded	Lac –	PM5	PA6	Mucoid	Lac +
BM6	BB3	Wrinkly	Lac +	PM6	PA7	Mucoid	Lac +
BM7	BB3	Wrinkly	Lac –	PM7	PA8	Mucoid	Lac –
BM8	BB4	Wrinkly	Lac +	PM8	PA9	Mucoid	Lac +
BM9	BB4	Wrinkly	Lac –	PM9	PA10	Mucoid	Lac –
BM10	BB5	Wrinkly	Lac +	PM10	PA11	Mucoid	Lac +
BM11	BB5	Wrinkly	Lac –	PM11	PA12	Mucoid	Lac –
BM12	BC1	Wrinkly	Lac +	PM12	PA13	Mucoid	Lac –
BM13	BC1	Tiny Mucoid	Lac +	PM13	PA14	Mucoid	Lac –
BM14	BC1	Wrinkly	Lac –	PM14	PA15	Mucoid	Lac +
BM15	BC1	Studded	Lac –	PM15	PA16	Mucoid	Lac +
BM16	BC3	Wrinkly	Lac +	PM16	PA17	Mucoid	Lac –
BM17	BC3	Tiny Mucoid	Lac +	PM17	PA18	Mucoid	Lac –
BM18	BC3	Studded	Lac –	PM18	PA19	Mucoid	Lac –
				PM19	PA20	Mucoid	Lac –

Table 1.1: Biofilm- and planktonic-adaptive mutants. Mutants isolated from biofilm (bead) and planktonic environments, their individual morphology, presence (+) or absence of (-) of the Lac marker, and the population from which it was isolated. Only a single mutant with of the same morphology and Lac background was isolated from each biofilm population, to prevent repeats.

The genomes of each of the eighteen biofilm-adapted mutants were individually sequenced and analyzed using breseq to identify the genetic basis of adaptation (Table S2). Breseq is a computational pipeline that predicts mutations in slightly divergent re-sequencing data aligned to a previously sequenced reference genome. It was specifically designed to accommodate microbial genomes up to 10Mb in size, and is an excellent tool for monitoring mutations over time in experimental populations. Short-read sequences were aligned using Bowtie 2 and individual base quality scores were re-calculated to include new information from the alignment, including the reference base and its position within the read. These re-calibrated estimated error rates were then used with the haploid SNP caller, which calculated the Bayesian posterior probability of all bases at each position in the alignment and recorded the base with the highest likelihood. Base substitutions were identified when the read alignment (RA) evidence quality score exceeded a specified cutoff for consensus mutations (E-value = 10). Single nucleotide polymorphism (SNP) mutations were called if a single base was affected; substitution mutations (SUB) were called if multiple substitutions occur together, or adjacent to insertions or deletions. Possible insertions and deletions (indels) are identified as candidate junctions and called if the position-hash score exceeded the individually calculated cut-off threshold. Given these stringent methods for identifying mutations, and an average coverage greater than 150 reads per base, we feel confident all true mutations were positively identified.

Genotype	Mutation				Description	Fitness (r)	Mutants
G1	wspHK	L135F	SNP	NS		1.305	BM13; BM17
G2	wspA	S285W	SNP	NS		1.183	BM3; BM7
G3	wspA	V389A	SNP	NS		1.543	BM16
G4	wspA	A452V	SNP	NS		1.389	BM1
G5	wspE	D652N	SNP	NS		0.991	BM2; BM4; BM9; BM11
G6	wspE	S726L	SNP	NS		1.453	BM6; BM8
G7	wspE	D733V	SNP	NS		0.207	BM18
G8	wspE	E565D	SNP	NS		0.910	BM5
	fabA-like	C134G	SNP	NS	3-hydroxylacyl ACP dehydratase		
G9	wspE	D696G	SNP	NS		0.884	BM10
	Cytochrome C	S119S	SNP	Syn	Cytochrome C/electron carrier		
G10	wspE	D696G	SNP	NS		1.742	BM12
	Lysine	H437R	SNP	NS	Lysine decarboxylase		
G11	wspA	Δ21bp	DEL			1.551	BM14
	exuT?	L14P	SNP	NS	Major Facilitator Transporter		
	MFT	M114F	SNP	NS	Major Facilitator Transporter		
G12	wspE	S654L	SNP	NS		0.152	BM15
	ABC	V176G	SNP	NS	ABC Transport		
	wrbA	D46E	SNP	NS	Multimeric flavoprotein		

Table 1.2: The genotypes of biofilm-adaptive mutants and their associated fitness effects, measured as selection rate constants.

Of these 18 mutants, 13 differed from the ancestral sequence by a single mutation (single-mutant), three differed by two mutations (double-mutant), and two mutants differed by three mutations (triple-mutant). Several of the 13 single mutants were also found to share the same mutation, and thus were grouped and classified by genotype for future fitness analyses (Table 1.2). Although the likelihood of an identical mutation occurring in parallel is generally very low, we determined that all mutants were independently derived. All mutants sharing a particular genotype were isolated from a separately evolved population. In one case, two mutants with the same mutation carried different Lac markers, further indicating that they did not evolve from the same lineage.

All 18 biofilm-adaptive mutants contained a mutation located in the *wsp* operon, which is known to directly influence levels of the messenger molecule cyclic diguanylate (cyclic-di-GMP) and has been characterized in *Pseudomonas aeruginosa* (Hickman *et al.*, 2005; Guvener and Harwood, 2007) Briefly, a membrane-associated protein, WspA, is constitutively activated by the addition of a methyl group from WspC, a methyltransferase. When activated, WspA causes autophosphorylation of WspE, a histidine kinase that in turn phosphorylates the di-guanylate cyclase, WspR. Once WspR is activated, it produces cyclic-diGMP by joins two GTP molecules together. Phosphorylated WspE also activates WspF, a methylesterase that removes a methyl group from WspA, forming a feedback loop and resetting the Wsp pathway. Cyclic-diGMP production by the Wsp pathway has also been associated with biofilm-specific adaptation in experimentally evolved *Burkholderia* populations (Traverse *et al.*, in press). All mutations reported here were located within three genes Wsp genes - a *wsp*-associated

histidine kinase (here referred to as wspHK), *wspA*, and *wspE* (Table S1 for details). Two mutants with a "tiny mucoid" morphology shared the same *wspHK* mutation, and all "studded" mutants contained a different *wspE* mutation. Secondary and tertiary mutations were also identified in genes likely influencing fatty acid synthesis (FabA-like) and polyamine synthesis (lysine decarboxylase), as well as major facilitator and ABC transporter proteins.

To test whether each isolated mutant is adaptive, and directly measure the magnitude of that fitness benefit, each mutant was competed against the ancestor in their selective environment (Table 1.2). All biofilm-adaptive mutants were more fit than their ancestor, with a mean fitness of 1.15/day for single-mutant genotypes, and an overall mean fitness of 1.11/day for all biofilm genotypes. Mutants isolated from, and competed in, the planktonic environment were also found to have a mean fitness increase of 1.1/day (Figure 1.3).

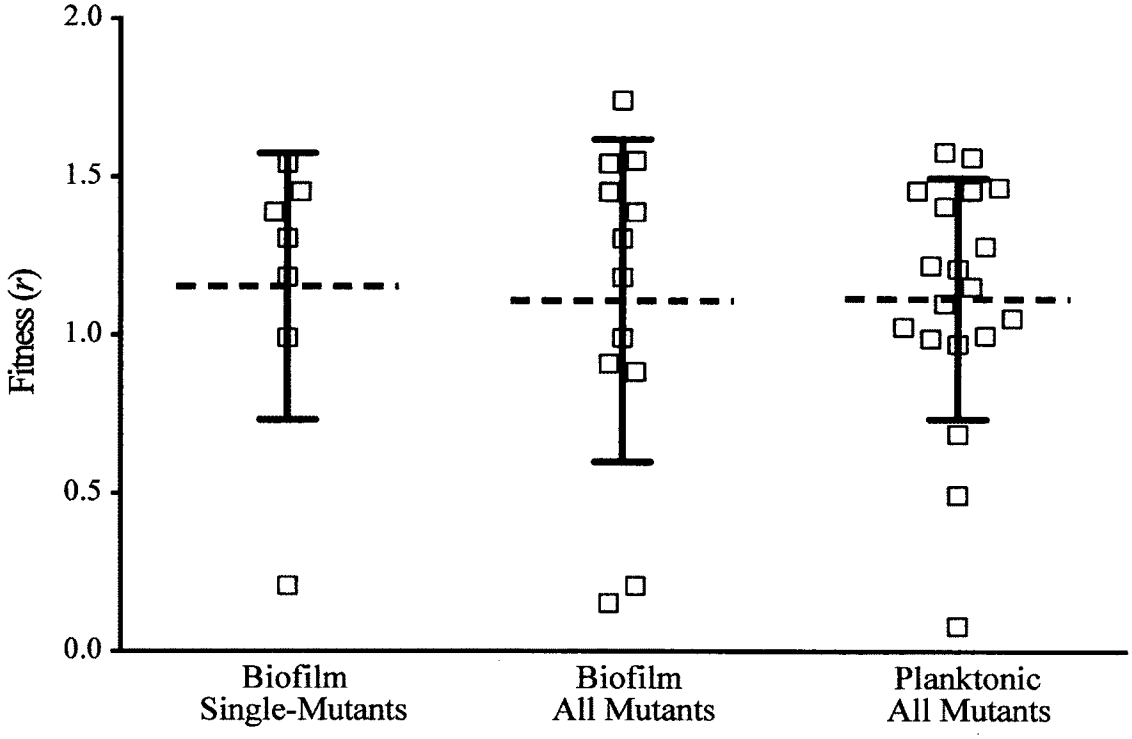


Figure 1.3: The fitness effects of beneficial mutations. The mean fitness effect, including 95% confidence intervals, of single-mutant biofilm-adaptive genotypes, all biofilm-adaptive genotypes, and planktonic-adaptive mutants.

The increased fitness of all biofilm-adaptive mutants is likely the result of increased cyclic-diGMP production by WspR. Mutations in *wspE* may lead to constitutive phosphorylation of WspR or prevent the pathway from being reset by WspF, both increasing c-diGMP. The mutation in *wspHK* may be performing a similar function by increasing phosphorylation of WspR. The membrane associated protein, WspA, directly regulates WspE activity. Mutations in *wspA* may result in constitutive autophosphorylation of WspE, again increasing c-diGMP and fitness. Interestingly, there wasn't a significant difference in fitness values among genotypes containing a single mutation and those with double or triple mutations. There was also no correlation between fitness and location of the mutation within the *wsp* operon (Figure 1.4). Genotypes G9 and G10, which share a common *wspE* mutation, have very different fitness mean fitness values ($r = 0.884$ and $r = 1.742$, respectively). Genotype G9 contained a synonymous secondary mutation in a cytochrome C electron carrier gene, which likely did not impact fitness (Peris *et al.*, 2010). A secondary mutation in a lysine decarboxylase further separates G10, and is probably influencing polyamine synthesis. The effects of both mutations in G10 may be additive or synergistic, but together are significantly higher than G9.

