

University of New Hampshire
University of New Hampshire Scholars' Repository

Doctoral Dissertations

Student Scholarship

Winter 2014

Causes and Consequences of Diversity within Experimental Biofilms of *Pseudomonas aeruginosa*

Kenneth Mark Flynn
University of New Hampshire, Durham

Follow this and additional works at: <https://scholars.unh.edu/dissertation>

Recommended Citation

Flynn, Kenneth Mark, "Causes and Consequences of Diversity within Experimental Biofilms of *Pseudomonas aeruginosa*" (2014).
Doctoral Dissertations. 2165.
<https://scholars.unh.edu/dissertation/2165>

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.

CAUSES AND CONSEQUENCES OF DIVERSITY WITHIN EXPERIMENTAL
BIOFILMS OF *PSEUDOMONAS AERUGINOSA*

BY

KENNETH MARK FLYNN

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Microbiology

December, 2014

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Microbiology by:

Dissertation Director, Cooper, Vaughn S., Associate Professor of Microbiology and Genetics

Sullivan, Elise R., Clinical Assistant Professor

Bolker, Jessica A., Associate Professor of Biological Sciences

Cooper, Timothy F., Associate Professor of Ecology and Evolution, University of Houston

O'Toole, George A., Associate Professor of Microbiology and Immunology, Dartmouth College

On September 12th 2014

Original approval signatures are on file with the University of New Hampshire Graduate School.

ACKNOWLEDGEMENTS

Over the past six years I have received support and encouragement from many individuals that deserve mention here. First, I would like to thank Dr. Vaughn Cooper for his mentorship and willingness to support me in all my endeavors both in and out of the laboratory. I would also like to thank my dissertation committee of Elise Sullivan, Jessica Bolker, George O'Toole and Tim Cooper for their support and invaluable input on all aspects of my research from PowerPoint presentations to manuscript drafts. In addition, Dr. Joe Harrison provided countless hours of his time to teach me how to perform genetic manipulations with *Pseudomonas* strains.

During data collection and writing, Gabrielle Dowell spent months of her time keeping me on task and listening to me talk about my research. I can say with confidence that this document would not have existed today without her input and encouragement. Wendy Carlson, a laboratory technician when I first started, carried out the original experimental evolution experiment that gave rise to this body of work. Megan McLaughlin, an undergraduate researcher, helped me with numerous day-to-day activities and laboratory experiments. I have also learned so much from the diverse group of graduate students that helped me start Feed the Data Monster. I would also like to thank my parents, Karen Sealund and Peter Flynn, for their love and support. Finally, thank you to Dr. Frank Fekete and Dr. Brooke Jude for giving me the initial push to get started during my time at Colby College.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
LIST OF TABLES	VI
LIST OF FIGURES.....	VII
ABSTRACT	X

CHAPTER	PAGE
I. INTRODUCTION	1
II. THE EVOLUTION OF RARE, SUBDOMINANT SPECIALISTS DRIVES COMMUNITY COMPOSITION AND FUNCTION IN EXPERIMENTAL PSEUDOMONAS AERUGINOSA BIOFILMS	12
INTRODUCTION	13
RESULTS	15
DISCUSSION	29
MATERIALS AND METHODS	34
III. FROM GENOMICS TO ECOLOGY AND BACK: THE EVOLUTION AND MAINTENANCE OF VAST DIVERSITY IN PSEUDOMONAS AERUGINOSA BIOFILMS	39
INTRODUCTION	40
RESULTS	41
DISCUSSION	59
MATERIALS AND METHODS	62
IV. EVOLUTION OF HYPERMUTATION IN EXPERIMENTAL PSEUDOMONAS BIOFILMS AS AN ESCAPE FROM CLONAL INTERFERENCE	69
INTRODUCTION	70

RESULTS	71
DISCUSSION	84
MATERIALS AND METHODS	87
V. BIOFILM ADAPTATION THROUGH MODIFYING THE TIMING OF DISPERSAL THROUGH FINE-TUNING OF POLYPHOSPHATE PRODUCTION.....	91
INTRODUCTION	92
RESULTS	93
DISCUSSION	101
MATERIALS AND METHODS	104
LIST OF REFERENCES	110
APPENDIX A CHAPTER 2 SUPPLEMENTAL MATERIAL	123
APPENDIX B CHAPTER 3 SUPPLEMENTAL MATERIAL	134
APPENDIX C CHAPTER 4 SUPPLEMENTAL MATERIAL	144

LIST OF TABLES

CHAPTER II:

Table 1. Characteristics of representative morphotypes from the B1 population isolated following 540 generations of biofilm adaptation	21
--	----

CHAPTER III:

Table 1. Low frequency mutations correspond with specific isolates following selection	52
Table 2. Convergence suggests mutational targets resulting in ecological separation	53

CHAPTER IV:

Table 1. Expected versus observed mutational counts calculated based on the mutation rate and time in generations	76
Table 2. The expected minimum number of generations needed to accumulate the observed number of mutational calls for individual isolates.....	79
Table 3. Comparison of mutational targets identified through convergence.....	82
Table 4. Hierarchical clustering of the variance associated with functional mutations among B1 540 generation isolates.....	82

LIST OF FIGURES

CHAPTER II:

Figure 1. Evolution of morphological diversity among three replicate biofilm populations.....	17
Figure 2. Replicate biofilm-evolved populations respond differently to diversity.....	20
Figure 3. Community fitness requires diversity.....	23
Figure 4. The D morphotype facilitates the attachment of other types partly defined by variation in c-di-GMP pools.....	25
Figure 5. Altering c-di-GMP levels in individual B1 morphotypes affects fitness as a whole.....	28

CHAPTER III:

Figure 1. Mutation frequencies in metagenomes and sequenced isolates	42
Figure 2. Biofilm structure preserves immense genetic variation.	46
Figure S1. Methods defining key differences across sequencing approaches.	48
Figure 3. Frequencies of mutational calls in different samples of the B1 population and associated cohorts.....	50
Figure 4. Characterization of mutants containing mutations in the anaerobic respiration transcriptional regulator, <i>anr</i>	54
Figure 5. Population genetic structure of a biofilm population over long-term and short-term evolution.....	58

CHAPTER IV:

Figure 1. Replicate biofilm-evolved communities of <i>Pseudomonas aeruginosa</i> PA14 exhibit extensive genetic variation consistent with hypermutation.	73
Figure 2. Population genetic structure of experimental biofilm	

communities through time each characterized by hypermutation.....	75
 Figure 3. Adaptive lineages accumulate more synonymous substitutions as adaptation proceeds.	80
 CHAPTER V:	
Figure 1. The <i>ppk</i> T443A substitution is a beneficial mutation that fixes in conjunction with hypermutation.	95
Figure 2. The fitness effect of <i>ppk</i> T443A mutant depends on growth conditions.....	97
Figure 3. T443A <i>ppk</i> mutants are phenotypically consistent with PolyP pooling.	99
Figure 4. The T443A substitution alters the protein conformation of the Ppk active site.	100
Figure 5. The proposed model of how enhanced activity of the polyphosphate kinase (Ppk) influences cellular physiology.	104
 APPENDICES:	
Figure A.1. Evolution of morphological diversity among three replicate planktonic populations.	124
Figure A.2. Hypermutation repeatedly evolves only in replicate biofilm populations.....	125
Figure A.3. Isolates of individual morphotypes tend to be relatively more fit than the complete community.....	126
Figure A.4. Variation in the timing of biofilm formation and attachment amongst B1 isolates.	127
Figure A.5. Ecotype physiology is mediated by c-di-GMP levels.....	128
Figure A.6. Reduced c-di-GMP levels in the subdominant specialist P type also disproportionately affects community dynamics.	129
Figure A.7. Altering c-di-GMP levels in an individual B3 morphotype does not affect the community as a whole.	130

Figure A.8. Sonication of <i>Pseudomonas aeruginosa</i> PA14 biofilms does not reduce cell viability.	131
Figure B.1. Comparison of the distribution of mutational types identified using both of our whole-population sequencing strategies.....	134
Figure B.2. Bead selection has a limited effect on mutational clusters identified in B1 after 100 generations.	135
Figure B.3. Bead selection has a limited effect on mutational clusters identified in B1 after 264 generations.	136

ABSTRACT

CAUSES AND CONSEQUENCES OF DIVERSITY WITHIN EXPERIMENTAL BIOFILMS OF *PSEUDOMONAS AERUGINOSA*

by

Kenneth Mark Flynn

University of New Hampshire, December, 2014

Currently, we do not understand how much biodiversity may be maintained in any environment, especially not in the structured environments of biofilms, which constitute the dominant mode of microbial life. Although maintenance is associated with the complex spatial structure of biofilm architecture and resulting ecological opportunities, how both the abiotic and biotic environment defines available niches remains poorly understood. Here, we experimentally evolved three replicate populations founded by *Pseudomonas aeruginosa* strain PA14 for 540 generations under conditions favoring a regular cycle of biofilm formation and dispersal. Utilizing a variety of laboratory and genomic sequencing approaches, we characterize the population genetic structure of these hypermutator populations and identify key genotypic innovations that allowed for the evolution of rare biofilm specialists creating strong interdependency between isolates, the maintenance of distinct ecological units driven by trade-offs associated with anaerobic growth, the repeated evolution of hypermutation and modification of the 'stick-or-swim' decision cascade via polyphosphate production and the sensing of free phosphate. Overall, this body of work illustrates the vast diversity harbored in biofilms and establishes methods to identify novel targets of adaptation to structured environments amidst a very high noise to signal ratio.

CHAPTER I

INTRODUCTION

Evolutionary biology aims to define mechanisms that explain how the vast biodiversity of life on our planet evolved. Ultimately, evolutionary biologists not only wish to use the knowledge of these mechanisms to better understand history but also make predictions about how organisms are going to adapt in the future. If underlying principles can be used to predict the adaptation of organisms in any context, better management practices could be developed for a variety of applications; these include protecting endangered species, preventing the spread of antibiotic resistant strains of bacteria, treating chronic infections and improving bioremediation of disturbed environmental sites. However, making such predictions is difficult and currently remains outside of our reach. To traverse this gap, we need to develop methods that improve our ability to study evolutionary processes and elucidate new sources of variation currently absent from predictive models of adaptation.

Observing ‘evolution-in-action’

Given that evolutionary processes occur over such long time scales, direct experimentation and observation of such processes is often impossible. Despite this difficulty, a variety of approaches do exist; these include: retrospectively examining the adaptive history of organisms through the use of fossil records, mathematical modeling

and simulation, and experimental microbial evolution (EME) studies. Briefly, I will summarize these approaches with an emphasis on the recent successes of EME, in part due to the use of next-generation sequencing (NGS) technology and bioinformatic analyses of genomic data.

Retrospective approaches to understand evolutionary processes. Quantitative study of the evolution of macroorganisms has typically focused on unusual situations or case studies that mimic ideal experiments with no time constraints. For example, several fish species allow researchers to examine if evolution would proceed exactly the same if we could 'replay life's tape' (Gould, 1990). As populations of *Astyanax mexicanus* cavefish indigenous to Northeast Mexico became trapped in underwater caves, independent populations experienced similar environment conditions such as perpetual darkness and nutrient limitation. As a result, these populations display strong evidence of parallel evolution as they converge on similar adaptive phenotypes such as loss of vision, pigmentation and even sleep loss (Duboué et al., 2011). Independently isolated stickleback populations originating from oceanic populations likewise display convergence. Repeatedly adapting to the freshwater environment selects for ancient alleles either present at low frequency in the source marine population or new alleles created from *de novo* mutational changes resulting in changes to body structure and bony armor plating as they adapt to new freshwater conditions (Bell, 1995; McKinnon et al., 2004).

Recent advances in NGS technology also deepen our understanding of the genomic changes associated with parallel changes in phenotype and behavior. Sequencing of genomic DNA of multiple individuals from independently isolated

populations adapting to similar conditions allow for researchers to connect genotype to phenotype. For example, phenotypic convergence of the Mexican cavefish described above also display convergence at the genomic level (Badic et al., 2013). These studies also inform why genetic variation persists; genomics of stickleback populations suggests negative frequency dependent selection favors persistent variation of bony armor phenotypes (Marchinko et al., 2014).

Studying the evolution of macroorganisms nonetheless presents several limitations. As mentioned before, observing *de novo* evolution is difficult due to the fact that evolutionary processes occur over long time scales. Second, statistical power is often restricted by a small sample size of individuals due to logistics, cost and feasibility. Counting, tracking, monitoring and quantifying phenotypes for a large numbers of individuals in multiple populations is often difficult or even impossible. Despite rapid advances with NGS, eukaryotic genomes are complex and numerous difficulties associated with genomic assembly and mutation identification still exist. Lastly, genetic tools are typically lacking, making the creation of isogenic mutants to verify associated phenotypes impossible. Digital organisms and experimental evolution with microbes present alternatives that allow evolutionary biologists to overcome these limitations at the cost of reduced application to macroorganisms such as animals, plants and other multicellular organisms.

Computer simulation or *de novo* evolution. The evolution of artificial life or computer programs may seem like a simulation but these digital organisms behave very similarly to biological life. Self-replication, mutation, and variable genomic structure recreating higher-order epistatic interactions allow digital organisms to experience *de novo*

evolution. Placed in competition with one another, the ease of tracking and monitoring digital lineages allow researchers using artificial life to directly test hypotheses and examine evolutionary processes on a scale not possible with biological life (Maynard Smith, 1992).

A recent study with digital organisms tested millions of different genotypes to support the notion of diminishing returns epistasis during the accumulation of multiple beneficial alleles (Lenski et al., 1999). Even in bacterial systems, examining all possible combinations of this many mutations and interactions is difficult. Few studies have managed to achieve similar numbers. Only one study utilizing biological life managed to generate an impressive 225 genotypes of *E. coli* to examine an abundance of positive and negative epistatic interactions (Elena and Lenski, 1997).

Using digital organisms to study biological evolution is controversial; digital life may as well be classified as alien (Lenski et al., 1999). Despite this controversy, the biggest caveat is that the evolution of digital organisms can only inform how biological systems evolve and not simultaneously inform how they function. While there may be considerable interest in programs that can evolve to solve complex, computation challenges, microbial systems may be the best alternative providing the best middle ground between maximum statistical power and relevance.

Experimental microbial evolution studies. Laboratory experiments with microorganisms have been used very successfully to study evolutionary processes over the last 25 years. Microorganisms provide ideal candidates for studying evolution for three primary reasons: 1) microorganisms grow quickly, 2) frozen “fossil” records that

can be reanimated can be kept easily, 3) many genetically tractable microorganisms are available. Compared to most macroscopic organisms, the fast growth of microbes allows for numerous generations of growth in a single day. This decrease in generation time enables researchers to observe evolution in real time in as little as one week.

Additionally, these attributes allow researchers to examine long-term adaptation more easily. For example, in a long-term evolution experiment (LTEE) studying the dynamics of adaptation and its genetic basis in *E. coli*, 12 populations have been propagated for more than 60,000 generations and continue to adapt to a simple defined environment (Cooper, 2014; Lenski, 1991). Although the rate of increase slows, fitness of these populations continues to increase with no limit in sight despite adapting to the same laboratory conditions for 50,000 generations (Wiser et al., 2013). Combined with NGS technology, such EME studies enable researchers to link a molecular mechanism to an observed adaptive response. For example, the genetic basis and the individual fitness effects of many of the genetic changes during the LTEE continue to be identified (Blount et al., 2012). In an extension to this work coupled with genetic manipulation techniques, researchers have begun examining all possible interactions of many mutations that help shape the genomic basis of adaptation. One study examined the nature of epistasis among the first five adaptive mutations that arose and fixed in one replicate population was found to be characterized by negative epistasis where subsequent beneficial mutations confer less and less of a benefit (Khan et al., 2011). Going a step further, another study found that these interactions also fluctuate in response to different environmental conditions (Flynn et al., 2013). All in all, these studies reveal the intricacies of adaptation and the distribution of beneficial effects that

selection acts on. These capabilities allow EME studies to continually improve our understanding of how long-term adaptation proceeds.

How is the vast amount of diversity maintained in microbial populations?

Currently, we do not understand how much biodiversity may be maintained in any environment, especially not in the structured environments of biofilms, which constitute the dominant mode of microbial life. The wide success of the biofilm lifestyle is associated with the intimate connection between the complex spatial structure associated with biofilm architecture and maintenance of biodiversity (Besemer et al., 2012; Parsek & Singh, 2003; Periasamy & Kolenbrander, 2009; Yildiz & Visick, 2009). Metabolic activities of neighboring cells produce gradients of nutrients, wastes and signaling molecule concentrations that may alter local forces of selection and create distinct ecological opportunities (Poltak & Cooper, 2011). Additional forces such as endogenous oxidative stress associated with the anaerobic and resource-starved conditions of the biofilm lifestyle may also create opportunities (Dietrich et al., 2013). However, defining available niches with any certainty remains difficult or even impossible despite knowledge of the environment *a priori*.

Whole-population sequencing of experimental populations comprised of a single bacterial species has revealed how complex seemingly ‘simple’ laboratory environments truly are. For example, the classic model of adaptation assumes simple patterns of periodic selection and transient polymorphism. However, sequencing reveals that adaptation to a simple experimental laboratory microcosm involves the co-occurrence of alternative contenders and harbor rare, low frequency variation (Dettman et al., 2012; Lang et al., 2013). Although these patterns may simply reflect an abundance of

alternative adaptive solutions due a high mutation supply rate (Drake 1991), these patterns may reflect a poor understanding of available ecological opportunities (Wang et al., 2010).

A case study: dynamic chronic lung infections of persons with cystic fibrosis

(CF). Chronic infections and associated biofilm communities are characterized by high biodiversity (Lieberman et al., 2014; Mena et al., 2008; Smith et al., 2006). For example, these persistent biofilms can harbor numerous bacterial species such as *Staphylococcus aureus*, *Burkholderia* spp. and *Pseudomonas aeruginosa* (Korgaonkar et al., 2013). The presence of specific species also correlates with clinical outcomes; although *P. aeruginosa* is more common, infection with members *Burkholderia cepacia* complex (BCC) associate with poor patient outcomes (Tomlin et al., 2004). Within a single species, diversity repeatedly manifests in independent infections as the development of important phenotypes such as mucoid conversion, hypermutation, rugose small colony variants (RSCVs), loss of virulence factor production and changes in cell surface virulence determinants (Rodriguez-Rojas et al., 2011; Starkey et al., 2009). Combined, this biodiversity contributes to the resilience of biofilms in chronic lung infections and makes them difficult to treat (Rakhimova et al., 2008).

Despite considerable work identifying the genetic basis of adaptation, why evolution during chronic infections repeatedly selects for these specific traits remains difficult to discern. One explanation is that these traits reflect adaptation to specific niches commonly found within host-associated biofilms. For example, selection for modified LPS structure is thought to be a consequence of the host immune response during infection (Ernst et al., 2007; Moskowitz & Ernst, 2010). However, mutations in

LPS-related genes were also repeatedly favored within experimental biofilm populations of *B. cenocepacia* (Traverse et al., 2013). Likewise, variation in the production of the sticky, polymer known as extracellular polysaccharide (EPS) typically thought to be provide physical protection may serve additional functions (Colvin et al., 2011). For example, tracking individually-marked *Pseudomonas aeruginosa* cells and Psl EPS trails through fluorescent marking and staining revealed a structural role of this EPS molecule allowing organization in initial colony formation and cell-to-cell interactions (Zhao et al., 2013). Lastly, the appearance of mutants exhibiting an elevated mutation rate is often thought to be a consequence of antibiotic use (Mena et al., 2008). However, the importance of hypermutation during chronic adaptation likely extends beyond antibiotic resistance (Oliver & Mena, 2010). For example, an elevated mutation rate can drive the evolution of virulence, persistence, transmissibility, and lung function decline during chronic lung infections.

Why do chronic infections remain unpredictable? Despite the considerable work characterizing the variation in traits associated with adaptation during chronic infections and why this diversity might evolve, clinical studies still report unpredictable patterns of phenotypic and genetic diversification (Ashish et al., 2013). Individual infections tend to be highly unique; large differences in the abundance of certain resident species or isolates with specific phenotypes are not uncommon. Furthermore, these communities are highly dynamic through time with the majority of variation occurring within rather than between infections.

Closing the gap: the importance of studying long-term adaptation to the biofilm lifestyle.

A multitude of EME studies have used NGS data to identify genomic diversity associated with adaptation, relying on either metagenomic sequencing or genomic re-sequencing of representative isolates. Both approaches leverage convergence of mutational calls, allowing us to more accurately distinguish errors from actual beneficial variation. The choice of approach largely depends on experimental design; temporal sampling can be used to increase statistical power to detect mutations, despite less biological replication (Lang et al., 2013; Lang et al., 2011). Sampling all the genetic variation present also provides insight into rare genetic variation that may never achieve fixation (Lang et al., 2013). In contrast, full genome re-sequencing of known colony phenotypes allows genotypes to be precisely linked to phenotypes of interest (Barrick et al., 2009; Tenaillon et al., 2012). However, what kind of design would be required to sequence ecology and link genotypes to niches?

Although biofilms preserve genetic variation, thanks to ecological complexity inherent to spatial structure (Besemer et al., 2012), few studies have examined the extent to which this diversity persists. Previous work in our lab examining genomic adaptation associated with biofilms of *Burkholderia cenocepacia* HI2424 (Poltak & Cooper, 2011) found that colony phenotypes belie genetic complexity, and occasionally represent relics of presumably extinct lineages (Traverse et al., 2013). However, these findings were only possible through a combination of both whole-population and isolate sequencing with pre-existing knowledge of the ecology of the system.

Although we should be able to observe distinct genetic lineages if strong ecological interactions evolve (Herron & Doebeli, 2013), the evolution of these types of interactions may be the exception not the rule. For example, if the evolution of biofilm specialization is as easy as some EME studies of shorter time frames suggest (McElroy et al., 2014), the rapid turnover of genotypes could make deciphering ecological interactions and associated lineages difficult especially if associated niches occur on a small, or micro-, environmental scale. The repeated evolution of similar phenotypes can lead to incorrect assumption about their adaptive history.

Causes and consequences of diversity within experimental biofilms of

Pseudomonas. In an effort to improve our understanding of the evolution and maintenance of diversity, we experimentally evolved three populations founded by *Pseudomonas aeruginosa* strain PA14 under conditions favoring a regular cycle of biofilm formation and dispersal for approximately 600 generations. Each replicate biofilm-evolved population quickly gave rise to immense phenotypic and genotypic variation driven, in part, by the early fixation of hypermutation. As such, this work sought to decipher this bewildering amount of diversity utilizing a combination of classic microbiology techniques, whole-population and isolate sequencing and genetic manipulation. This work illustrates the current limitations of our ability to measure diversity harbored in biofilms and define ecological niches. Our approach builds upon previous work in these fields by testing the limits of microbiology and bioinformatics to decipher the exceptional population dynamics of these populations and creates suggestions for future studies studying the maintenance of diversity.

Here, this body of work is broken up into four experimental chapters. First, we characterize the phenotypic diversity present in replicate biofilm-evolved populations and identify the evolution of rare, specialist genotypes that generate strong interdependency between community members. Second, we characterize the population genetic structure of one of these populations in detail and identify key genotypic innovations that define the evolution and maintenance of distinct, co-existing genetic lineages. Third, we examine what aspects of the population genetic structure enabled the repeated evolution of hypermutation under our biofilm conditions. Lastly, we describe a non-synonymous substitution and the associated molecular mechanism that defines a key adaptive innovation in the biofilm lifestyle by altered regulation of intracellular levels of polyphosphate.

CHAPTER II

THE EVOLUTION OF RARE, SUBDOMINANT SPECIALISTS DRIVES COMMUNITY
COMPOSITION AND FUNCTION IN EXPERIMENTAL *PSEUDOMONAS*
AERUGINOSA BIOFILMS

INTRODUCTION

Biofilms, or aggregates of microbial cells attached to surfaces, represent one of the most successful forms of life for bacteria found in a wide variety of different environments (Besemer et al. 2012; Periasamy & Kolenbrander 2009). This success is attributed to the great amount of biodiversity, from species to genes, harbored in biofilms (Lee et al. 2014; Yarwood et al. 2007). In such biofilms, the architecture and metabolic activities of neighboring cells produce gradients of nutrients, wastes and signaling molecules that may alter local forces of selection, and create distinct ecological opportunities (Poltak et al. 2010). These opportunities enable both the creation of new variation through diversification and persistence of pre-existing variation. For example, biofilms associated with chronic lung infections of persons with Cystic Fibrosis (CF) harbor distinct colony morphotypes differing in antibiotic resistance, motility, and biofilm formation (Lieberman et al. 2011; Smith et al. 2006; Kirisits et al. 2005). This vast biodiversity often results in biofilm communities that are nearly impossible to eradicate (Drenkard & Ausubel 2002; Høiby et al. 2010; Goerke & Wolz 2010).

While biodiversity enables populations to rapidly respond to changing conditions, the evolution of biotic interactions amongst this diversity can also have a profound effect on how microbial communities adapt. The adaptive radiation of *Pseudomonas fluorescens* in static culture vials is largely driven by resource competition for oxygen (Spiers et al. 2003). However, this selective pressure also leaves ample room for strong biotic effects of specialization; successful colonization of the air-liquid interface by wrinkly-spreader specialists not only alters pre-existing niches, but creates new

ecological opportunity by modifying oxygen gradients through structure (Koza et al. 2011). These alternative outcomes may depend on the extent to which organisms shape their environment and construct new niches: some organisms may fail to thrive except in the presence of others, whereas others may exist independently of biotic inputs (Odling-Smee et al. 2003). These biotic inputs are also subsequent to change as participating members adapt further, and alter the fitness landscape of the community as a whole (Andrade-Domínguez et al. 2014).

Experimental microbial evolution (EME) offers the opportunity to study the adaptive process in real time, and when conducted in an environment favoring surface attachment one can begin to probe the evolution of biofilm diversity. While EME studies over shorter durations have had success identifying forces driving diversification and incipient interactions among emerging types (for example, see McElroy et al. 2014), longer experiments allowing more generations of reproduction are necessary to study the eventual consequences of diversity on long-term adaptation and community function. For example, short-term EME studies repeatedly observe phenotypes commonly observed during chronic lung infections such as mucoid conversion, hypermutation, rugose small colony variants (RSCVs), loss of virulence factor production and changes in cell surface virulence determinants during chronic lung infections (Moskowitz & Ernst 2010; Oliver et al. 2000; Rau et al. 2010). However, how well adaptive events over relatively short-term scales inform on the long-term adaptation by pathogens during chronic infections remains poorly understood.

To assess how diversity and biotic interactions influenced the long-term adaptation of *Pseudomonas aeruginosa*, we experimentally evolved three replicate

populations founded by strain PA14 for 540 generations under conditions favoring a regular cycle of biofilm formation and dispersal described previously (Poltak & Cooper 2011). Similar to short-term EME studies, each population evolved extensive diversity of different colony morphologies likely reflecting distinct ecological strategies and varied in EPS production, timing of surface attachment, and swarming motility. To our surprise, however, the distribution and persistence of this variation varied greatly across replicate biofilm-evolved populations over the long-term. Disrupting individual ecological strategies through modified levels of a secondary messenger molecule, c-di-GMP, in one focal population revealed the evolution of strong interdependency between isolates. In conclusion, this study emphasizes the importance of subdominant specialist types that can have disproportional effects on the community as a whole.

RESULTS

Experimental biofilm selection recapitulates some features of isolates from CF infections. Prior studies of biofilm adaptation in a defined experimental system with *Burkholderia* revealed that novel colony morphologies were linked to distinct ecological strategies within the community (Poltak & Cooper 2011). Although heritable colony types might belie even greater genetic diversity (Traverse et al. 2013), we reasoned that patterns of morphological diversity would illuminate adaptive dynamics. Plating samples of evolved communities at early, middle and late time points (approximately 100, 275, and 540 generations) from three replicate biofilm populations (B1, B2, B3) revealed distinct, heritable colony phenotypes whose composition and frequency fluctuated throughout the experiment (Figure 1). Two to seven colony types were observed across replicate biofilm populations in no consistent pattern, contrasting more predictable

patterns of diversification from a common ancestral clone seen in prior studies (39, 41, 52). Although similar phenotypes appeared in replicate populations (i.e. small colony variants (SCVs)), auto-aggregative variants, increased production of EPS as evidence by Congo Red uptake and reduced production of phenazines), these phenotypes appeared at different times throughout the experiment, occurred independently of one another, and varied in their ability to persist (Figure 1). However, highly mucoid colony variants that overproduce EPS, and are commonly observed during later stages of chronic infections, were not observed in this study. Meanwhile, the three replicate planktonic-evolved populations (P1, P2, P3) were less diverse, and more consistent among replicates and through time (Figure A.1).

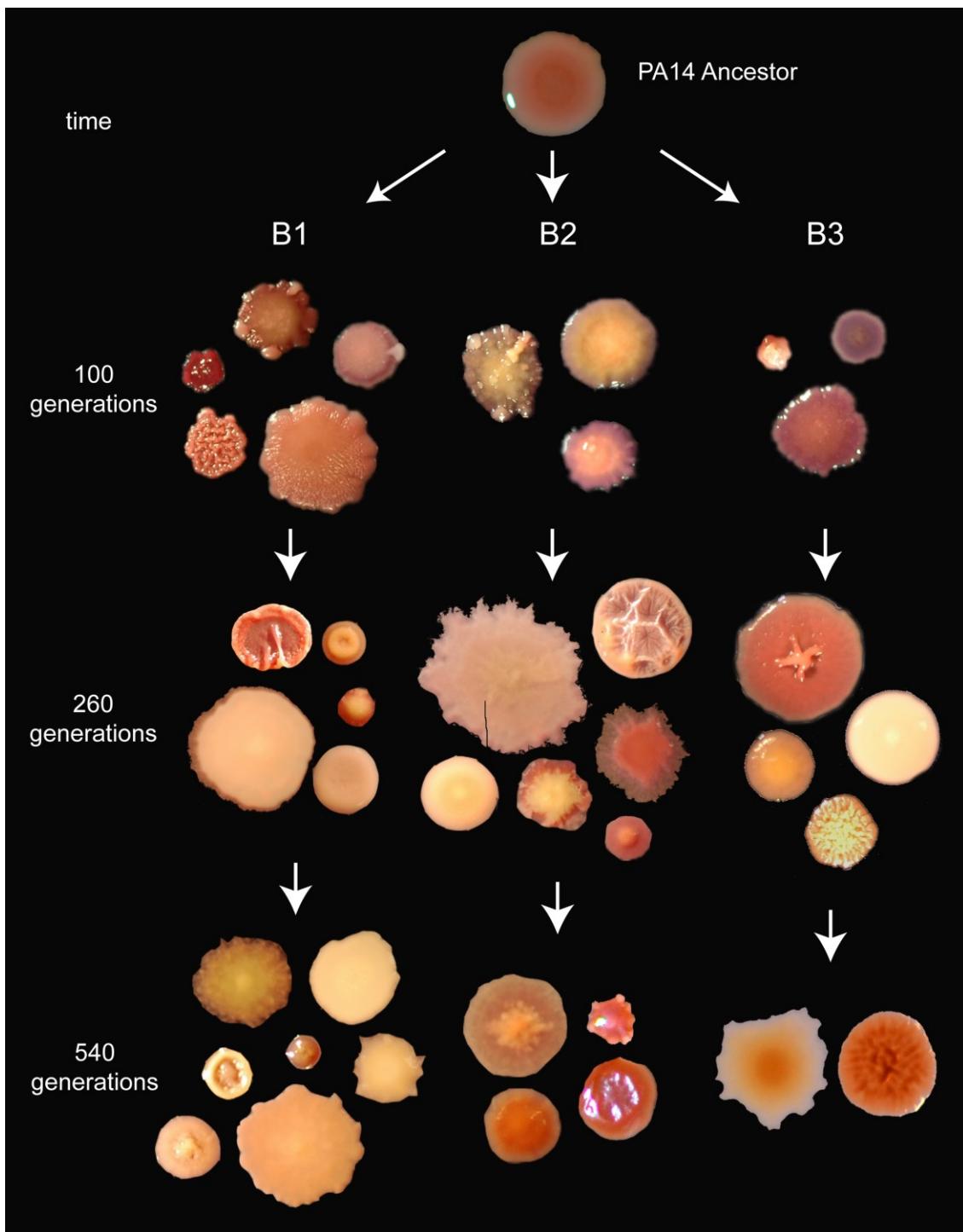


Figure 1. Evolution of morphological diversity among three replicate biofilm populations.

Aliquots from 100, 260 and 540 generations were grown on 1% tryptone supplemented with Coomassie blue and Congo red. Number and letter designations (B1-3) refer to the replication population numbers.

Increased phenotypic diversity among biofilm isolates coincided with approximately 50-fold greater rates of mutation (Figure A.2, Table S1) a phenomenon also commonly reported during chronic infections (Oliver et al. 2000). Genome sequences of evolved clones and complete communities were conducted, as previously described, to determine the likely causes of these mutator phenotypes. We will report the population-genomic dynamics of these communities elsewhere (Chapter 3), but these studies revealed substitutions in the *mutS* (T112P, which evolved independently in populations B1 and B2) and *mutL* genes (D467G, population B3) which fixed by day 44 (B1), 25 (B2), and 90 (B3), and caused the hypermutation phenotype. The *mutS* substitutions occurred at a highly conserved threonine, which has been shown in *E. coli* to be functionally indistinguishable from Δ *mutS*, and has a lower affinity for heteroduplex DNA (Junop et al. 2003). Planktonic-evolved populations retained the wild-type mutation rate and no mutations expected to influence the mutation rate were found in isolates from planktonic-evolved populations (Figure A.2).