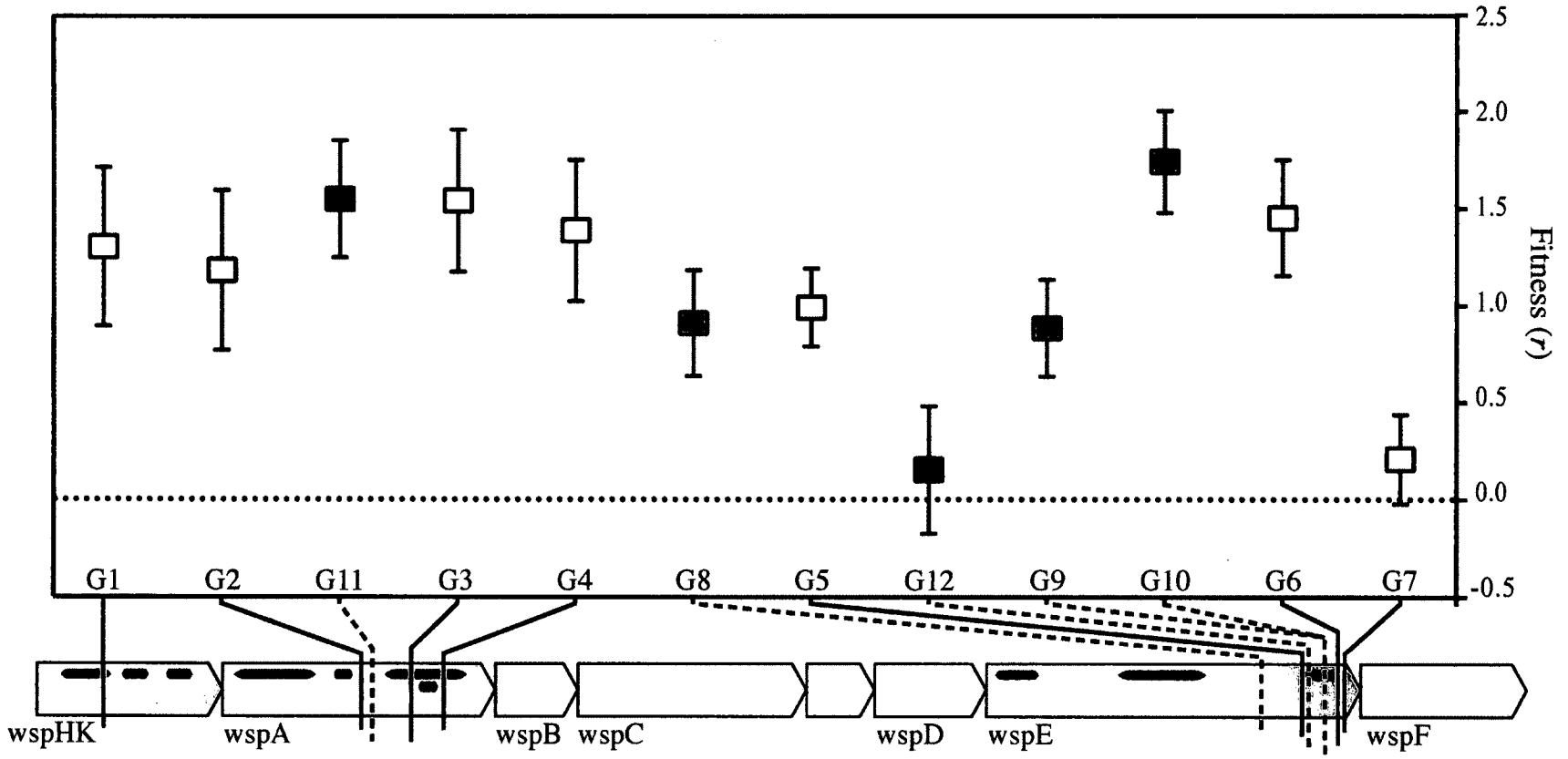


Figure 1.4: Biofilm-adaptive genotypes and their corresponding fitness values. Genotypes are ordered along the x-axis based upon the location of their *wsp* mutation, with related mean fitness (squares) and 95% confidence intervals indicated on the y-axis. Single-mutants (G1-G7) are indicated by solid lines and light gray squares, double-mutants (G8-G10) are indicated by dashed lines and medium gray squares, and triple-mutants (G11-G12) are indicated by dashed lines and black squares. Putative active sites and conserved domains identified by NCBI are highlighted in red. As shown, the greatest variation in fitness appears among *wspE* mutant genotypes; however, there is no correlation between average fitness and the number of mutations, *wsp* genotype, nor the location of individual *wsp* mutations within putative active sites and conserved domains.

The fitness effects of single-mutant (SM) biofilm-adaptive genotypes, all biofilm-adaptive genotypes and planktonic-adaptive mutants were used to determine whether the effects of beneficial mutations are exponentially distributed. Fitness values were normalized to the lowest measured value to account for the loss of smaller benefit mutations as a result of drift or competition, then used to estimate the overall shape of each distribution. The likelihood-ratio test rejected the null hypothesis that the fitness effects of biofilm-adaptive single-mutants (LRT=9.707, P=0.0003), all biofilm-adaptive mutants (LRT=12.749, P=0.0006), and planktonic-adaptive mutants (LRT=24.623, P=0.0000) are exponentially distributed (Beisel *et al.*, 2007). Instead, the right tail of a distribution of fitness effects is best described by a Weibull-shape, with a clustering of larger benefit mutations and a truncated tail. This model was supported by the shape parameter that was individually calculated for each of the three distributions ($\kappa \approx -1$). Each data set favored a similarly shaped distribution of slightly different widths, with a scale parameter (τ) of 1.34 for SM biofilm genotypes, 1.59 for all biofilm genotypes, and 1.50 for planktonic mutants. Frequency histograms and individual fitness values of biofilm-adaptive genotypes (Figure 1.5) and planktonic-adaptive mutants (Figure 1.6) support this visible clustering of fitness effects.

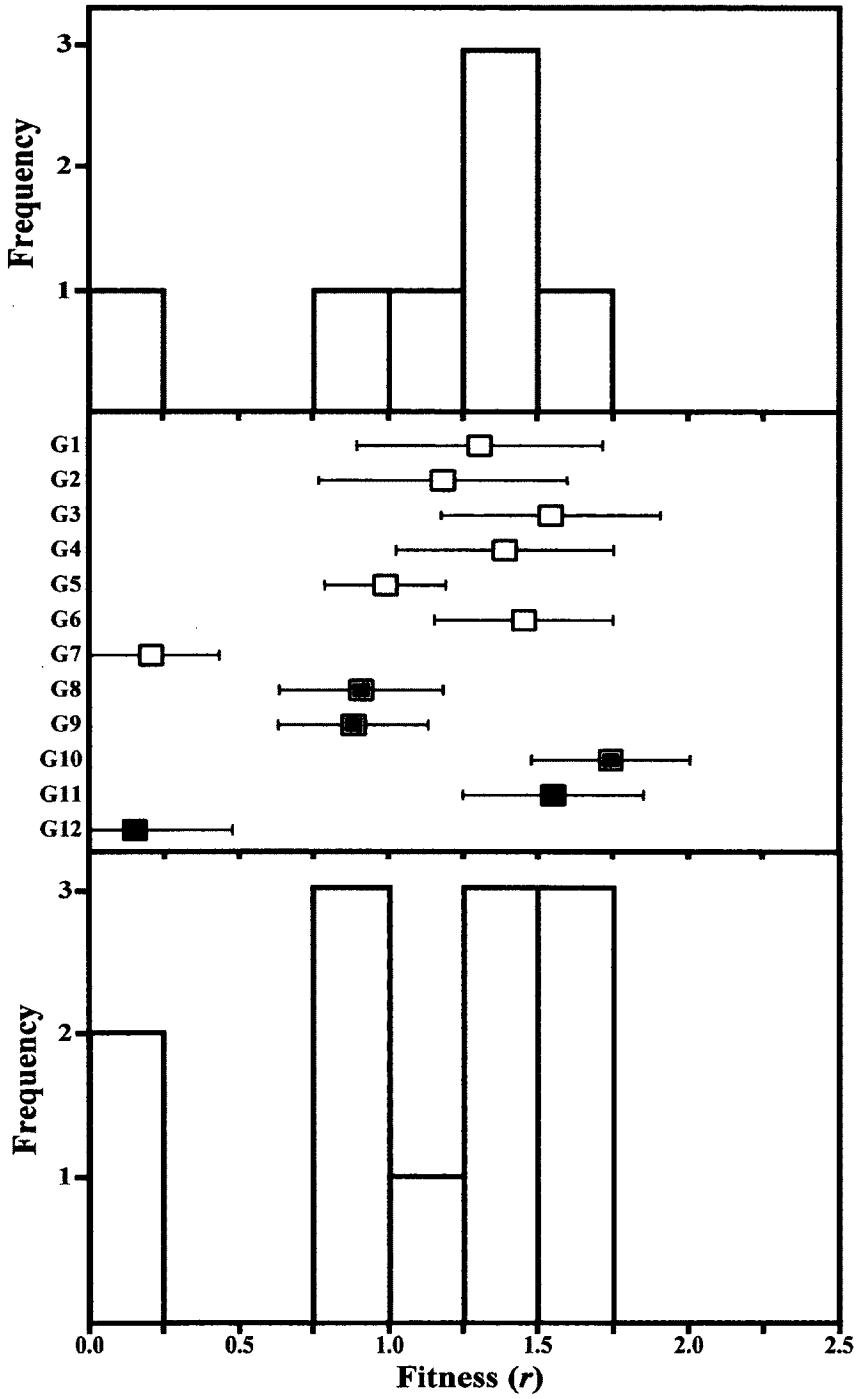


Figure 1.5: The distribution of fitness effects of biofilm-adaptive genotypes.
Top: The distribution of single-mutant fitness effects, G1-G7.
Middle: Individual fitness values, including 95% confidence intervals, for each biofilm-adaptive genotype.
Bottom: The distribution of all biofilm-adaptive fitness effects, including single mutant; double mutant, G8-G10; and triple mutants, G11-G12.