Phenotypic diversity between replicate populations reflects differences in community function. To determine if changes in composition over time influenced community function, we compared properties of the complete community with properties of each constituent morphotype, both as clonal populations and as predicted sums-of-parts given additive relationships (Loreau & Hector 2001). The first phenotype we measured was fitness in the biofilm model itself, determined by competition against the PA14 ancestral strain for adhesion and persistence on the plastic bead. Individual isolates tended to be more fit than complete diverse mixtures (Figure A.3) and demonstrated increased productivity (CFU/mL) on a bead surface (Table S2), which is

indicative of negative effects of competition among isolates (rendering the whole less than the sum of parts) or perhaps facilitative effects of the mixed community on the ability of the ancestor to colonize and persist. Biofilm biomass of distinct phenotypes from each population, and the complete communities, were also quantified. In contrast with the fitness measurements, biomass of the B1 population after 540 generations, and B2 after 260 and 540 generations of adaptation, significantly exceeded the expected, scaled sum of the output of individual morphotypes, whereas mixed B3 communities were always worse than expectations from individual constituents alone (Figure 2, Table S3).

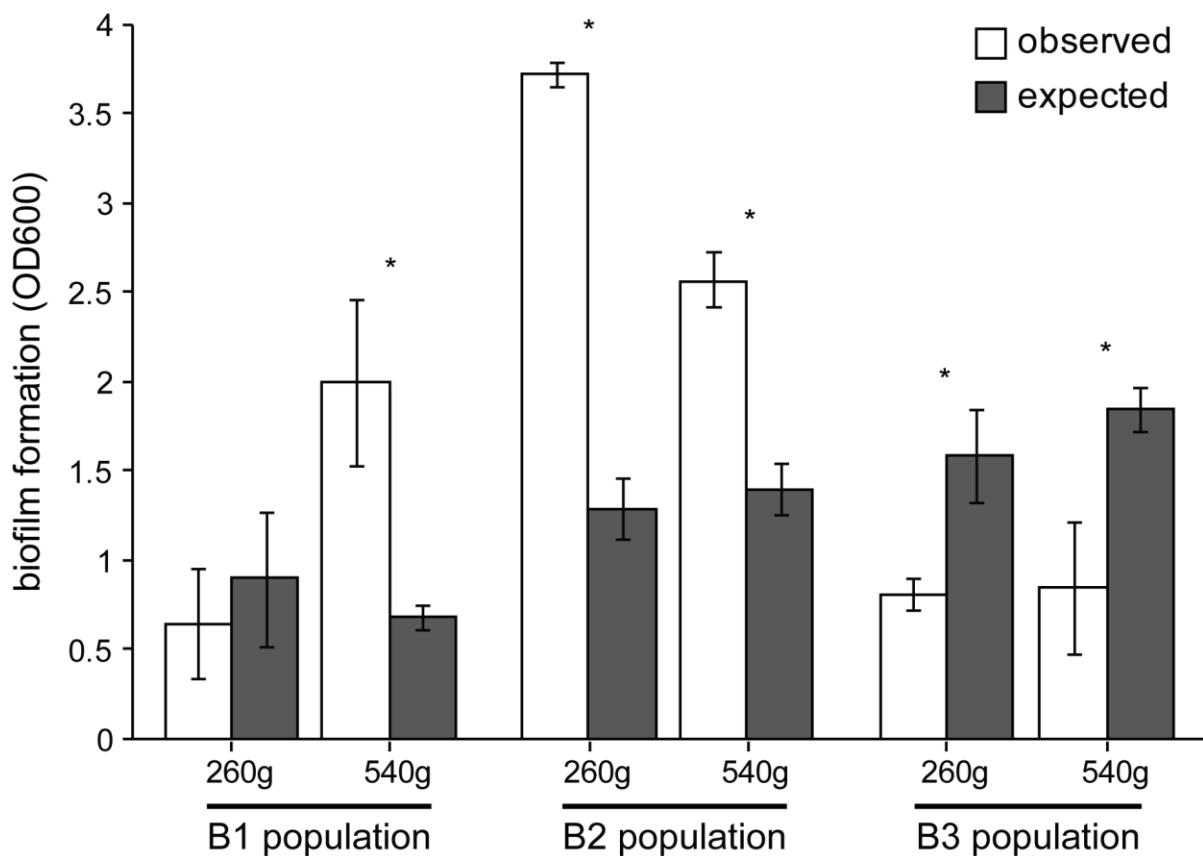


Figure 2. Replicate biofilm-evolved populations respond differently to diversity. Biofilm formation of the mixed communities (white) after 260 and 540 generations was compared to an expected value based on the scaled sum of individual morphotypes (grey). Asterisks denote a significant difference between observed and expected values ($P < 0.05$) based on Student's *t*-tests.

Rare variants disproportionately contribute to community fitness. We focused on the B1 population at the end point of the experiment (Figure 1) to begin to assess how diversity contributed to biofilm adaptation. Colony morphologies within this population were the most diverse, which raised questions in both population genetics and microbial ecology. First, how did so many distinct lineages rise to detectable frequency in a large population, and second, how can such diversity persist in a relatively simple environment? This study tests the hypothesis that these heritable variants defined different ecological strategies within the community that contributed individually and collectively to biofilm function.

Table 1. Characteristics of representative morphotypes from the B1 population isolated following 540 generations of biofilm adaptation.

Colony Morphology *	Wildtype	Apollo	Demeter	Hera	Minerva	Olympus	Poseidon	Vulcan
Name	Abbrev.	W/T	A	D	H	M	O	P
Growth‡	N/A	Planktonic, late attachment	Biofilm, early attachment	Biofilm, delayed attachment	Planktonic, delay/late attachment	Planktonic, late attachment	Biofilm, delayed attachment	Biofilm, late attachment
SCV?	No	No	No	Yes	Yes	Yes	Yes	No
Biofilm formation (95% C.I.)	0.214 (0.061)	0.106 (0.024)	0.798 (0.153)	0.561 (0.042)	0.871 (0.135)	0.162 (0.059)	0.252 (0.060)	0.471 (0.094)
Frequency (95% C.I.)	0	0.094 (0.048)	0.199 (0.163)	0.046 (0.027)	0.492 (0.143)	0.014 (0.040)	0.020 (0.014)	0.135 (0.044)

* Pictures were scaled to their relative sizes on a plate. See Figure 1 for bigger pictures.

‡ Growth patterns are based on biofilm formation patterns of individual types (Figure A.4, 5) and timing of attachment of individual types to a bead grown in M63 (Figure A.6). Types were categorized based on when the majority of attachment occurred: early <4 hours, delayed 4-12 hours, late >12 hours.

We explored the ecological mechanisms by which fitness and biomass of these evolved biofilm communities correlates with diversity though assays where each one of these seven representative isolates was subtracted from the mixed B1 community. Although individual isolates from B1 after 540 generations tended to be more fit than the diverse community as a whole (Figure A.3), we found the complete community was still more fit relative to the ancestor compared to constructed communities lacking one of the seven representative morphotypes (selection rate, r , = 3.124 ± 0.289 [95% confidence interval, C.I.]). Communities lacking the M, P, D or V morphotypes lost up to 40% of their fitness advantage over their ancestor (Figure 3, Table S4). Surprisingly, this decrease in fitness did not correlate with biofilm productivity or fitness of an individual type alone (Table 1). For example, removing minority members P or V, which contributed less than 15% to the total cellular yield of the community, reduced fitness by nearly as much as removing the most abundant (~49%) constituent, M.

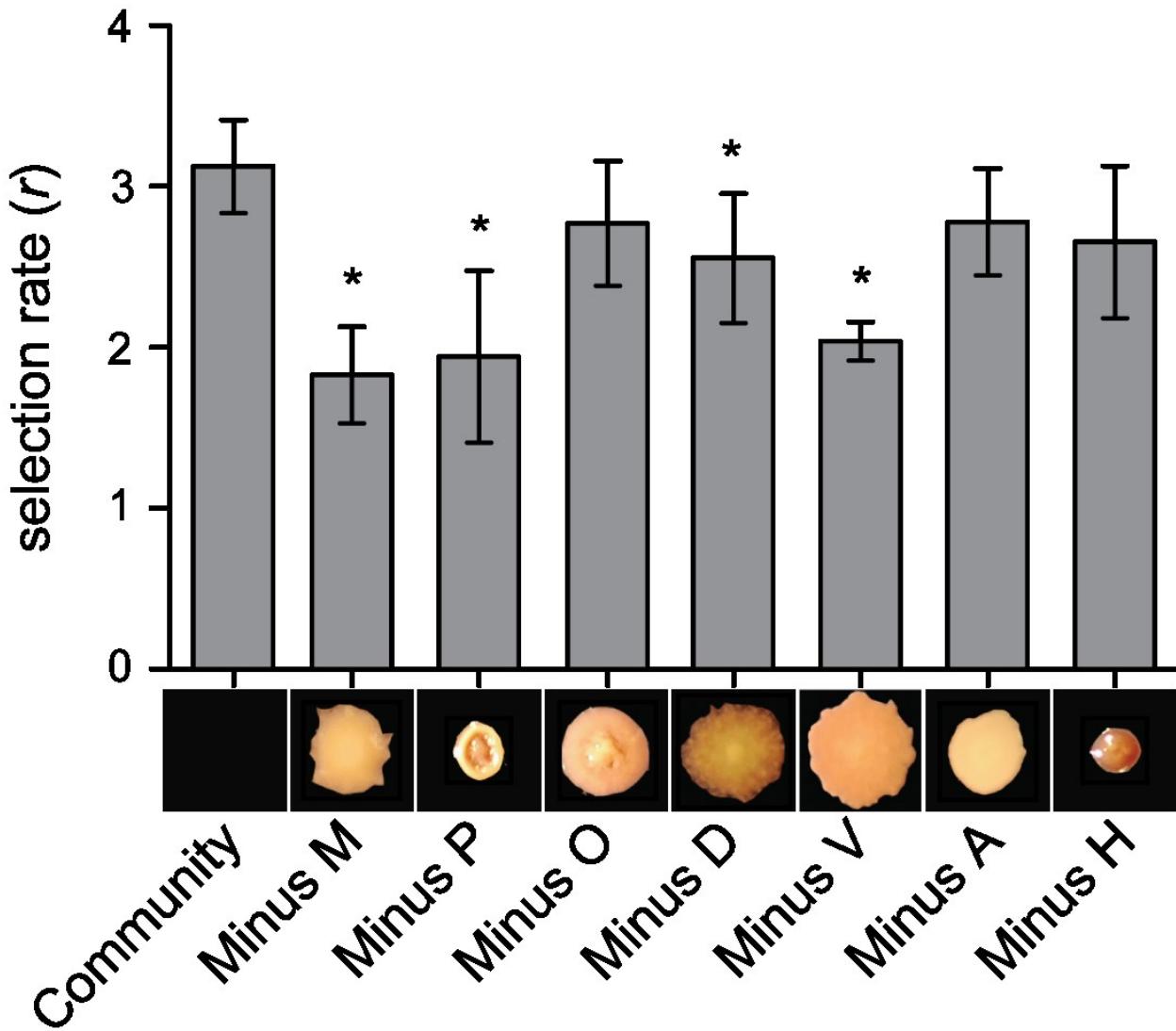


Figure 3. Community fitness requires diversity.

Incomplete communities lacking each of the seven members were constructed and competed versus the ancestor, as was the complete reconstructed community. Post-hoc analysis was performed using Dunnett's test with the complete community as the control; asterisks denote $P < 0.05$. Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan (Table 1).

Variation in c-di-GMP levels contributes to phenotypic diversity in B1. We hypothesized that the unique timing of biofilm production (Figures S4a) and attachment (Figure A.4b) may allow the D type to occupy a distinct ecological niche that strongly influences community fitness and composition. Unlike the other morphotypes, D

displayed enhanced biofilm productivity and attachment early during the growth cycle (Figure A.4), so D may facilitate attachment by other types by providing a binding surface (a form of niche construction) and/or possibly by external signaling. To test this hypothesis, we assayed biofilm formation of pair-wise combinations of morphotypes in which D was allowed to establish a biofilm prior to adding the other type (Figure 4a). When the M and H types and the ancestor were grown in combination with the D type, these combinations produced more robust biofilms than expected from either clone growing alone (Figure 4b).

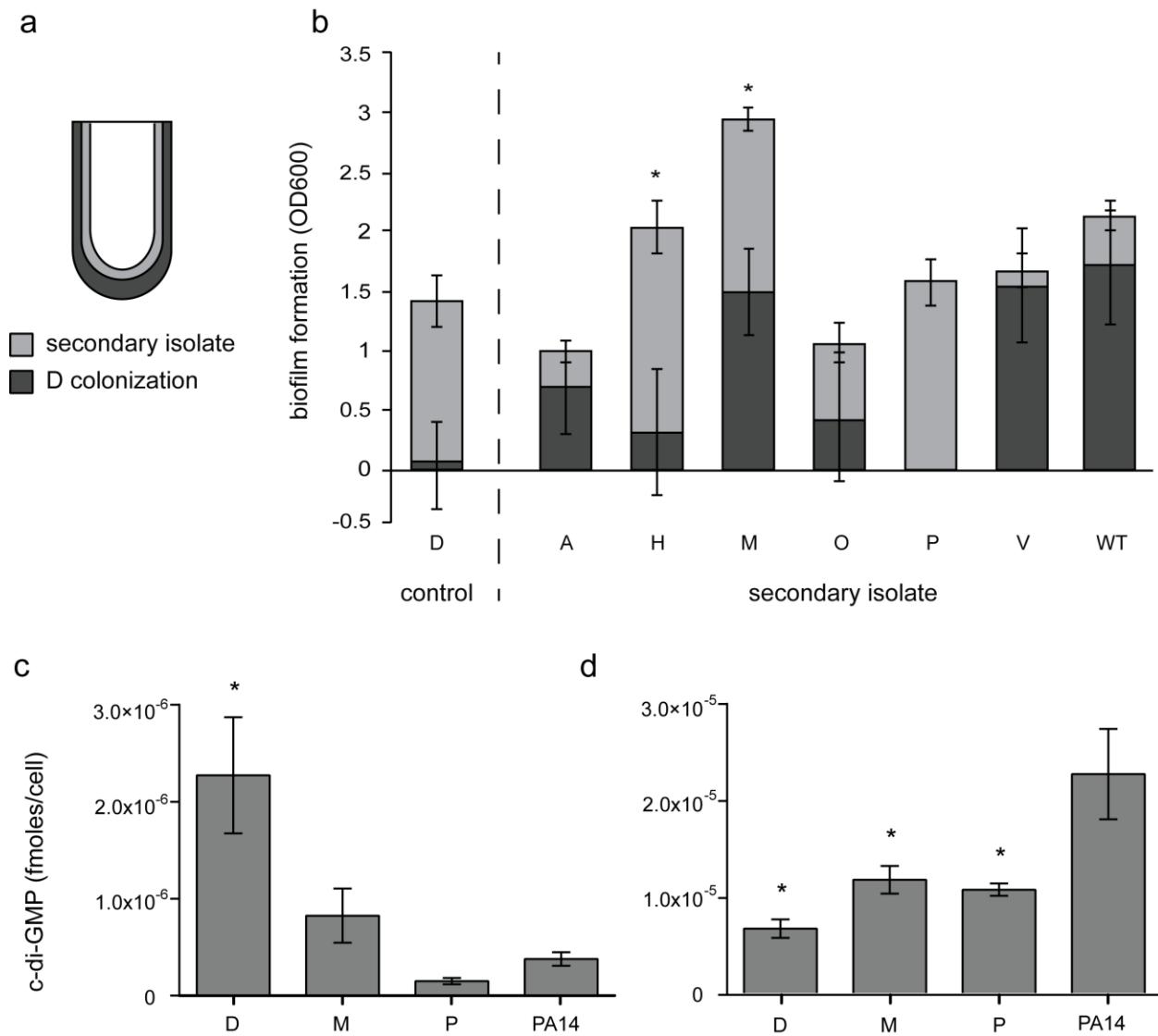


Figure 4. The D morphotype facilitates the attachment of other types partly defined by variation in c-di-GMP pools.

A facilitative interaction between D (Demeter, Table 1) and the rest of the community was suspected based on other biofilm and timing of attachment data (Figure A.4). a. This interaction was assessed by pair-wise biofilm production assays where D was allowed to colonize a surface for four hours before the addition of a secondary isolate. b. Biofilm formation of individual morphotypes alone (grey) and the improvement of total biofilm formation with D colonization (white). Asterisks signify combinations that are more productive than both the morphotype grown alone and D with more D added (D^2) to control for greater cell density. c, d. C-di-GMP was extracted from cells of three focal variants: M, D, and P. Cells were grown for 17 hours in M63 media supplemented with a bead. C-di-GMP was extracted from cells taken from (b) the planktonic phase or (c) from biofilms in triplicate. The amount of c-di-GMP was standardized by cellular yield (CFU/mL) determined by plating both the planktonic phase and biofilm-associated cells. Error bars are SEM and asterisks denote significant differences from other genotypes

($P < 0.05$). Post-hoc analysis was performed using two-way student t-tests. A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan, WT = PA14 ancestor (Table 1).

To probe the molecular mechanism underlying early colonization, c-di-GMP was extracted from planktonic and biofilm-grown D, M, P, and ancestral cells, following precedent that increased concentrations of the second messenger c-di-GMP can produce colony variants similar to those reported here (Meissner et al. 2007; Starkey et al. 2009). Planktonically growing D cells had significantly higher levels of c-di-GMP compared to M, P, and the ancestor (Figure 4c), whereas biofilm-associated D cells had significantly less c-di-GMP (Figure 4d). Although some subtle differences in c-di-GMP concentrations may exist, we found no significant differences between M, P, or the ancestor with either biofilm or planktonically grown cells in these assays. These results suggest that high levels of this regulatory molecule in the planktonic phase stimulate early surface attachment but then levels drop.

Altered physiology of a single ecotype disrupts community composition and function. To assess how selection for biofilm specialists with variation in a second messenger molecule could influence community composition and function, we experimentally manipulated levels of c-di-GMP in individuals and examined the consequences on the community as a whole. We reasoned that altering c-di-GMP levels in a single ecotype should not only disrupt its function, but also the balance and function of an entire community if ecological interactions are present. Communities were constructed with a single member expressing either the phosphodiesterase (PDE) *bifA*, effectively reducing the intracellular levels of c-di-GMP, or the diguanylate cyclase (DGC) *sadC*, effectively increasing c-di-GMP levels. When induced, these enzymes altered biofilm production, colony morphology, and pigment secretion (pyoverdine,

pyocyanin) (Figure A.5). Community fitness and composition of these altered communities were compared to control assemblies in which expression of these enzymes were not induced.

Amidst complex and asymmetric responses by the community in terms of fitness and composition following these manipulations, four overall patterns emerged. First, manipulating c-di-GMP in the most abundant M type exhibiting moderate biofilm formation and delayed attachment during the growth cycle (Figure A.5), affected other types present to the detriment of the community (Figure 5, M+*pbifA*, 1.498 ± 0.562 [95% C.I.], and M+*psadC*, 1.316 ± 0.142 [95% C.I.] compared to unmanipulated, 3.389 ± 0.502 [95% C.I.]). Reduced c-di-GMP in M caused the D morphotype to become nearly undetectable. Increased c-di-GMP in the M type reduced not only its relative abundance but also that of O and H types; however, the A type increased drastically (Fig. 5b).

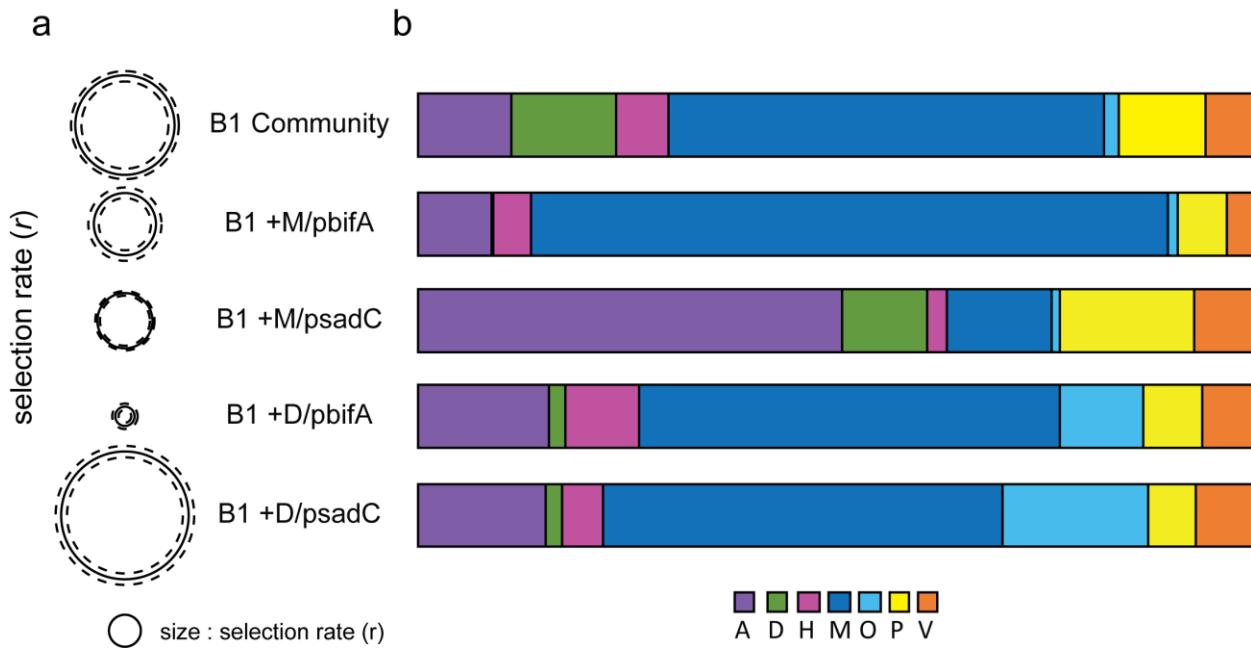


Figure 5. Altering c-di-GMP levels in individual B1 morphotypes affects fitness as a whole.

Communities were constructed with certain members expressing the phosphodiesterase *bifA*, effectively reducing intracellular levels of c-di-GMP, or the diguanylate cyclase *sadC*, a phosphodiesterase effectively increasing c-di-GMP levels. The effect of altering c-di-GMP in a single member on the community as a whole was assessed through fitness (a) and composition (b) relative to a natural B1 community. a. The size of the circle correlates with fitness of each community versus the PA14 ancestor. Dotted line circles represent 95% confidence intervals. Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan, WT = PA14 ancestor (Table 1).

Second, increasing or decreasing c-di-GMP in the D type tended to increase the relative abundance of O and H variants (Figure 5). Reducing c-di-GMP in D also reduced community fitness, but surprisingly increasing c-di-GMP in D increased community fitness ($t_{12} = 3.557$, $P = 0.004$) while producing a large increase in O at the expense of D despite D being important for community fitness (Figure 3). Third, altered physiologies of other minority members besides the D type also disproportionately influence the community. Reducing c-di-GMP levels in a minority member, P (representing ~5% of the population) to disfavor its biofilm lifestyle (Figure A.5) lowered

community fitness more than removing P from the community (Figure 3, $r = 1.102 \pm 0.255$ [95% C.I.] vs. 2.328 ± 0.685 [95% C.I.]; $t_6 = 5.34$, $P = 0.002$). At the same time, reduced c-di-GMP in P also enriched for the A and H types at the expense of the M and P types (Figure A.6). Lastly, manipulating c-di-GMP levels in the B type of the B3 community after 260 generations, devoid of a benefit of mixture when forming a biofilm at this time point (Figure 2), did not negatively impact the community as a whole (Figure A.7).

Although the rise of some genotypes associating with the fall of others may be explained by their occupation of complementary niches, we emphasize that these shifts were almost always costly to the community as a whole. For example, although the A type could fill the vacancy left by the M type, the lower capacity of A to form robust biofilms (Figures S4) may explain the reduced community output. Because nearly every disturbance altered community function or composition, the evolution of the community was thus influenced by interactions between types and their unique characteristics or functions. In agreement with this explanation, we found that similar disruptions in different replicate populations did not always have an effect on the community as a whole (Figure A.7). For example, disruption of the B isolate from the B3 population after 260 generations had little effect on other types.

DISCUSSION

Although increased biodiversity and rapid phenotypic diversification appears to be inherent to biofilms, the long-term consequences of this diversity on adaptation are not well understood. At its conception, the present study sought to identify this gap in understanding to experimentally evolve *Pseudomonas aeruginosa* PA14 for over 500

generations under conditions favoring a regular cycle of attachment and dispersal from a polystyrene bead. By comparison, most studies examining biofilm adaptation rely on models like drip-flow reactors that emphasize attachment and persistence to a surface enabling only short-term adaptation (for example, McElroy et al. 2014). By integrating dispersal into our model, our design allows not only to observe the origin of diversity patterns of diversification, but also how variation associated with biofilms resolves over much longer time frames. *Pseudomonas* lung infections of persons with Cystic fibrosis (CF) are characterized by immense genetic and phenotypic diversity in a variety of medically important traits. However, does this variation in medically important traits evolve repeatedly or reflect persistent variation associated with initial colonization and diversification? Distinguishing between these two possibilities requires an experimental evolution model that allows for sufficient time for fixation and extinction events to occur.

Adaptation to the biofilm lifestyle itself may promote variation in medically important traits in the absence of a host (Rau et al. 2010; Oliver et al. 2000). For example, hypermucoid phenotypes are repeatedly observed in biofilms of *P. fluorescens* Pf-01 presumably in response to adaptation to a crowded biofilm (Kim et al. 2014). Likewise, prior work in our lab with experimentally evolved biofilm populations of *Burkholderia cenocepacia* HI2424 (Traverse et al. 2013) has shown that variation in LPS structure also repeatedly evolves in the absence of the host immune response (Moskowitz & Ernst 2010; Ernst et al. 2007). In agreement with these findings, we also observe colony phenotypes typically associated with chronic infections including SCVs, altered motility, and increased EPS production. We also find that the production of virulent factors pyoverdine and pyocyanin are also reduced. Curiously, however,

variation in these traits never appear to fix, and we find little convergence across the replicate biofilm-evolved populations (Figure 1).

Adapting populations are expected to be relatively homogenous with transient periods of diversity that is eliminated by strong purifying selection, but this expectation is most certainly flawed. Genomic re-sequencing of experimental evolution studies performed over the past decade has revealed how complex adaptation to relatively simple, homogenous, environments truly are (Dettman et al. 2012). For example, a recent study examining genetic diversity present during the well-studied Lenski Long-Term Evolution Experiment (LTEE) (Wiser et al. 2013; Barrick et al. 2009) highlights the presence of a staggering amount of genetic diversity persisting as co-occurring competing lineages (Leiby & Marx 2014). Although several studies suggest adaptation to the biofilm lifestyle may proceed through few adaptive paths (McDonald et al. 2009; McElroy et al. 2014; van Ditmarsch et al. 2013), biofilms are certainly more complex with diversity being maintained through the plethora of distinct ecological niches (Poltak and Cooper, 2010). Although the lack of parallelism of colony phenotypes across replicate biofilm populations evolved under identical conditions (Figure 1, S1) may appear stochastic, we found that disrupting the biofilm community produced repeatable and directed patterns of change (Figure 5b).

Prior studies from our lab with biofilm-evolved populations of *B. cenocepacia* saw the evolution of unique colony types that underwent ecological succession as they modified the abiotic environment, and constructed new niches (Poltak & Cooper 2011). Specialist ecotypes were able to invade vacant niches produced by metabolic cross-feeding, but also enhanced biofilm architecture, producing a synergistic community

(Traverse et al. 2013). In contrast, we found strong antagonism between members of these *P. aeruginosa* communities in terms of fitness (Figure A.3). This discrepancy could result from insufficient time for co-evolution among *P. aeruginosa* variants or their greater diversity, which may require further time to evolve neutrality or synergy. In agreement with this explanation, some populations do exhibit a benefit of mixture for biofilm formation (Figure 2), which correlates with the occurrence of the hypermutation phenotype in the B1 and B2 replicate populations (Figure A.2). Combined, these data suggest that hypermutation accelerated the evolution of cooperation in this context. In the B1 population, this cooperation takes the form of enhanced biofilm formation of the community due to the evolution of biofilm specialists like the D type that facilitate the attachment of others (Figure 4b), in part, due to modified c-di-GMP regulation (Figure 4c,d).

Previous work has identified variation in cyclic-di-guanylate cyclase variation as an important mechanism of adaptation in biofilms (Meissner et al. 2007; Bobrov et al. 2014). Here, we find additional evidence that functionally important biofilm ecotypes are partly defined by variation in intracellular levels of this molecule (Figure 4c,d). Although we found few significant differences between the co-evolving ecotypes, more subtle variation in c-di-GMP likely exists due to local tuning of c-di-GMP regulation and physical co-localization of enzymes with GGDEF and EAL domains within a cell (Lindenberg et al. 2013; Hengge 2009).

This study also highlights how both abiotic and biotic environments define selective pressures that bacteria face in structured environments, and potentially alter adaptive outcomes. Since much research has focused on how a single species adapt to

simple conditions (for example, LTEE), biotic influences on adaptation are less understood (Lawrence et al. 2012). Examining the co-evolution of two species in a structured environment (Hillesland & Stahl 2010) revealed genetic adaptation in one member to exploitative the dominant species demonstrating an evolutionary response in the presence of the other species. Similarly, the evolution of the facilitative D type in this system (Figure 4) likely influenced the adaptation of the other types in the community; otherwise, its removal would have only minimally affected other variants (Figure 5). Disproportionate effects of removing single variants were not limited to the D type (Figure 5 and Figure A.6), which demonstrates the substantial influence of the biotic environment in these biofilms. Furthermore, we emphasize that shifts in response to these disruptions were nearly always costly to the community as a whole (Figure 5a). For example, although the A type could fill the vacancy left by the M type, the lower capacity of A to form robust biofilms (Figure A.4) may explain the reduced community output. We suggest these biotic constraints on community assembly and the contingent adaptive history of each population more adequately explain the lack of parallelism across replicate biofilm-evolved populations adapted under identical conditions (Figure 1).

In summary, the evolution of rare, numerically underrepresented biofilm specialists can create strong interdependency through biotic interactions that can ultimately influence adaptive outcomes in structured environments. It is worth noting here that the evolution of synergistic effects of diversity in these three populations correlates with the timing of mutator evolution (Figure 2). The mutator phenotype fixed first in population B2 and benefits of mixture on biomass soon followed; later on, the rise

of the mutator phenotype in population B1 also led to benefits of mixture. In contrast, the *mutL* allele only became dominant late during the evolution of population B3, which continued to show negative effects of mixture. Why mutators evolve remains largely unsolved, but is thought to result due to limiting beneficial variation, and we plan to test this explanation in the context of our experimental system. In agreement with this hypothesis, the fixation of hypermutation was never observed amongst replicate biofilm-evolved lines of *B. cenocepacia* adapted under similar conditions (Traverse et al. 2013). Regardless, the mechanistic cause of the positive correlation between mutation rate and the evolution of synergy remains a topic for a future report. In particular, we predict that increased genetic variation may facilitate the evolution of rare genotypes that produce synergistic biofilm specialists and plan to dissect the eco-evolutionary dynamics of these events through next generation sequencing technology.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *Pseudomonas aeruginosa* PA14 was used as the founder strain for this study. *Escherichia coli* SM10 λpir was used to conjugate plasmids into PA14. *E. coli* SM10 λpir cells were grown overnight in lysogeny broth (LB) medium (1.0% w/v tryptone, 0.5% w/v yeast extract, 1.0% w/v NaCl) at 37°C with shaking. *P. aeruginosa* PA14 and derived isolates were grown overnight from freezer stock in a mixture of 1 ml tryptic soy broth (TSB) and 4 ml M63 media (base consisting of 15 mM (NH₄)₂SO₄, 22 mM KH₂PO₄, 40 mM K₂HPO₄ supplemented with 40 mM Galactose, 1 mM MgSO₄, 25 μM FeCl₂, and 0.4% w/v Arginine) with a polystyrene bead. All beads were aseptically added to sterile 15 mL centrifuge tubes and sonicated in 1.5 mL PBS to remove all biofilm cells before passage or plating. Sonicating was

determined to have no effect on the viability of biofilm-grown cells, and allowed for more accurate assessments of bead growth in terms of yield and reproducibility (Figure A.8).

When necessary, the following antibiotic concentrations were used: 10 µg/ml of gentamicin and 20 µg/ml nalidixic acid for *E. coli* and 80 µg/ml gentamicin for *P. aeruginosa*.

Experimental evolution. Three replicate ancestral populations of *P. aeruginosa* PA14 were grown on 7mm polystyrene beads suspended in 5 mL of M63 media in a test tube for ninety days or ~550 generations. Populations were selected for reversible surface attachment by daily transfer of the bead to a new test tube where cells were required to adhere to a new bead in order to persist, using a method described previously (Poltak & Cooper 2011). Three replicate planktonic lines were also maintained as a control by transferring 50 µL of media daily under the same growth conditions lacking a bead for attachment.

Characterization of morphotypes. Biofilm-evolved and planktonic-evolved replicate populations were examined for phenotypic diversity by repeat plating on 1% tryptone supplemented with 20 µg/mL Coomassie blue and 40 µg/mL Congo red allowing variation in colony morphology to be more apparent due to dye uptake after growth at 37°C for 24 hours and an additional 48 hours at room temperature. The plates with less than 200 CFUs were used for isolate identification and phenotype scoring since plate density influenced colony phenotype development. Representative morphotypes from population and each time point were isolated and grown in monoculture, and those isolates with heritable, distinguishable phenotypes determined through passaging in M63, were used for further characterization (Figure 1, S1 and S2). Isolates were also

screened for variation in specific phenotypes such as biofilm formation, colony size (SCVs), and production of secondary metabolites to confirm differences (Table 1, data not shown). Isolates were classified as SCVs by direct comparison to the PA14 ancestral colony phenotypes plated together. Since the B1 population was selected as the focal population for further study, we tested if the seven morphotypes used in this study were representative of the diversity present in the natural B1 replicate biofilm-evolved population. We found no significant difference in the fitness of an artificially constructed B1 community containing these seven members added together at their observed frequencies (Figure 1, Table 1) and the natural population (fitness of the natural community, selection rate (r) = 0.421, compared to the constructed community, r = 0.327; t_6 = 2.213, P = 0.075).

Biofilm production assay. Biofilm production was quantified as described previously (O'Toole & Kolter 1998) after four, eight, and twenty-four hours. To assess if the attachment of one type facilitated the attachment of another, one type was allowed to establish a biofilm in wells of a 96-well plate for four hours before the addition of a second type, standardized using optical density. Biofilms were stained with 1.0% crystal violet (CV) and solubilized with 95% ethanol after a full 24 hours of growth.

To determine if biofilms formed by the complete B1 community after 24 hours were more robust than the sum of the constituent parts, a predicted 24 hour community biofilm value was calculated from the biofilm capacity of the individual types alone, scaled for the frequency at which each type was observed in the community (Table 1) and then summed. A similar approach was performed to determine if the community was more productive in cellular yield, CFU/mL.

Fitness assays. *P. aeruginosa* wild-type strain PA14 derived competitors and PA14^{+lacZ} were also grown overnight separately and then equal amounts of competitors harvested from a bead were added to 5 mL of M63 media with a new polystyrene bead at a 1:100 dilution. CFU/mL was determined for each competitor at the time of inoculation and after 24 hours off of the bead by plating onto 1% tryptone supplemented X-gal. Fitness was calculated as the selection rate, r, defined as the difference in $\ln(\Delta\text{CFU}/\text{mL})$ over a 24 hour time period, which has units of inverse time. Since fitness is defined as the ability to attach to a bead, less fit competitors can produce fewer CFUs on a bead after 24 hours than added initially. Selection rate, rather than relative fitness, allowed for fitness to be accurately determined given the low fitness of the PA14 ancestor in some contexts (Travisano & Lenski 1996). For a more detailed explanation, see <http://myxo.css.msu.edu/ecoli/srvsrf.html>.

Change in frequency assays. Communities were reconstructed by adding them together in their natural relative frequency (defined as the observed frequencies in the unmanipulated 90d community, Table 1) with substituted complements harboring a vector of interest *graciously provided by the George O'Toole laboratory at Dartmouth College. Plasmids were introduced in PA14 by electroporation as described previously (Choi et al. 2006) or through biparental mating with E. coli SM10*. For expression of *bifA*, pMQ80-His-*bifA*⁺ (Merritt et al. 2007; Kuchma et al. 2007), or *sadC*, *psadC* (Merritt et al. 2007; Kuchma et al. 2007), arabinose was added to the media at a final concentration of 0.5%. Change in relative abundance was calculated as the difference between the observed frequency of a morphotype in a mixed community on a bead after 24 hours

and the expected frequency based on the inoculated frequencies (modeled after observed morphotype frequencies, Table 1).

Extraction of cyclic-di-GMP. Cyclic-di-GMP was extracted using a nucleotide extraction method described previously using liquid chromatography-mass spectrometry (LC-MS) (Newell et al. 2011). Briefly, cells isolates were suspended in 250 µL of extraction buffer containing 40:40:20 methanol, acetonitrile, and water with 0.1N formic acid. Extraction mixtures were incubated at 20°C for one hour. After the incubation, 200 µL of the supernatant was transferred to fresh, cold tubes containing 8 µL of 15% ammonium bicarbonate. Samples were dried using a gas manifold and re-suspended in 100 µL of the mobile phase buffer the same day as quantification due to the instability of c-di-GMP in mass spectrometry mobile phase buffer (10 mM tributylamine and 15 mM acetic acid in 97:3 water/methanol). Extractions were performed after 17 hours from both planktonic phase and biofilm-associated cells grown under evolution conditions (see above). Cellular yield before each extraction was also determined to allow for standardization of the results following quantification.

CHAPTER III

FROM GENOMICS TO ECOLOGY AND BACK:
THE EVOLUTION AND MAINTENANCE OF VAST DIVERSITY IN *PSEUDOMONAS AERUGINOSA* BIOFILMS

INTRODUCTION

We do not understand how much genetic variation for fitness may be maintained in any environment, especially not in the structured environments of biofilms, which constitute the dominant mode of microbial life. Large population sizes provide adapting bacterial populations ample access to beneficial variation that is maintained by the complex spatial structure associated with biofilm architecture (Drake 1991). Ecological niches created by the spatial and temporal structure of biofilms is predicted to facilitate variation that would otherwise go extinct in more homogenous environments (Besemer et al., 2012; Parsek & Singh, 2003; Periasamy & Kolenbrander, 2009; Yildiz & Visick, 2009). For example, oxygen gradients commonly selects for variation in redox chemistry and extrapolymeric saccharide (EPS) production due to the endogenous oxidative stress and resource-starved conditions associated with anaerobic conditions (Dietrich et al., 2013) . However, defining these available niches with any certainty remains difficult or even impossible despite knowledge of the environment or organisms involved *a priori*.

Genomic sequencing of experimentally evolved microbial populations has revealed that presumably simple homogenous environments are even more complex than previously thought. Adapting bacterial populations harbor immense genetic diversity despite appearing seemingly homogeneous (Dettman et al., 2012) . Only a few mutational changes may be required to cause a single bacterial species to act like two species, persisting as co-occurring genetic lineages (Herron & Doebeli, 2013) . However, the evolution of distinct ecological units requires some sort of trade-off for long-term maintenance (Koeppel et al., 2013, Shapiro & Polz, 2014). Whether or not adaptation to a structured environment readily creates these trade-offs remains unclear.

Here, we characterize the exceptional population genetic structure of an experimentally evolved hypermutator biofilm community described previously. Utilizing a combination of whole-population sequencing and full genome re-sequencing of representative clones (Traverse et al., 2013), our findings illustrate our inability to fully appreciate the vast diversity harbored in biofilms. Despite this complexity, we identified distinct genotypic clusters associated with a preference for bead attachment. We previously found that these populations evolved unique patterns of community assembly, owing in part to interactions with early-attaching biofilm specialists (Chapter 2). Combined with genotype information, we identify mutations involved in the regulation of anaerobic respiration to associate with a preference for bead attachment and creation of distinct ecological units.

RESULTS

We used a combination of genome sequencing of representative clones and complete population samples (metagenomes) to identify the mutations that underlie global adaptation and lineage diversification within three replicate *P. aeruginosa* biofilm populations (B1, B2, B3) evolved for 540 generations. Combined, these approaches generated 20.2 gigabase pairs (Gbps) of whole-population and 54.4 Gbps of isolate sequencing data. From these data, we identified 1,183 mutations among the 40 isolates selected for sequencing (Table B.1) and 8,931 mutations from surveys of the total genetic diversity present within each replicate population at six different time points: 107, 150, 264, 400, 450 and 540 generations (Table B.2). We previously found that defects in mismatch repair (MMR) fixed in each replicate biofilm population (Chapter 2).

Consistent with a mutational bias towards transitions in MMR- genotypes, most mutations identified with both approaches were transitions (96.5%, Table B.1, 2).

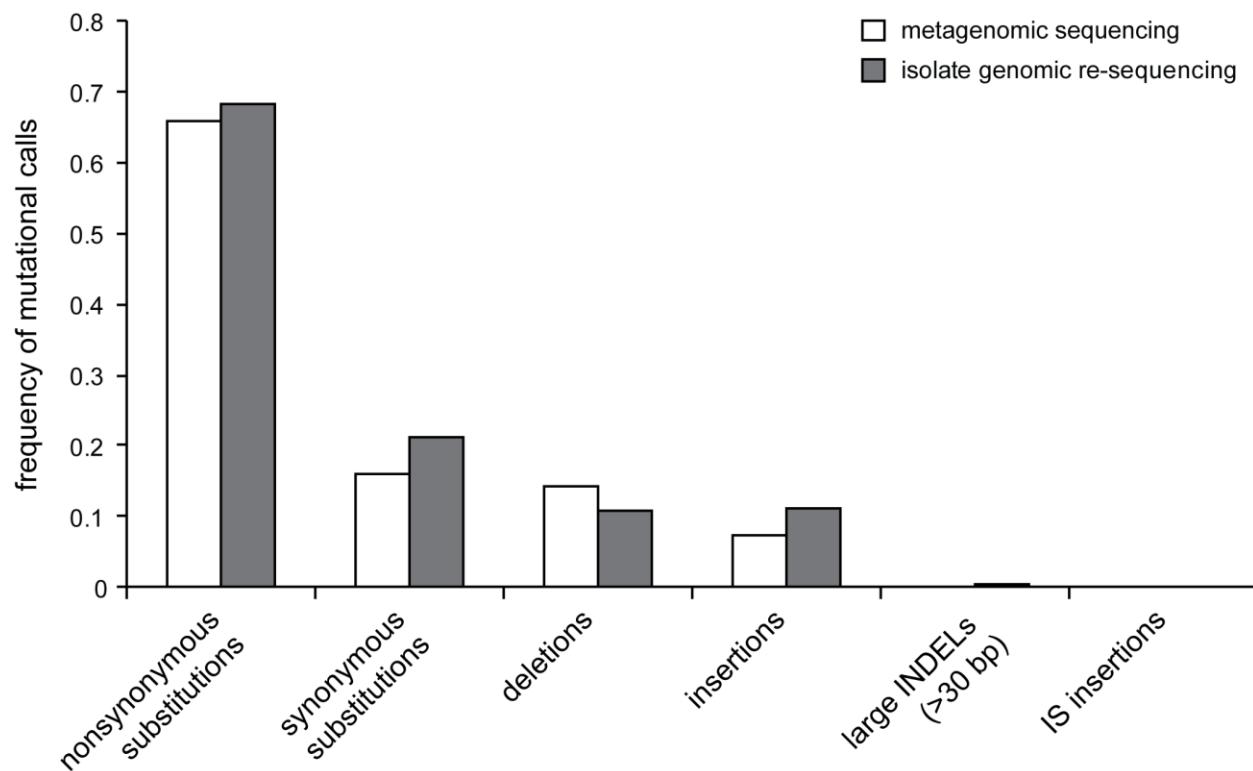


Figure 1. Mutation frequencies in metagenomes and sequenced isolates

In both data sets, mutations were predominantly non-synonymous single-nucleotide polymorphisms. For metagenomes ($n = 397$, Table B.2): non-synonymous, 718; synonymous: 223; deletions: 111; insertions, 117; large INDELS, 2; IS insertions, 0. For isolate sequencing ($n = 1,183$, Table B.1): non-synonymous, 252; synonymous: 61; deletions: 54; insertions, 27; large INDELS, 0; IS insertions, 0.

Based on the understanding that colonies originate from a single bacterial cell, high-confidence mutational calls from isolates provided a complete genomic snapshot of cells present in these biofilm populations at a given time. Previous work suggested that colony phenotype reflect distinct ecological strategies in these biofilm communities (Chapter 2). As such, the clones selected for sequencing were isolated based on morphology and expected to represent the array of ecological strategies that evolved in

each population. Twenty-six clones were isolated and sequenced from the B1 population and an additional nine and five isolates were sequenced from the B2 and B3 replicate populations, respectively. Meanwhile, mutational calls from the metagenomes represent the total genetic variation present in a population at a given time. To better eliminate false positives caused by sequencing or alignment errors, we only retained those mutational calls supported by repeated sampling of evolved populations over time to increase our confidence of some low-frequency mutational calls and generate a list of high-confidence adaptive trajectories of mutations. In addition, matching calls from sequenced isolates added greater certainty. Ultimately, we defined 397 high-confidence mutational trajectories across all three biofilm-evolved populations (Table B.3). Combined, these data allowed us to identify the adaptive history of the distinct ecological strategies represented by the isolates sequenced.

Isolates harbor many mutations unseen with whole-population sequencing. High-confidence mutational calls identified in isolates should also be present within the metagenomic samples from which they were derived, given sufficient sequencing depth for resolution. Much to our surprise, only 54.75% of these known isolate mutations were also identified in metagenomes despite 338x coverage on average across samples; all forty isolates harbored many shared and unique mutations unseen with whole-population sequencing (Figure 2a). As such, these biofilm populations may be much more diverse than we originally anticipated and these patterns may simply reflect inadequate sampling depth of the genetic variation present in each population.

Alternatively, one simple methodological explanation for the identification of the large number mutations unseen in the metagenome is that these mutations actually

arose in isolates after isolation from the evolved population but prior to genomic DNA extraction. Individual isolates experienced four additional growth cycles amounting to ~40 generations: a day under conditions mimicking the original EME study, a day of morphological development on solid media and two days in rich media. With this additional growth, we expect approximately 7.2 additional unique mutations (0.18 mutations per genome per generation (Heilbron et al., 2014) over 40 generations), but instead saw an average of 16.7 unique mutations. In addition, these unique mutations also exhibit a ratio of non-synonymous to synonymous mutational changes (dN/dS) consistent with a signal of positive selection after correcting for the observed bias towards transitions (Figure 1, unique isolate calls, 2.96, compared to the neutral expectation of 1.75). Taken together, many of the mutations specific to isolates were likely favored by selection yet undetected by whole-population sequencing.

We examined the possibility of a sampling bias by focusing on the twenty clones isolated after 540 generations from the B1 population. This collection of isolates was comprised of nine isolates selected based on morphology including several previously characterized rare variants (lettered names correlating with phenotype: A, D, H, M, O, P and V, see Chapter 2), and an additional 11 isolates selected at random (isolates numbered 101 through 109, 111 and 112). The phylogenetic relationship between these isolates revealed that known biofilm specialists such as the D, P and V types did, in fact, harbor more unique mutations on average, manifesting as longer branch tip lengths (Figure 2b). Although common phenotypes were more likely to be closely related, interesting exceptions were observed. For example, the M isolates, originally selected as being representative of common variants, and the 107 isolate were found to be more

genetically similar to rare variants despite exhibiting commonly observed colony phenotypes (Figure 2b).

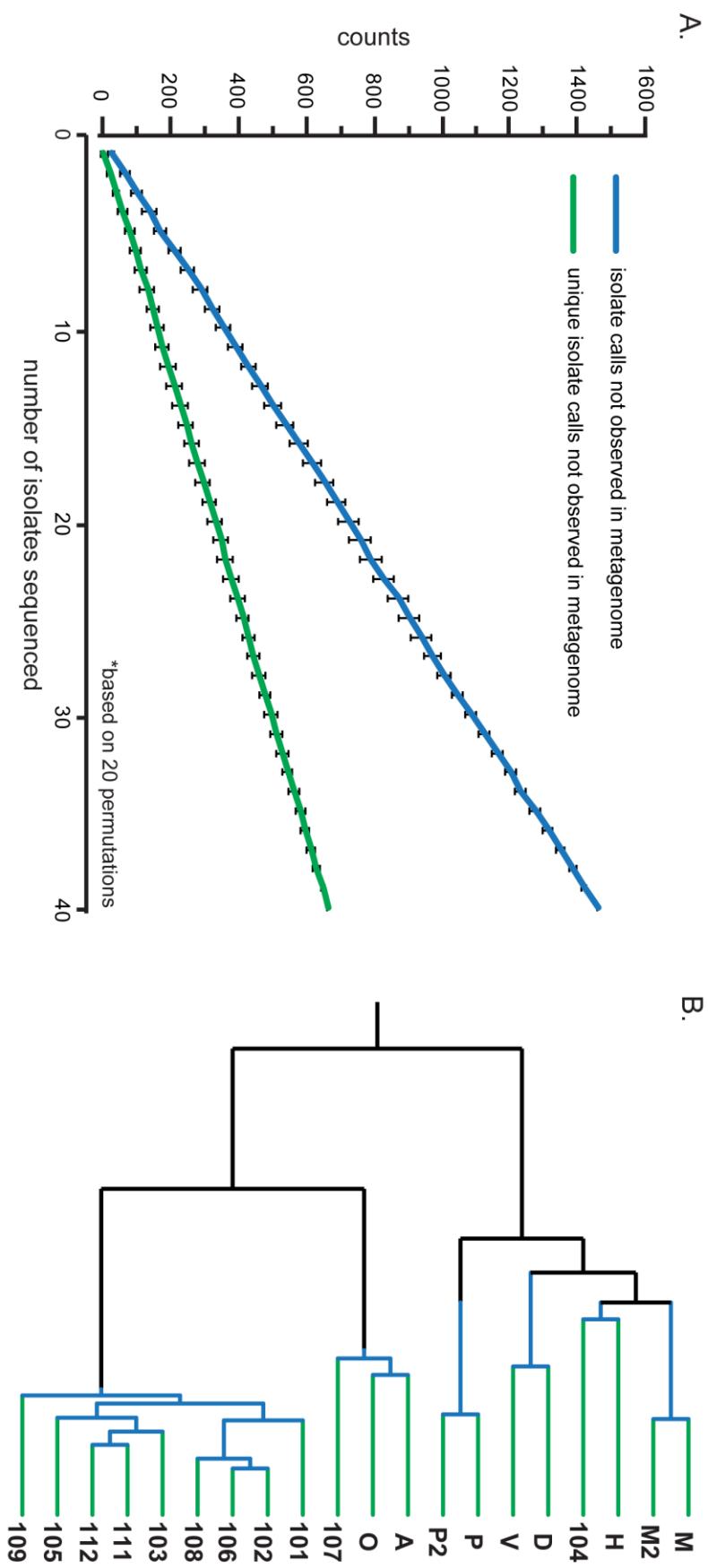


Figure 2. Biofilm structure preserves immense genetic variation.

Genome sequences of forty individual isolates from three evolved biofilm populations revealed many mutations unobserved in metagenomes. A. Unobserved isolate mutations were classified into two categories: mutational calls not observed in the metagenome (blue), but observed in multiple clones, and mutational calls observed only once in a single isolate (green). Error bars represent 95% confidence intervals calculated from randomizations of the data set with 20 permutations. B. Phylogenetic relationship among isolates from population B1 after 540 generations of evolution. Branches are colored as in panel A: unobserved shared clonal mutations are in blue and unique mutations are in green. The phylogenetic tree was created using Ward's method using binary data.

Previous work demonstrated that these biofilm communities could respond quickly to perturbation, displaying patterns of rapid ecological turnover related to niche-filling over the course of a single day (Chapter 2). Since our isolate sequencing approach focused on cells attached to a plastic bead, we hypothesized that selection for bead attachment over the course of a single day of growth (approximately 7 generations) was strong enough to significantly enrich for these rare genotypes that were undetected in the community metagenome. To more accurately identify patterns of enrichment, we grouped individual mutations with similar trajectories through time utilizing a hierarchical clustering analysis approach (See Methods). This analysis combined the 397 high-confidence mutations identified during our original whole-population sequencing approach into 12 collections of mutations with shared adaptive history referred to as mutational cohorts (Table B.4).

To test our hypothesis that selection for bead attachment over the course of a single day enriches for specific genotypes, we subjected mixed samples of the replicate B1 population after 107, 264 and 540 generations to a day of bead selection before performing additional whole-population sequencing. If insufficient read depth is solely responsible, we expected that additional sampling would uniformly improve our ability to identify rare genotypes associated with these 12 cohorts. Meanwhile, a bias towards mutations associated with specific cohorts would suggest an effect of bead selection. This approach provided an additional 2.7 Gbps of whole-population sequencing data and 128.8X additional read depth to test for an effect of bead selection on our ability to observe rare genotypes (Figure A.1, Table B.4, Figure B.1).

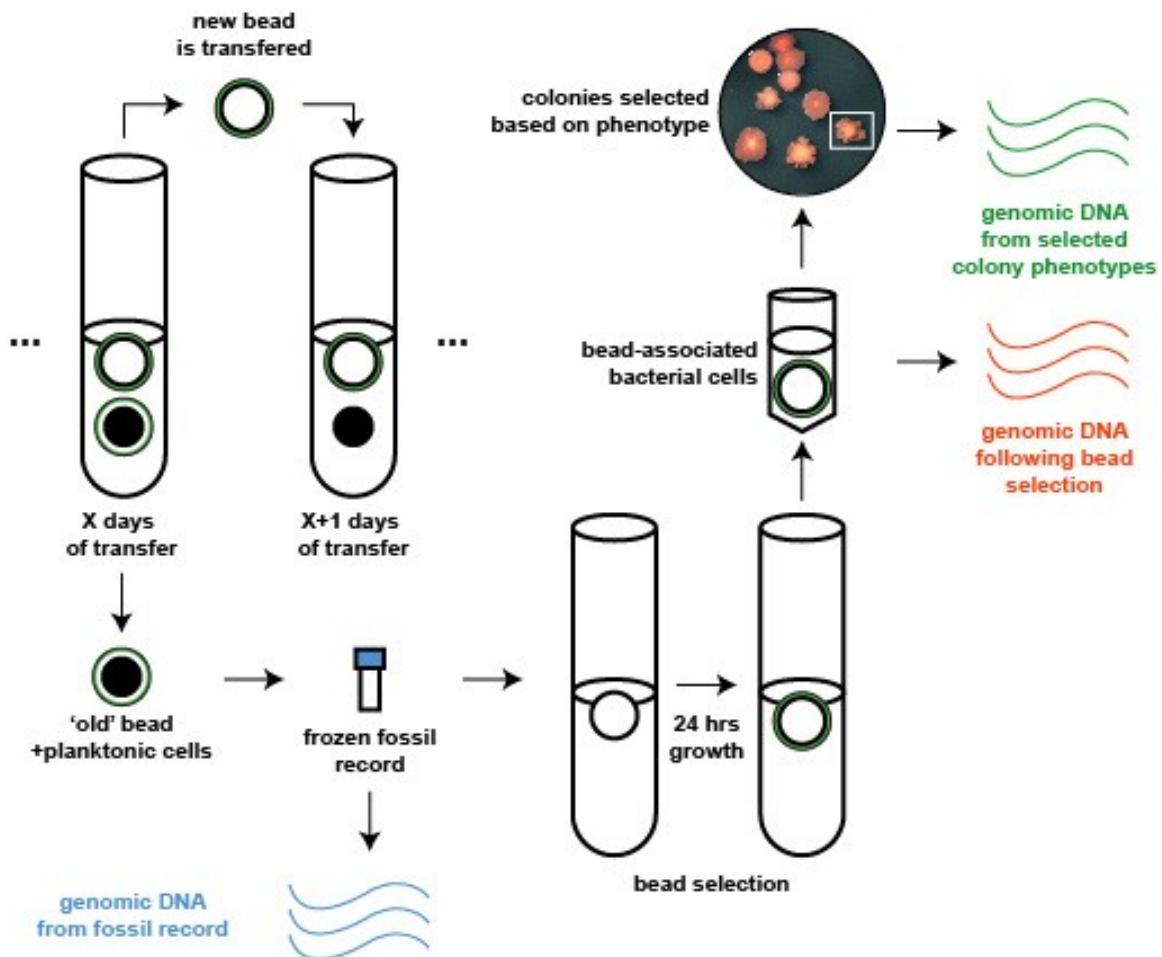


Figure S1. Methods defining key differences across sequencing approaches.

Genomic DNA was extracted in three different ways. First, frozen aliquots taken throughout the duration of the original EME study (Chapter 2) were constituted overnight in rich media before DNA was extracted. Second, samples were initially propagated the same but allowed an additional day of growth under conditions mimicking the original selection conditions and genomic DNA was harvested from bead-associated cells (blue). With our isolate sequencing approach, bead-associated cells were plated on solid media and allowed at least 48 hours to develop morphological differences. Subsequently, colony morphologies of interest were selected and regrown in rich media before DNA was extracted.

Bead selection enriches for specific mutational clusters with shared adaptive

history. Our additional sequencing of the B1 population generated a second opinion of the allele frequencies for previously identified mutations and the rare mutations. Consistent with inadequate sampling, we only observed slight changes in allele frequency for 129 mutations previously identified in both the B1 metagenomes after 107

and 264 generations (Figure B.2 & B.3, Table B.4). Likewise, we identified new mutations at both time points: 30 mutations in the 107 generations metagenome and 66 in the 264 metagenome. Additional sampling of the B1 population after 540 generations also saw limited change in allele frequency for mutations at or near fixation (green cluster, Figure 3a). In contrast, however, 23 mutations belonging to a single cohort did exhibit enrichment consistent with a bias due to bead selection (Figure 3a). This mutational cluster was identified at 39.5% allele frequency following bead selection, increasing from our estimate of 19.8% with our initial sequencing. This increase in frequency by these 23 mutations also coincided with an observable change in allele frequency of another cohort consisting of 21 mutations, decreasing from 69.6% to 51.2% (Figure 3b).

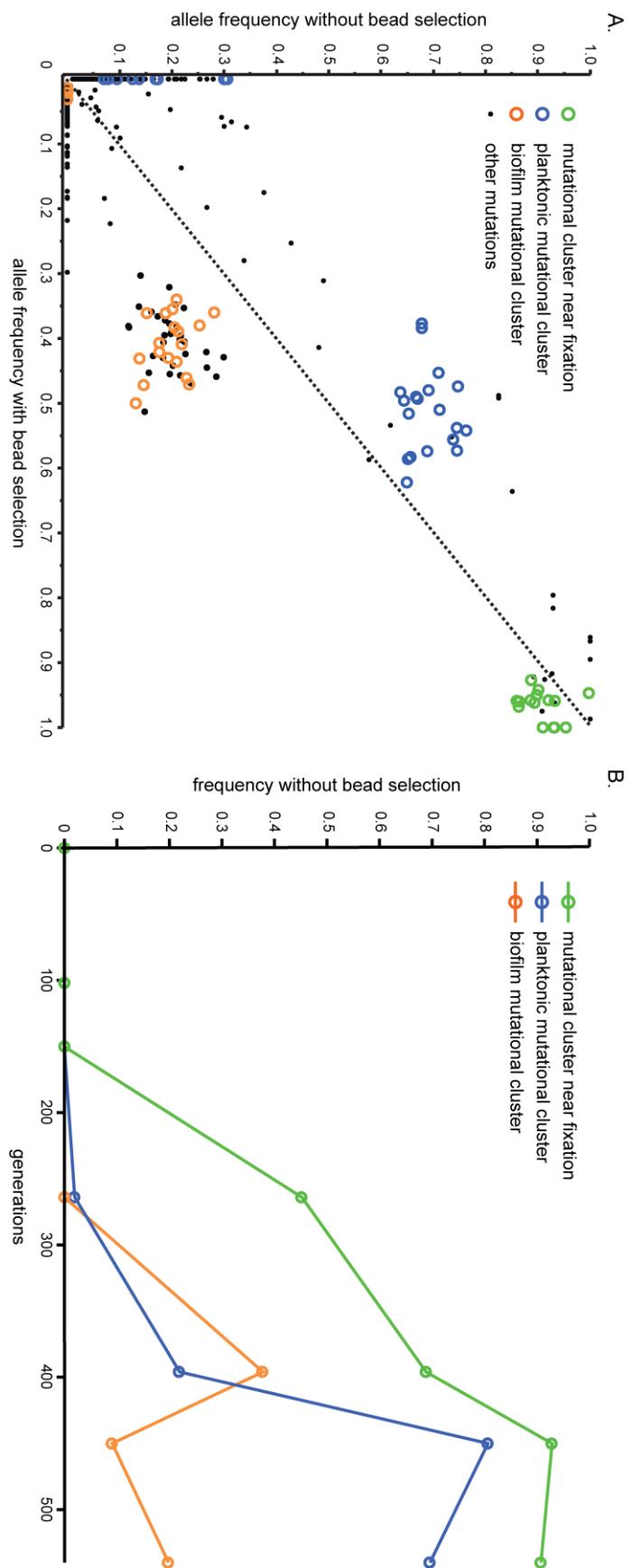


Figure 3. Frequencies of mutational calls in different samples of the B1 population and associated cohorts.

A. Linear regression of observed allele frequency from whole-population sequencing utilizing two different preparation methods. Points represent a total of 534 mutational calls, 128 of which were identified in both data sets. Mutations below the line are enriched following bead selection while mutations above are enriched during planktonic growth. B. Examination of three clusters identified visually reveals that these mutations belong to particular lineages with shared adaptive history

In an attempt to associate these changes in allele frequency to information on the ecology of the system (Chapter 2), we sought to identify which mutations associated with known isolates and, by extension, specific phenotypes and ecological strategies in two ways. First, we determined which isolates associated with the two mutational cohorts described above (Figure 3). Of the 23 mutations increasing following bead selection, the majority were shared among nine isolates: 101-103, 105, 106, 108, 109, 111 and 112. Likewise, the 21 mutations shown to decrease belonged to the O and H types. Second, we examined the 406 rare mutations only observed once in either metagenome sample. Although the majority of these low-frequency mutations are more likely to be the result of error, 51 of these singleton mutations were verified with isolate sequencing: 28 mutations following bead selection and 33 mutations without. Similar to above, we found that these mutations tended to associate with particular isolates (Table 1). For example, we only observed mutations specific to known early biofilm specialists, D and V, following bead selection (Chapter 2). Likewise, 13 mutations associated with the O and H types were only observed without bead selection.

Table 1. Low frequency mutations correspond with specific isolates following selection.

isolate(s)	planktonic selection	bead selection	total
101	0	2	2
103	0	1	1
104	1	1	2
105	0	6	6
107	5	0	5
101, 102, 106, 108	0	1	1
101, 111	1	0	1
101-103, 105, 106, 108-112	1	0	1
102, 106, 108	0	2	2
109, 110	0	8	8
131, M, M2, H	2	0	2
131, M, M2, P, P2, H, 104	3	0	3
D, V	0	3	3
H	3	0	3
O	6	0	6
P, P2	1	1	2
P, P2, 104	1	0	1
P2, 104	1	0	1
P2, O, 105, 108-110	0	1	1

Identifying potential targets of bead selection responsible for the observed patterns of change. In an effort to identify the molecular mechanism responsible for this response to bead selection, we examined the mutational clusters described above (Figure 3). In particular, we sought to identify potential targets of selection through convergence at the gene and operon level. Of the 44 mutations present (Table B.4), only two loci were repeatedly found to be mutated in samples at multiple time points and replicate populations: *anr* and *flgB* (Table 2). Of the five *anr* mutations identified, one of these mutations associated with each of the co-existing lineages identified above (Figure 3).

Table 2. Convergence suggests mutational targets resulting in ecological separation.

functional unit*	class	description	group	count	X^2	P
Anaerobic growth	Operational unit	Anaerobic respiration	biofilm-evolved populations	19	4.5907	< 0.05
			metagenomic sampling	7	6.6312	< 0.05
<i>anr</i>	gene	transcriptional regulator Anr	biofilm-evolved populations	5	179.7047	< 0.05
			metagenomic sampling	4	363.6257	< 0.05
flagellar biosynthesis	Operational unit	flagellar basal body structural components	biofilm-evolved isolates	15	5.2136	< 0.05
			metagenomic sampling	7	6.6312	< 0.05
<i>flgB</i>	gene	flagellar basal body rod protein	biofilm-evolved populations	1	11.73360248	< 0.05
			metagenomic sampling	1	39.86110076	< 0.05

In total, five independently derived *anr* mutations were observed in the B1 and B2 replicate populations. Mapping these mutations to the *anr* coding sequence revealed that mutations were clustered in two predicted protein domains (Figure 4a), ligand-binding, and helix-turn-helix/cyclic-AMP binding (HTH-CRP). Given the importance of Anr as a global transcriptional regulator controlling numerous pathways needed for anaerobic respiration, we hypothesized that these mutations could influence the ability of derived isolates to utilize an alternative terminal electron acceptor in the absence of oxygen. To test this hypothesis, we grew replicate cultures of representative isolates containing each of the five identified *anr* mutations under anaerobic conditions with the addition of nitrate, a common alternative electron acceptor for Pseudomonads. In partial agreement with our hypothesis, three out of the five isolates were unable to grow anaerobically. Interestingly, where the mutation occurred correlated with distinct phenotypic consequences; mutations occurring near the ligand binding site resulted in a complete loss of mutants to grow anaerobically while isolates containing *anr* mutations occurring in the HTH-CRP domain retained the ability to grow anaerobically (Figure 4b). We verified these phenotypes were not the result of secondary mutations through the

creation of isogenic mutants containing either the I119T or the G180D *anr* mutation that exhibited the same pattern.