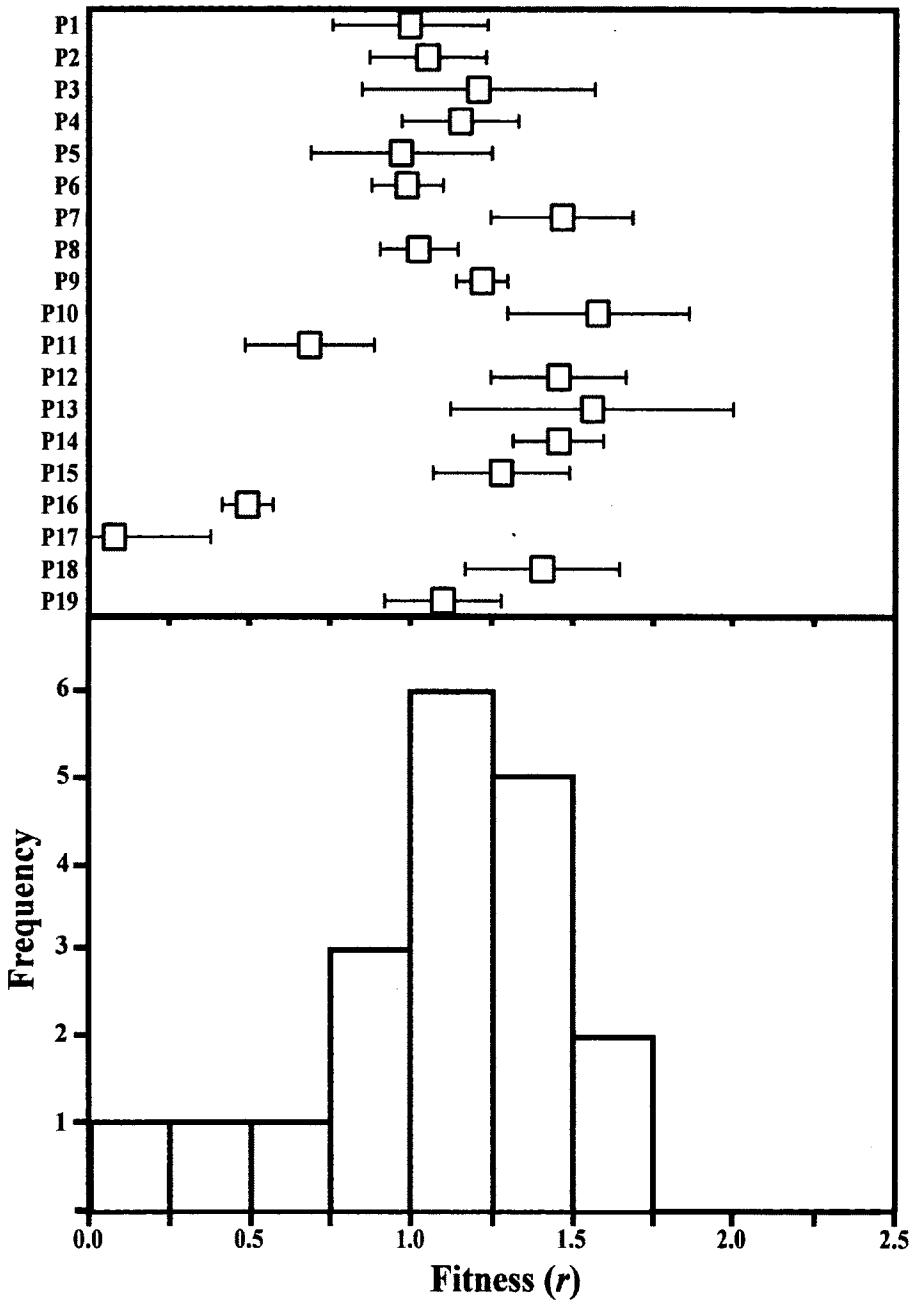


Figure 1.6: The distribution of planktonic-adaptive fitness effects. Top: The distribution of all planktonic-adapted fitness effects, P1-P19. Middle: Individual fitness values, including 95% confidence intervals, for each planktonic-adaptive mutant.

This work suggests that mutations of large fitness benefit may not be as rare as once believed, at least in some systems, and that the distribution of these fitness effects need not fit an exponential model. Although the clustered distribution of fitness effects may have been molded by selection – that is, beneficial mutants could have effects better resembling an exponential distribution if sampled prior to selection – we suggest otherwise, because these mutants were sampled long before they influenced average properties of the population and while multiple mutations were contending. Thus selection had not yet sorted among beneficial alleles to generate the clustered effects, so the distribution itself must be an inherent property of adaptation by the ancestral genotype in each of the environments.

We expected that the beneficial mutations isolated from each of the selective environment would vary, and were surprised by the congruence in fitness values. A population that is generally better adapted to the selective environment is expected to have fewer beneficial mutations available to it because their ancestor's initial fitness is much closer to the theoretical "optimum" fitness in that environment, and a larger number of mutations will be deleterious (Patwa and Wahl, 2008). Variation in observed fitness effects should be greater, therefore, when ancestral fitness is lower. In this case, ancestral fitness in the biofilm environment was much lower ($\overline{m}$ = 2.04) than that in the planktonic environment ($\overline{m}$ = 7.54), suggesting that there may be greater variation among biofilm-adaptive fitness effects (Martin and Lenormand, 2006; Vale *et al.*, 2012).

To determine whether there was significant difference between biofilm and planktonic variation in measured fitness effects, a Forsythe-Brown test for homogeneity of variance was performed. This test was chosen because a Shapiro-Wilke test concluded

that both the planktonic-adaptive and single-mutant biofilm-adaptive data sets were not normally distributed (P=0.047, P=0.048 respectively) although the data set containing all biofilm-adaptive mutants was not quite significant (P=0.133, see appendix for test statistics). The Forsythe-Brown test determined that there was a significant difference in variation between biofilm-adaptive and planktonic-adaptive fitness effects (F_{151}=9.233, P=0.003). It also found that there was no significant difference between the fitness effects of single-mutant biofilm adaptive genotypes and all biofilm-adaptive genotypes (F_{111}=0.290, P=0.591). Although the mean fitness values were very similar, this significant difference in variation indicates that the mutants themselves may be specific to the selective environment (Elena and Sanjuán, 2007).

The similarity in biofilm-adaptive and planktonic-adaptive distributions was also surprising. Previous theoretical work using extreme value theory derived an exponential distribution on the assumption that ancestral fitness in that environment was high (Gillespie, 1984). Orr also derived a exponential distribution when focusing on mutants of a single gene, and comparing the individual change in fitness (ΔW) between the wild-type allele and one the few predicted beneficial alleles containing a single point mutation (Orr, 2003). Moreover, he found that this distribution remained the same regardless of the wild-type allele's exact fitness ranking among beneficial alleles. However, this model still rested on the assumption of a well adapted ancestor, and later work suggested it may not be the best model of adaptation to novel environments. A larger subset of mutations is predicted to be beneficial in a novel environment where the ancestor is not well adapted (Barrett *et al.,* 2006). Thus, we expected that the selective environment, and ancestral fitness in that environment, would dictate the shape of the distribution.

Although our ancestral genotype was naïve to both selective environments, we expected that a higher planktonic ancestral fitness would correlate to an exponentially distributed set of beneficial fitness effects because it met the assumption of high initial fitness specified by extreme value theory. However, neither of our observed distributions were exponential, and the parallelism between them suggests that mutants of similar effect may be available to a given genotype, regardless of ancestral fitness in the selective environment, and shape of the distribution may be independent of ancestral fitness. The spectrum of effects of beneficial mutations available to natural selection is expected to determine the rate by which a population adapts (Orr, 2003; Patwa and Wahl, 2008). If a similar distribution is available regardless of the environment, it appears that the dynamics by which adaptation proceeds may not be defined by the availability of beneficial mutations, after all. Rather, the extent to which beneficial mutants co-occur and compete (clonal interference) may more strongly influence adaptive dynamics.

Overall, the non-exponential distribution of beneficial effects has been supported by a few empirical studies. Rokyta *et al.* (2008) determined that the distribution of beneficial fitness effects in a DNA virus are best modeled by a Weibull-like domain of attraction, rather than exponential. In this case, the fitness effects of several unique beneficial mutations were measured as progeny produced after 24 hours, instead of direct competition with the ancestor. The distribution of effects was estimated using the same likelihood-ratio framework described above, which accounts for the loss of small benefit mutations. This finding was in stark contrast, however, to previous work confirming an exponential distribution of beneficial effects conferring antibiotic resistance in *Escherichia coli* and *Pseudomonas aeruginosa* (Kassen and Bataillon, 2006; MacLean

and Buckling, 2009). This disparity between distributions is likely the result of variation in available beneficial mutations and large differences in initial, ancestral fitness. MacLean and Buckling found that the fitness effects of rifampicin resistant *P. aeruginosa* mutants were exponentially distributed at low levels of the antibiotic. However, the fitness effects measured at high rifampicin levels, where ancestral fitness was much lower, were no longer exponentially distributed (MacLean and Buckling, 2009). The mutants tested were collected only at high levels of rifampicin, so the distributions may actually highlight a difference in the shape and distribution of pleiotropic fitness effects. It also emphasizes the difficulty in determining a singular model of adaptation because the environment plays a heavily deterministic role.

The distribution of beneficial fitness effects is likely dictated by the internal, genetic environment as well as the external environment. As organisms adapt and population fitness increases, the availability of beneficial mutations is expected to decrease, and the measured effect of those mutations diminishes (Lenski *et al.*, 1991; Lenski and Travisano, 1994; Cooper and Lenski, 2000; de Visser *et al.*, 1999; de Visser and Lenski, 2002). It has long been assumed that the scarcity of beneficial variation ultimately limits the rate of further adaptation, however there is little empirical evidence to delineate whether large benefit mutations are altogether absent, or just no longer beneficial on an adapted background. Ongoing work focusing on this relationship between genotype and environment may help to better define the dynamics of adaptation by directly measuring the effects of all available beneficial mutations, regardless of genetic background. By sampling distributions from several populations differing in

selective environment, as well as genetic background, we hope to paint a broader picture of the forces governing the dynamics of adaptation.

CHAPTER II

THE PLEIOTROPIC FITNESS EFFECTS OF BENEFICIAL MUTATIONS

Introduction

Mutations that increase fitness by producing a beneficial phenotype may simultaneously alter a number of other phenotypes that are not directly selected. This phenomenon, known as pleiotropy, is the result of a single gene influencing multiple phenotypes, and is thought to be common (Fisher, 1958; Mayr, 1963). How these indirect effects of adaptation influence fitness when conditions and selection pressures change is poorly understood. Environmental variation and pleiotropy are thought to play a role in maintaining diversity, particularly within a heterogeneous environment (Lynch and Gabriel, 1987; Futuyma and Moreno, 1988; MacLean *et al.*, 2004; Knight *et al.*, 2006). Pleiotropic fitness effects are also expected to influence several aspects of evolution, and may help to explain why certain beneficial alleles ultimately prevail. It has been theorized that the magnitude of fitness benefit or trade-off is directly associated with that in the original, selective environment (Lande, 1983). If this is true, mutations that greatly increase fitness may experience an equally large deficit if environmental settings change. Under these more complex, potentially realistic conditions, the spectrum of pleiotropic effects may better predict survival than direct fitness alone.

Recent studies attempting to quantify the pleiotropic effects of beneficial mutations have, in large part, focused on consecutive beneficial mutations evolving on a common genetic background that have been isolated over time from experimentally evolved populations (Travisano and Lenski, 1996; Cooper, 2002; Ostrowski *et al.*, 2005; Barrett *et al.*, 2005). The majority of this work has shown a general trend toward positive pleiotropy, in which mutants that had adapted to a single resource tended to be more fit than their ancestor when directly competed in alternative carbon sources (Travisano and Lenski 1996; Ostrowski 2005), although a cost of adaptation has also been seen (Cooper and Lenski, 2000; Cooper, 2002; MacLean et al., 2005). They also found that the magnitude of fitness increase was greatest in environments most similar to the selective environment, although the same trend was not always observed for those few mutants with deleterious pleiotropic effects (Travisano and Lenski 1996; Ostrowski 2005; Cooper and Lenski 2000, 2001, Cooper 2002). These studies have mainly focused on the pleiotropic effects of mutants that were already well adapted to their environment, while little has been reported about how pleiotropy may influence the very first steps of adaptation. By studying the pleiotropic effects of several mutants differing from a common ancestor by a single mutation, we hope to better understand the role of pleiotropy in adaptation. Given prior findings, we expected that the majority of pleiotropic effects would be positive although some mutants would experience fitness costs in environments more divergent from their original one. Additionally, we were interested in how the initial selective value compared to the magnitude of fitness gain or loss in alternative environments, and predicted a positive correlation based on previous

studies that found a significant positive correlation among only mostly favorable fitness effects in several environments (Ostrowski *et al.*, 2005).

We evolved replicate populations founded by a single strain of *Burkholderia cenocepacia* in a structured environment selecting for biofilm formation. Populations founded from that same ancestor were also evolved in a liquid, planktonic environment (Poltak and Cooper, 2010). Isolates containing a single beneficial mutation were collected and competed against the ancestor to measure their direct fitness effects. We then quantified the pleiotropic effects of these mutants by directly competing them against the ancestor in alternative environments. In comparing different aspects of the biofilm environment alone, we also examined what niche the mutants had adapted to occupy within a heterogeneous environment containing surface structure.