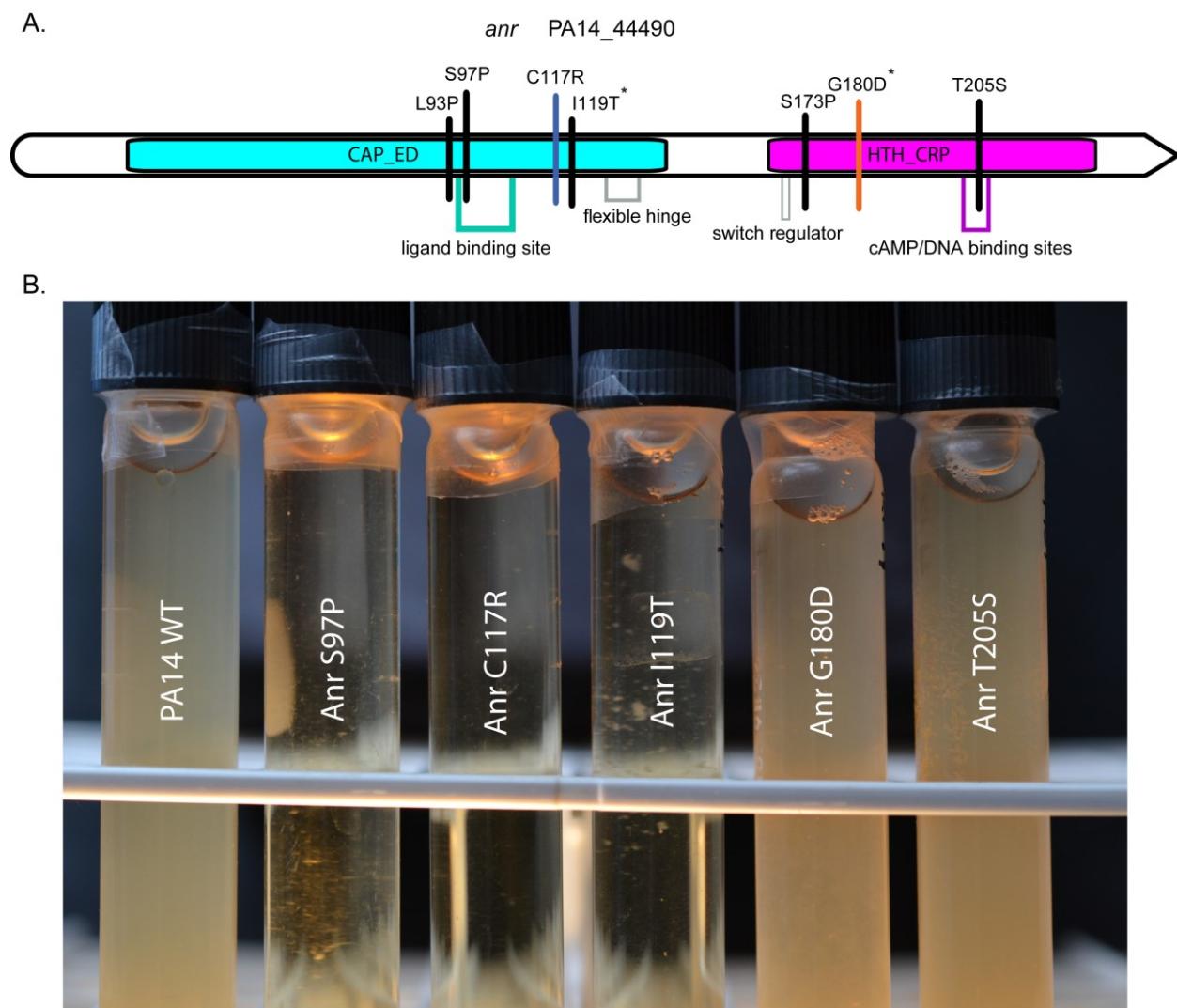


Figure 4. Characterization of mutants containing mutations in the anaerobic respiration transcriptional regulator, *anr*.

A total of five mutations in *anr* produced variable growth phenotypes under anaerobic conditions. A. The position of the five identified *anr* mutations overlaid onto functional protein domains with vertical bars as well as two additional mutations found at low frequency in the raw sequencing data. The blue and orange bars correlate with cohorts highlighted in Figure 3: increased with bead selection (orange), T205S *anr*, and decreased with bead selection (blue), C117R. B. *Anr* mutants with a mutation in the CAP_ED domain are unable to respire anaerobically using nitrate as an alternative terminal electron acceptor. Asterisks denote mutations that influenced anaerobic growth in isogenic constructs.

Whole-population sequencing over short timescales reveals micro-geographical

niches within biofilms. The long-term maintenance of distinct ecological groups originating from a single population requires some sort of fitness trade-off. Given the potential of the two types of *anr* phenotypes, anaerobic-capable or not, we sought to build a detailed snapshot of the population genetic structure of B1 and verify that these mutations associated with distinct co-existing lineages. Revisiting the 12 mutational cohorts identified to share similar trajectories through time, we systematically sorted and nested each cohort based on its maximum allele frequency and probability of occurrence at each time point. For example, two cohorts were nested if they are observed at greater than 50% at the same time point or continued to follow similar patterns of increase or decrease through time (See Appendix B for access to scripts that automate this process). This analysis produced a detailed image of the B1 population genetic structure characterized by the evolution of four distinct clades or collections of nested lineages (Figure 5a).

During the first 100 generations of adaptation, the B1 population saw the evolution of two competing lineages that slowly increased in frequency over the next 50 generations. Some time after 150 generations, one of these lineages acquired the T112P *mutS* SNP (green) and eventually outcompeted the alternative lineage (red) by 264 generations. The global success of this hypermutator lineage was likely attributed to the rapid accumulation of secondary mutational changes on this background, producing three new lineages. Since these lineages remained segregated up until the end of the experiment, we classified these three lineages as clades: pink clade, purple clade and lime green clade. By 264 generations, these clades already represented a

majority of the entire population and dominated the population for the remainder of the experiment (pink clade, ~5%, purple clade, 20% and lime green clade, ~40%).

The purple clade remained at low frequency, slowly decreasing in size to ~5% allele frequency by 450 generations. By 400 generations, however, the pink clade increased in allele frequency to ~10% and persisted until the end of the experiment. Interestingly, this subtle increase from 5% to 10% in allele frequency coincided with the occurrence of one of the observed *anr* mutations, the *anr* S97P SNP that was unable to grow anaerobically (Figure 4). Around the same time, the lime green clade continually increased in frequency from 264 until 450 generations, representing the majority at ~80%, in part due to the success of two new lineages on this background (orange and blue). Similar to the pink clade, this success coincided with the occurrence of two *anr* mutations; the orange *anr* T205S SNP lineage and blue *anr* C117R SNP lineage, previously shown to rapidly respond to bead selection (Figure 3). Both lineages co-existed within the lime green clade until the end of the experiment while accumulating additional mutations (light red, light purple and light blue lineages).

Despite the observed stability of these three major clades (pink, purple, lime green) during the long-term experiment, we wondered how much the population genetic structure fluctuated as a whole in response to bead selection. Revisiting our additional whole-population sequencing data following a single day of bead selection, we generated an alternative view of the population genetic structure present at 540 generations (Figure 5b). We reasoned that this analysis demonstrates the extent of fluctuations due to the daily bead transfer and associated forces of selection that may favor planktonic growth initially and biofilm attachment later (Chapter 2). Surprisingly,

we observed shifts in allele frequency as great as 23%, comparable to shifts previously observed over the past hundred generations of adaptation (Table B.4). In particular, the orange T205S *anr* lineage, which retained anaerobic growth capacity (Figure 4), increased in frequency. Consequently, eight out of the twenty isolates harbored this mutation. Meanwhile, lineages containing *anr* mutations without anaerobic capacity decreased: the pink clade, S97P, and the blue *anr* C117R lineage. Taken together, these patterns suggest a potential ecological trade-off between bead attachment and planktonic growth due to anaerobic growth conditions or stress.

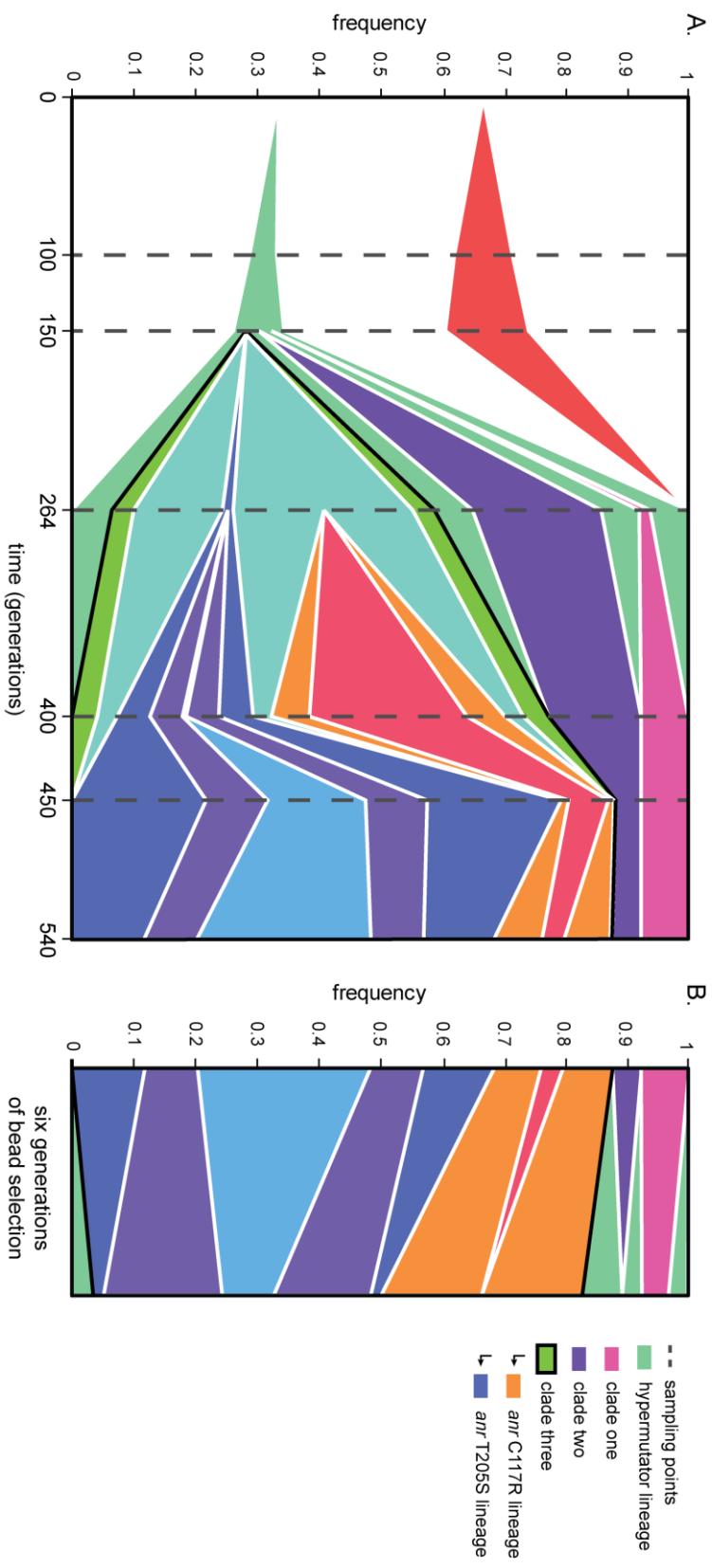


Figure 5. Population genetic structure of a biofilm population over long-term and short-term evolution.

A. Frequencies of major haplotypes in population B1 during long-term evolution. Following the fixation of a MMR-genotype (green), three distinct clades emerged: the pink clade, purple clade and lime green clade. The success of the lime green clade coincided with the occurrence of two additional lineages on this background, each harboring a NS SNP in *anr*: *anr* C117R, blue, and *anr* T205S, orange. Vertical dotted lines denote time points where whole-population sequencing was performed. **B.** Although the population genetic structure appears relatively stable between 400 and 540 generations, examining fluctuations over the course of a single day reveals dynamics undetected with our original whole-population sequencing approach.

DISCUSSION

It is now commonplace to study the dynamics of adaptation using NGS methods, and especially for studies using experimental microbial evolution. There are two common approaches: genome sequencing of representative clones and whole-population sequencing of the total genetic variation present. Both approaches leverage data from multiple samples (among temporal series or clones) to increase the confidence of mutational calls. The choice of approach may largely depend on experimental design; temporal sampling can be used to increase statistical power, despite less biological replication (Lang et al. 2013; Lang et al. 2011). However, sequencing of isolates with defined traits of interest allows genotypes to precisely linked to phenotypes of interest that is otherwise difficult with whole-population sequencing (Barrick et al., 2009; Tenaillon et al., 2012). To achieve the benefits of both methods, we used a combination of approaches and multiple sample preparation methods to resolve the extraordinary mutational diversity found in these biofilm populations of *P. aeruginosa* that became mutators. We found that both isolate and whole-population sequencing defined similar patterns of mutational variation (Figure 1), but both were required in tandem to decipher likely adaptive targets residing in rare lineages.

Although biofilms may predictably preserve genetic variation owing to spatial structure and the ecological interactions that consequently arise (Besemer et al., 2012; Parsek & Singh, 2003; Periasamy & Kolenbrander, 2009; Yildiz & Visick, 2009), few studies have examined the extent to which this diversity persists. Much to our surprise, a substantial number of lineages with unique beneficial mutations existed below our limit of detection using metagenomic sequencing (Figure 2). Although a sampling bias may,

in part, explain why genetic variation goes unseen (Figure 2b, Table 1), this study illustrates the vast amount of diversity present in these diverse communities and our inability to measure the total genetic variation truly present in structured communities. For example, we found that 54.75% of mutations identified with isolate sequencing were not observed with whole-population sequencing despite >300X average read coverage combined. Similar to previous work in our laboratory examining genomic adaptation associated with biofilms of *Burkholderia cenocepacia* HI2424 (Poltak & Cooper, 2011), colony phenotypes occasionally represent relics of presumably extinct lineages (Traverse et al., 2013). Taken together, these findings suggest that biofilms harbor a staggering amount of persistent genetic variation even amidst what appear to be global fixation events.

Can we sequence ecology? Our findings suggest that a single sequencing approach is insufficient to do so. Examining the genetic structure of biofilm-evolved populations in the absence of ecological information reveals population dynamics that are nearly indistinguishable from more simple systems characterized by periodic selection (Lang et al., 2013; Wong et al., 2012). We found that identifying targets of selection through convergence alone becomes exceedingly difficult as complexity increases. For example, the B1 population is characterized by the co-occurrence of numerous genetic lineages that rapidly accumulate many mutations due to hypermutation obscuring the signal of adaptive mutations and convergence (Table B.4). We were only able to generate a testable model explaining the co-occurrence of two genetic lineages (Figure 3) in combination with some knowledge of the ecology of the B1 population *a priori* (Chapter 2). Even with this progress, we have certainly only

scratched the surface of the ecological complexity within these biofilms with the presence of many other genetic lineages with no obvious ecological trait or below our level of detection in the B1 population alone (Figure 5).

Colony phenotypes and associated ecological strategies evolve repeatedly on a variety of different genetic backgrounds and persist. If the evolution of biofilm specialization is as easy as some EME studies of shorter time frames and models of bacterial speciation suggest (Koeppel et al., 2013; McElroy et al., 2014), the rampant re-evolution of ecological strategies from any genetic background would certainly camouflage any genetic signal of ecology. Ecological stability and, by extension, the observation of ecology requires physical, mutational, or fitness barriers to prevent mixing (Herron & Doebeli, 2013). With previous work in our lab, generalist lineages repeatedly invaded specialist niches despite the evolution of synergy and, coupled with mutations improving competition for resources, readily displaced specialist lineages (Traverse et al., 2013). These findings support a model where highly diverse biofilm populations will still be largely defined by global adaptive events.

Understanding adaptation to structured environments over longer time scales will thus require knowledge of how diversity segregates into distinct ecological units (Shapiro & Polz, 2014). For example, distinct genotypic clusters of free-swimming Vibrios associated with an ecological trade-off based on food particle size preference (Polz et al., 2006). Here, we originally identified two distinct genotypic clusters exhibiting enrichment within bead-associated samples (Figure 3). In particular, mutational clusters that increased in response to bead selection retained the ability to grow anaerobically (Figure 4). Examining how this potential trade-off influenced the population genetic

structure of B1, we found that *anr* mutations associated with three distinct lineages (Figure 5a). Furthermore, a single day of bead selection resulted in large fluctuations in the population genetic structure of B1 (Figure 5b). Taken together, we argue these patterns illustrate micro-environment variation that occurs daily in our system with bead transfers. Once a bead is transferred, selection likely favors free-swimming genotypes initially but eventually favors those genotypes that can either handle the stressors inherent to the biofilm lifestyle. Furthermore, the B2 population also saw similar *anr* mutations further emphasizing the importance of this subtle target of selection.

In conclusion, this study emphasizes the need for better experimental design and bioinformatic approaches to accurately identify rare, likely important, adaptive variation. Despite only working with a single species of bacteria, the amount of persistent genetic variation makes identifying important beneficial variation challenging. Although there is a growing trend of mixed sample sequencing, deciphering the adaptation of microbial populations, especially with species-rich communities, will require some combination of both metagenomics, genomic sequencing of representative isolates, and laboratory bench work to fully understand long-term adaptation to a structured environment. In addition, this study highlights how this combination can also lead to testable hypotheses of adaptive mechanisms associated with a response to micro-scale selection.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Replicate biofilm populations of *Pseudomonas aeruginosa* PA14 were adapted during an EME study under conditions described previously (Chapter 2). To achieve anaerobic growth conditions, 7mL of tryptic soy broth supplemented with 100 mM KNO₃ was inoculated in 10mL screw cap

tubes with the addition of 1 mL of mineral oil to eliminate the exchange of oxygen at the air-liquid interface. To ensure an absence of growth was due to the inability of cells to utilize nitrate as a terminal electron acceptor, an aliquot of the inoculum was spot plated on agar plates to assess viability.

Preparation of samples for genome sequencing. Samples were sequenced in four different experimental blocks varying in design and preparation: a) 23 clones from both biofilm and planktonic-evolved populations (B1, B2, B3, P1, P2 and P3) b) 20 clones from the B1 population, c) 27 metagenome samples isolated after 107, 150, 264, 396, 450 and 540 generations from all three replicate biofilm-evolved populations and 396 and 540 generations from all three replicate planktonic-evolved populations and d) three metagenome samples isolated after 107, 264 and 540 generations from the B1 population following a day of bead selection. Re-extracted genomic DNA from the ancestral PA14 strain from our laboratory was included in experimental blocks a and b to both verify the exact DNA sequence of our ancestral strain, and to help eliminate false positives introduced by sequencing and alignment error. Blocks c and d did not include our ancestral strain since leveraging mutation identification across time points allowed us to eliminate false positives. For blocks a-c, genomic DNA was extracted from 1 mL of overnight cultures utilizing Qiagen's DNeasy Blood and Tissue kit. For block d, genomic DNA was extracted using a traditional phenol, chloroform extraction method for library construction in-house. Sample quality was assessed with both NanoDrop and gel electrophoresis before samples were processed for library construction.

Library construction was sequencing was performed as follows: a) genomic DNA was extracted using a traditional phenol, chloroform extraction method and purified

using library construction and sequencing was performed at University of Vermont's DNA and Microarray Core Facility using NuGen's Ovation library construction system to generate 100bp paired-end reads for reference alignment, b) library construction was conducted through the Michael Lynch laboratory at the Indiana University utilizing services available at Center for Genomics and Bioinformatics (CBG) before samples were sequenced on Illumina's HiSeq2500 at the Hubbard Center for Genomic Studies (HCGS) at the University of New Hampshire, c) library construction for metagenomic samples was performed using the Nextera XT DNA Sample Prep Kit, and also sequenced at the HCGS, d) libraries were constructed in-house using a modified version of a protocol made available by the University of Massachusetts' Medical Center Deep Sequencing facility (<http://www.umassmed.edu/nemo/>). Briefly, genomic DNA was fragmented (<800bp) using Invitrogen's Nebulizer Kit with sterilized 0.2µm air filters and tubing at 40psi. The resulting fragments were blunted and ligated with adaptors using Epicentre's END-IT DNA repair kit and FAST LINK KIT (#LK11025) before sequencing at the Tufts University Core Facility (TUCF).

Mutation identification. The BRESEQ software package (<http://www.barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing>) was used to align Illumina reads to the *Pseudomonas aeruginosa* PA14 genome, and also called polymorphisms for individual isolates and mixed populations (Deatherage & Barrick, 2014). All reads that covered a particular site were extracted via BRESEQ's graphical output of aligned reads providing evidence, and mutational events were manually validated by examining this graphical output. This method of validation has been previously demonstrated to effectively screen for errors in mutation identification.

Polymorphic mutational events were identified utilizing BRESEQ's experimental polymorphism prediction feature. Briefly, breseq tests that the reads aligned at a given position in the genome supports a model of mutational variation through calculating the maximum likelihood percentage with each base to predict polymorphism at 0.1% resolution. Statistical support is based on the probability that the observed alignment could be generated by chance using a likelihood-ratio test. Since the empirical error model and quality re-calibration used by breseq does not fully capture some second-order sources of variation in mutational error rates, false-positive polymorphism predictions were overcome through only retaining mutations identified across multiple time points or samples for subsequent analysis.

Estimating the ratio of non-synonymous to synonymous substitutions for mismatch repair (MMR) mutants. The content of the *Pseudomonas aeruginosa* PA14 genome was used to calculate the rate at which non-synonymous and synonymous substitutions would occur by chance. Codon usage for PA14 was downloaded from the codon usage database summarizing NCBI-GenBank information (<http://www.kazusa.or.jp/codon/>). Using this information, the neutral ratio of non-synonymous to synonymous substitutions (dN/dS) was estimated to be 2.96. MMR-deficit mutants have been shown to enrich for transition (AT->GC) over transversion (AC->CG, AT->TA) substitutions (Miller, 1996). Given that transitions favor synonymous substitutions, we adjusted the expected dN/dS ratio to reflect the observed rate of transition substitutions in our data set of 96% (Figure 2) with an expected dN/dS ratio of 1.75.

Clustering of mutational trajectories and assessing population genetic structure

parameters. Mutations by only retaining mutations that were: i) are present at multiple time points and ii) achieved greater than 10%. This data set was used to create Fisher-Müller diagrams (Figure 3) using MATLAB scripts provided by Katya Kosheleva and Michael Desai of Harvard University. The scripts automate the process of defining mutational cohorts and drawing Fisher-Müller diagrams of population-genetic dynamics (Lang et al., 2013). The scripts were modified to accommodate a variety of data structures and experimental designs. To summarize this method briefly, mutations were first grouped into clusters of mutations or cohorts based on mutational trajectories through time. At each time point, cohorts were sorted by their maximum allele frequency and nested based on the probability of occurrence. For example, a trajectory can occur on the wildtype background or on the background of another cohort at higher frequency. Cohorts were nested if they are observed at greater than 50% or follow similar patterns of increase or decrease. Finally, this information was used to produce a Fisher-Müller diagram for each replicate population.

Assessing the convergence of adaptive targets. Targets of selection were identified based convergence. After the identification of mutational clusters of interest, genes or operons observed to be mutated multiple times were selected for further study (For example, Table B.4).

Mutant construction with two-step allelic exchange. *P. aeruginosa* PA14 was grown at 37 °C using Tsoy broth, LB broth or in Vogel–Bonner minimal medium (VBMM; 0.2 g MgSO₄•7H₂O, 2.0 g citric acid, 3.5 g NaNH₄HPO₄•4H₂O, 10 g K₂HPO₄, pH 7.0) optionally supplemented with 1.5% Bacto agar for semi-solid media. *Escherichia coli*

strains were routinely cultured at 37 °C using Tsoy broth. Antibiotics were added as follows: for *E. coli*, gentamicin (Gm) at 5µg/ml to maintain plasmids, Gm at 10µg/ml for selection and 15µg/mL naladixic acid for selection against donor strains and for *P. aeruginosa*, Gm at 30µg/ml for plasmid retention or Gm at 60µg/ml for selection.

Isogenic mutants were generated using methods described elsewhere (Zhao et al., 2013). Briefly, PCR products were amplified using primers (Table B.5) that targeted 1 kb regions surrounding the I119T *anr*, G180D *anr*, R376C *pslE*, and W588* *pi/Y1* mutations tailed with attB1 or attB2 sequences as described in the Gateway Cloning Technology Manual (Invitrogen). To control against the occurrence of additional mutations throughout the cloning process, PCR products were always amplified with the Phusion high-fidelity polymerase (NEB) and gel extractions were performed with the addition of 1mM guanosine to protect DNA from the mutagenic effects of UV. Gateway BP Clonase II Enzyme Mix (Invitrogen) allowed for *in vitro* recombination between resulting PCR products and the pDONRPEX18Gm plasmid derived from pEX18Gm (Hoang, Karkhoff-Schweizer, Kutchma, & Schweizer, 1998). Plasmid constructs were confirmed to contain only the intended mutations by Sanger sequencing using the M13 universal primers. Next, plasmids were introduced into Z-competent DH5α *E. coli* cells (Zymo Research) for plasmid maintenance and storage. Plasmid constructs were harvested overnight cultures of the resulting DH5α using the PureYield Plasmid Miniprep System (Promega) and electroporated into S17.1 *E. coli* cells. Next, the plasmid was introduced into PA14 utilizing biparental mating with *E. coli* S17.1. Transformants were selected using 60 ug/mL gentamicin. Electroporations and biparental mating protocols were performed by adapting standard methods surrounding

the use of the mini-Tn7 system with *Pseudomonas* (Choi, Kumar, & Schweizer, 2006).

Finally, constructed mutants were confirmed to contain the mutant allele using Sanger sequencing on a collection of eight to sixteen potential mutants.

CHAPTER IV

EVOLUTION OF HYPERMUTATION IN EXPERIMENTAL *PSEUDOMONAS* BIOFILMS
AS AN ESCAPE FROM CLONAL INTERFERENCE

INTRODUCTION

The evolution of an elevated mutation rate, or hypermutation, is commonly observed with bacteria in laboratory and clinical settings (Jayaraman, 2011; Mena et al., 2008). Since hypermutation is expected to have no direct benefit on organismal fitness (Shaver et al., 2002), fixation of this trait in large populations dominated by selection requires that it be linked to at least one other beneficial mutation, in a process called genetic hitchhiking (Taddei et al., 1997). Despite extensive theoretical and experimental work to understand conditions where genetic hitchhiking can allow hypermutation to fix, the prevalence of these conditions in more complex environments are largely unknown.

Many plausible scenarios favoring the genetic hitchhiking of hypermutation have been proposed: mutational biases (Couce et al., 2013), fluctuating environmental conditions (Travis & Travis, 2002), strong clonal interference (Tenaillon et al., 1999), and simple co-occurrence with a single mutation (Wylie et al., 2009). Yet it is also evident that we still have much to understand about the evolution of mutators. Studies utilizing similar experimental conditions report varying frequencies of hypermutation (Gaut et al., 2012; Maharjan et al., 2006; Notley-McRobb et al., 2002; Sniegowski et al., 1997), which suggests that specific organismal traits, environmental conditions, or unknown dynamics within the population-genetic environment influence the evolution of bacterial mutation rates.

Pseudomonas lung infections of persons with cystic fibrosis (CF) are characterized by immense genetic and phenotypic diversity (Ashish et al., 2013). Although typically explained as a direct result of antibiotic therapy (Maciá et al., 2006), hypermutation may facilitate genomic adaptation to the biofilm lifestyle (Mena et al.,

2008). Association with a surface creates gradients of nutrients, waste, and signaling molecules that alter local forces of selection and create distinct ecological opportunities (Hall-Stoodley, Costerton, & Stoodley, 2004). Genetic variation created by hypermutation may enable rapid adaptation to these opportunities and other micro-environmental variation inherent to a structured environment (Couce et al., 2013; Torres-Barceló et al., 2013).

Here, we examine how the population genetic structure both influenced and changed with the ultimate success of hypermutation in three experimental biofilm populations of *P. aeruginosa* PA14 described previously (Chapter 3). Using a combination of both metagenomic sequencing and genome sequencing of representative clones (Traverse et al., 2013), we are able to identify functional targets of selection despite a very high noise (genetic variance caused by hypermutability) to signal (selected mutations) ratio. In particular, we find dynamics consistent with the explanation that hypermutation represents a 'quick fix' to escape rampant clonal interference that evolved following extensive phenotypic diversification and ecological segregation. Furthermore, the association of genetic variation in traits observed during chronic infections to various lineages in these model populations provide additional insight into the unpredictable patterns of adaptation common among CF infections infected with *P. aeruginosa*.

RESULTS

Previously, we utilized a combination of deep sequencing and isolate full genome re-sequencing to examine the genomic basis of adaptation and ecological differentiation within experimental biofilm communities of *P. aeruginosa* PA14 (Chapter 3, Table A1-3).

That data set allowed us to examine why the population genetic history of the three replicate biofilm-evolved populations (B1, B2, B3) differed from control populations evolved under planktonic conditions (P1, P2, P3).

Hypermutation drives adaptive genomic variation within replicate biofilm

populations. Previously we reported a strong bias for transition substitutions (96.5%) consistent with the early occurrence of hypermutation in replicate biofilm populations (Chapter 3). The B1 and B2 populations, and eventually B3, saw the fixation of hypermutation via non-synonymous substitutions elevating the rate of mutation approximately 50-fold due to a loss of 80.1% in MMR activity in each biofilm-evolved population. The same SNP in *mutST112P* evolved independently in populations B1 and B2, and *mutL* D467G evolved in B3. The number of mutations identified in each is also consistent with the order by which hypermutation fixed: after 540 generations of adaptation, we identified 187 mutations in the B2 population followed by 143 in the B1 population, with only 19 in the B3 population (Table B.2). Insertion and deletion mutations were much less common; only 190 deletions and insertions were observed across all populations. In agreement with previous findings (Chapter 2), fewer mutational changes were identified in the planktonic-evolved populations due to retention of the wild-type mutation rate throughout the experiment (Table C.1).

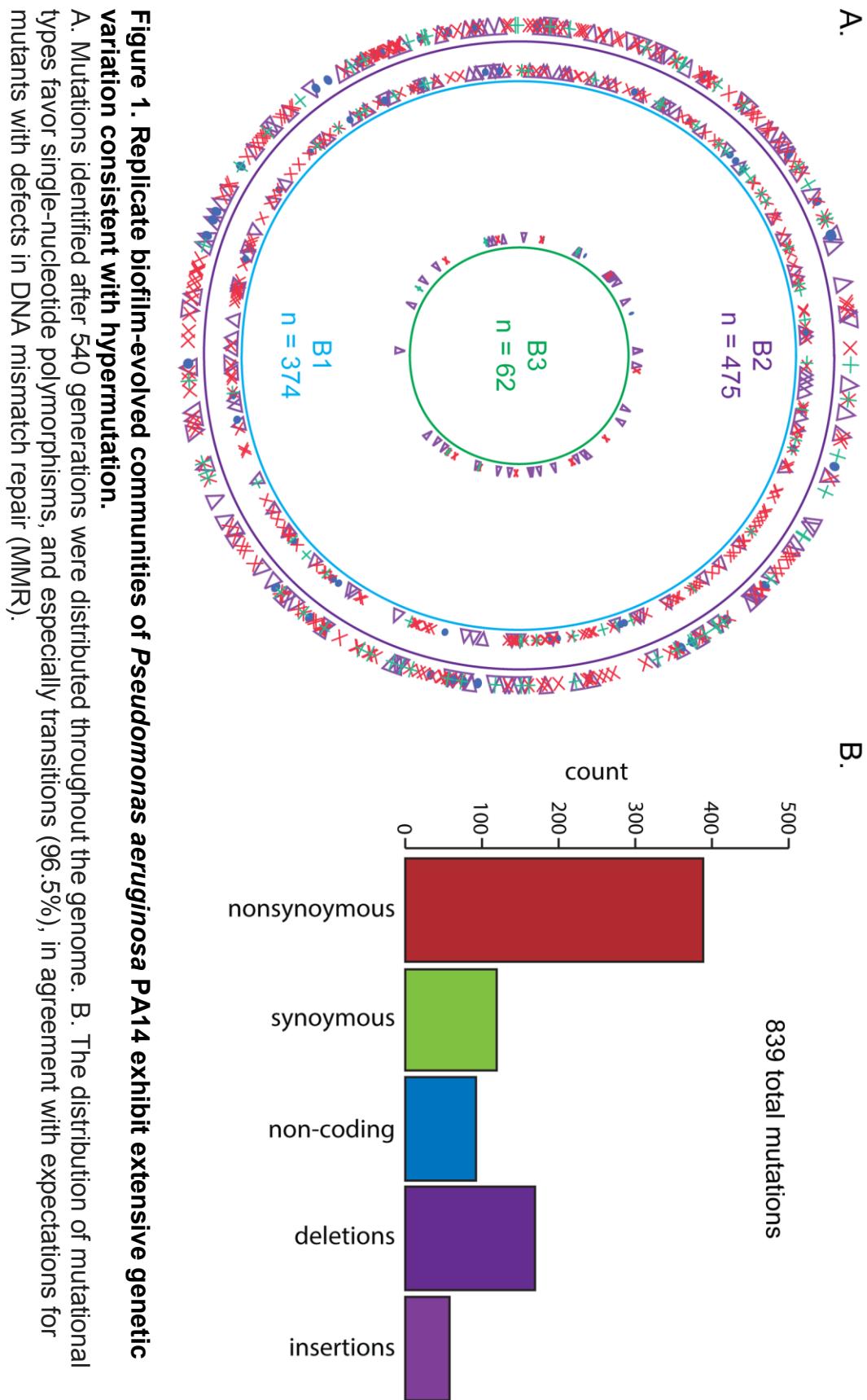
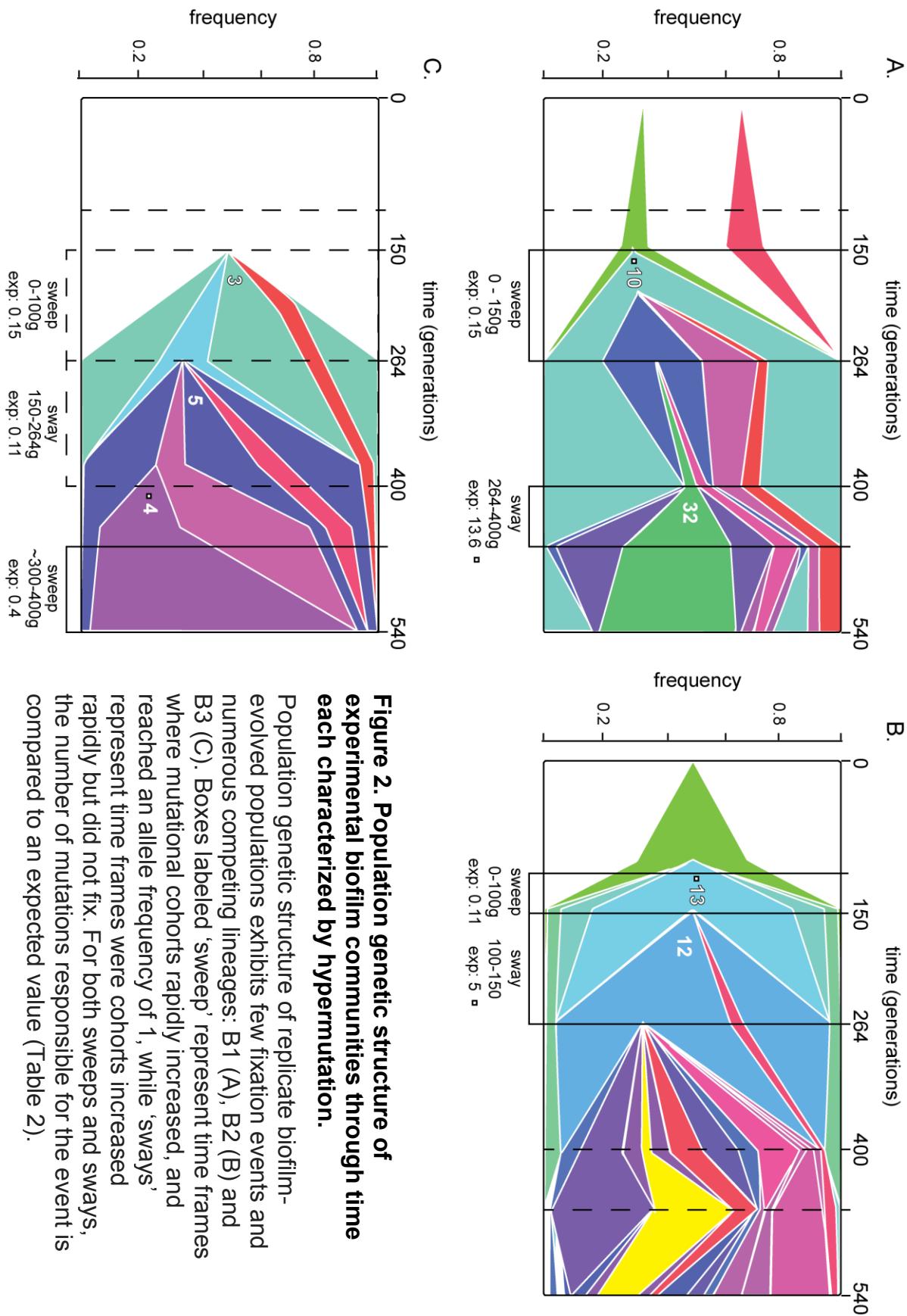


Figure 1. Replicate biofilm-evolved communities of *Pseudomonas aeruginosa* PA14 exhibit extensive genetic variation consistent with hypermutation.

A. Mutations identified after 540 generations were distributed throughout the genome. B. The distribution of mutational types favor single-nucleotide polymorphisms, and especially transitions (96.5%), in agreement with expectations for mutants with defects in DNA mismatch repair (MMR).

Populations are defined by numerous competing lineages and few selective sweeps. Although linked sets of mutations that define haplotypes did undergo periods of rapid change in frequency, genetic sweeps that purged genetic diversity from each replicate population were relatively uncommon (Figure 2). Of the sweeps that did occur, three of four associated with mutator alleles (Figure 2a,b). In population B3 one sweep occurred independently of any mutations related to MMR between 150 and 264 generations (Figure 2c).



Successful biofilm lineages accumulate more mutations faster. Previous work examining the long-term maintenance of genetic variation in the B1 population illustrated the presence of a vast amount of low-frequency variation in these biofilms (Chapter 3). Based on this precedent, we hypothesized that hypermutation repeatedly fixed as the indirect effect of strong selection to acquire multiple beneficial mutations caused by clonal interference among many competing lineages. To test this hypothesis, we compared the number of mutations associated with genetic sweeps or sways in each replicate biofilm population to the expected number of mutations that would accumulate on average in a mutator genotype (Table 2). Although estimates of the genome-wide mutation rate may suffer from inaccuracy (Baer et al., 2007), a mutation accumulation experiment with hypermutator *P. aeruginosa* PAO1 reported a rate of 0.18 mutations per genome per generation (Heilborn et al., 2014). Based on observed loss of 80.1% in MMR activity (Chapter 2), expected values were calculated assuming 0.144 mutations per genome per generation. Consistent with our hypothesis, each of the genotypes associated with the seven events described above harbored many more mutations than expected (Figure 2, Table 1).

Table 1. Expected versus observed mutational counts calculated based on the mutation rate and time in generations.

population	event	time (generations)	expected*	observed
B1	sweep	150	0.15	8
B1	sway	136	32	13.6
B2	sweep	110	0.11	13
B2	sway	50	5	12
B3	sweep	100	0.15	3
B3	sway	114	0.11	5
B3	sweep	100	0.4	4

* calculated assuming 0.144 mutations per genome per generation

Although the mutator lineages that achieved fixation did so with many mutations (B1, 8, B2, 13 and B3, 4, Table 1) one haplotype fixed with only three mutations in the absence of hypermutation (Figure 2c, Table C.3). This haplotype contained three mutations likely to have large effects on protein expression and function: i) a 12 bp deletion in *orfH*, ii) a premature stop in *ddl*, and a frameshift mutation in PA14_66380. Interestingly, all three genes are associated with traits exterior to the cell that selection could act on: *orfH* encodes UDP-N-acetyl-D-mannosaminuronate dehydrogenase involved in polysaccharide production, *ddl* encodes D-alanine-D-alanine ligase implicated in cell wall assembly and PA14_66380 is predicted to create a potassium/proton antiporter. Thus, we predict that the rapid acquisition of multiple beneficial mutations, rather than hypermutation itself, drives global adaptation in our biofilm system.

Global success in biofilms is constrained as adaptation proceeds. In contrast, successful genotypes identified within planktonic-evolved populations reach frequencies above 40% with only two to three mutations (Figure C.1), and raise the question why mutators did not evolve to high frequency in planktonic populations. One possibility is that the effective population sizes (N_e) of biofilm and planktonic populations differed, which influences the supply rate of beneficial mutations. However, planktonic and biofilm selection regimes exhibited no significant difference in the size of the population bottleneck of the daily transfer (bead bottleneck, $1.36 \times 10^8 \pm 0.51 \times 10^8$, compared to planktonic, $1.79 \times 10^8 \pm 0.69 \times 10^8$, $t_{10} = 0.959$, $P = 0.3502$). Nevertheless these calculations may not capture the complex dynamics of structured communities. For example, we typically assume that all genetic lineages experience the same number of

generations per day, even despite genetic variation for fitness. However, the evolution of ecologically distinct genotypic clusters within biofilms due to micro-environmental separation may reduce the N_e of these subpopulations and potentially also limit the number of generations of growth. These subpopulations may also harbor multiple competing lineages that produce clonal interference. In fact, we find evidence of both dynamics within these biofilms as they adapted.

Colonies isolated after 540 generations in populations B1 that contained more unique mutations that were undetected in the whole population sequencing (which is biased towards higher-frequency mutations) tended to harbor fewer total mutations than the mean of all isolates sequenced (linear regression between unobserved versus total mutations, $F_{20} = 11.035$, $P = 0.0036$). Put another way, these isolates experienced as many as 200 fewer generations compared to the average number predicted, and could be viewed as evolutionary relicts with mutations that do not represent the population (Table 2).

Table 2. The expected minimum number of generations needed to accumulate the observed number of mutational calls for individual isolates.

isolate	mutations unobserved with deep sequencing	total mutations identified	minimum generations needed*	distance from grand mean^
M	43	61	423.05	204.43
D	54	62	429.99	197.49
P	57	77	534.02	93.46
M2	61	77	534.02	93.46
P2	85	108	749.01	-121.53
V	74	82	568.69	58.79
A	39	101	700.47	-72.99
H	62	81	561.76	65.72
O	47	110	762.88	-135.40
101	38	98	679.66	-52.18
102	34	93	644.98	-17.50
103	39	98	679.66	-52.18
104	53	74	513.21	114.27
105	40	99	686.60	-59.12
106	31	90	624.18	3.30
107	45	104	721.27	-93.79
108	35	94	651.92	-24.44
109	47	106	735.14	-107.66
111	24	84	582.57	44.91
112	36	95	658.85	-31.37

* minimum generations needed was calculated based on 0.144 mutations/ genome/generation.[^]subtraction from the grand mean of 627.49 generations.

If clonal interference caused mutator lineages to prevail, then the lineages that succeeded to reach high frequency should be enriched in functionally significant mutations. By comparison, unsuccessful lineages or isolates from rare subpopulations should not show this bias, and the overall enrichment in nonsynonymous mutations should also decline over time as subpopulations stabilize. To this end, we quantified the ratio of non-synonymous to synonymous substitutions (dN/dS) to measure enrichment in functionally significant mutations. Consistent with clonal interference, early hypermutator lineages represented by early metagenomic calls (metagenome at 107, 150, and 264 generations) exhibited the highest dN/dS overall (4.25, Figure 3). As

expected, this ratio declined in a stepwise manner as adaptation proceeded; late metagenome samples saw stabilization (3.25) while mutations only identified in individual isolates, likely representing rare genetic lineages (Chapter 3), exhibited the lowest dN/dS ratios around 3.

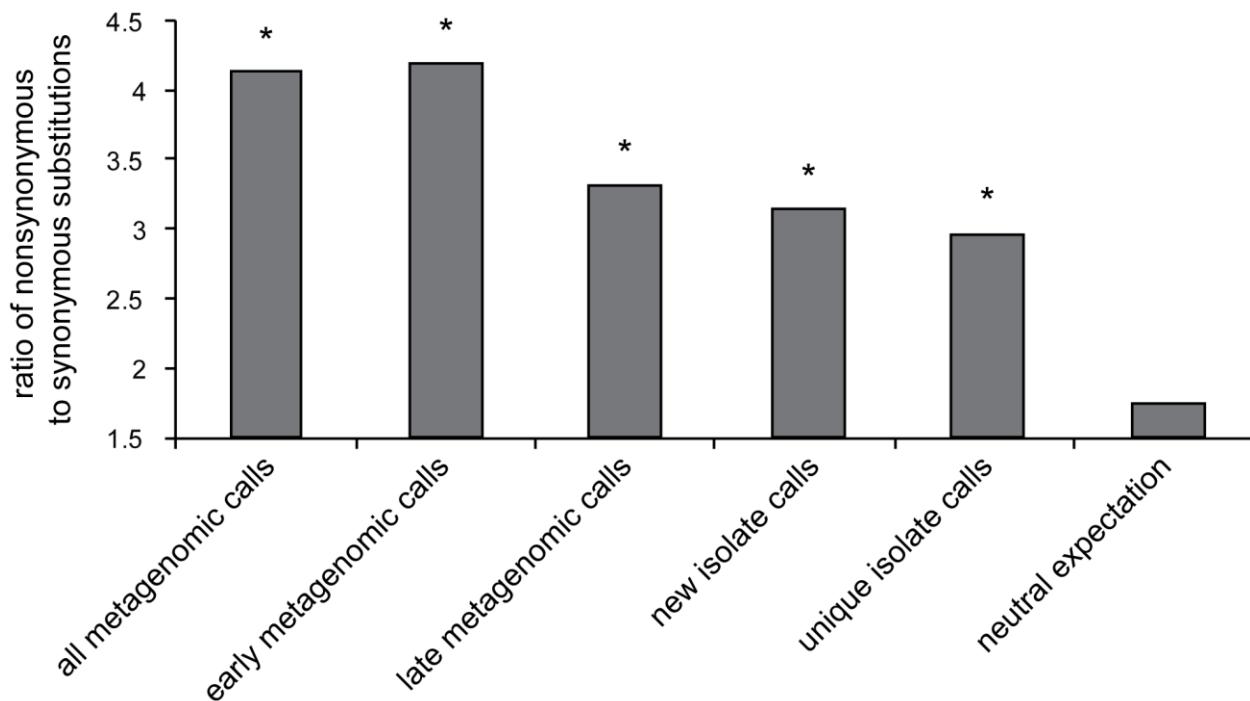


Figure 3. Adaptive lineages accumulate more synonymous substitutions as adaptation proceeds.

Replicate biofilm populations display a declining ratio of non-synonymous to synonymous substitutions through time. A bias towards the number of non-synonymous to synonymous mutational calls always remains greater than the neutral expectation. The neutral expectation was calculated as 2.96 based on genomic content, and corrected to 1.75 given the observed bias towards transitional mutations that favor synonymous substitutions.

Population genetic structure of replicate biofilms populations revealed limited convergence at the gene level. Given the large amount of genetic variation present, distinguishing driver mutations from passengers is difficult with the observed mutational counts. Of the mutational clusters demonstrating large shifts in allele frequency between

time points, only six of the 75 mutations demonstrated convergence at the operon level (Table C.1-3). In an effort to identify additional adaptive targets, we assessed convergence at numerous levels of cellular organization; the observed number of mutational hits in genes, operons, and functional categories was compared to expected values assuming neutral patterns of mutation (See Methods). We identified significant enrichment in mutations affecting eight functional categories (Table 3). These included mutations related to cell wall, LPS, capsule and polysaccharide production, which include all three mutations implicated in the first successful lineage observed in the B3 population (Table 1). In addition, we identified many categories known to be important for biofilm formation such as flagellar biosynthesis and glycosylation, GacA/GacS signal transduction, *wsp* genes and type IV pili.

Table 3. Comparison of mutational targets identified through convergence.

functional unit*	class	group	count	X^2	P
Anaerobic growth	Operational unit	biofilm-evolved populations	19	4.5907	< 0.05
		metagenomic sampling	7	6.6312	< 0.05
<i>anr</i>	gene	biofilm-evolved populations	5		
		metagenomic sampling	4		
Cell wall / LPS / capsule	Functional category	biofilm-evolved populations	30	122.6268	< 0.05
		metagenomic sampling	10	0.2122265	
DNA mismatch repair genes	Protein complex	biofilm-evolved populations	4	7.1806	< 0.05
		metagenomic sampling	10	19.2541	< 0.05
flagellar biosynthesis	Operational unit	biofilm-evolved isolates	15	5.2136	< 0.05
		metagenomic sampling	7	6.6312	< 0.05
GacA/GacS/Rsm signaling	Signaling pathway	biofilm-evolved populations	5	19.5763	< 0.05
		metagenomic sampling	1	1.1037	
glycogen synthesis and degradation	operon	biofilm-evolved populations	4	7.5385	< 0.05
		metagenomic sampling	0		
Motility & Attachment	Functional category	biofilm-evolved populations	30	153.4073	< 0.05
		metagenomic sampling	9	2.1122	
Nucleotide biosynthesis and metabolism	Functional category	biofilm-evolved populations	22	125.5950	< 0.05
		metagenomic sampling	10	4.6551	< 0.05
Pel / Psl polysaccharides	<i>pel</i> / <i>psl</i> operons	biofilm-evolved populations	8	2.3922	
		metagenomic sampling	2	0.1992	
Purine conversions	Operational unit	biofilm-evolved populations	9	5.1201	< 0.05
		metagenomic sampling	6	3.8842	
Related to phage, transposon, or plasmid	Functional category	biofilm-evolved populations	15	41.8374	< 0.05
		metagenomic sampling	10	9.0396	< 0.05
RNA polymerase	Protein complex	biofilm-evolved populations	13	52.8759	< 0.05
		metagenomic sampling	5	1.8675	
Transcription, RNA processing and degradation	Functional category	biofilm-evolved populations	19	163.27832	< 0.05
		metagenomic sampling	7	2.7028443	
Transcriptional regulators	Functional category	biofilm-evolved populations	93	101.0600	< 0.05
		metagenomic sampling	37	15.1669	< 0.05
Two-component regulatory systems	Functional category	biofilm-evolved populations	56	35.5878	
		metagenomic sampling	22	8.7192	< 0.05
Type IV pili	Operational unit	biofilm-evolved populations	16	18.6960	< 0.05
		metagenomic sampling	8	10.1085	
wrinkly spreader phenotype	operon	biofilm-evolved populations	3	3.9057	< 0.05
		metagenomic sampling	1	1.1037	

Ongoing work in our lab examining the effect of selected mutations suggests they may, in fact, be adaptive. For example, four mutations in the anaerobic transcriptional regulator, *anr*, were previously implicated as associated with ecological variation, or niches varying in oxygen availability, and also with the evolution of major lineages (Chapter 3). The T443C *ppk* mutation was one of eight mutations found to fix in the B1 population (Table C.1). Despite conferring a large fitness benefit (Chapter 5), no

additional mutations in the *ppk* gene were observed across all replicate populations. However, we did observe significant enrichment for mutations related to nucleotide biosynthesis and metabolism, of which the *ppk* gene is a member.

Functional clustering of mutations was sometimes only observed among isolates and not observed in the metagenome. For instance, in population B1 isolates after 540 generations (n=20), most variation between isolates could be explained by mutations in three functional categories (Table 3). Mutations affecting nucleotide biosynthesis and metabolism explained 55.6% of the variance, mutations occurring in motility and attachment explained 11.7%, and transcriptional regulator genes explained 7.0% of the variation between isolates (Table 4). As an example, we observed several mutations in flagellar biosynthesis and type IV pili among isolates, but only two mutations in *flgB* and *fliF* achieved greater than 50% frequency in the B1 population.

Table 4. Hierarchical clustering of the variance associated with functional mutations among B1 540 generation isolates.

PseudoCAP category*	sum of squares	df	mean squares	f model	percent	P
MA	229.57	1	229.57	9.652	11.71%	0.003
NBM	1090.42	1	1090.42	45.846	55.64%	0.001
TR	136.95	1	136.95	5.758	6.99%	0.006
Other	50.99	1	50.99	2.144	2.60%	0.118
Residuals	451.9	19	23.78		23.06%	
Total	1959.83	23			100%	

* MA = Motility and attachment, NBM = nucleotide biosynthesis and metabolism, TR = transcriptional regulators based on classifications from the Pseudomonas.com database.

In combination, these results suggest that evolutionary convergence affecting similar functions occurred among replicate biofilm populations, but hypermutation obscures these patterns. We previously found that numerous genetic lineages likely persisted below the level of detection by whole-population sequencing within these diverse communities (Chapter 3). The abundance of rare yet adaptive variation thus

hindered our ability to identify adaptive targets using a criterion of strict genetic convergence.

DISCUSSION

Here, we report that hypermutation in long-term evolved *P. aeruginosa* populations repeatedly fixed under biofilm, but not planktonic, conditions (Figure 1, Table C.1). Several hypotheses to explain the success of hypermutation have been proposed. Recent evidence suggests that hypermutation itself may confer a fitness benefit in response to endogenous oxidative stress (Torres-Barceló et al., 2013). Although additional work will be required, preliminary results cause us to disfavor this explanation. Without a direct fitness, hypermutators must overcome the large mutation supply rate imparted on non-mutator by large population sizes to successfully invade these populations (Drake 1991). The co-occurrence of mutations enabling social cheating or negative frequency dependency commonly could enable hypermutator genotypes to overcome this disadvantage when rare (Lujan et al., 2011). However, we do not observe any of the commonly observed social cheat genotypes, nor do we find that artificially created social cheat genotypes can invade more readily (preliminary data not shown, we plan to improve this analysis further). Lastly, hypermutation may enable the rapid accumulation of multiple beneficial alleles on a single background to overcome interference competition among competing lineages and limits of small N_e within subpopulations. Of these three hypotheses, our evidence best fits this one.

With no means to combine co-occurring beneficial mutations (Vos & Didelot, 2009), periodic selection favors genotypes that accumulate multiple beneficial mutations within adapting populations of asexual organisms (Lang et al., 2013). Likewise, we

observe similar patterns within the population genetic structure of these diverse hypermutator biofilm communities (Figure 2); lineages appear to succeed globally through rapidly accumulating many mutations (Table 1). We present evidence that high rates of diversification and cladogenesis inherent to these adapting biofilm populations may limit N_e and increase competition between co-occurring lineages enough to enable hypermutator lineages to invade.

Several studies have suggested that adaptation to the biofilm lifestyle may be predictable, proceeding through a few mutations in a small number of mutational targets (Junop et al., 2003, Traverse et al., 2014). In contrast, we find convergence at the gene or operon level to be the exception, not the rule (Tables C.1-3). The high level of competition among long-lived adaptive lineages may explain this discrepancy. First, clonal interference may create less predictable adaptive outcomes. Co-occurring beneficial variation limits the relative benefit of any one mutation, lengthens fixation times, and may favor mutations of large benefit (de Visser & Rozen, 2006). In support of this explanation, dN/dS declines through time suggesting that these large benefit mutations may become limiting as adaptation proceeds (Figure 3). Second, the recurrent evolution of new persistent lineages may further increase the amount of clonal interference and obscure any signal of convergence through reducing N_e (Figure 2, Table 4). The combination of small N_e and high mutation rate increases genetic variance and likely also explains the lack of mutational convergence.

Although convergence at the gene or operon level was limited, we did identify convergence for mutations in traits or functional groups implicated in adaptation seen during chronic lung infections. For example, rugose small colony variants (RSCVs)

rapidly evolve in many biofilm adaptation models and often associated with motility and attachment mutations (Gupta et al., 2014). Interestingly, enrichment for mutations associated with RSCs were only observed with isolate sequencing (Table 5, Chapter 3). If medically important genetic variation tends to associate with rare, specialist lineages, the observed combination of clonal interference and small N_e may explain the large amount of variation seen amongst CF patients (Ashish et al., 2013).

The genomic environment of the PA14 genome may also inherently favor hypermutation. Although these findings implicate something inherent to the nature of selection experienced during adaptation to a structured environment (Figure 1, Table C.1), previous research in our laboratory utilizing the exact same selection regime with another potential cystic fibrosis pathogen, *Burkholderia cenocepacia* HI2424, did detect hypermutation over 1050 generations of adaptation to the same selection regime (Traverse et al., 2013). The trade-offs associated with the mechanisms of ecological separation among lineages in PA14 biofilms may be more extreme, may further subdivide the population, and result in a larger reduction of N_e (Conibear et al., 2009). In support of this explanation, *Burkholderia* generalist genotypes were observed to maintain a large percentage of the total population size (~70%), whereas most evolved PA14 lineages persisted at levels below 10% (Chapter 3). However, key experimental differences between the EME studies may also explain these patterns. For example, this study utilized different growth medium with additional supplements that may enable greater genetic diversity by alternative metabolic strategies. In fact, we now know that these supplements are very important for the evolutionary success of the *ppk* SNP observed in the replicate B1 population (Chapter 5).

Although more direct tests are needed to eliminate the possibility of a direct fitness benefit of hypermutation or social cheating, hypermutators should not succeed due to increased clonal interference and accumulation of deleterious mutations. Our findings suggest that the spatially complex environment of biofilms alleviates these shortcomings and creates ideal conditions for hypermutators to succeed. It is also worth evaluating the generality of these findings among different strains of *P. aeruginosa* and *B. cenocepacia*, which can more directly test the roles of the genomic history of these organisms and, by extension, the distribution of mutational effects on mutator evolution. Perhaps most importantly, this study demonstrates that the evolution of mutator phenotypes in biofilm-forming organisms need not be the result of a particular external stress like antibiotic selection, but rather may evolve as the side effect of strong competition among neighboring lineages that become starved for the next beneficial mutation required to prevail.

MATERIALS AND METHODS

Genome sequencing and mutation identification among experimentally evolved PA14 populations. Replicate biofilm populations of *Pseudomonas aeruginosa* PA14 were previously adapted during a EME study under conditions described previously (Chapter 2). Genome sequencing was carried out in four experimental blocks described previously (Chapter 3). The BRESEQ software package was used to align Illumina reads to the *Pseudomonas aeruginosa* PA14 genome, and also called polymorphisms for individual isolates and mixed populations (Deatherage & Barrick, 2014). In total, we identified 1,183 mutations among 40 isolates (Table B.1) and defined 397 high-confidence mutational trajectories across all three replicate populations (Table B.1, 3).

Calculations based on mutation rate. To estimate the number of mutations associated with a genetic sweep or sway in each replicate biofilm population, we assumed a mutation rate of 0.144 mutations per genome per generation. Based on the reported rate of 0.18 mutations per genome per generation for a hypermutator *P. aeruginosa* PAO1 strain (ref), we extrapolated the mutation rate for our PA14 laboratory hypermutator strain with the observed 80.1% loss in MMR activity (Chapter 2). To estimate the number of generations distinct genetic lineages experienced throughout the duration of the EME study, the observed number of mutational calls within individual clones isolated after 540 generations was used to generate an expected value (expected number of generations = observed mutational count / mutation rate).

Estimates of effective population size among experimental conditions. Under planktonic conditions, replicate populations were sub-cultured daily with a 1:100 dilution factor while biofilm conditions relied on the daily transfer of a 7mm polystyrene bead (Poltak and Cooper, 2010). To estimate the size of the population bottleneck experienced under both selection regimes, we measured the cellular yield or CFUs/mL of replicate cultures reconstituted from both biofilm-evolved and planktonic-evolved 540 generation frozen stocks in 5mL of M63 growth medium. The bottleneck was calculated as the natural log transformed ratio of the total cellular yield of an entire 24 hour culture to the daily inoculum size, which was either 50 μ L of planktonic cells or total cells attached to a bead. Ne was extrapolated from this value as the product of the bottleneck size and cellular yield of replicate cultures. Cells were harvested off beads and enumerated as described previously (Chapter 2).

Assessing convergence of adaptive targets. To identify adaptive convergence, the observed number of mutational hits in a particular genomic target was compared to an expected value based on the target size in basepairs and total number of mutational calls observed. Adaptive targets of interest were based on single genes, operons and higher levels of cellular organization spanning multiple regions of the genome. For functional classes, each gene in the PA14 genome was grouped into one of 26 functional classes based on PseudoCAP Functional Class annotation information defined at <http://www.pseudomonas.com> (Winsor et al., 2011). We also defined additional classes called operational units that represent collections of genes with a common function but tended to be more specific. For example, the PseudoCAP annotation information for DNA mismatch repair genes places them into the much larger DNA replication, recombination, modification and repair class. However, selection can likely act on subsets of these larger groups of genes. As such, defining these additional operational units allowed us to define targets of selection at intermediate levels of cellular organization.

Given high rates of diversification, we hypothesized that individual isolates or groups of isolates may experience selection differently. For example, biofilm specialists may experience more oxidative stress inherent to anaerobic conditions and, by extension, stronger selection for redox chemistry. To examine this possibility, we enumerated the number of mutational hits occurring in each of PseudoCAP functional class for each individual isolate and generated an Euclidean distance matrix using the statistical package in R (<http://www.r-project.org>) (see refs). Next, we compared the functional dissimilarity among isolates to identify adaptive targets only enriched in

individual clones or groups of isolates. The functional and genetic relationship between isolates was then visualized using Ward's minimum variance method. We also assessed the genetic dissimilarity among isolates using a binary distance matrix using the dist.clust function of the ade4 package (<http://pbil.univ-lyon1.fr/ADE-4/home.php?lang=eng>). Since we defined genetic similarity on the presence or absence of mutational changes in an isolate's genome, we used this comparison to distinguish between patterns of enrichment caused by selection for functional traits and other possibilities such as variation in adaptive history.

Hierachical analysis of the variance was performed using the adonis function of the vegan package (<http://CRAN.R-project.org/package=vegan>) to estimate the variance contributing to observed relationships between isolates. P values are based on the twenty permutations within the adonis function. We defined genetic dissimilarity between isolates as the presence of , a binary distance matrix was calculated using the dist.clust function of the ade4 package (<http://pbil.univ-lyon1.fr/ADE-4/home.php?lang=eng>) based on the presence or absence of each mutation.

CHAPTER V

BIOFILM ADAPTATION THROUGH MODIFYING THE TIMING OF DISPERSAL
THROUGH FINE-TUNING OF POLYPHOSPHATE PRODUCTION

INTRODUCTION

Microbes commonly live in close quarters with one another in spatial structures known as biofilms associated with both abiotic and biotic surfaces. This complex spatial structure of biofilms enables persistence by providing a variety of benefits (Besmer et al., 2012; Parsek & Singh, 2003; Periasamy & Kolenbrander, 2009; Yildiz & Visick, 2009). The biodiversity harbored within biofilms allows communities to rapidly respond to changing conditions. Furthermore, biofilm-associated cells often exhibit enhanced resistance to numerous stressors such as antimicrobials and the host-immune response (Folkesson et al., 2012; Gupta, Liao, Petrova, Cherny, & Sauer, 2014; Soto, 2013). The co-occurrence of a diverse collection of microbial species can also facilitate genetic exchange of niche-specific genes and subsequent adaptation (Molin & Tolker-Nielsen, 2003). Despite the inherent benefits, the decision to make the lifestyle switch from a free-living cell to a surface-associated member of a diverse community does not come lightly.

The biofilm lifestyle exposes participants to endogenous oxidative stress, nutrient deprivation and strong competition for available resources. As a result, this 'swim-or-stick' decision is tightly regulated process involving a complex network of signal transduction and regulation pathways in response to environmental cues (Belas, 2014; Fazli et al., 2014). However, deciphering associated pathways involved in this decision has been challenging. For example, although the secondary messenger molecule c-di-GMP is a known key player in this decision for many bacterial species, microbial genomes harbor upwards of 50 proteins with active domains predicted to interact with or bind c-di-GMP (Hengge, 2009, Boyd & O'Toole, 2012). Understanding of how these

networks evolve and how bacteria can modify these networks in response to selection remains unclear but nevertheless paramount for our understanding.

Here, we characterize a non-synonymous substitution in the polyphosphate kinase (*ppk*) gene that occurs in one experimentally evolved biofilm population of *Pseudomonas aeruginosa* PA14 and eventually fixes as part of a mutator haplotype (Chapter 3). Although the experimental selection method favored a regular cycle of surface attachment, biofilm formation, and dispersal (Poltak & Cooper, 2011), mutants containing this *ppk* substitution outcompeted the parental PA14 strain in the absence of biofilm selection. We find that this mutant is not more fit due to an intrinsically faster growth rate but rather through fine-tuning of the precise moment when growth is the fastest, subsequently enabling more timely surface colonization. This study defines a molecular mechanism whereby cells can modify the 'stick-or-swim' decision via polyphosphate production and subsequent activation of the Pho regulatory network responding to free phosphate. Furthermore, our results demonstrate the utility of experimental evolution to understand the new ecological opportunities that drive adaptation in structured biofilm environments.

RESULTS

The ppkT443A mutation is one of eight mutations to fix in conjunction with hypermutation. The first population-genetic sweep in the replicate B1 evolved biofilm population involved nine mutations associated with global success of this lineage (Figure 1a). Previous work demonstrated that numerous adaptive lineages co-occurred within each population (Chapter 3) and that successful genetic lineages required an excess of adaptive mutations to be successful (Chapter 4). Consequently, we sought to

examine the potential fitness benefit and genetic consequence of the seven mutations that were linked to a loss-of-function SNP in *mutS* that elevated the mutation rate. An excess of mutations affecting genes involved in nucleotide biosynthesis and metabolism throughout the experiment (Chapter 3, Table 3) lead us to examine the role of the *ppk* T443A mutation in more detail. The single, engineered *ppk*T443A mutation on the PA14 background revealed that this mutation alone confers a significant fitness benefit under the selective conditions experienced during the original EME study. We reasoned this benefit could be the result of adaptation to the biofilm lifestyle or some aspect of our growth conditions. However, we found that this fitness benefit was not biofilm-dependent; the *ppk* mutant was able to exclude the ancestor in M63 media with or without a bead for attachment (Figure 1b). Additionally, we found that this fitness benefit was lost during direct competition in LB (Figure 1b) or KT10 π (data not shown) suggesting *ppk*T443A represents an adaptation to our growth medium.