Prior work has suggested that adaptation to a heterogeneous environment comprised of two highly contrasting physical niches, such as the structured surface and liquid phases of our experimental biofilm environment, would ultimately favor two optimal phenotypes- each able to occupy one of the niches (Futuyma and Moreno, 1988; Gillespie and Turelli, 1989; Kassen, 2002; MacLean *et al.*, 2005; Jasmin and Kassen, 2007). Mutations that are beneficial in one patch are, therefore, less likely to be beneficial in the other (Via and Lande, 1985). Because only cells attached to the plastic bead are transferred, we expect that our collection of mutants will contain biofilm specialists capable of exploiting this niche, and are predicted to have the largest fitness benefit in a structured environment. As a result of this specialization, however, mutants may be more likely to experience negative trade-offs and reduced niche breadth (Cooper, 2002; Jasmin and Kassen, 2007). We hypothesized that biofilm-adaptive mutants would

be more likely to pay a pleiotropic fitness cost in the absence of a surface for biofilm growth because of their specific adaptation.

Materials and Methods

Experimental Evolution and Collection of Mutants

Mutants of *Burkholderia cenocepacia* strain HI2424 were isolated from one of two environmental conditions as previously explained (Chapter 1). Briefly, replicate populations of the ancestral genotype were experimentally evolved in two environments: a liquid, planktonic environment and a physically structured, biofilm environment. Variants with a suspected beneficial mutation were collected, and their fitness benefit was measured by head-to-head competition with the ancestor in their selective environment. Altogether, 18 biofilm-evolved and 19 planktonic-evolved beneficial mutants were collected and used to assay fitness in alternative environments.

All isolates were maintained at -80°C; culture conditions remained the same except when noted. Mutant and ancestor monocultures were always preconditioned for 24 hours in the alternative environment in which they were to be competed.

Fitness Assays in Alternative Environments

Fitness in alternative environments was measured using the method for direct competition previously described (Chapter I). The fitness of each biofilm evolved mutant was measured in the opposite selective environment (planktonic), as well as bead fitness in lower nutrient environments, alternative carbon sources, and general stress environments. For planktonic fitness, biofilm evolved mutants were competed for 24 hours in 5mL of 3% galactose minimal media (GMM) without a bead. Competitions were

created by adding 25 µl each of monoculture ancestor and mutant to fresh media, then sampled by diluting to 10^{-4} in phosphate buffered saline (PBS) and plating 100 µl on ½ strength tryptic soy agar containing Xgal (Tsoy-Xgal). Cultures were incubated for 24 hours, then sampled by diluting in PBS, and plating 50µl at 10^{-5} and 100µl at 10^{-6}.

To assay fitness in lower nutrient environments, biofilm evolved mutants were separately competed in 1% GMM and 0.3% GMM. Competitions were created, sampled, and plated using the exact method for biofilm fitness previously described. Mutant fitness in alternative carbon sources was also measured using the same method for creating, sampling, and plating competitions, but the minimal media was supplemented with either 3% fructose (3% FMM) or 1% trehalose (1% TMM) instead of galactose.

Finally, fitness of all biofilm mutants was measured under the general stress conditions of low iron and low oxygenation. Low iron levels were achieved by adding 200 µM bathophenanthroline disulfate, an iron chelator, to the 3% GMM used. Competitions were then created and sampled as previously explained. Fitness in low levels of oxygen was assayed using 3% GMM and incubating media cultures in an orbital shaker at 80 RPM. Cultures were immediately sampled at 2^4, 10^5, and 2^5; then again at 10^5, and 2^5 after 24 hours.

The fitness of planktonic evolved mutants was measured in the opposite selective environment and in a low oxygen environment. Mutants were competed in 1% GMM with a 7mm polystyrene bead using the same method to measure biofilm fitness. Low oxygen fitness was determined using the previous protocol for measuring planktonic fitness with a few alterations. All cultures were incubated in a standing rack without shaking to produce an oxygen-depleted environment. Competition cultures were created

after a 1:10,000 dilution from overnight culture and sampled at 10^2. After 48 hours, competition cultures were sampled at 10^6.

Mutant and ancestor colony forming units (CFUs) were counted for each replicate and used to calculate fitness and overall yield. Fitness values (selection rate constants) were calculated for each environment as the difference in Malthusian parameters, or difference in log yield, over the amount of time assayed (Lenski *et al.*, 1991).

Statistical Analyses

Fitness in alternative environments (indirect effects) was compared to fitness in the selective environment (direct effects), as follows. Absolute values of fitness (means of at least three-fold replication) from each foreign environment (planktonic growth, 1% GMM, 0.3% GMM, 3% fructose minimal media, 1% trehalose minimal media, iron-limited, and oxygen-limited) were averaged for each mutant. This grand mean value, referred to as the "pleiotropic index", was then regressed against direct fitness values using JMP 9. These regressions evaluated the correlation between fitness in the selective environment and the extent of fitness gain or trade-off in alternative environments, using $p<0.05$ as a statistical criterion.

Swimming Motility

Mutant and ancestor swimming ability was tested on tryptone-swim plates (1% tryptone; 0.3% agar; 0.5% NaCl) with threefold replication. Plates were inoculated with 50µl of overnight culture, then incubated at 37°C for 20 hours. The diameter of each colony was then measured in millimeters.

Results and Discussion

Adaptation to the selective environment

A single ancestral strain was evolved under a regime selecting for increased biofilm production requiring adherence to, then dispersal from, a polystyrene bead that was transferred to fresh media every 24 hours. From this environment, we isolated 18 mutants that were assumed to differ from the ancestral strain by a single mutation because they so rapidly rose to a detectable frequency within a large population of average mutation rate. Each of these different mutations was confirmed to be beneficial in the environment from which they were isolated, having increased fitness when directly competed against the ancestral strain (3% galactose with a bead, $\bar{r}_{12} = 1.11$). This benefit in the selective environment was expected because natural selection is inherently shortsighted, favoring phenotypes that are immediately beneficial. However, natural environments may be more variable and complex, ultimately influencing the success and overall adaptability of single mutations that likely influence multiple phenotypes. To better understand how beneficial mutations may influence alternative phenotypes, and subsequently fitness, we competed each of these 18 mutants against the ancestor under different environmental conditions.

Planktonic and low nutrient fitness

In studying the overall pleiotropic fitness effects of adaptive mutations, we were interested in determining how mutants had functionally adapted to their selective environment. Specifically, we hoped to identify whether mutants that evolved in an environment containing two very different niches – biofilm growth or planktonic growth -- had adapted to only one niche specifically as previously reported (Jasmin and Kassen,

2007), rather than both. Fitness in the planktonic (liquid) niche was measured by removing the polystyrene bead, and fitness in both environments were compared. Adaptation to both niches should produce equivalent fitness in each. However, we predicted that mutants had only adapted to the biofilm portion of their environment and would therefore pay a fitness cost when it was removed. Fitness was also measured in lower levels of galactose (1% and 0.3%) to determine whether mutants had solely adapted to the provided carbon source at a specified concentration. If this were the case, fitness should decrease when galactose levels are significantly reduced, regardless of whether or not a bead is present.

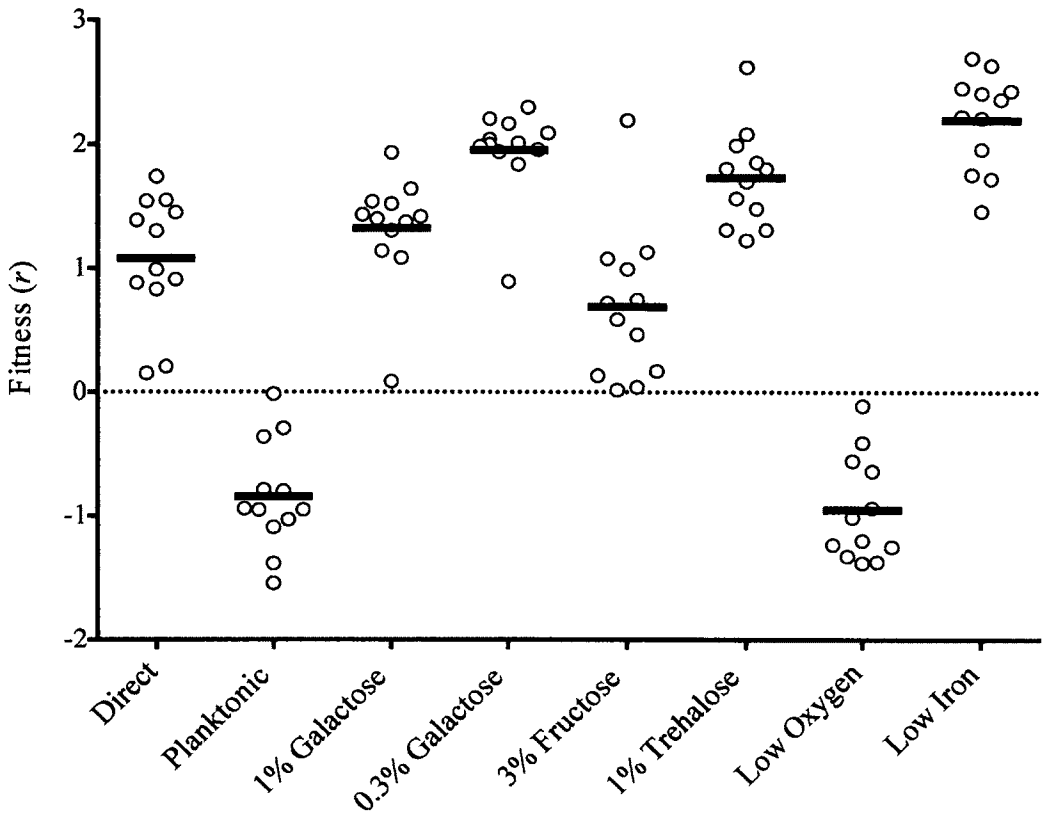


Figure 2.1: The fitness effects of all biofilm-adaptive genotypes were measured in their selective environment (3% galactose) and alternative environments, including: an unstructured environment (planktonic); lower concentrations of the original carbon source (1% galactose and 0.3% galactose); alternative carbon sources (3% fructose and 1% trehalose) and general stress environments (low oxygen and low iron).

As predicted, we found that most mutants were actually significantly less fit ($\bar{r}_{12}$ = -0.845, SD = 0.441) in the planktonic phase than their ancestor, and overall planktonic fitness was significantly lower (t_{22} = -10.045, $P < 0.0001$) than that in the selective environment (Figure 2.1). However, mutants did not experience fitness costs at lower concentrations of galactose. All mutants were more fit than their ancestor at 1% galactose ($\bar{r}_{12}$ = 1.324, SD = 0.448) and 0.3% galactose ($\bar{r}_{12}$ = 1.955, SD = 0.357). Whereas fitness at 1% galactose did not significantly differ from that in the selective environment (t_{22} = -1.098, P = 0.284), mutants were significantly more fit at 0.3% galactose than at the 3% galactose of the selective environment (t_{20} = -4.706, P = 0.0001). This increased mean fitness (Figure 2.1) may result from relaxed osmotic pressure at lower sugar concentrations. The high, initial galactose concentration (3% weight by volume; approximately 31mg/ml) may have been stressful, if not toxic, which could cause mutant fitness to increase as sugar concentrations decrease. This hypothesis is supported by the distributions of individual fitness effects (Figure 2.2) that follow a similar pattern at all concentrations of galactose, as the mean shifts slightly toward larger benefit values. The similar distribution of mutant effects across environments supports the hypothesis that biofilm mutants adapted specifically to surface growth rather than all aspects of the selective environment, and that beneficial mutations may not be adaptive under all alternative conditions.

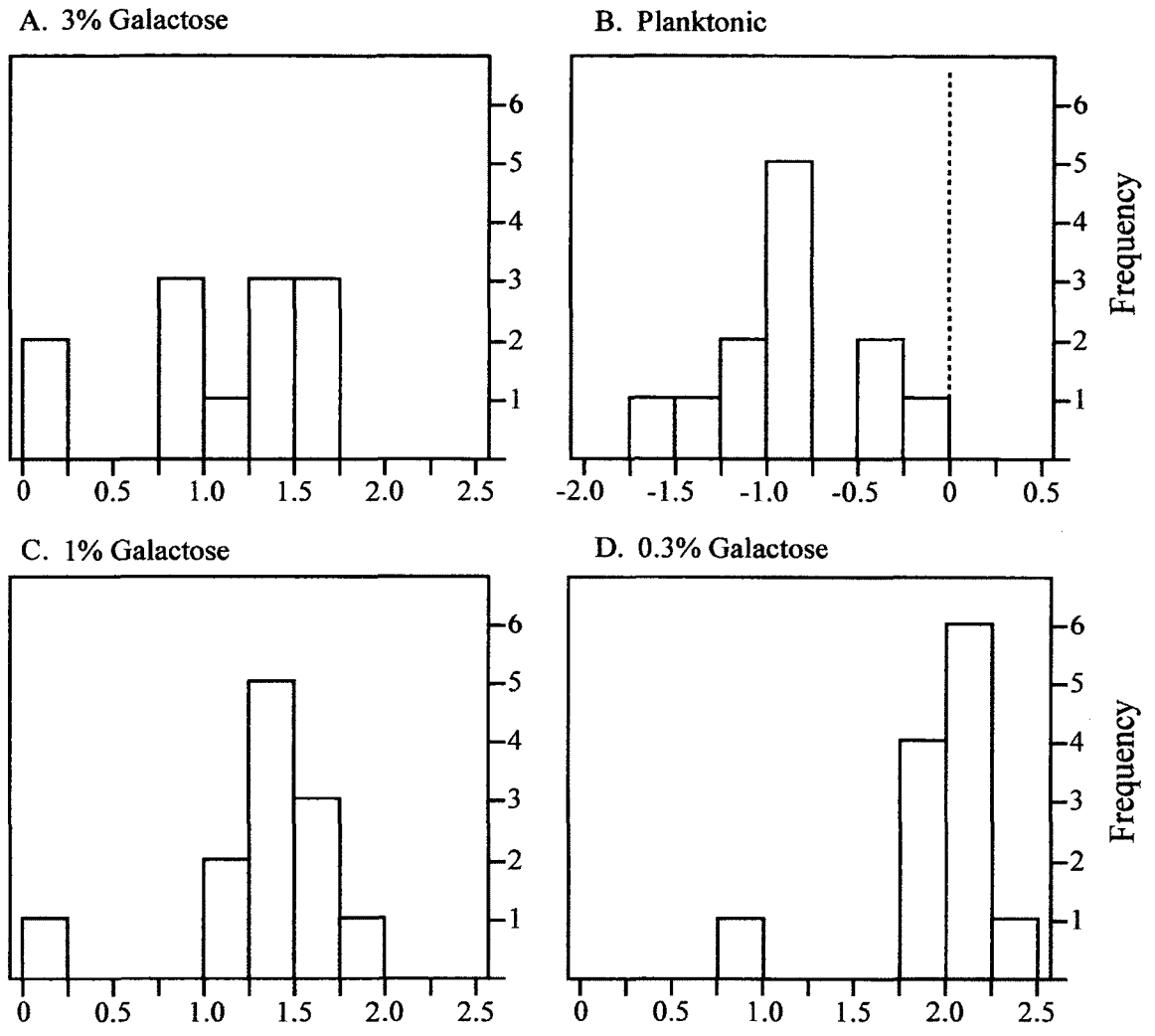


Figure 2.2: Frequency distributions of biofilm mutant fitness effects. A) Selective biofilm environment (3% galactose). B) Planktonic environment; dashed line denotes fitness of the ancestor. C) Reduced galactose (1%), exhibiting a similar distribution to the selective environment. D) Reduced galactose (0.3%).

Fitness in alternative carbon sources

Alternative carbon sources were chosen based upon their mechanisms of transportation across the cellular membranes, with the prediction that mutants would tend toward positive pleiotropy and fitness would be highest in resources most similar to the galactose (Lin, 1987; Travisano and Lenski, 1996; Ostrowski *et al.*, 2005). Galactose is transported across the outer membrane through the porin OmpF, and passes through the inner membrane through a non-phosphotransferase system (PTS). Fructose also crosses the outer membrane through the OmpF porin, but unlike galactose, is transported across the inner membrane by the PTS. Trehalose is a resource most dissimilar to galactose, and as a disaccharide, requiring the larger porin LamB to pass through the outer membrane. It is then transported across the inner membrane by the PTS (Travisano and Lenski, 1996). Trehalose is also known to sustain membrane integrity in times of dehydration and its synthesis has been shown to protect the cell from stressful conditions such as extreme temperatures or osmotic pressures (Crowe et al., 1984; Kandror et al., 2002).

As expected, the mean fitness effect of all mutants competed in fructose was positive ($\bar{r}_{12} = 0.673$, SD = 0.357), indicating that mutations were generally beneficial (Figure 2.1). The average benefit in fructose was actually lower than that in the selective environment, although the difference was not considered significant ($t_{21} = -1.899$, $P = 0.0712$) (Figure 2.3). This distribution is consistent with our predictions that most mutations would be beneficial in an alternative carbon source, yet the measurable effects would be less than in galactose.

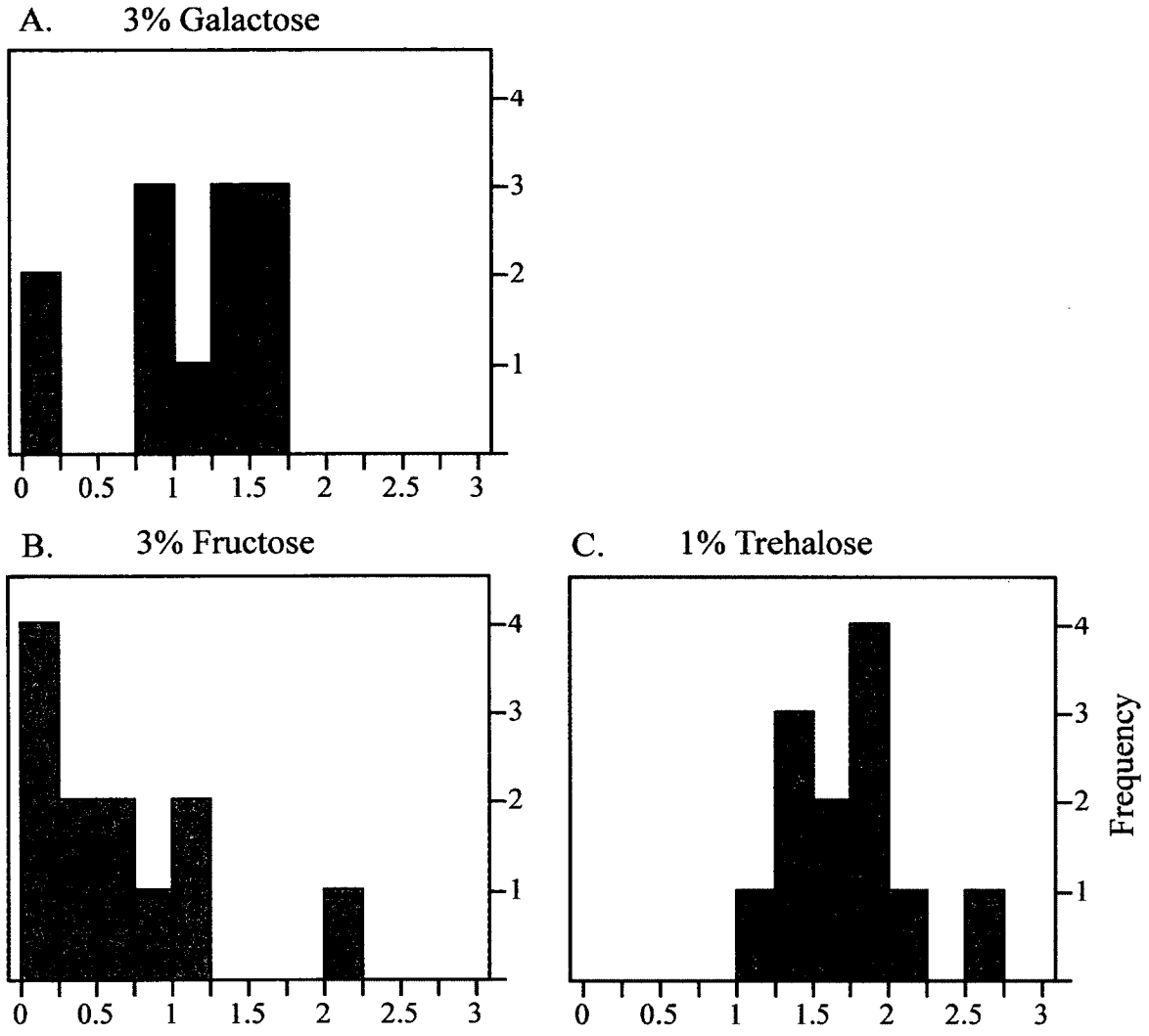


Figure 2.3: Frequency distributions of mutant fitness effects in alternative carbon sources. A) Fitness in the selective environment, 3% galactose. B) Fitness effects in 3% fructose, a sugar similar to galactose in structure but differently transported across cell membranes, are generally positive, yet the mean value is lower than that in the selective environment, as predicted. C) Fitness effects in 1% trehalose, a disaccharide sugar that is most dissimilar to galactose in terms of both structure and cellular uptake, are generally greater than those in the selective environment.

Mean fitness in trehalose was expected to be even lower than that in fructose because its cellular uptake and structure is less similar to galactose. However, we found that mean fitness in trehalose ($\bar{r}_{12} = 1.732$, SD = 0.395) was significantly greater than that in fructose and galactose ($t_{21} = 3.344$, $P = 0.0031$). This result was initially surprising, not only because it did not fit with our predictions, but because previous work focusing on targets of selection reported an opposite trend. Travisano and Lenski (1996) reported that *E. coli* mutants selected in glucose, an OmpF/PTS sugar, were less fit in trehalose than in other OmpF/PTS sugars. This suggested that OmpF transport was likely a target for selection (Travisano and Lenski, 1996). However, those *E. coli* mutants had evolved for 2,000 generations in a homogenous environment and differed from their ancestor by multiple mutations, whereas the genotypes tested here evolved for a short period of time under selection for biofilm formation and contain, at most, three adaptive mutations. These differences in the selective environment may explain the opposite responses in trehalose. For example, biofilm adapted mutants may have a fitness edge in trehalose if it plays a role in biofilm development. A recent study of *Klebsiella pneumoniae* found that expression of genes involved in trehalose metabolism were elevated during phases of biofilm formation (Wu *et al.*, 2011). Transposon mutants lacking these genes produced lower levels of biofilm and capsular polysaccharide, suggesting that trehalose is an important component in biofilm development (Wu *et al.*, 2011). If biofilm-specific mutants make better use of the provided trehalose to form robust biofilms than their ancestor, they will prove more fit. However, the underlying mechanism for these *Burkholderia* mutants in this system remains uncertain.