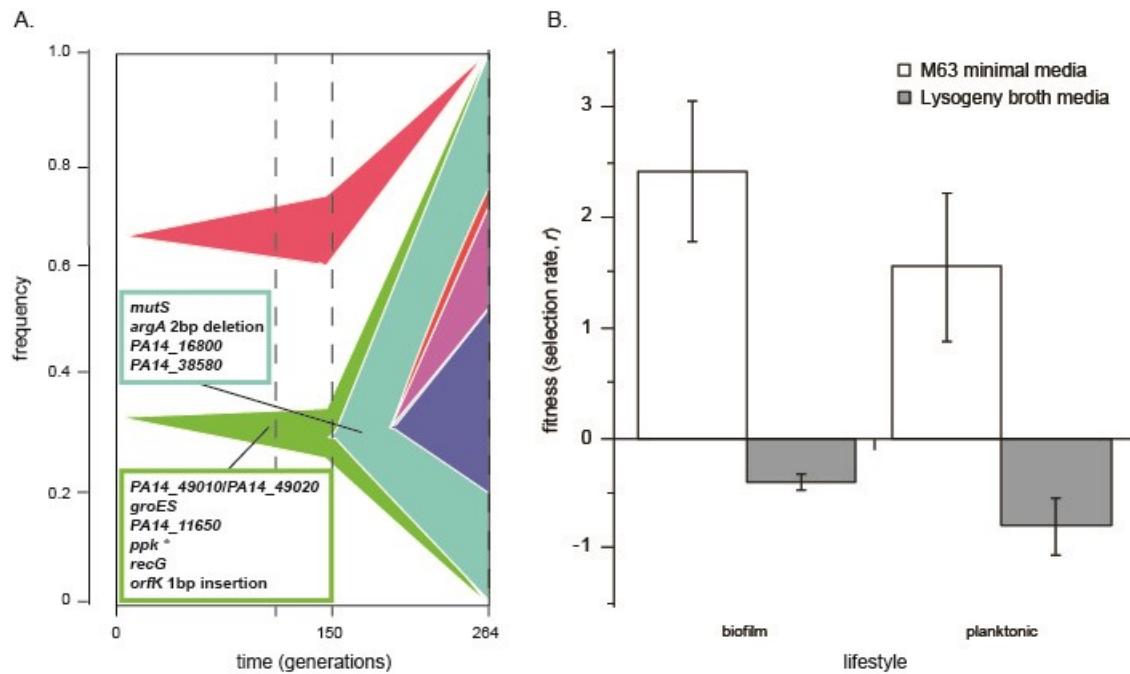


Figure 1. The *ppk* T443A substitution is a beneficial mutation that fixes in conjunction with hypermutation.

A. The *ppk* T443A mutation is one of nine mutations that gradually increase in allele frequency before a non-synonymous substitution in *mutS* occurs and fixes by 264 generations. B. Direct fitness competitions of *ppk* T443A against the wild-type PA14 genotype in minimal media (M63, selective conditions, white bars) or rich media (LB, grey bars) with or without a bead surface for attachment (biofilm versus planktonic). Fitness is expressed as the difference in absolute fitness, r . Error bars represent 95% confidence intervals of the means.

The evolutionary advantage of the *ppk* T443A mutant depends on the selective conditions. We hypothesized that the evolutionary benefit of the *ppk* mutation represents an adaptation to some component of our selective media, M63 minimal media. Several components of our selective media could be responsible: the phosphate level (60mM), galactose (40mM) and arginine (23mM). To isolate the effect of available phosphate and galactose on the fitness effect of the *ppk* mutant, we assayed fitness of the *ppk* mutant in KT10 π growth medium supplemented with equivalent levels of phosphate and galactose found in M63 separately and in combination. Although intracellular phosphate is the substrate for the Ppk enzyme, we found that the fitness

advantage of the mutant was greatest in the presence of galactose, relies more on the additional sugar not present in the other media tested (Figure 2a). In particular, the *ppk* mutant was the most fit when phosphate was limiting while galactose was available.

Since other *Pseudomonas* spp. preferentially utilize amino acids over sugars through catabolite repression (Moreno, Martínez-Gomariz, Yuste, Gil, & Rojo, 2009), we predict this adaptation alters the influence of catabolite repression to better utilize the available galactose. We tested this prediction by comparing growth curves of the *ppk* mutant and the WT PA14 strain in M63 media containing both sugar and amino acids. In fact, the *ppk* mutant exhibited a slower rate of growth (V_{max} : *ppk* T443A, 0.071 ± 0.006 , compared to wild-type, 0.084 ± 0.006 , $t_{14} = 3.363$, $P = 0.0046$) and displayed a delayed growth pattern in which maximum growth was achieved much later (Figure 2b).

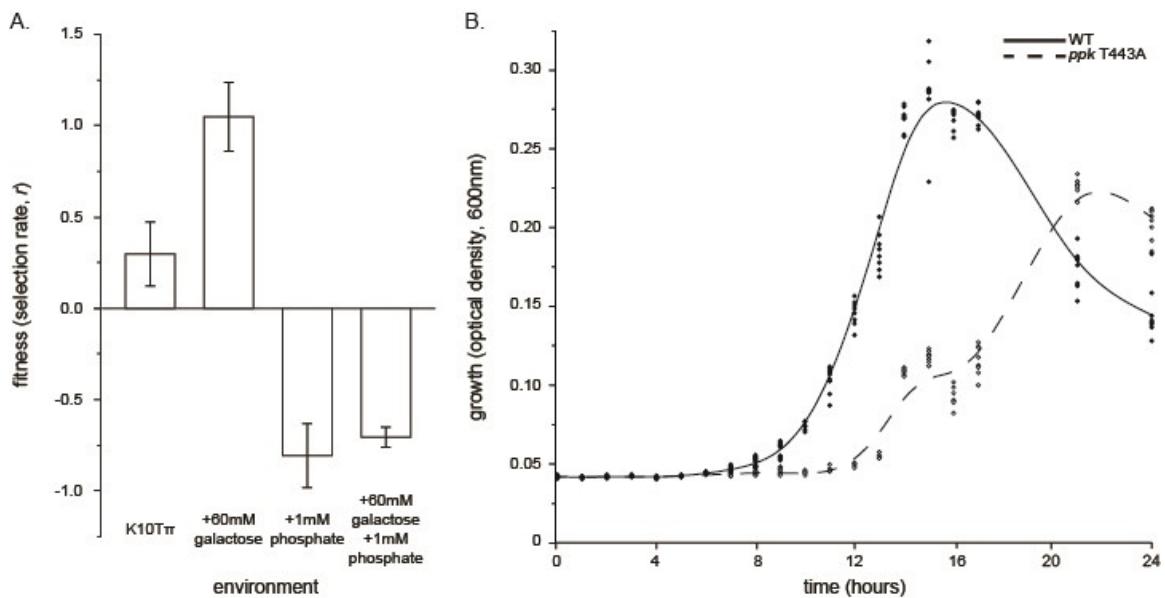


Figure 2. The fitness effect of *ppk* T443A mutant depends on growth conditions. A. The selective advantage of the *ppk* T443A substitution examined in K10T π minimal media optionally supplemented with 60mM Galactose, 1mM phosphate or both. B. Growth curve dynamics of wild-type PA14 (solid line, filled circles) compared to *ppk*T443A (dotted line, open circles) in 5mL M63 minimal media used during the original EME experiment.

ppkT443A enhances enzyme activity. To test whether the observed patterns directly resulted from altered Ppk enzymatic activity, we quantified PolyP production. Since the Ppk enzyme controls the conversion of intracellular phosphate (Pi) into PolyP and the polyphosphatase (Ppx) enzyme degrades PolyP into organic phosphate, we reasoned that phenotypic comparison of mutants lacking Ppk or Ppx would reveal if the *ppk* T443A mutation represented a gain-of-function or loss-of-function mutation. This expectation is based on previous work with PolyP in *P. fluorescens* demonstrating that activation of the Pho regulon when Pi is limiting and results in biofilm dispersal through *rapA*-mediated degradation of c-di-GMP (Monds et al., 2006).

Following this reasoning, we hypothesized that mutants favoring the conversion of Pi into PolyP (Δppx) will generally exhibit less biofilm formation compared to wildtype

PA14 under phosphate-sufficient conditions. That is, reduced conversion of PolyP back into Pi or enhanced enzymatic activity of the Ppk enzyme would favor biofilm dispersal by reducing intracellular Pi levels. To test this hypothesis we compared the biofilm formation of the *ppk* mutant to wildtype, Δppk and Δppk under variable phosphate availability: K10T π media supplemented with 0, 1, 15 or 60mM phosphate. Consistent with our hypothesis, we find that the *ppk* mutant is more similar phenotypically to Δppk mutants lacking the Ppx enzyme (Figure 3). Overall, both *ppk*T443A and Δppk formed less biofilm than wildtype under all phosphate sufficient conditions (Figure 3a) and also displayed reduced colony size (Figure 3b). As a result, we expect the *ppk* mutant is experiencing some form of PolyP pooling similar to the Δppk mutant. Curiously, we also find that the observed biofilm phenotypes are even more exaggerated in the *ppk* mutant than mutants lacking Ppx entirely (biofilm formation: *ppk* T443A, 0.620 ± 0.082 , compared to Δppk , 1.161 ± 0.182 , $t_{14} = 6.392$, $P < 0.0001$). Lastly, *ppk* T443A mutant cells produced slightly higher, but not significant, levels of PolyP on average compared to wildtype PA14 cells (Figure 3c, relative fluorescent units (r.f.u.) *ppk*, 426.38 ± 223.48 , compared to wildtype, 277.38 ± 166.79 , $t_3 = -1.70$, $P = 0.144$).

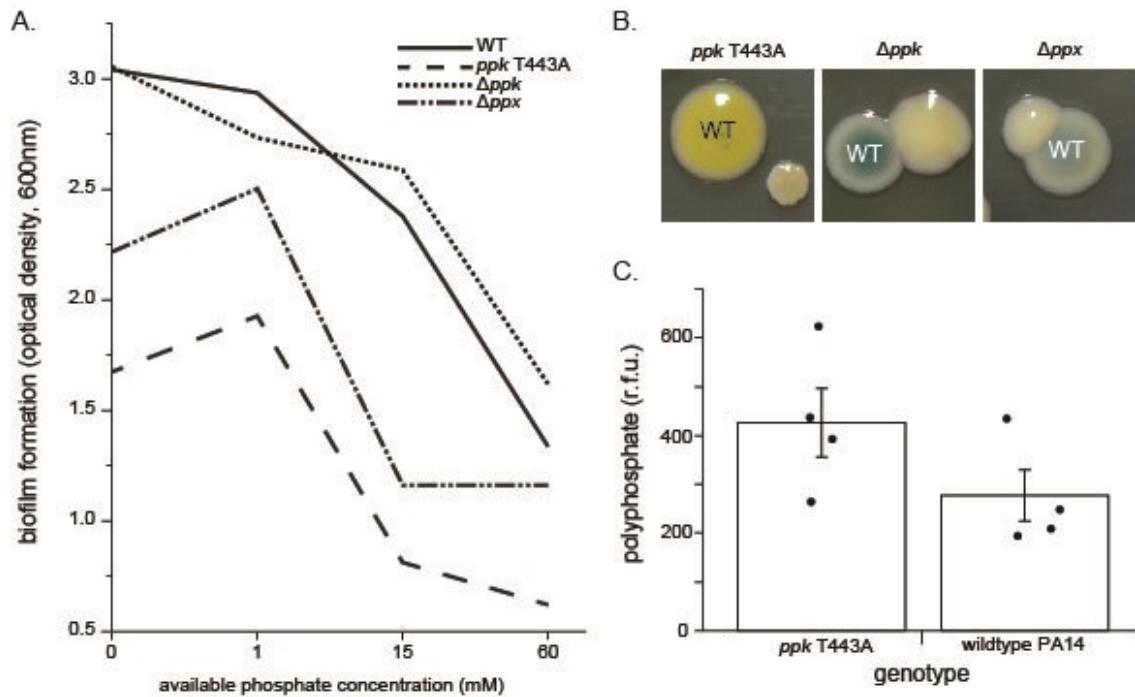


Figure 3. *T443A ppk* mutants are phenotypically consistent with PolyP pooling.
 A. Biofilm formation of *ppk*-related strains in K10T π media supplemented with 0, 1, 15 or 60mM phosphate. B. Observed colony phenotypes of *ppk*-related mutants in co-culture with the ancestral strain. Colonies labeled 'WT' are founded by the PA14 LacZ+ genotype. C. Polyphosphate extractions from batch cultures of T443A *ppk* mutants and wild-type PA14 cells grown in M63 minimal media. Values are relative fluorescent units (r.f.u.) and error bars represent standard error based on four-fold replication but were not significant based on a two-sample t-test.

The effect of the observed threonine to alanine amino acid change at position 443 of the protein was examined in several ways. First, we took advantage of the extensive research performed on this enzyme in *E. coli* and known crystallized protein structure (Zhu et al., 2005). In *E. coli*, the *ppk* mutation corresponds to the 397 amino acid position with no obvious connection to the active site. Second, we generated protein structures with or without *ppk*T443A using *in silico* methods to test for any potential effects to protein structure overall. Despite occurring distally from the active site, this analysis suggests that the introduction of the hydrophobic side chain accompanying the threonine to alanine substitution of *ppk*T443A may, in fact, cause a

conformational change within the Ppk active site around the 633 amino acid position (Figure 4).

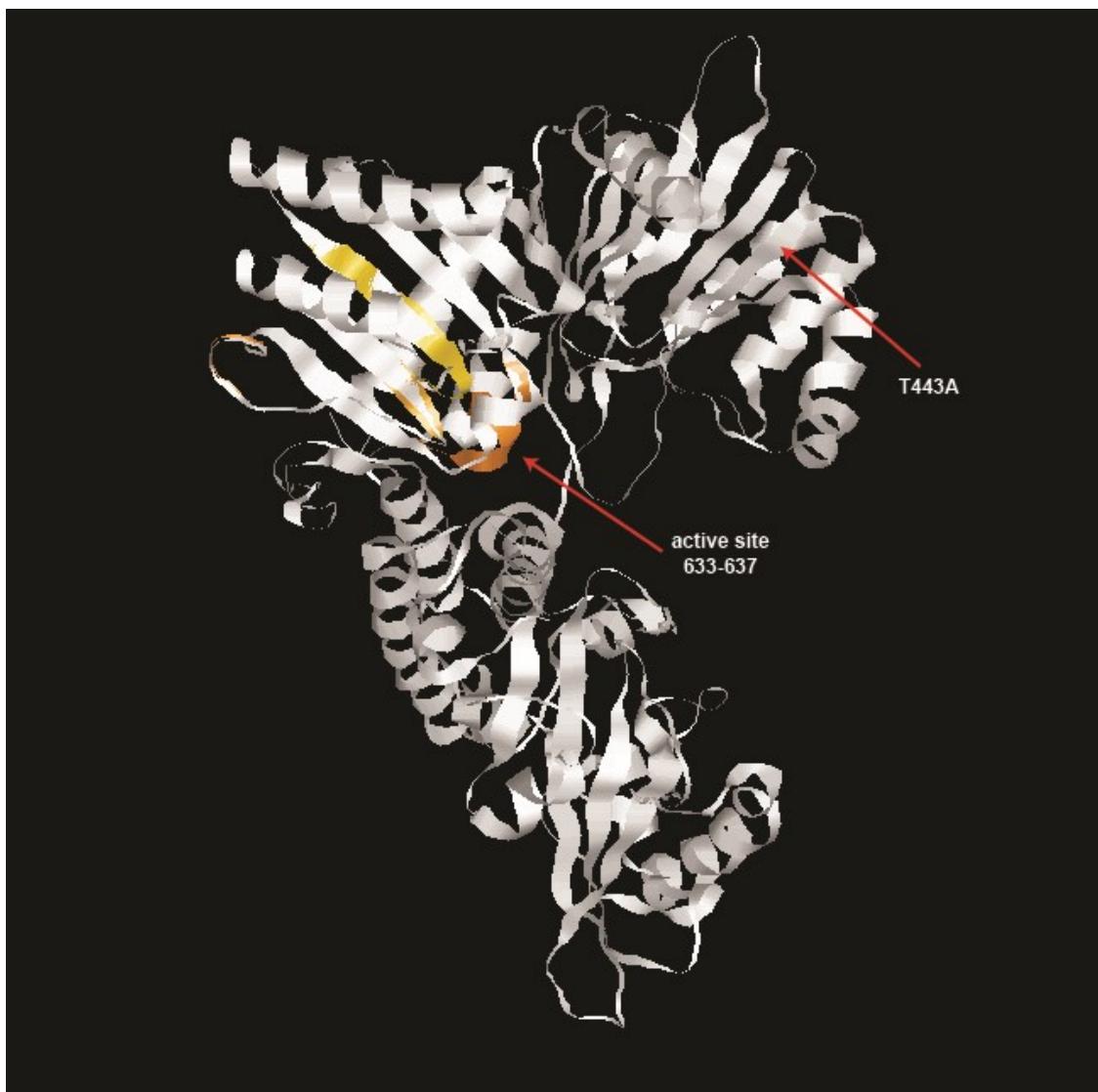


Figure 4. The T443A substitution alters the protein conformation of the Ppk active site.

The effect of the T443A substitution on predicted protein structure examined *in silico* generated using Phyre2 and superimposed using FATCAT. Arrows indicate the position of the active site present at positions 633 to 637 in relation to the location where the observed mutation occurred.

DISCUSSION

PolyP is a highly conserved polymer implicated in numerous aspects of microbial physiology such as metabolism, stress protection and even virulence (Brown & Kornberg, 2008). Previous work with *P. fluorescens* has demonstrated the importance of fine-tuning of PolyP production for organismal fitness in soil where complete loss of either Ppk or Ppx was detrimental (Silby, Nicoll, & Levy, 2009b). As a result, PolyP production is a tightly regulated process but the exact mechanism by which PolyP influences organismal fitness in soil remains unclear (Silby, Nicoll, & Levy, 2012). Here, we describe a gain-of-function substitution in the *ppk* gene that increases fitness of *P. aeruginosa* PA14 during adaptation to a regular cycle of surface attachment, biofilm formation, and dispersal. Although the T443A *ppk* mutation swept to fixation with a variety of additional substitutions (Figure 1a), we found that this mutation alone confers a fitness benefit dependent on our growth medium (Figure 1b). More specifically, we found that the fitness benefit of the T443A *ppk* mutation could only be recreated in different growth medium with the addition of both galactose and phosphate (Figure 2a). With an over abundance of galactose in the media, modified PolyP production may allow for better utilization of this available sugar that is not a preferred substrate for *Pseudomonas* (Moreno, Martínez-Gomariz, Yuste, Gil, & Rojo, 2009).

This study also highlights the utility of experimental evolution as a nuanced mutant screen to elucidate the intricacies of regulatory networks involved in the 'stick-or-swim' decision. Gain-of-function mutations are presumably rare in comparison to the wide array of mutational possibilities (for example, (Segrè, Murray, & Leu, 2006)). Despite no obvious connection to the active site of the Ppk enzyme (Zhu, Huang, Lee, &

Xu, 2005), our findings suggest this mutation may represent one of these rare gain-of-function mutations (Figure 3, 4). Future work verifying these findings and further examining the connection between PolyP, catabolite repression and the Pho regulon will be possible thanks to this *de novo* occurrence of this substitution.

The mechanism by which bacteria differentially metabolize substrates in a mixed-resource environment is often catabolite repression, which inhibits some enzymes to enable more rapid metabolism on preferred carbon sources. This regulatory system in Pseudomonads is largely controlled through the Catabolite Repression Control (Crc) protein, however, molecular mechanisms involved in its regulatory activity remains poorly understood (Morales et al., 2004; Moreno et al., 2009). Nonetheless, we offer several lines of evidence suggesting a possible connection between PolyP and catabolite repression. First, *Streptomyces* has been shown to utilize PolyP glucokinases, enzymes uses PolyP rather than ATP for the phosphorylation of sugar before entry into the Entner– Doudoroff pathway for glucose metabolism (Koide, Miyanaga, Kudo, & Eguchi, 2013). Although polyphosphate glucokinases have not been identified in Pseudomonads, we speculate PolyP-pooling could impact the activity of such proteins or related proteins if present within the PA14 genome. Lastly, Crc-mediated repression of sugar catabolism may be indirectly connected to the sensing of environmental phosphate levels. The sensing of available phosphate is largely achieved through the phosphotransferase system (PTS), which couples transport and phosphorylation of sugars (Monds et al., 2006). Activation of the PTS system coincides with activation of the alternative RNA polymerase, RpoN, which removes catabolite

repression through Crc (Rojo, 2010). A strong connection between sugar catabolism and efficient phosphate utilization is conceivable.

We propose a model in which the T443A *ppk* mutation stimulates expression of the Pho regulon through depleted intracellular levels of organic phosphate to remove catabolite repression (Figure 5). Although the majority of work characterizing the Pho regulon has been done in *P. fluorescens*, previous studies have demonstrated that PA14 has a functional *rapA* homolog, a well characterized phosphodiesterase or PDE, and functions similarly under the control of Pho regulon to effectively reduce the intracellular levels of c-di-GMP (George O'Toole, personal communication). phosphodiesterase (PDE) *bifA*, effectively reducing the intracellular levels of c-di-GMP. Consequently, our model builds upon the wealth of knowledge with *P. fluorescens*, as follows. Under phosphate-limiting conditions, the PstS system causes PhoR to phosphorylate PhoB, which in turn stimulates biofilm dispersal through RapA-mediated degradation of c-di-GMP, in turn, promoting biofilm dispersal. In support of this model, we find that T443A *ppk* mutant cells always display decreased levels of biofilm formation compared to wild-type cells, indicative of biofilm repression (Figure 3a). Although counterintuitive given that the experimental evolution selected for surface attachment, our results suggest that delayed attachment or enhanced biofilm dispersal may have been under stronger selection.

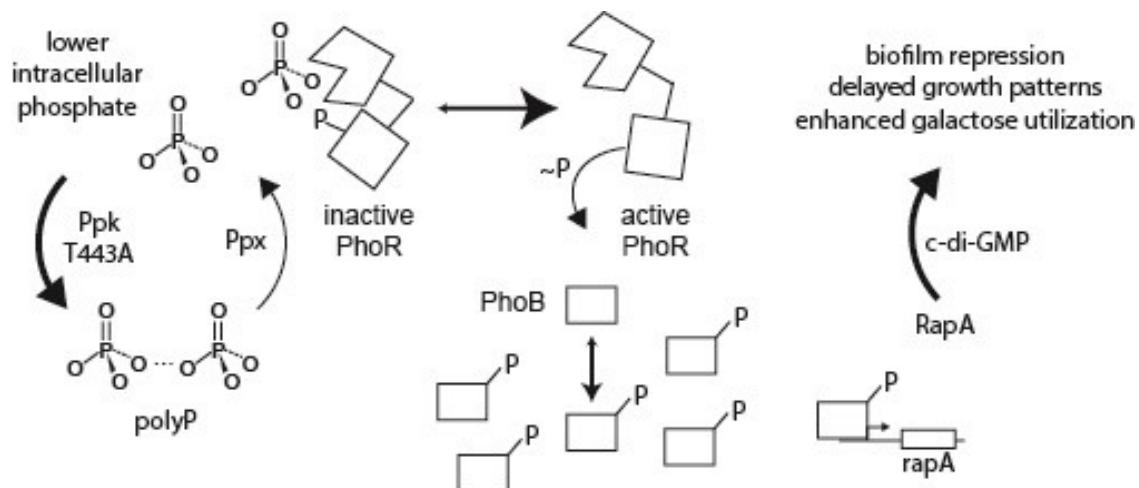


Figure 5. The proposed model of how enhanced activity of the polyphosphate kinase (Ppk) influences cellular physiology.

Enhanced activity of the Ppk enzyme stimulates the conversion of organic phosphate (PO_4) into PolyP ($(\text{PO}_4)_n$) depleting intracellular levels of PO_4 . The phosphotransferase system (Pts) senses intracellular levels of PO_4 resulting in the phosphorylation of PhoR and subsequently PhoB that activates the Pho regulon. Activation of the Pho regulon enhances biofilm dispersal via RapA activity and may indirectly influence catabolite repression.

In conclusion, this study highlights one way bacteria can modify the 'stick-or-swim' decision cascade during adaptation to the biofilm lifestyle. We hope that future work examining the connections between PolyP, catabolite repression, and the Pho regulon with EME, will elucidate more about the intricacies of secondary messenger molecules in the 'stick-or-swim' decision cascade and biofilm adaptation. Likewise, our results demonstrate the utility of EME to identify novel ecological strategies associated with adaptation in structured environments. Furthermore, we find that this work highlights the complexity of the 'stick-or-swim' decision and provides insight into why such complex regulatory mechanisms may have evolved to control it.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Replicate biofilm populations of *Pseudomonas aeruginosa* PA14 were adapted during an EME study under conditions

described previously (Flynn et al., unpublished). Selection for biofilm attachment to 7mm polystyrene beads was achieved through growth in M63 media (15mM (NH₄)₂SO₄, 22mM KH₂PO₄, 40mM K₂HPO₄, 1mM MgSO₄, 0.4% arginine, 25µM FeCl₂ 40mM Galactose). *P. aeruginosa* PA14 was also routinely grown at 37 °C using tryptic soy broth (TSoy), lysogeny broth (LB, 10g tryptone, 5g yeast extract, 5g NaCl), K10Tπ (50mM Tris-HCl, pH 7.4, 0.2% (wt/vol) Bacto tryptone, 0.15% (vol/vol) glycerol, 0.61mM MgSO₄), or in Vogel–Bonner minimal medium (VBMM; 0.2g MgSO₄•7H₂O, 2.0g citric acid, 3.5g NaNH₄HPO₄•4H₂O, 10g K₂HPO₄, pH 7.0) optionally supplemented with 1.5% Bacto agar for semi-solid media. K10Tπ was used to grow cells under various available Pi conditions. K10Tπ has been previously shown to contain 0.14mM Pi (Monds, Newell, Schwartzman, & O'Toole, 2006) and was periodically supplemented with 1, 15 or 60mM K₂HPO₄ yielding K10T-1, K10T-15, and K10T-60, respectively.

Assessment of growth dynamics through time. Growth dynamics of *ppkT443A* was compared with the ancestral strain in the selective media utilized during the original EME study. The optical density at 600nm of 100µL aliquots were measured every hour for the first 17 hours and after 21 and 24 hours of growth using a Tecan plate reader. Growth curve parameters such as Vmax were calculated based on the greatest change in optical density per hour.

Fitness assays. Head-to-head competition experiments were performed against a PA14 strain marked chromosomally at a neutral position with lacZ using a mini-Tn7 system (ref). First, competitors and the marked ancestral strain are grown overnight and allowed to acclimate to the competition conditions, separately. After 24 hours, cells are harvested from the beads utilizing methods described previously (See Chapter 2), and

introduced in mixture into 5mL of fresh media with a new polystyrene bead at a 1:100 dilution at the start of the competition. CFU/mL is determined for each competitor at the time of inoculation and off of the beads after 24 hours of growth by plating onto 1% tryptone supplemented with 80 μ g/mL X-gal. Fitness was calculated as the selection rate, r , defined as the difference in $\ln(\Delta\text{CFU}/\text{mL})$ over the course of the competition, which has units of inverse time. Since fitness is defined as the ability to attach to a bead, less fit competitors can produce fewer CFUs on a bead after 24 hours than added initially. Selection rate, rather than relative fitness, allows for fitness to be more accurately determined given the low fitness of the PA14 ancestor in some contexts (Travisano & Lenski 1996). For a more detailed explanation, see <http://myxo.css.msu.edu/ecoli/srvsrf.html>.

Construction of the *ppk* T443A, *ppk* and *ppx* deletion mutants. *P. aeruginosa* PA14 was grown at 37°C using TSoy, LB or VBMM during genetic manipulation. *Escherichia coli* strains were routinely cultured at 37 °C using TSoy or LB broth. Antibiotics were added as follows: for *E. coli*, gentamicin (Gm) at 5 μ g/ml to maintain plasmids, Gm at 10 μ g/ml for selection, and 15 μ g/mL naladixic acid for selection against donor strains and for *P. aeruginosa*, Gm at 30 μ g/ml for plasmid retention or Gm at 60 μ g/ml for selection.

The *ppk* allele was commercially synthesized (Genscript) and cloned into pMQ30 (X). Isogenic deletion mutants were generated using methods described elsewhere (Zhao et al., 2013). Briefly, PCR products were amplified using primers that targeted 500bp regions of the PA14 chromosome flanking the *ppk* and *ppx* genes and joined via splicing by overlapping extension (SOE) PCR. The upstream forward and downstream

reverse primers were used to generate deletion alleles that were tailed with attB1 or attB2 sequences as described in the Gateway Cloning Technology Manual (Invitrogen). To control against the occurrence of additional mutations throughout the cloning process, PCR products were always amplified with the Phusion high-fidelity polymerase (NEB) and gel extractions were performed with the addition of 1 mM Guanosine to protect DNA from the harmful effects of UV. Gateway BP Clonase II Enzyme Mix (Invitrogen) allowed for in vitro recombination between resulting PCR products and the pDONRPEX18Gm plasmid derived from pEX18Gm (Hoang, Karkhoff-Schweizer, Kutchma, & Schweizer, 1998). Plasmid constructs were confirmed to contain the mutation of interest without additional mutations with Sanger sequencing using the M13 universal primers. Next, plasmids were introduced into Z-competent DH5 α *E. coli* cells (Zymo Research) for plasmid maintenance and storage. Plasmid constructs were harvested overnight cultures of the resulting DH5 α using the PureYield Plasmid Miniprep System (Promega) and electroporated into S17.1 *E. coli* cells. Next, biparental matings with *E. coli* S17.1 and PA14 to introduce the new plasmid and select for transformants using 60ug/mL gentamicin. Electroporations and biparental mating protocols were performed by adapting standard methods surrounding the use of the mini-Tn7 system with *Pseudomonas* (Choi, Kumar, & Schweizer, 2006). Finally, mutants were confirmed to contain the deletion alleles using Sanger sequencing on a collection of eight to sixteen potential mutants.

Biofilm production assay. Biofilm production was quantified as described previously (O'Toole & Kolter 1998) after four, eight, and twenty-four hours. To assess if the attachment of one type facilitated the attachment of another, one type was allowed to

establish a biofilm in wells of a 96-well plate for four hours before the addition of a second type, standardized using optical density. Biofilms were stained with 1.0% crystal violet (CV) and solubilized with 95% ethanol after a full 24 hours of growth.

Quantification of PolyPhosphate from *Pseudomonas aeruginosa* cells.

Polyphosphate (PolyP) was extracted from *P. aeruginosa* PA14 in two ways. PolyP quantification has been successfully performed for *P. fluorescens* Pf-01 and *P. aeruginosa* PAO1 utilizing the affinity of PolyP to glassmilk (Silby, Nicoll, & Levy, 2009a). Following DNase and RNase treatment, quantification of the amount of bound PolyP is achieved through measuring the reaction of this complex with toluidine blue O dye in 40mM acetic acid at optical densities of 530 and 630nm after incubation, and the amount of PolyP is then expressed as the 530/630 ratio. However, we found that the sensitivity of this assay was incompatible with the low levels of PolyP inherent to the PA14 background.

To increase the sensitivity of our PolyP measurements, we adapted a method relying on the binding of PolyP to 4'-6-Diamidino-2-Phenylindole (DAPI) to quantify PolyP levels both *in vitro* and *in vivo* for marine microbes (Kulakova et al., 2011). Briefly, *P. aeruginosa* cells were grown at 37 °C in 50mL batch cultures of M63 containing 15mM or 60mM phosphate. Cultures were split and pelleted before freezing at -20 °C overnight. After thawing, pellets were washed with 50mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and 200µL was side aside for *in vivo* quantification before the remainder of the suspension was lysed with 100µL H₂SO₄ for 5 minutes at room temperature. Lysis mixtures were neutralized with the addition of 2M NaOH and 200µL of 1M Tris, pH 7.0, containing 0.6% neutral red. Neutral red was used

for quality control to ensure a proper pH and optimal PolyP binding; neutralized mixtures should be an orange-red color while pink (acidic) or yellow (basic) are problematic.

Samples were pelleted and treated with 600 μ L of 6M NaI before transferring the supernatant to EconoSpin spin columns for DNA (Epoch Life Sciences, Cat No. 1910-050). The columns were washed twice with buffer (10mM Tris, pH 7.5, 50% EtOH, 1mM EDTA, 100mM NaCl) and before transferring the columns to a fresh 1.5mL centrifuge tube. PolyP was eluted by washing 50 μ L of water over the columns twice to increase yield and diluted 1:100 in fresh tubes. Since microbes tend to utilize long-chains of PolyP (Kornberg, Rao, & Ault-Riché, 1999), 100 μ L of these samples were combined with 200 μ L of DAPI mixtures tailored to quantify long polymer chains of PolyP (150mM KCl, 20mM HEPES-KOH, pH 7.0, 10 μ M DAPI). PolyP binding was quantified by excitation at 420nm and measured at 550nm. Additionally, PolyP levels were quantified *in vivo* by mixing 100 μ L of 1:1,000 cells prior to lysis to 200 μ L of the same DAPI mixture described above and quantified in the same way. Values were expressed as relative fluorescent units (r.f.u.) calculated based on a standard curve generated using type 45 phosphate glass obtained from Sigma-Aldrich at 0.5, 1, 2.5, 5, 7.5, 10, and 50 μ g/mL. The molecular weight was calculated as described previously (Diaz & Ingall, 2010).