Fitness under general stress conditions

To determine whether adaptation to a biofilm lifestyle influenced fitness under stressful conditions, all 18 mutants were competed at reduced concentrations of oxygen and iron. Biofilms provide a protective environment for bacteria, and increased biofilm production is a common response when external conditions are stressful. Previous studies have shown that low iron levels actually induce biofilm formation in several bacterial species (Johnson *et al.* 2005; Alves *et al.*, 2010). Hence, we expected that biofilm-adaptive mutants would be more fit than their ancestor under low iron conditions, which proved to be true ($\bar{r}_{12} = 2.203$, SD = 0.403). Mutant fitness in low iron was also significantly greater than that in the selective environment ($t_{12} = 5.835$, $P < 0.0001$) and values were normally distributed (Figure 2.4). The low iron concentration was harmful to the ancestor, allowing much lower yield than that in the selective environment, where iron is more readily available (although not supplemented). However, mutant yield appeared unaffected by the low concentrations, suggesting that adaptation to a biofilm lifestyle may also provide an advantage when iron is scarce.

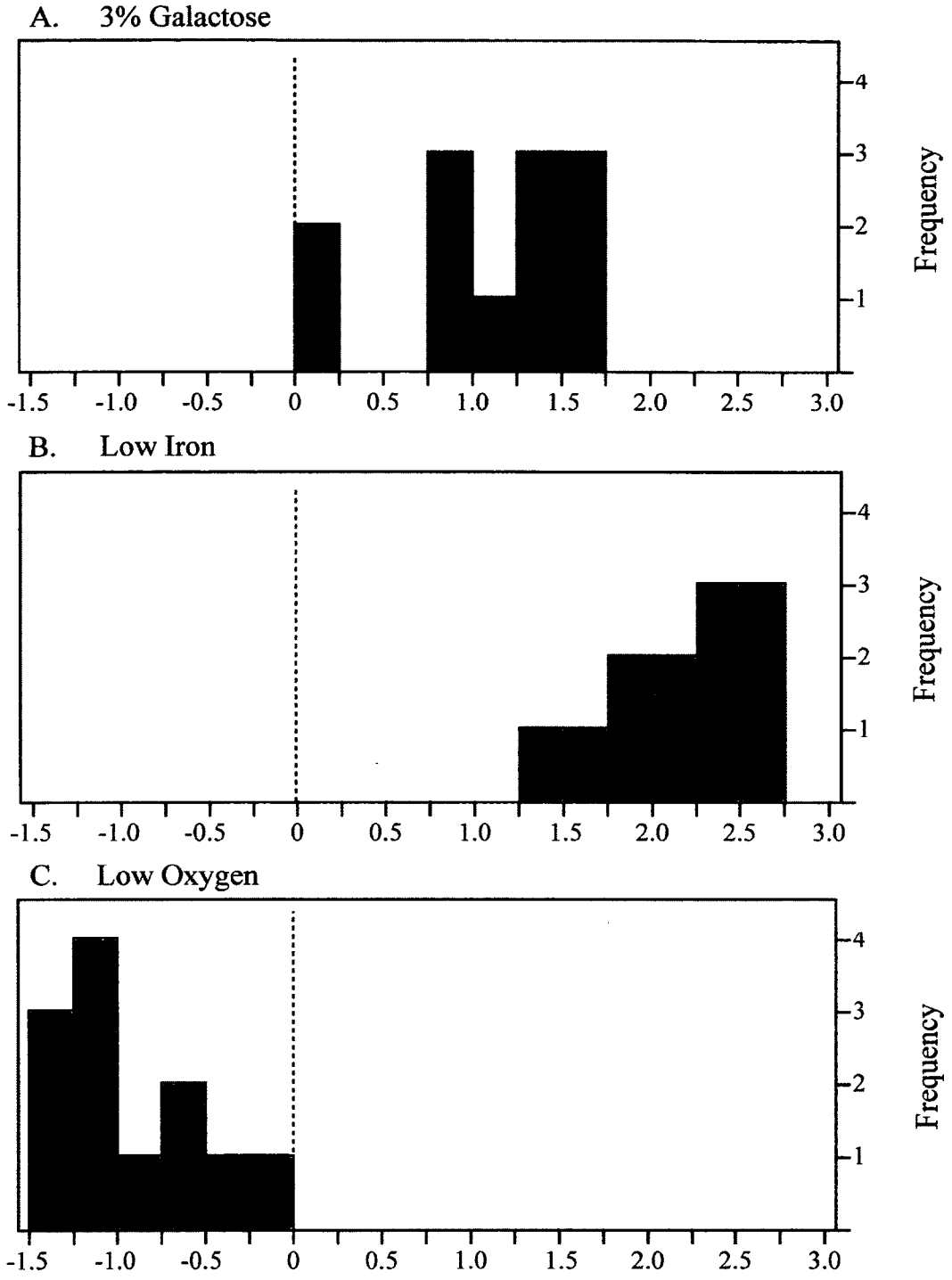


Figure 2.4: Frequency distributions of mutant fitness effects in general stress environments. A) Fitness in the selective environment, 3% galactose. B) Mutant fitness at iron concentrations was generally positive, with a mean fitness effect greater than that in the selective environment. B) Mutants were generally at a fitness disadvantage at a lower oxygen concentration, with a negative mean effect.

We also quantified fitness at low concentrations of oxygen, with the prediction that mutant fitness would again tend toward positive pleiotropy and be greater than the ancestor. Surprisingly, mutants were actually less fit than their ancestor ($\bar{r}_{12}$ = -0.962, SD = 0.416), and overall fitness was significantly lower than that in the selective environment (t_{21} = -10.906, $P < 0.0001$). Although total yield declined for both mutant and ancestor, we found that ancestral Malthusian fitness was similar to that in the selective environment, whereas Malthusian fitness of biofilm-evolved mutants was extremely low. We hypothesized that this low fitness may have been a result of the inability of biofilm-adapted mutants to occupy the planktonic phase near the air-liquid interface, where oxygen concentrations are likely higher than at the bottom of a non-shaking tube. Although the planktonic phase was never directly sampled, the ability of the ancestral genotype to swim to areas of higher oxygen concentration may have allowed it to grow to a higher density and still colonize the bead, whereas biofilm mutants were confined to areas of extremely low oxygen, restricting their overall growth. To further test this hypothesis, we measured the motility of all mutants alongside their ancestor (Figure 2.5). After 20 hours, we found that the mean swimming motility of biofilm-adaptive mutants (μ_{12}= 11.4mm, SD = 4.3mm) was significantly lower (t_{55} = -8.96, $P < 0.0001$) than that of the ancestor ($\bar{x}$ = 36.2cm). This, along with a directly measured fitness cost in the planktonic phase, supports our prediction that initial adaptation to a biofilm environment limits the ability to occupy the planktonic niche.

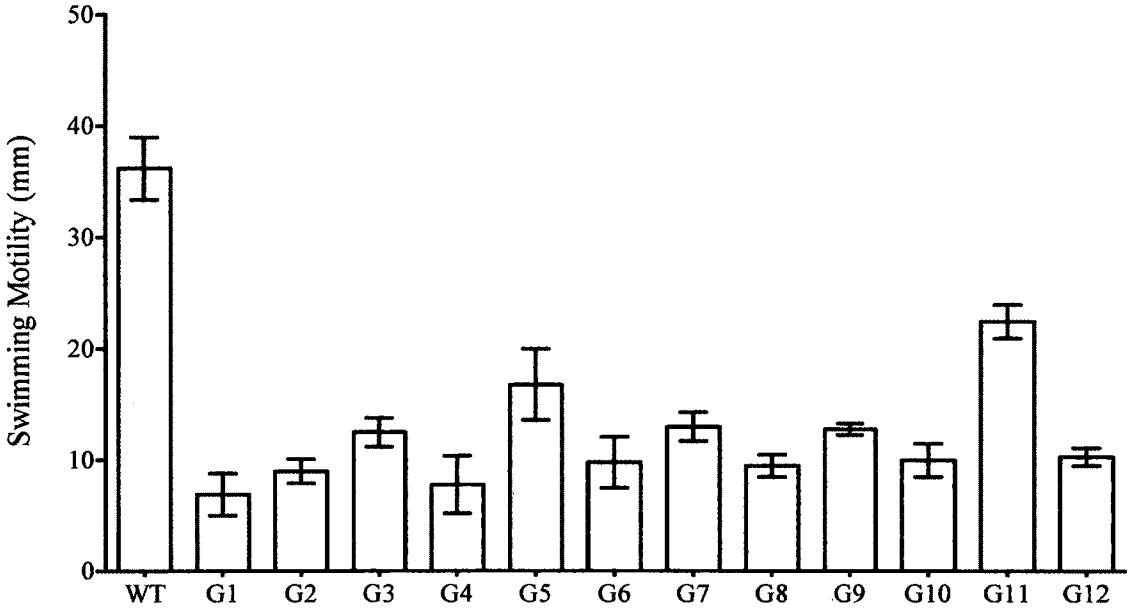


Figure 2.5: Swimming motility of biofilm-adaptive genotypes and their ancestor. All mutants appear to have lost some motility (measured in millimeters), relative to their ancestor (HI2424), as a result of adaptation to a structured environment. Error bars represent standard deviation.

The cost of adaptation

While exact fitness in alternative environments was not always predictable, the effects of biofilm-adaptive mutants were generally positive. Interestingly, all mutants tended to show a similar response in a given alternative environment, regardless of whether it was positive or negative. This symmetry contrasts with previous experiments that found that the costs of adaptation were not always predictable, and often differed even among mutants evolved under identical conditions (Travisano, 1997; Kassen, 2002; Ostrowski *et al.*, 2005). While attempting to define the influence of environmental variation on the distribution of mutational fitness effects (DMFE) in an RNA virus, Lalić *et al.* (2011) found that they were unable to predict the effect of a mutation given its effect in the original host. Another recent study concluded that the variation in DMFEs

for a single set of mutations measured in multiple hosts reflect the distinctive fitness landscapes of that host (Vale *et al.*, 2012). However, that work used viruses and focused on the distribution of all mutational fitness effects, including those that are neutral and deleterious. We are interested in defining how the fitness of mutants favored by a selective environment correlates to fitness effects.

Theory has long assumed that the magnitude of pleiotropic fitness effect would directly correlate to the magnitude of its initial fitness benefit (Lande, 1983), but limited empirical data exist and tradeoffs have been scarce (Ostrowski *et al.*, 2005). We used the complete array of fitness measurements reported here to calculate a pleiotropic index for each mutant, as the mean of absolute effects in each environment. This value, calculated for biofilm-adaptive single mutants as well as all biofilm-adaptive mutants, was then regressed against fitness in the selective environment.

We found a significant correlation ($r^2 = 0.657$, $F_6 = 9.568$, $P = 0.0271$) between direct effects and pleiotropic index for mutants with a single mutation (Figure 2.6 A). The strength of the correlation and its significance actually increased when all biofilm-adaptive mutants (single or multiple mutations) were included in the regression ($r^2 = 0.689$, $F_6 = 22.144$ $P = 0.0008$). Evidently, pleiotropic effects of secondary and tertiary mutations generate a similar pattern (Figure 2.6 B).

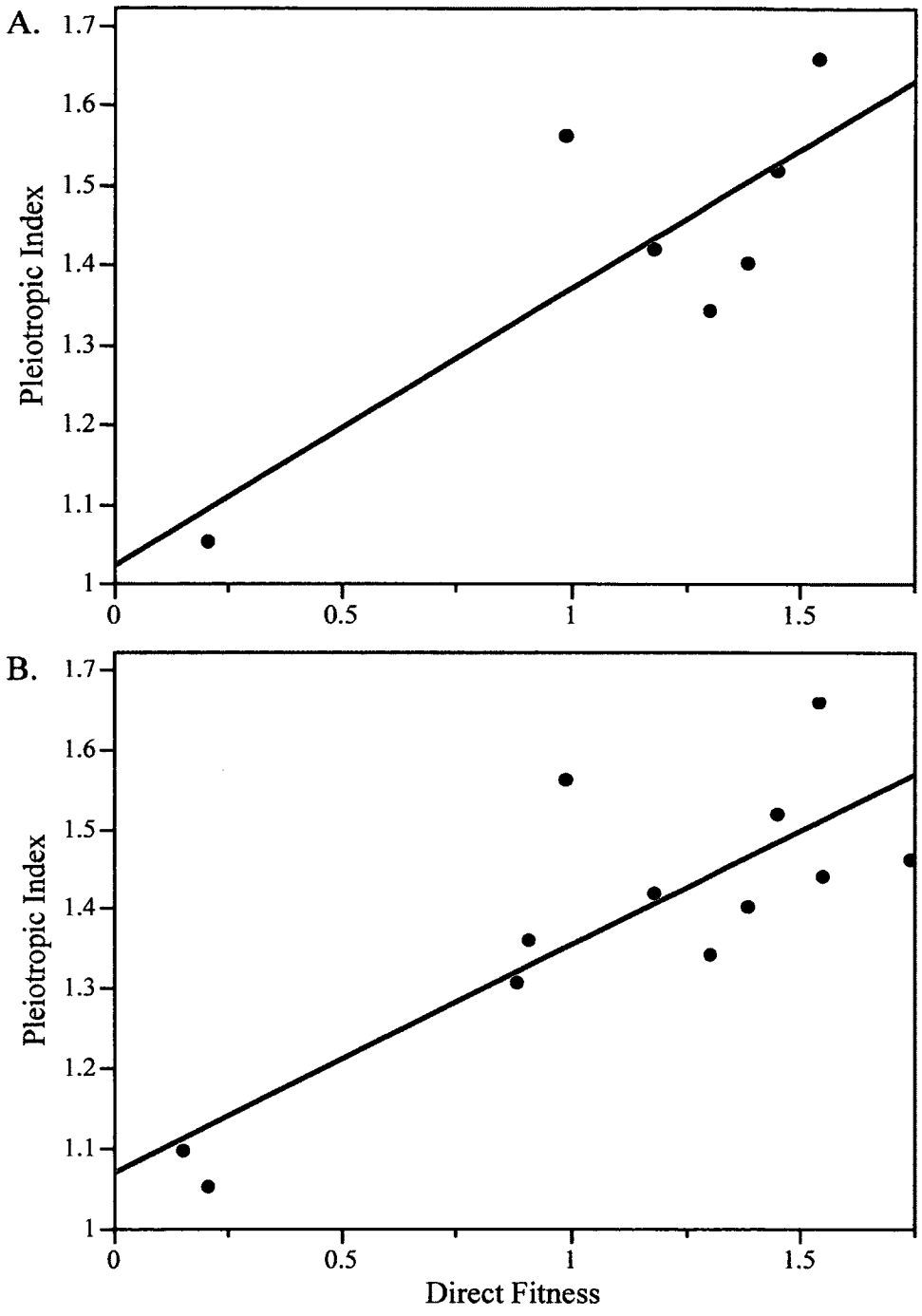


Figure 2.6: Pleiotropic Index (see text for definition) of biofilm-adaptive mutants with a single mutation (A) and all biofilm-adaptive mutants (B) strongly correlates with direct fitness effects in the selective environment.

While it is not always possible to anticipate the direction of fitness effects, this work suggests that the magnitude of indirect effect correlates well with the direct fitness effect in the selective environment, and may therefore be predictable. Defining the influence of mutations in alternative environments is crucial for developing models of adaptation and understanding the role of specialist phenotypes, particularly in heterogeneous environments. This work suggests that the fittest biofilm-adaptive variants will experience the largest trade-offs under fluctuating conditions, similar to the attachment and dispersal of biofilm communities. While the fate of mutants described here has not been characterized, prior work with similar biofilm mutants supports this conclusion. Previously, a single clone that was passaged under selection for biofilm formation diversified into a community with three morphologically distinct ecotypes. A representative of each ecotype was then separated from the community and evolved in a homogenous, liquid media. Biofilm ecotypes with the highest initial fitness experienced the greatest trade-off in motility, biofilm production, and fitness while adapting to the liquid environment, indicating that there is a cost associated with specialization (Ellis, 2011). While the original biofilm communities remained diverse, generalist ecotypes eventually dominated the population and produced new variants capable of invading the niches of specialist ecotypes (Poltak and Cooper, 2011). We believe that the pleiotropic fitness costs of specialization may directly impact the process of evolution by restricting the potential for further adaptation, and impeding long-term success within a population.

The role of pleiotropy in adaptation also remains relevant to the emergence of novel and multi-host pathogens (Gandon, 2004; Vale *et al.*, 2012; Yates *et al.*, 2006). The evolution of host specificity is dependent upon the distribution of all available

mutational fitness effects in heterogeneous environments (Pepin *et al.*, 2010). Lalic *et al.* characterized the distribution of mutational fitness effects in the *Tobacco etch potyvirus* (TEV) across multiple hosts, and found evidence that the virus could easily broaden its host range and adapt to new hosts (2011). By specifically restricting our study to beneficial mutations, we hope to better understand how adaptive mutations interact with their environment and entirely novel hosts, what role genotype-by-environment interactions in pathogen adaptation, and whether those interactions are predictable (Dennehy, 2009; Pepin *et al.*, 2010).

LIST OF REFERENCES

Alves, J. R., Pereira, A. C. M., Souza, M. C., Costa, S. B., Pinto, A. S., Mattos-Guaraldi, A. L., Hirata, R., Rosa, A. C. P. and Asad, L. M. B. O. (2010) Iron-limited condition modulates biofilm formation and interaction with human epithelial cells of enteroaggregative *Escherichia coli* (EAEC). *Journal of Applied Microbiology.* **108**:246-255.

Barrett, R.D.H., M'Gonigle, L.K. and Otto, S.P. (2006) The distribution of beneficial mutant effects under strong selection. *Genetics.* **174**:2071-2079.

Barrett, R. D. H., MacLean, R. C. and Bell, G. (2005) Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. *American Naturalist.* **166**:470-480.

Barrick, J.E. and Knoester, D.B. (2010) Breseq: Identifying More Mutations From Short-Read Microbial Resequencing Data. (Available at: http://www.barricklab.org/breseq)

Barrick, J.E., Yu, D.S., Yoon, S.H., Jeong, H., Oh, T.K., Schneider, D., Lenski, R.E. and Kim, J.F. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli. Nature.* **461**:1243-1247.

Beisel, C.J., Rokyta, D.R., Wichman, H.A. and Joyce, P. (2007) Testing the extreme value domain of attraction for distributions of beneficial fitness effects. *Genetics.* **176**:2441-2449.

Bull, J.J., Badgett, M.R., Wichman, H.A., Huelsenbeck, J.P., Hillis, D.M., Gulati, A., Ho, C. and Molineau, I.J. (1997) Exceptional convergent evolution in a virus. *Genetics.* **147**:1497-1507.

Cooper, V.S. and Lenski, R.E. (2000) The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature.* **407**:736-739.

Cooper, V.S. (2002) Long-term experimental evolution in *Escherichia coli.* X. Quantifying the fundamental and realized niche. *BMC Evolutionary Biology.* **2**:1-9.

Crowe, J. H., Crowe, L. M. and Chapman, D. (1984) Preservation of Membranes in Anhydrobiotic Organisms - the Role of Trehalose. *Science.* **223**:701-703.

Dennehy, J. J. (2009) Bacteriophages as model organisms for virus emergence research. *Trends in Microbiology,* **17**:450-457.

Elena, S.F. and Sanjuán, R. (2007) Virus evolution: insights from an experimental approach. *Annual Review of Ecology, Evolution, and Systematics.* **38**:27-52.

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

Ellis, C.N. (2008) The experimental evolution of host adaptation of the emerging pathogen *Burkholderia cenocepacia.* (Master's thesis). University of New Hampshire, Durham, NH.

Fisher, R.A. (1930, 1958) The Genetical Theory of Natural Selection. Oxford: Oxford University Press.

Futuyma, D.J. and Moreno, G. (1988) The evolution of ecological specialization. *Annual Review of Ecology and Systematics.* **19**:207-233.

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

Gillespie, J.H. (1983) A simple stochastic gene substitution model. *Theoretical Populations Biology.* **23**:202-215.

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

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

Gillespie, J. H. and Turelli, M. (1989) Genotype-Environment Interactions and the Maintenance of Polygenic Variation. *Genetics.* **121**:129-138.

Guvener, Z.T., and Harwood, C.S. (2007) Subcellular location characteristics of the *Pseudomonas aeruginosa* GGDEF protein, WspR, indicate that it produces cyclic-di-GMP in response to growth on surfaces. *Molecular Microbiology* **66**:1459-73.

Haldane, J.B.S. (1927) The mathematical theory of natural and artificial selection, Part V: Selection and Mutation. *Mathematical Proceedings of the Cambridge Philosophical Society.* **23**:838-844.

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

Hickman, J.W., Tifrea, D.F., and Harwood, C.S. (2005) A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proceedings of the National Academy of Sciences of the United States of America.* **102**: 14422-7.

Hill, W.G. and Robertson, A. (1966) The effect of linkage on limits to artificial selection. *Genetical Research.* **8**:269-294.

Isles, A., Maclusky, I., Corey, M., Gold, R., Prober, C., Fleming, P. and Levison, H. (1984) *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *The Journal of Pediatrics.* **104**:206-10.

Jasmin, J. N. and Kassen, R. (2007) On the experimental evolution of specialization and diversity in heterogeneous environments. *Ecology Letters.* **10**:272-281.

Johnson, M., Cockayne, A., Williams, P. H. and Morrissey, J. A. (2005) Iron-responsive regulation of biofilm formation in *Staphylococcus aureus* involves fur-dependent and fur-independent mechanisms. *Journal of Bacteriology.* **187**:8211-8215.

Kandror, O., DeLeon, A. and Goldberg, A. L. (2002) Trehalose synthesis is induced upon exposure of Escherichia coli to cold and is essential for viability at low temperatures. *Proceedings of the National Academy of Sciences of the United States of America.* **99**:9727-9732.

Kassen, R. (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology.* **15**:173-190.

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

Kimura, M. and Ohta, T. (1969) The average number of generations until fixation of a mutant gene in a population. *Genetics.* **61**:763-771.

Knight, C. G., Zitzmann, N., Prabhakar, S., Antrobus, R., Dwek, R., Hebestreit, H., and Rainey, P.B. (2006) Unraveling adaptive evolution: how a single point mutation affects the protein coregulation network. *Nature Genetics.* **38**: 1015–1022.

Korona, R., Nakatsu, C.H., Forney, L.J. and Lenski, R.E. (1994) Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proceedings of the National Academy of Sciences of the United States of America.* **91**:9037-9041.

Lalić, J., Cuevas, J. M. and Elena, S. F. (2011) Effect of Host Species on the Distribution of Mutational Fitness Effects for an RNA Virus. *PLoS Genetics.* **7**:e1002378.

Lande, R. and Arnold, S. J. (1983) The Measurement of Selection on Correlated Characters. *Evolution.* **37**:1210-1226.

Lang, G.I., Botstein, D. and Desai, M.M. (2011) Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics.* **188**:647-661.