LIST OF REFERENCES

- Ashish, A., Paterson, S., Mowat, E., Fothergill, J. L., Walshaw, M. J., & Winstanley, C. (2013). Extensive diversification is a common feature of *Pseudomonas aeruginosa* populations during respiratory infections in cystic fibrosis. *Journal of Cystic Fibrosis : Official Journal of the European Cystic Fibrosis Society*, 12(6), 790–3. doi:10.1016/j.jcf.2013.04.003
- Baer, C. F., Miyamoto, M. M., & Denver, D. R. (2007). Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Reviews. Genetics*, 8(8), 619–31. doi:10.1038/nrg2158
- Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., ... Kim, J. F. (2009). Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*, 461(7268), 1243–7. doi:10.1038/nature08480
- Belas, R. (2014). Biofilms, flagella, and mechanosensing of surfaces by bacteria. *Trends in Microbiology*, 22(9), 517–527. doi:10.1016/j.tim.2014.05.002
- Bell, M. A. (1995). Intraspecific Systematics of *Gasterosteus Aculeatus* Populations: Implications for Behavioral Ecology. *Behaviour*. doi:10.1163/156853995X00496
- Besemer, K., Peter, H., Logue, J. B., Langenheder, S., Lindstrom, E. S., Tranvik, L. J., & Battin, T. J. (2012). Unraveling assembly of stream biofilm communities. *ISME J.* doi:ismej2011205 [pii] 10.1038/ismej.2011.205
- Blount, Z. D., Barrick, J. E., Davidson, C. J., & Lenski, R. E. (2012). Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature*. doi:nature11514 [pii] 10.1038/nature11514
- Boyd, C. D., & O'Toole, G. A. (2012). Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. *Annual Review of Cell and Developmental Biology*, 28, 439–62. doi:10.1146/annurev-cellbio-101011-155705
- Bradic, M., Teotónio, H., & Borowsky, R. L. (2013). The population genomics of repeated evolution in the blind cavefish *Astyanax mexicanus*. *Molecular Biology and Evolution*, 30(11), 2383–400. doi:10.1093/molbev/mst136
- Brown, M. R. W., & Kornberg, A. (2008). The long and short of it - polyphosphate, PPK and bacterial survival. *Trends in Biochemical Sciences*, 33(6), 284–90. doi:10.1016/j.tibs.2008.04.005

Choi, K. H., Kumar, A., & Schweizer, H. P. (2006). A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods*, 64(3), 391–397. doi:S0167-7012(05)00158-2 [pii] 10.1016/j.mimet.2005.06.001

Colvin, K. M., Gordon, V. D., Murakami, K., Borlee, B. R., Wozniak, D. J., Wong, G. C., & Parsek, M. R. (2011). The pel polysaccharide can serve a structural and protective role in the biofilm matrix of *Pseudomonas aeruginosa*. *PLoS Pathog*, 7(1), e1001264. doi:10.1371/journal.ppat.1001264

Conibear, T. C. R., Collins, S. L., & Webb, J. S. (2009). Role of mutation in *Pseudomonas aeruginosa* biofilm development. *PLoS One*, 4(7), e6289. doi:10.1371/journal.pone.0006289

Cooper, V. S. (2014). The origins of specialization: insights from bacteria held 25 years in captivity. *PLoS Biology*, 12(2), e1001790. doi:10.1371/journal.pbio.1001790

Couce, A., Guelfo, J. R., & Blázquez, J. (2013). Mutational spectrum drives the rise of mutator bacteria. *PLoS Genetics*, 9(1), e1003167. doi:10.1371/journal.pgen.1003167

De Visser, J. A., & Rozen, D. E. (2006). Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. *Genetics*, 172(4), 2093–2100. doi:genetics.105.052373 [pii] 10.1534/genetics.105.052373

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology (Clifton, N.J.)*, 1151, 165–88. doi:10.1007/978-1-4939-0554-6_12

Dettman, J. R., Rodrigue, N., Melnyk, A. H., Wong, A., Bailey, S. F., & Kassen, R. (2012). Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular Ecology*, 21(9), 2058–77. doi:10.1111/j.1365-294X.2012.05484.x

Diaz, J. M., & Ingall, E. D. (2010). Fluorometric quantification of natural inorganic polyphosphate. *Environmental Science & Technology*, 44(12), 4665–4671. doi:10.1021/es100191h

Dietrich, L. E. P., Okegbe, C., Price-Whelan, A., Sakhtah, H., Hunter, R. C., & Newman, D. K. (2013). Bacterial community morphogenesis is intimately linked to the intracellular redox state. *Journal of Bacteriology*, (January). doi:10.1128/JB.02273-12

Dinamarca, M. A., Aranda-Olmedo, I., Puyet, A., & Rojo, F. (2003). Expression of the *Pseudomonas putida* OCT plasmid alkane degradation pathway is modulated by two different global control signals: evidence from continuous cultures. *Journal of Bacteriology*, 185(16), 4772–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC166476/>

Duboué, E. R., Keene, A. C., & Borowsky, R. L. (2011). Evolutionary convergence on sleep loss in cavefish populations. *Current Biology : CB*, 21(8), 671–6. doi:10.1016/j.cub.2011.03.020

Elena, S. F., & Lenski, R. E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. *Nature*, 390(6658), 395–8. doi:10.1038/37108

Ernst, R. K., Moskowitz, S. M., Emerson, J. C., Kraig, G. M., Adams, K. N., Harvey, M. D., ... Miller, S. I. (2007). Unique lipid a modifications in *Pseudomonas aeruginosa* isolated from the airways of patients with cystic fibrosis. *The Journal of Infectious Diseases*, 196(7), 1088–1092.

Fazli, M., Almblad, H., Rybtke, M. L., Givskov, M., Eberl, L., & Tolker-Nielsen, T. (2014). Regulation of biofilm formation in *Pseudomonas* and *Burkholderia* species. *Environmental Microbiology*. doi:10.1111/1462-2920.12448

Flynn, K. M., Cooper, T. F., Moore, F. B.-G., & Cooper, V. S. (2013). The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. *PLoS Genetics*, 9(4), e1003426. doi:10.1371/journal.pgen.1003426

Folkesson, A., Jelsbak, L., Yang, L., Johansen, H. K., Ciofu, O., Høiby, N., & Molin, S. (2012). Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews. Microbiology*, 10(12), 841–51. doi:10.1038/nrmicro2907

Gaut, B. S., Long, A. D., Gaut, R. L., McDonald, P., Bennett, A. F., Rodriguez-Verdugo, A., & Tenaillon, O. (2012). The Molecular Diversity of Adaptive Convergence. *Science*. doi:10.1126/science.1212986

- Gupta, K., Liao, J., Petrova, O. E., Cherny, K. E., & Sauer, K. (2014). Elevated levels of the second messenger c-di-GMP contribute to antimicrobial resistance of *Pseudomonas aeruginosa*. *Molecular Microbiology*, 92(3), 488–506. doi:10.1111/mmi.12587
- Hall-Stoodley, L., Costerton, J. W., & Stoodley, P. (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol*, 2(2), 95–108. doi:10.1038/nrmicro821
- Heilbron, K., Toll-Riera, M., Kojadinovic, M., & Maclean, R. C. (2014). Fitness Is Strongly Influenced by Rare Mutations of Large Effect in a Microbial Mutation Accumulation Experiment. *Genetics*, genetics.114.163147–. doi:10.1534/genetics.114.163147
- Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nature Reviews Microbiology*, 7(4), 263–273. doi:10.1038/nrmicro2109
- Herron, M. D., & Doebeli, M. (2013a). Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. *PLoS Biology*, 11(2), e1001490. doi:10.1371/journal.pbio.1001490
- Herron, M. D., & Doebeli, M. (2013b). Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. *PLoS Biology*, 11(2), e1001490. doi:10.1371/journal.pbio.1001490
- Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., & Schweizer, H. P. (1998). A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene*, 212(1), 77–86. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9661666>
- Jayaraman, R. (2011). Hypermutation and stress adaptation in bacteria. *Journal of Genetics*, 90(2), 383–391. doi:10.1007/s12041-011-0086-6
- Junop, M. S., Yang, W., Funchain, P., Clendenin, W., & Miller, J. H. (2003). In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair*, 2(4), 387–405. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12606120>

Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., & Cooper, T. F. (2011). Negative epistasis between beneficial mutations in an evolving bacterial population. *Science (New York, N.Y.)*, 332(6034), 1193–6. doi:10.1126/science.1203801

Kim, W., Racimo, F., Schluter, J., Levy, S. B., & Foster, K. R. (2014). Importance of positioning for microbial evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 111(16), E1639–47. doi:10.1073/pnas.1323632111

Koeppel, A. F., Wertheim, J. O., Barone, L., Gentile, N., Krizanc, D., & Cohan, F. M. (2013). Speedy speciation in a bacterial microcosm: new species can arise as frequently as adaptations within a species. *The ISME Journal*, 7(6), 1080–91. doi:10.1038/ismej.2013.3

Koide, M., Miyanaga, A., Kudo, F., & Eguchi, T. (2013). Characterization of polyphosphate glucokinase SCO5059 from *Streptomyces coelicolor* A3(2). *Bioscience, Biotechnology, and Biochemistry*, 77(11), 2322–4. doi:10.1271/bbb.130498

Korgaonkar, A., Trivedi, U., Rumbaugh, K. P., & Whiteley, M. (2013). Community surveillance enhances *Pseudomonas aeruginosa* virulence during polymicrobial infection. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3), 1059–64. doi:10.1073/pnas.1214550110

Kornberg, A., Rao, N. N., & Ault-Riché, D. (1999). Inorganic polyphosphate: a molecule of many functions. *Annual Review of Biochemistry*, 68, 89–125. doi:10.1146/annurev.biochem.68.1.89

Kulakova, A. N., Hobbs, D., Smithen, M., Pavlov, E., Gilbert, J. A., Quinn, J. P., & McGrath, J. W. (2011). Direct Quantification of Inorganic Polyphosphate in Microbial Cells Using 4'-6-Diamidino-2-Phenylindole (DAPI). *Environmental Science & Technology*, 45(18), 7799–7803. doi:10.1021/es201123r

Lang, G. I., Botstein, D., & Desai, M. M. (2011). Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics*, 188(3), 647–61. doi:10.1534/genetics.111.128942

Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., & Desai, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, 500(7464), 571–4. doi:10.1038/nature12344

- Lenski, R. E. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*, 138, 1315–1341.
- Lenski, R. E., Ofria, C., Collier, T. C., & Adami, C. (1999). Genome complexity, robustness and genetic interactions in digital organisms. *Nature*, 400(6745), 661–4. doi:10.1038/23245
- Lieberman, T. D., Flett, K. B., Yelin, I., Martin, T. R., McAdam, A. J., Priebe, G. P., & Kishony, R. (2014). Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nature Genetics*, 46(1), 82–7. doi:10.1038/ng.2848
- Lujan, A. M., Macia, M. D., Yang, L., Molin, S., Oliver, A., & Smania, A. M. (2011). Evolution and adaptation in *Pseudomonas aeruginosa* biofilms driven by mismatch repair system-deficient mutators. *PLoS One*, 6(11), e27842. doi:10.1371/journal.pone.0027842 PONE-D-11-17239 [pii]
- Maciá, M. D., Borrell, N., Segura, M., Gómez, C., Pérez, J. L., & Oliver, A. (2006). Efficacy and potential for resistance selection of antipseudomonal treatments in a mouse model of lung infection by hypermutable *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, 50(3), 975–983.
- Maharjan, R., Seeto, S., Notley-McRobb, L., & Ferenci, T. (2006). Clonal adaptive radiation in a constant environment. *Science (New York, N.Y.)*, 313(5786), 514–517. doi:10.1126/science.1129865
- Marchinko, K. B., Matthews, B., Arnegard, M. E., Rogers, S. M., & Schluter, D. (2014). Maintenance of a genetic polymorphism with disruptive natural selection in stickleback. *Current Biology : CB*, 24(11), 1289–92. doi:10.1016/j.cub.2014.04.026
- McElroy, K. E., Hui, J. G. K., Woo, J. K. K., Luk, A. W. S., Webb, J. S., Kjelleberg, S., ... Thomas, T. (2014). Strain-specific parallel evolution drives short-term diversification during *Pseudomonas aeruginosa* biofilm formation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), E1419–27. doi:10.1073/pnas.1314340111
- McKinnon, J. S., Mori, S., Blackman, B. K., David, L., Kingsley, D. M., Jamieson, L., ... Schluter, D. (2004). Evidence for ecology's role in speciation. *Nature*, 429(6989), 294–8. doi:10.1038/nature02556

- Mena, A., Smith, E. E., Burns, J. L., Speert, D. P., Moskowitz, S. M., Perez, J. L., & Oliver, A. (2008). Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *J Bacteriol*, 190(24), 7910–7917. doi:JB.01147-08 [pii] 10.1128/JB.01147-08
- Miller, J. H. (1996). Spontaneous mutators in bacteria: insights into pathways of mutagenesis and repair. *Annual Review of Microbiology*, 50, 625–43. doi:10.1146/annurev.micro.50.1.625
- Molin, S., & Tolker-Nielsen, T. (2003). Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Current Opinion in Biotechnology*, 14(3), 255–61. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12849777>
- Monds, R. D., Newell, P. D., Schwartzman, J. A., & O'Toole, G. A. (2006). Conservation of the Pho regulon in *Pseudomonas fluorescens* Pf0-1. *Applied and Environmental Microbiology*, 72(3), 1910–24. doi:10.1128/AEM.72.3.1910-1924.2006
- Morales, G., Linares, J. F., Beloso, A., Albar, J. P., Martínez, J. L., & Rojo, F. (2004). The *Pseudomonas putida* Crc global regulator controls the expression of genes from several chromosomal catabolic pathways for aromatic compounds. *Journal of Bacteriology*, 186(5), 1337–44. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=344427&tool=pmcentrez&rendertype=abstract>
- Moreno, R., Martínez-Gomariz, M., Yuste, L., Gil, C., & Rojo, F. (2009). The *Pseudomonas putida* Crc global regulator controls the hierarchical assimilation of amino acids in a complete medium: evidence from proteomic and genomic analyses. *Proteomics*, 9(11), 2910–28. doi:10.1002/pmic.200800918
- Moskowitz, S. M., & Ernst, R. K. (2010). The role of *Pseudomonas* lipopolysaccharide in cystic fibrosis airway infection. *Sub-Cellular Biochemistry*, 53, 241–253.
- Notley-McRobb, L., Seeto, S., & Ferenci, T. (2002). Enrichment and elimination of mutY mutators in *Escherichia coli* populations. *Genetics*, 162(3), 1055–1062.
- Oliver, A., & Mena, A. (2010). Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 16(7), 798–808. doi:10.1111/j.1469-0691.2010.03250.x

- Parsek, M. R., & Singh, P. K. (2003). Bacterial biofilms: an emerging link to disease pathogenesis. *Annu Rev Microbiol*, 57, 677–701.
doi:10.1146/annurev.micro.57.030502.090720
- Periasamy, S., & Kolenbrander, P. E. (2009). Mutualistic biofilm communities develop with *Porphyromonas gingivalis* and initial, early, and late colonizers of enamel. *J Bacteriol*, 191(22), 6804–6811. doi:JB.01006-09 [pii] 10.1128/JB.01006-09
- Poltak, S., & Cooper, V. (2011). Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *The ISME Journal*, 5(3), 369–378. doi:10.1038/ismej.2010.136
- Polz, M. F., Hunt, D. E., Preheim, S. P., & Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1475), 2009–21. doi:10.1098/rstb.2006.1928
- Rodriguez-Rojas, A., Oliver, A., & Blazquez, J. (2011). Intrinsic and Environmental Mutagenesis Drive Diversification and Persistence of *Pseudomonas aeruginosa* in Chronic Lung Infections. *J Infect Dis*. doi:jir690 [pii] 10.1093/infdis/jir690
- Rojo, F. (2010). Carbon catabolite repression in *Pseudomonas* : optimizing metabolic versatility and interactions with the environment. *FEMS Microbiology Reviews*, 34(5), 658–84. doi:10.1111/j.1574-6976.2010.00218.x
- Segrè, A. V., Murray, A. W., & Leu, J.-Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS Biology*, 4(8), e256.
doi:10.1371/journal.pbio.0040256
- Shapiro, B. J., & Polz, M. F. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends in Microbiology*, 22(5), 235–47.
doi:10.1016/j.tim.2014.02.006
- Shaver, A. C., Dombrowski, P. G., Sweeney, J. Y., Treis, T., Zappala, R. M., & Sniegowski, P. D. (2002). Fitness evolution and the rise of mutator alleles in experimental *Escherichia coli* populations. *Genetics*, 162(2), 557–566. Retrieved from
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12399371

- Silby, M. W., Nicoll, J. S., & Levy, S. B. (2009a). Requirement of polyphosphate by *Pseudomonas fluorescens* Pf0-1 for competitive fitness and heat tolerance in laboratory media and sterile soil. *Applied and Environmental Microbiology*, 75(12), 3872–3881.
- Silby, M. W., Nicoll, J. S., & Levy, S. B. (2009b). Requirement of polyphosphate by *Pseudomonas fluorescens* Pf0-1 for competitive fitness and heat tolerance in laboratory media and sterile soil. *Applied and Environmental Microbiology*, 75(12), 3872–81. doi:10.1128/AEM.00017-09
- Silby, M. W., Nicoll, J. S., & Levy, S. B. (2012). Regulation of polyphosphate kinase production by antisense RNA in *Pseudomonas fluorescens* Pf0-1. *Applied and Environmental Microbiology*, 78(12), 4533–7. doi:10.1128/AEM.07836-11
- Smith, E. E., Buckley, D. G., Wu, Z., Saenphimmachak, C., Hoffman, L. R., D'Argenio, D. A., ... Olson, M. V. (2006). Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A*, 103(22), 8487–8492. doi:0602138103 [pii] 10.1073/pnas.0602138103
- Smith, J. M. (1992). Evolutionary biology. Byte-sized evolution. *Nature*, 355(6363), 772–3. doi:10.1038/355772a0
- Sniegowski, P. D., Gerrish, P. J., & Lenski, R. E. (1997). Evolution of high mutation rates in experimental populations of *E. coli*. *Nature (London)* , 387, 703–705.
- Soto, S. M. (2013). Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence*, 4(3), 223–9. doi:10.4161/viru.23724
- Starkey, M., Hickman, J. H., Ma, L., Zhang, N., De Long, S., Hinz, A., ... Parsek, M. R. (2009). *Pseudomonas aeruginosa* rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. *J Bacteriol*, 191(11), 3492–3503. doi:JB.00119-09 [pii] 10.1128/JB.00119-09
- Taddei, F., Radman, M., Maynard-Smith, J., Toupiance, B., Gouyon, P. H., & Godelle, B. (1997). Role of mutator alleles in adaptive evolution. *Nature*, 387(6634), 700–702. doi:10.1038/42696
- Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., & Gaut, B. S. (2012). The molecular diversity of adaptive convergence. *Science (New York, N.Y.)*, 335(6067), 457–61. doi:10.1126/science.1212986

Tenaillon, O., Toupancre, B., Le Nagard, H., Taddei, F., & Godelle, B. (1999). Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics*, 152(2), 485–493. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10353893

Tomlin, K. L., Clark, S. R., & Ceri, H. (2004). Green and red fluorescent protein vectors for use in biofilm studies of the intrinsically resistant *Burkholderia cepacia* complex. *J Microbiol Methods*, 57(1), 95–106. doi:10.1016/j.mimet.2003.12.007 S016770120300349X [pii]

Torres-Barceló, C., Cabot, G., Oliver, A., Buckling, A., & Maclean, R. C. (2013). A trade-off between oxidative stress resistance and DNA repair plays a role in the evolution of elevated mutation rates in bacteria. *Proceedings. Biological Sciences / The Royal Society*, 280(1757), 20130007. doi:10.1098/rspb.2013.0007

Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R., & Cooper, V. S. (2013). Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc Natl Acad Sci U S A*, 110(3), E250–9. doi:1207025110 [pii] 10.1073/pnas.1207025110

Travis, J. M. J., & Travis, E. R. (2002). Mutator dynamics in fluctuating environments. *Proceedings. Biological Sciences / The Royal Society*, 269(1491), 591–597.

Vos, M., & Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *The ISME Journal*, 3(2), 199–208. doi:10.1038/ismej.2008.93

Wang, L., Spira, B., Zhou, Z., Feng, L., Maharjan, R. P., Li, X., ... Ferenci, T. (2010). Divergence involving global regulatory gene mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome Biology and Evolution*, 2, 478–87. doi:10.1093/gbe/evq035

Wiser, M. J., Ribeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science (New York, N.Y.)*, 342(6164), 1364–7. doi:10.1126/science.1243357

Wong, A., Rodrigue, N., & Kassen, R. (2012). Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*. *PLoS Genetics*, 8(9), e1002928. doi:10.1371/journal.pgen.1002928

Wylie, C. S., Ghim, C.-M., Kessler, D., & Levine, H. (2009). The fixation probability of rare mutators in finite asexual populations. *Genetics*, 181(4), 1595–1612.

Yildiz, F. H., & Visick, K. L. (2009). Vibrio biofilms: so much the same yet so different. *Trends Microbiol*, 17(3), 109–118. doi:S0966-842X(09)00023-7 [pii] 10.1016/j.tim.2008.12.004

Zhao, K., Tseng, B. S., Beckerman, B., Jin, F., Gibiansky, M. L., Harrison, J. J., ... Wong, G. C. L. (2013). Psl trails guide exploration and microcolony formation in *Pseudomonas aeruginosa* biofilms. *Nature*, 497(7449), 388–91. doi:10.1038/nature12155

Zhu, Y., Huang, W., Lee, S. S. K., & Xu, W. (2005). Crystal structure of a polyphosphate kinase and its implications for polyphosphate synthesis. *EMBO Reports*, 6(7), 681–7. doi:10.1038/sj.embor.7400448

APPENDICES

Appendix A Chapter 2 Supplemental material.

Figures A.1-8 and Tables A.1-3

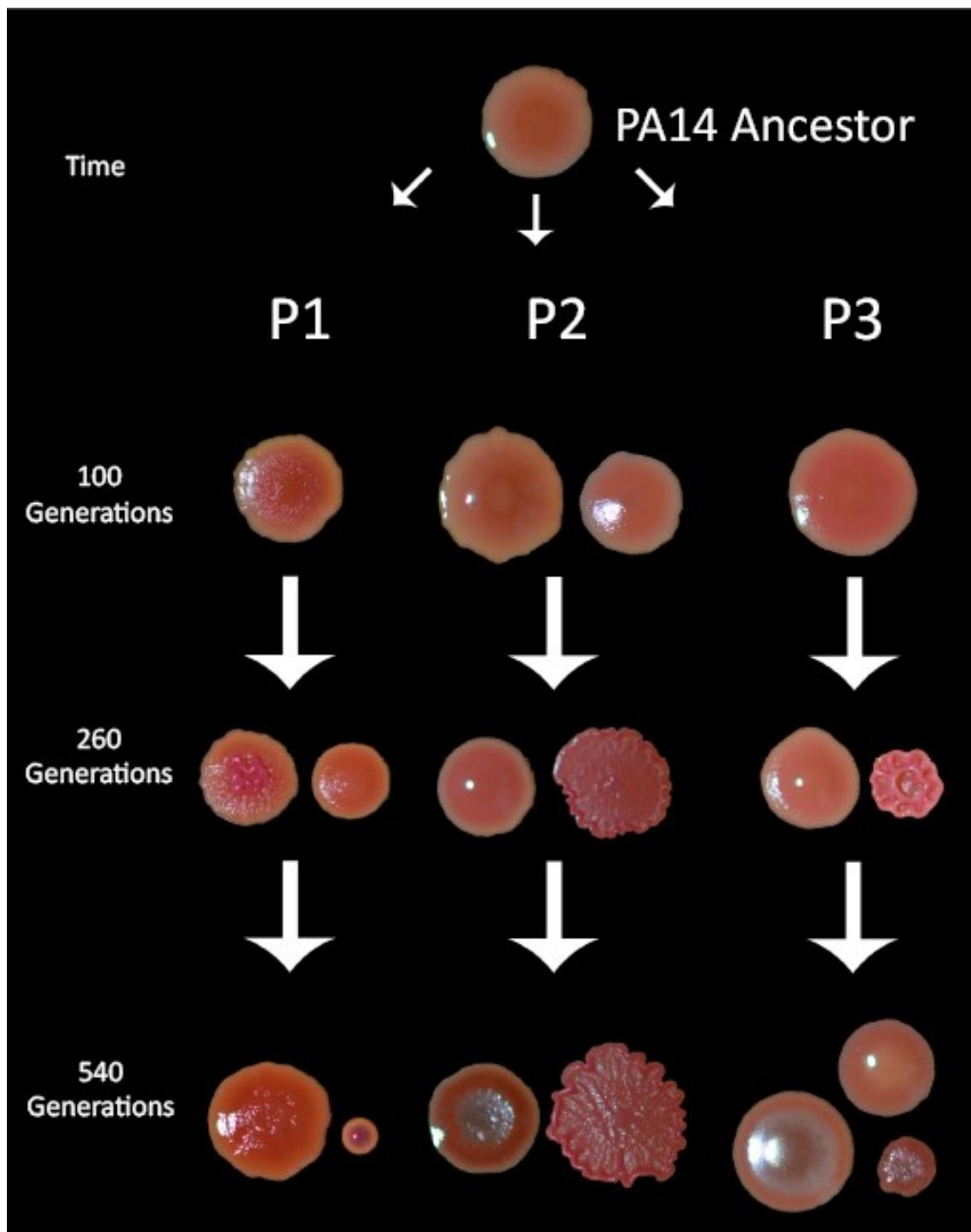


Figure A.1. Evolution of morphological diversity among three replicate planktonic populations.

Aliquots from different time points were grown on 1% tryptone supplemented with Coomassie blue and Congo red. Number and letter designations (B1-3) refer to the replication population numbers.

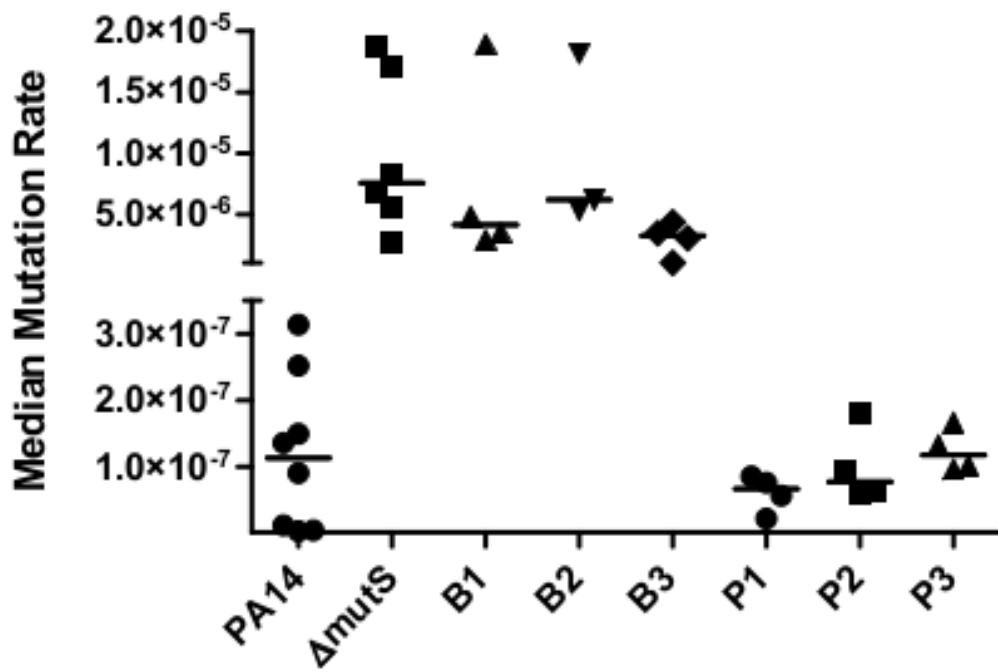


Figure A.2. Hypermutation repeatedly evolves only in replicate biofilm populations.

Fluctuation tests with representative isolates from replicate populations were performed to approximate the mutation rate through the rate spontaneous resistance mutants arise in response to 1 $\mu\text{g/mL}$ ciprofloxacin and 100 $\mu\text{g/mL}$ gentamicin (data not shown). Horizontal bars represent the media rate of mutation.

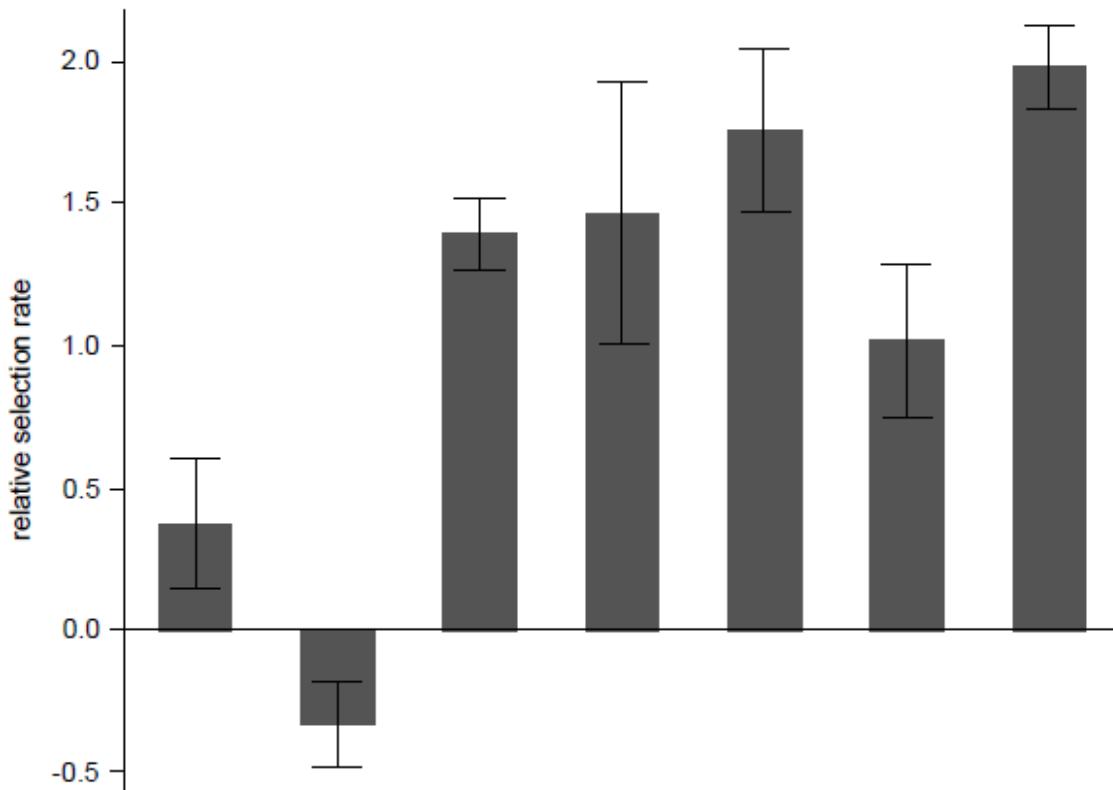


Figure A.3. Isolates of individual morphotypes tend to be relatively more fit than the complete community.

Individual isolates were competed head-to-head against the PA14 ancestor and standardized compared to the competitive ability of the complete community. An experimental mean of one represents equivalency to the complete community. Four out of the seven types present in B1 at the end of the experiment were significantly more fit than the entire community ($P < 0.05$). Included for convenience are the *t*-statistics (*t*) and degrees of freedom (d.f.) for post-hoc analyses. Numbers in parentheses represent 95% confidence intervals (CI) for relative selection rate values. Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan (Table 1).

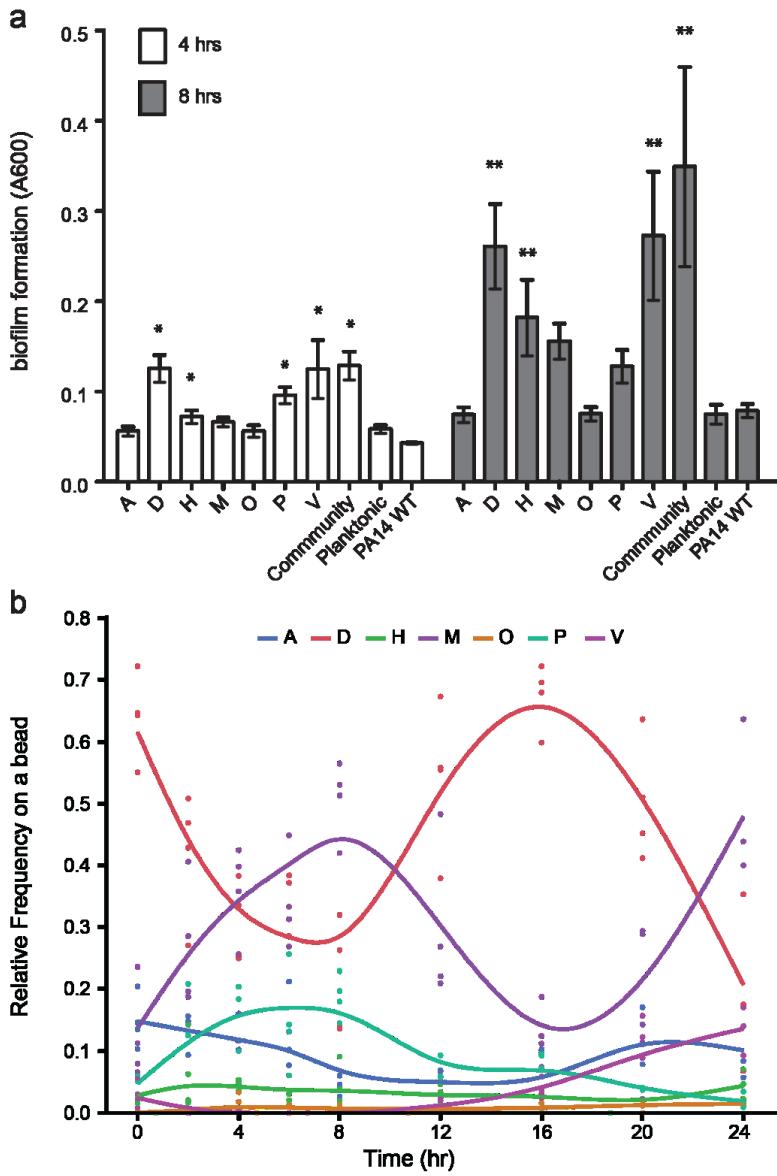


Figure A.4. Variation in the timing of biofilm formation and attachment amongst B1 isolates.