Lenski, R.E. (1991) Quantifying fitness and gene stability in microorganisms. *Biotechnology.* **15**:173-192.

Lenksi, R.E. and Travisano, M. (1994) Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences of the United States of America.* **91**:6808-6814.

Lin, E.C.C. (1987) *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. Washington, D.C.: American Society for Microbiology.

LiPuma, J. J., Spilker, T., Coenye, T. and Gonzalez, C. F. (2002) An epidemic *Burkholderia cepacia* complex strain identified in soil. *Lancet.* **359**:2002-3.

Lynch, M. and Gabriel, W. (1987) Environmental tolerance. *American Naturalist.* **129**:283-303.

MacLean, R. C. (2005) Adaptive radiation in microbial microcosms. *Journal of Evolutionary Biology,* **18**:1376-1386.

MacLean, R.C. and Buckling, A. (2009) The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa. PLOS Genetics.* **5**:e1000406.

MacLean, R.C., Rainey, P.B. and Bell, G. (2004) The evolution of a pleiotropic fitness trade-off in *Pseudomonas fluorescens. Proceedings of the National Academy of Sciences of the United States of America.* **101**:8072-8077.

Mahenthiralingam, E., Bischof, J., Byrne, S. K., Radomski, C., Davies, J. E., Av-Gay, Y. and Vandamme, P. (2000) DNA-Based diagnostic approaches for identification of *Burkholderia cepacia* complex, *Burkholderia vietnamiensis, Burkholderia multivorans, Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. *Journal of clinical microbiology.* **38**:3165-73.

Martin, G. and Lenormand, T. (2006). The fitness effect of mutations across environments: a survey in light of fitness landscape models. *Evolution.* **60**:893-907.

Mayr, E. (1963) Animal species and evolution. Cambridge: Belknap Press of Harvard University Press.

Nei, M., Maruyama, T. and Chaeraborty, R. (1975) The bottleneck effect and genetic variability in populations. *Evolution.* **29**:1-10.

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

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

Patwa, Z. and Wahl, L.M. (2008) The fixation probability of beneficial mutations. *Journal of the Royal Society Interface.* **5**:1279-1289.

Pepin, K. M., Lass, S., Pulliam, J. R. C., Read, A. F. and Lloyd-Smith, J. O. (2010) Identifying genetic markers of adaptation for surveillance of viral host jumps. *Nature Reviews Microbiology.* **8**:802-813.

Perfeito, L., Fernandes, L., Mota, C., Gordo, I. (2007) Adaptive mutations in bacteria: high rate and small effects. *Science.* **317**:813-815.

Perfeito, L., Gordo, I., and Campos, P.R.A. (2006) The effect of spatial structure in adaptive evolution. *European Physical Journal B.* **51**:301-306.

Poltak, S.R., and Cooper, V.S. (2011) Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *ISME Journal.* **5**:369-378.

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Available at: http://www.R-project.org)

Reik, R., Spilker, T. and Lipuma, J. J. (2005) Distribution of *Burkholderia cepacia* complex species among isolates recovered from persons with or without cystic fibrosis. *Journal of Clinical Microbiology.* **43**:2926-8.

Rokyta, D.R., Beisel, C.J., Joyce, P., Ferris, M.T., Burch, C.L., Wichman, H.A. (2008) Beneficial fitness effects are not exponential for two viruses. *Journal of Molecular Evolution.* **67**:368-376.

Rozen, D.E., de Visser, J.A. and Gerrish, P.J. (2002) The fitness effects of fixed beneficial mutations in microbial populations. *Current Biology.* **12**:1040-1045.

Sanjuán, R., Moya, A. and Elena, S.F. (2004) The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proceedings of the National Academy of Sciences of the United States of America.* **101**:8396-8401.

Speert, D. P., Henry, D., Vandamme, P., Corey, M. and Mahenthiralingam, E. (2002) Epidemiology of *Burkholderia cepacia* complex in patients with cystic fibrosis, Canada. *Emerging Infectious Diseases.* **8**:181-7.

Springman, A. C., Jacobs, J. L., Somvanshi, V. S., Sundin, G. W., Mulks, M. H., Whittam, T. S., Viswanathan, P., Gray, R. L., Lipuma, J. J. and Ciche, T. A. (2009) Genetic diversity and multihost pathogenicity of clinical and environmental strains of *Burkholderia cenocepacia*. *Applied and Environmental Microbiology.* **75**:5250-60.

Traverse, C.C. (2012) Evolutionary genomics and ecological interactions of biofilm adaptation. (Master's thesis). University of New Hampshire, Durham, NH.

Travisano, M. and Lenski, R. E. (1996) Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. *Genetics.* **143**:15-26.

Wilke, C.O. (2004) The speed of adaptation in large asexual populations. *Genetics.* **167**:2045-2053.

Vale, P.F., Choisy, M. Froissart, R., Sanjuán, R., and Gandon, S. (2012) The distribution of mutational fitness effects of phage ϕX174 on different hosts. *Evolution.* **66**:3495-3506.

Vanlaere, E., Baldwin, A., Gevers, D., Henry, D., De Brandt, E., LiPuma, J. J., Mahenthiralingam, E., Speert, D. P., Dowson, C. and Vandamme, P. (2009) Taxon K, a complex within the *Burkholderia cepacia* complex, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. *International Journal of Systematic and Evolutionary Microbiology.* **59**:102-11.

Vanlaere, E., Lipuma, J. J., Baldwin, A., Henry, D., De Brandt, E., Mahenthiralingam, E., Speert, D., Dowson, C. and Vandamme, P. (2008) *Burkholderia latens* sp. nov., *Burkholderia diffusa* sp. nov., *Burkholderia arboris* sp. nov., *Burkholderia seminalis* sp. nov. and *Burkholderia metallica* sp. nov., novel species within the *Burkholderia cepacia* complex. *International Journal of Systematic and Evolutionary Microbiology.* **58**:1580-90.

Via, S. and Lande, R. (1985) Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution.* **39**:505-522.

Wu, M. C., Lin, T. L., Hsieh, P. F., Yang, H. C. and Wang, J. T. (2011) Isolation of genes involved in biofilm formation of a *Klebsiella pneumoniae* strain causing pyogenic liver abscess. *PLOS One.* **6**:e39805.

Yates, A., Antia, R. and Regoes, R. R. (2006) How do pathogen evolution and host heterogeneity interact in disease emergence? *Proceedings of the Royal Society B-Biological Sciences.* **273**:3075-3083.

APPENDIX

	Number of Total Reads	Average Coverage
BM1	33,500,574	435
BM2	36,668,563	476
BM3	30,991,427	402
BM4	15,446,128	201
BM5	25,323,922	329
BM6	25,767,153	335
BM7	27,238,621	354
BM8	24,797,574	322
BM9	22,792,877	296
BM10	41,754,928	542
BM11	18,004,355	234
BM12	12,734,313	165
BM13	19,118,577	248
BM14	23,739,497	308
BM15	39,444,785	512
BM16	34,317,080	446
BM17	16,066,953	209
BM18	35,337,581	459

Table S1: Sequencing statistics for each isolate.

Mutant	Morph		Position	Effect	Locus Tag	Annotation	aa change	nt change	Ref Codon	Var Codon
B1	W	Chrom2	647404	NS	Bcen2424_3786	wspA	A452V	C-->T	GCC	GTC
B2	W	Chrom2	652772	NS	Bcen2424_3791	wspE	D652N	G-->A	GAC	AAC
B3	W	Chrom2	646903	NS	Bcen2424_3786	wspA	S285W	C-->G	TCG	TGG
B4	W	Chrom2	652772	NS	Bcen2424_3791	wspE	D652N	G-->A	GAC	AAC
B5	ST	Chrom1	3065571	NS	Bcen2424_2768	3-hydroxylacyl-(ACP) dehydratase-like	C134G	T-->G	TGC	GGC
		Chrom2	652513	NS	Bcen2424_3791	wspE	E565D	G-->T	GAG	GAT
B6	W	Chrom2	652995	NS	Bcen2424_3791	wspE	S726L	C-->T	TCG	TTG
B7	W	Chrom2	646903	NS	Bcen2424_3786	wspA	S285W	C-->G	TCG	TGG
B8	W	Chrom2	652995	NS	Bcen2424_3791	wspE	S726L	C-->T	TCG	TTG
B9	W	Chrom2	652772	NS	Bcen2424_3791	wspE	D652N	G-->A	GAC	AAC
B10	W	Chrom2	838824	NS	Bcen2424_3791	wspE	D696G	A-->G	GAC	GGC
		Chrom2	2243712	Syn	Bcen2424_5193	cytochrome C – electron carrier	S119S	C-->T	TCG	TCA
B11	W	Chrom2	652772	NS	Bcen2424_3791	wspE	D652N	G-->A	GAC	AAC
B12	W	Chrom2	652,905	NS	Bcen2424_3791	wspE	D696G	A-->G	GAC	GGC
		Chrom2	899966	NS	Bcen2424_4010	response regulator receiver domain	H437R	A-->G	CAC	CGC
B13	TM	Chrom2	645194	NS	Bcen2424_3785	wspHK- histidine kinase	L135F	C-->T	CTT	TTT
B14	W	Chrom2	646969	DEL	Bcen2424_3786	wspA	Δ21bp			
		Chrom2	1073395	NS	Bcen2424_4159	major facilitator transporter-D-galacturonate and D-glucoronate	L14P	T-->C	CTC	CCC
		Chrom3	548243	NS	Bcen2424_6354	major facilitator transporter-arabinose transport	I114F	A-->T	ATT	TTT
B15	ST	Chrom1	1783269	NS	Bcen2424_1618	ABC - polar amino acid transport	V176G	T-->G	GTG	GGG
		Chrom2	652779	NS	Bcen2424_3791	wspE	S654L	C-->T	TCG	TTG
		Chrom3	288735	NS	Bcen2424_6131	wrbA- multimeric flavoprotein	D46E	C-->A	GAC	GAA
B16	W	Chrom2	647215	NS	Bcen2424_3786	wspA	V389A	T-->C	GTG	GCG
B17	TM	Chrom2	645194	NS	Bcen2424_3785	wspHK- histidine kinase	L135F	C-->T	CTT	TTT
B18	ST	Chrom2	653016	NS	Bcen2424_3791	wspE	D733V	A-->T	GAT	GTT

Table S2: Sequencing results for biofilm-adaptive mutants.

Shapiro-Wilk Test	Biofilm Single Mutants	Biofilm All Mutants	Planktonic Mutants
Count	7	12	19
Test Statistic (W)	0.807	0.894	0.899
P-value	0.048*	0.133	0.047*

Table S3: Results of a Shapiro-Wilk Test to determine goodness of fit to a normal distribution.

		Biofilm Single v. Biofilm All		Biofilm Single v. Planktonic		Biofilm All v. Planktonic	
Brown-Forsythe Test	F-ratio	0.290		4.953		9.233	
	P-value	0.591		0.028*		0.003*	
Levene's Test	F-ratio	0.316		5.229		9.915	
	P-value	0.575		0.024*		0.002*	
DF Numerator		1		1		1	
DF Denominator		111		132		151	
		Single	All	Single	Planktonic	All	Planktonic
Count		47	66	47	87	66	87
Standard Deviation		0.515	0.549	0.515	0.409	0.549	0.409
Mean Absolute Difference to Mean		0.419	0.451	0.419	0.303	0.451	0.303
Mean Absolute Difference to Median		0.416	0.448	0.416	0.301	0.448	0.301

Table S4: Results of Brown-Forsythe and Levene's Test for homogeneity of variance.