Isolates from the B1 population after 540 generations of adaptation were examined for variation in the timing of maximum biofilm formation (a) and attachment to a bead (b). a. Biofilm formation after four (white bars) and eight (grey bars) hours of growth. Asterisks denote significance based on post-hoc Tukey tests, $P < 0.05$. b. Replicate cultures were destructively sampled at nine different time points: 0, 2, 4, 6, 8, 12, 16, 20 and 24 hours. Six out of seven morphotypes except H significantly fluctuated in their relative abundance in the community based on CFU/mL on a bead. Asterisks represent significant one-way ANOVAs of abundance through time, $P < 0.05$. Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan (Table 1).

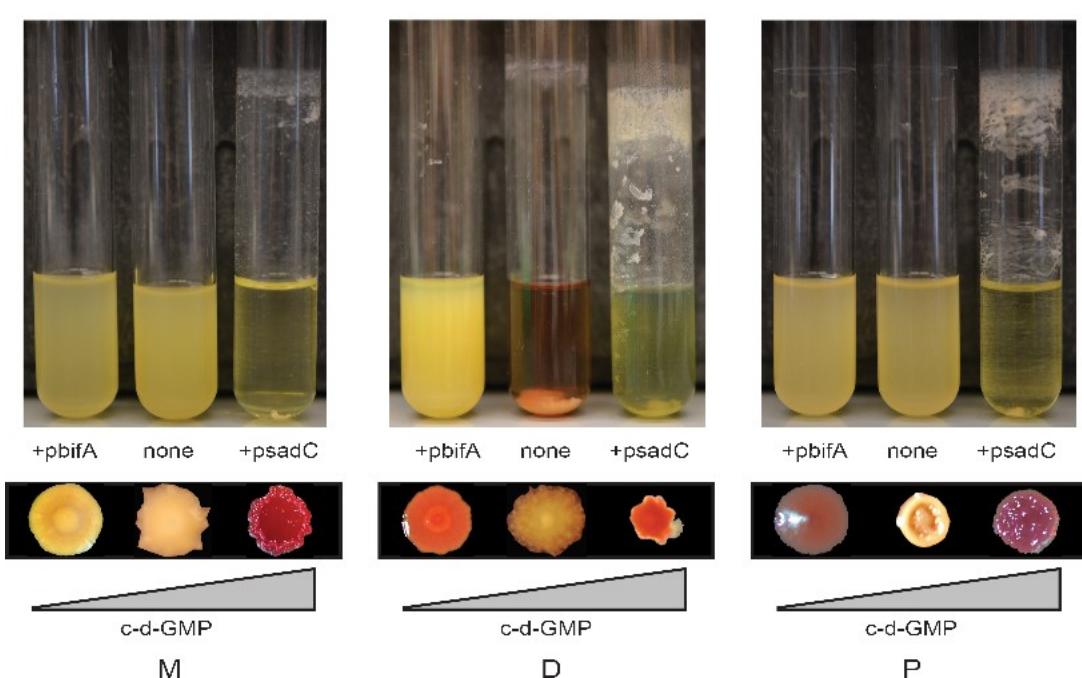
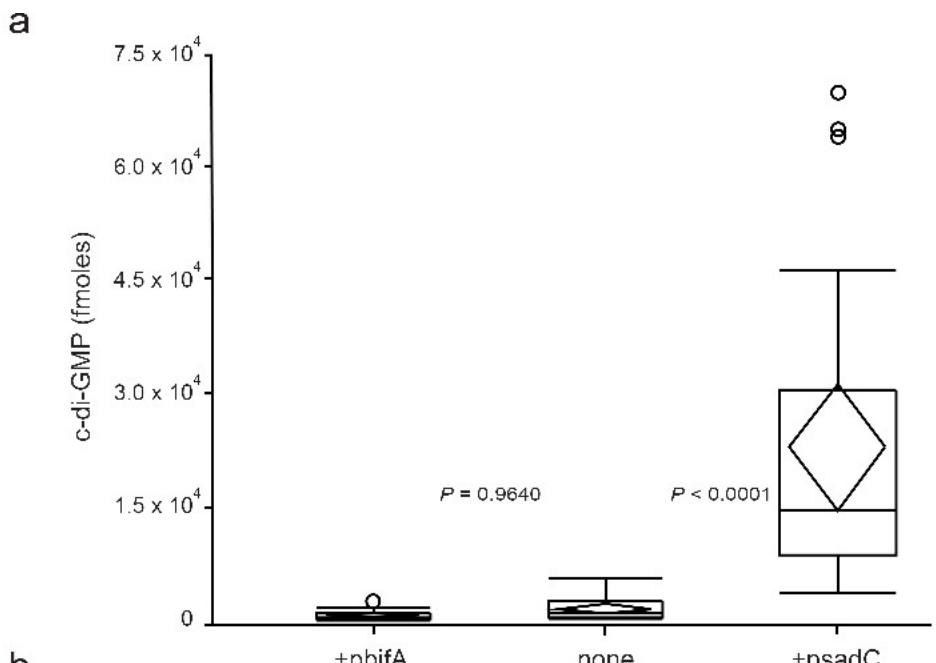


Figure A.5. Ecotype physiology is mediated by c-di-GMP levels.

Inducible plasmids containing *bifA*, a phosphodiesterase, or *sadC*, a diguanylate cyclase, were introduced into focal variants (D, M, or P, not shown, see Fig. S5) from the evolved biofilm population. a. Concentrations of c-di-GMP of cells from induced cells were quantified by LC-MS (41). *P* values between box plots are based on post-hoc Tukey tests. b. Effects of altered levels of c-di-GMP were determined from growth patterns in overnight cultures of LB and changes in colony morphology.

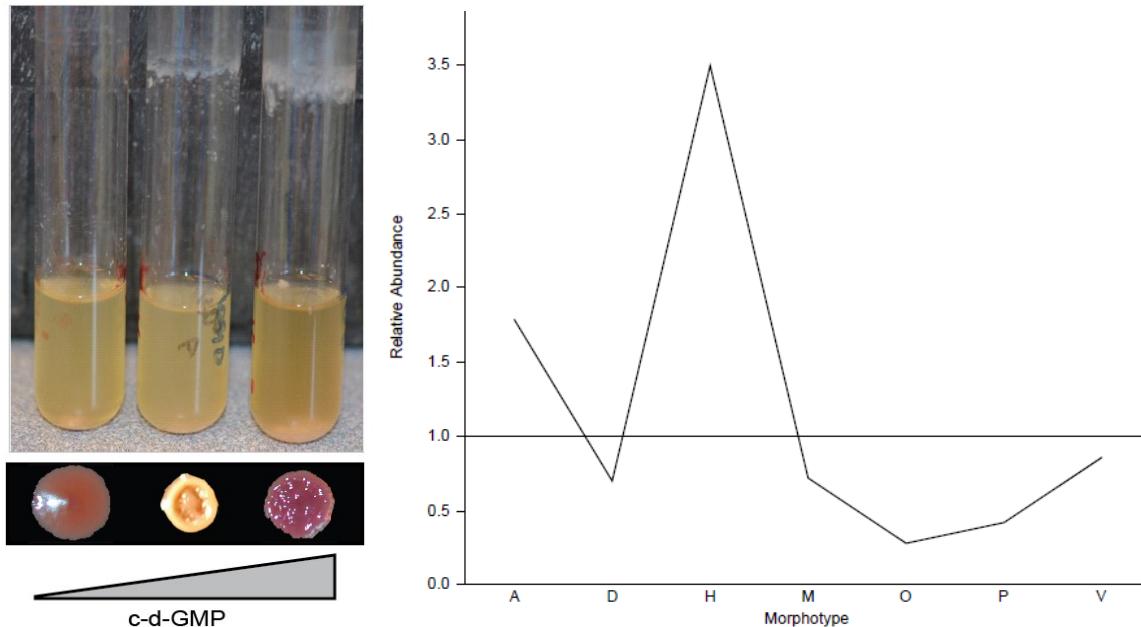


Figure A.6. Reduced c-di-GMP levels in the subdominant specialist P type also disproportionately affects community dynamics.

Plasmids containing *bifA*, a phosphodiesterase, or *sadC*, a diguanylate cyclase, under control of the P_{BAD} promoter (1, 2) were introduced into the P type from the evolved B1 biofilm population (left). Expression was induced with the addition of arabinose at a final concentration of 0.5%. The consequence of altering levels of c-di-GMP was determined by observing growth patterns in overnight cultures of LB. Communities were constructed with the P type expressing the phosphodiesterase *bifA* (1), effectively reducing intracellular levels of c-di-GMP (right). Consequences of supplementation of P expressing *bifA* (dashed line) on community composition relative to a natural B1 community (solid line). Values represent fold changes in relative frequency.

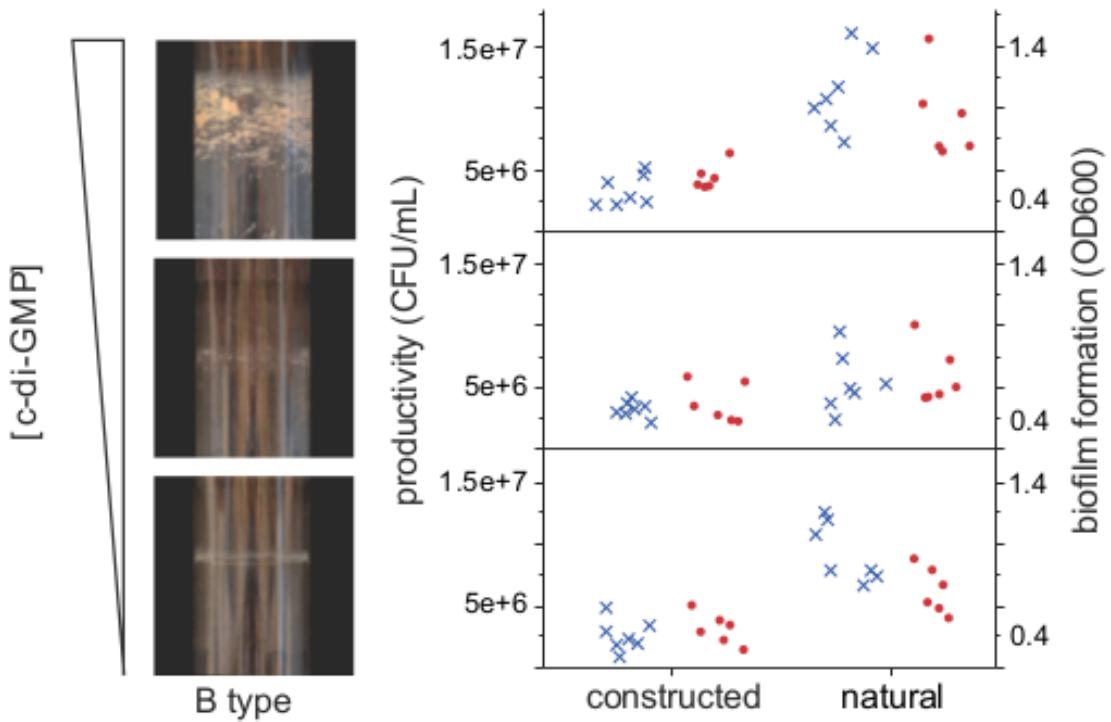


Figure A.7. Altering c-di-GMP levels in an individual B3 morphotype does not affect the community as a whole.

Communities were either constructed or supplemented with the B type expressing the phosphodiesterase *bifA* (bottom, left), effectively reducing intracellular levels of c-di-GMP, or the diguanylate cyclase *sadC* (top, right), a phosphodiesterase effectively increasing c-di-GMP levels. The effect of altering c-di-GMP in a single member on the community as a whole was assessed through productivity or CFU/mL on a bead after 24 hours (blue x's) and biofilm formation (red dots). Regardless of how the B type's physiology was disrupted (left), community function was unaffected. Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan, WT = PA14 ancestor (Table 1).

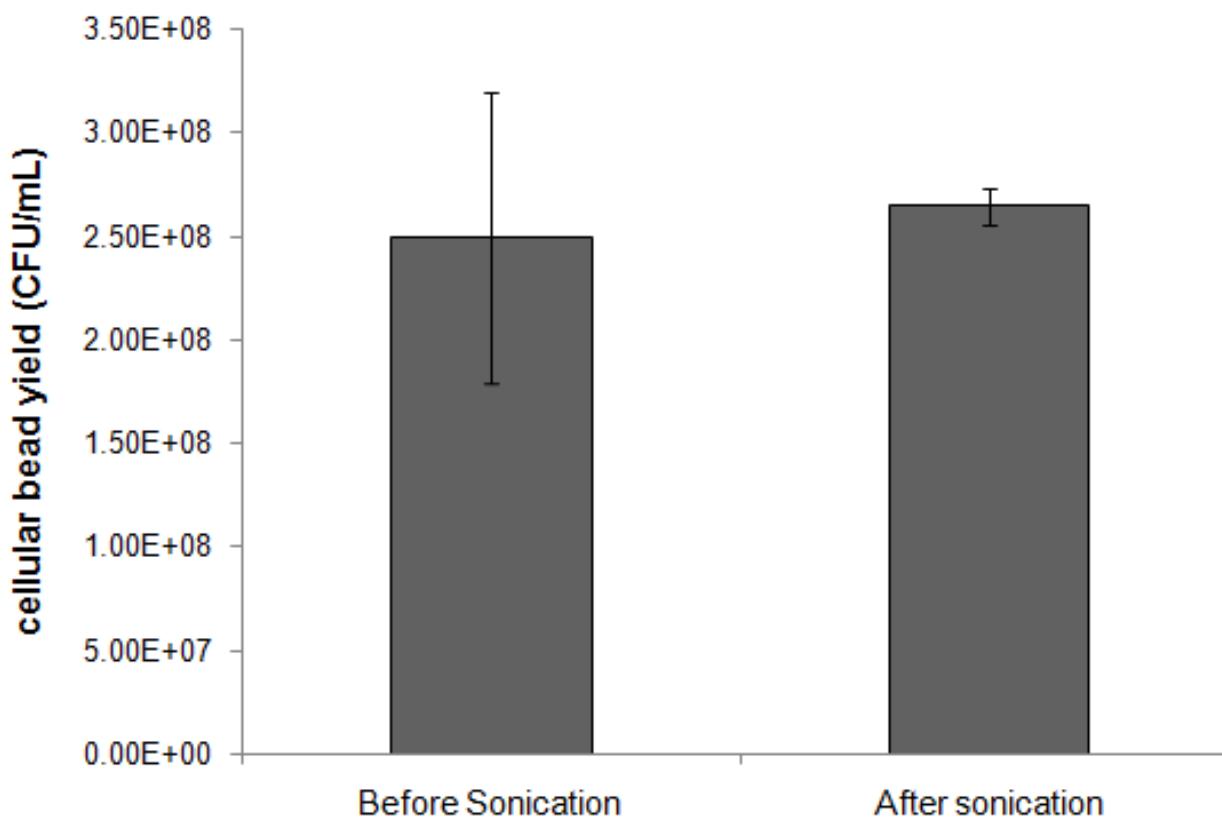


Figure A.8. Sonication of *Pseudomonas aeruginosa* PA14 biofilms does not reduce cell viability.

Replicate biofilms were sampled from polystyrene beads extracted after 24 hours of growth in M63. Samples were plated before and after being sonicated in 1.5 mL PBS for 10 seconds.

Table A.1. Comparisons of the mutation rate for ciprofloxacin resistance between 540 generations isolates, a *mutS* deletion strain and the ancestral PA14 strain. Significance was based on two-tailed t-tests.

Type	Mutation rate	t	d.f.	P
Biofilm-evolved	4.92x10 ⁻⁶	-2.39495	26	0.0244*
Planktonic-evolved	9.43x10 ⁻⁸	-0.69584	19	0.4954
<i>mutS</i> deletion	2.20x10 ⁻⁵	2.598907	15	0.0210*
Ancestral PA14	1.20x10 ⁻⁷	NA	NA	NA

Table A.2. Comparisons of yield (CFU/mL) of morphotypes grown in monoculture (mono) or community mixture (mix).

Morphotype*	Yield (mix)	Yield (mono)	t	d.f.	P
A	3.50x10 ⁶	3.66x10 ⁷	6.539	8	0.0002*
V	1.54x10 ⁶	1.30x10 ⁸	11.23	8	<0.0001*
P	3.54x10 ⁶	2.61x10 ⁷	2.977	8	0.0177*
M	2.01x10 ⁷	5.42x10 ⁷	3.362	8	0.0099*
H	2.16x10 ⁶	7.34x10 ⁷	10.91	7	<0.0001*
O	7.40x10 ⁵	5.01x10 ⁷	4.612	8	0.0017*
D	2.78x10 ⁶	3.50x10 ⁷	4.691	8	0.0016*

* Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan (Table 1).

Table A.3. Comparisons between expected observed biofilm formation for biofilm-evolved populations after 260 and 540 generations of adaptation. Significance was based on two-tailed t-tests.

Population	Time (gen)	Expected biofilm	Observed biofilm	t	d.f.	P
B1	260	0.90412	0.65325	-1.01694	15	0.3264
B1	540	0.69379	2.00795	5.465871	15	<0.0001*
B2	260	1.29909	3.72430	25.41343	15	<0.0001*
B2	540	1.40639	2.56886	11.26207	15	<0.0001*
B3	260	1.59892	0.82009	-5.51479	15	<0.0001*
B3	540	1.85427	0.85654	-4.99305	15	0.0002*

Table A.4. The selection rate or fitness of various communities constructed from 90day isolates. Post hoc comparisons to identify homogeneous groupings were conducted using Tukey's test.

Construct*	N	Selection rate (95% CI)	Homogeneous subsets
Complete community	5	3.124 (0.29)	A
Minus M	5	1.829 (0.30)	C
Minus P	4	1.942 (0.53)	C
Minus O	5	2.763 (0.39)	AB
Minus D	5	2.553 (0.40)	B
Minus V	5	2.039 (0.12)	C
Minus A	5	2.779 (0.33)	AB
Minus H	5	2.654 (0.47)	B

* Letter abbreviations refer to a specific B1 isolate: A = Apollo, D = Demeter, H = Hera, M = Minerva, O = Olympus, P = Poseidon, V = Vulcan (Table 1).

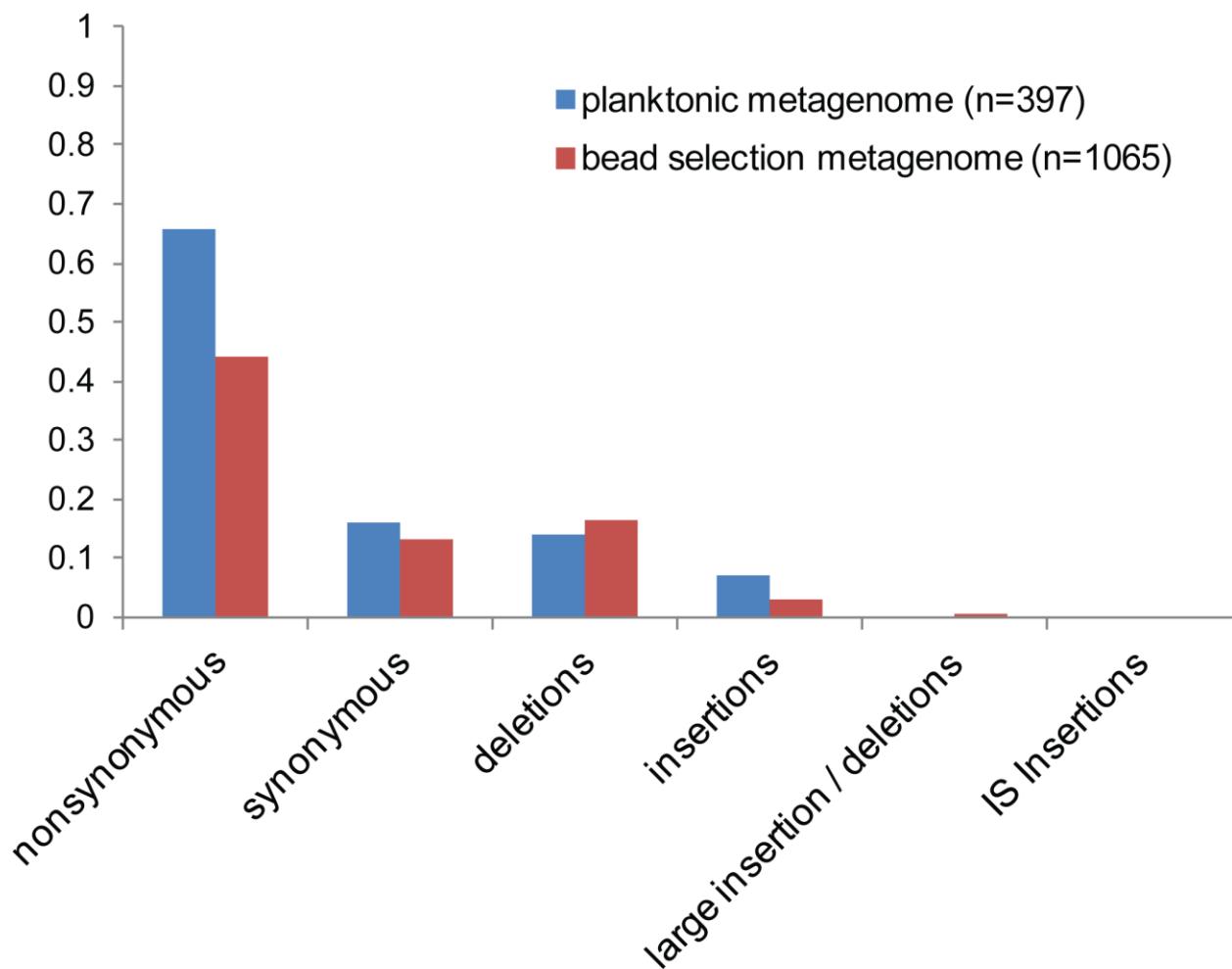


Figure B.1. Comparison of the distribution of mutational types identified using both of our whole-population sequencing strategies.

Whole-population sequencing with or without a day of bead selection identified similar distributions of mutational changes. Counts are expressed as the observed frequency in each data set given the large difference in sample size (without bead selection: n=396, bead selection: n=1065). As expected, transitions (96.5%) were the most abundant in agreement with the fixation of defects in DNA mismatch repair (MMR) in all replicate biofilm populations (Chapter 2).

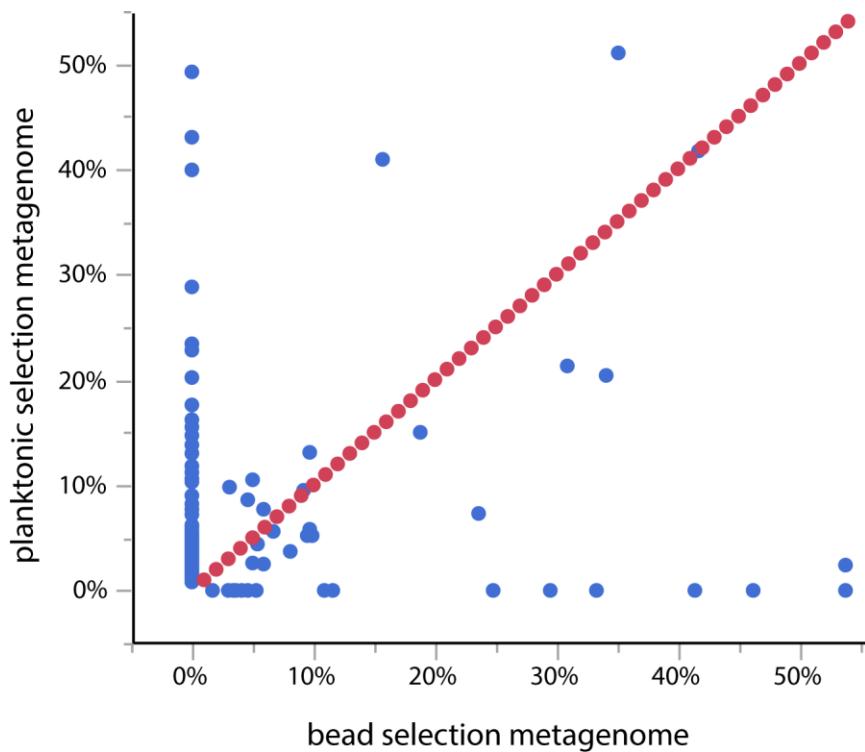


Figure B.2. Bead selection has a limited effect on mutational clusters identified in B1 after 100 generations.

Linear regression of observed allele frequencies of mutational calls identified with two different preparation methods. Mutations below the line are enriched following bead selection while mutations above are enriched during planktonic growth.

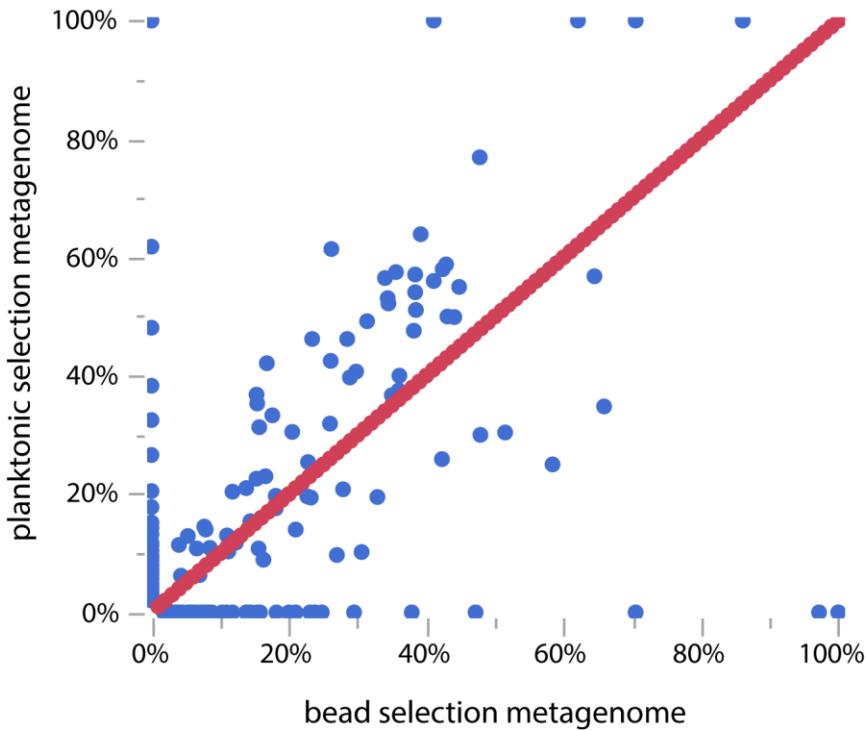


Figure B.3. Bead selection has a limited effect on mutational clusters identified in B1 after 264 generations.

Linear regression of observed allele frequencies of mutational calls identified with whole-population sequencing utilizing two different preparation methods. Mutations below the line are enriched following bead selection while mutations above are enriched during planktonic growth.

Tables B.1-3 summarize the mutational calls identified in the various next-generation sequencing approaches utilized during this study (See Figure S1 for a visual explanation). Table B.1 summarizes 1,183 mutational calls identified with our full genome re-sequencing of forty colony isolates, respectively. Table B.2 summarizes the 8,931 raw mutational calls identified with our whole-population sequencing of all three replicate biofilm-evolved populations (B1, B2, B3) after 107, 150, 264, 396, 450 and 540 generations while Table B.3 summarizes the resulting 397 high-confidence mutational trajectories. Lastly, Table B.4 summarizes the additional mutational calls identified with additional whole-population sequencing of the replicate B1 population samples adapted for 107, 264 and 540 generations following a single day of bead selection or six generations of growth under conditions described previously (Chapter 2). These data are made available online given their large size.

Table B.1. Mutational calls identified using full genomic re-sequencing of forty colony phenotypes isolated from all three replicate biofilm-evolved populations (B1, B2, B3) after 107, 264 or 540 generations with bead selection.

<https://drive.google.com/file/d/0B1ba0QiKxpZVW9oQjINQ0p3b0k/edit?usp=sharing>

Table B.2. All mutational calls identified using whole-population sequencing of all three replicate biofilm-evolved populations (B1, B2, B3) after 107, 150, 264, 396, 450 and 540 generations without bead selections.

<https://drive.google.com/file/d/0B1ba0QiKxpZV1hiNkZwb0NiWk0/edit?usp=sharing>

Table B.3. High-confidence mutational trajectories identified using whole-population sequencing of all three replicate biofilm-evolved populations (B1, B2, B3) after 107, 150, 264, 396, 450 and 540 generations without bead selections.

<https://drive.google.com/file/d/0B1ba0QiKxpZSk0tZnM0b09NNnM/edit?usp=sharing>

Table B.4. Mutational calls identified using whole-population sequencing of all three replicate biofilm-evolved populations (B1, B2, B3) after 107, 264 and 540 generations with bead selection.

<https://drive.google.com/file/d/0B1ba0QiKxpZVnZaNkpqaUtxdjq/edit?usp=sharing>

Detailed snapshots of the population genetic structure through time were created utilizing a hierarchical clustering analysis approach (Chapter 3). Leveraging multiple time points, mutations were combined into mutational cohorts based on the similarity of their trajectories through time. Table B.4 summarizes the mutational calls used in the creation of the Fisher-Muller diagram depicted in Chapter 3, Figure 5. This creation of these figures was automated using MATLAB scripts provided by the Michael Desai Lab at Harvard University. The scripts were written by Katya Kosheleva and modified to accommodate different data sets and unique experimental designs. These scripts will be made publicly available for use at the following location:

<https://drive.google.com/folderview?id=0B1ba0QiKxpZcXlhQjNkTTlwZ1k&usp=sharing>

Table B.5. A summary of the 136 mutational calls and their membership within 12 mutational clusters used to infer the population genetic of the replicate B1 population.

<https://drive.google.com/file/d/0B1ba0QiKxpZZHozVW1zaER5ZDg/edit?usp=sharing>

Appendix C Chapter 4 Supplemental Material.

Table C.1. Mutational calls identified with whole-population sequencing across all three replicate planktonic-evolved populations (P1, P2, P3) after 396 and 540 generations.

<https://drive.google.com/file/d/0B1ba0QiKxpZSjJsSHIyaVZCbjA/edit?usp=sharing>

Table C.2. Mutational clusters in B1 displaying large shifts in allele frequency through time.

replicate population	adaptive event*	period of fastest increase	type	mutation	involved mutations
B1	sweep	264 → 396	INS	FS INS	orfK
			SNP	E16K	groES
			SNP	L38P	recG
			SNP	G316S	PA14_11650
			SNP	T443A	ppk
			NC	T→C	PA14_49010, PA14_49020
			SNP	A42T	PA14_16800
			INS	FS INS	PA14_38580
			SNP	A294V	PA14_31470
			SNP	V224A	PA14_06030
B1	sway	264 → 396	NC	1 bp INS	PA14_50870, PA14_50880
			SNP	N6D	PA14_19450
			SNP	W95*	PA14_48210
			SNP	V282A	PA14_71750
			SNP	G521D	rpoB
			SNP	V1632A	pvdJ
			SNP	F25F	PA14_23670
			SNP	F125L	PA14_04050
			SNP	A306V	wbpM
			SNP	S246L	PA14_06450
			SNP	R256H	PA14_42950
			SNP	V82A	PA14_41730
			SNP	Y32C	PA14_28790
			DEL	1 bp DEL	PA14_50650
			NC	A→G	oprH, napE
			SNP	A889A	PA14_58250
			SNP	R145K	apt
			SNP	A120A	carB
			SNP	Q18R	PA14_69980
			SNP	A192A	PA14_29260
			SNP	V152I	PA14_19310
			SNP	V223A	PA14_68060
			SNP	T450T	betT1
			SNP	G145S	PA14_10910
			SNP	D95G	PA14_02990
			SNP	W384*	PA14_66380
			SNP	A751V	rcsC
			SNP	A641A	PA14_48160

*'sweeps' represent time frames where mutational cohorts rapidly increased, and reached an allele frequency of 1, while 'sways' represent time frames where cohorts increased rapidly, but do not achieve fixation.

Table C.3. Mutational clusters in B2 displaying large shifts in allele frequency through time.

replicate population	adaptive event*	period of fastest increase	type mutation	involved mutations
B2	sweep	102 → 150	SNP T46T	<i>PA14_56890</i>
			DEL FS DEL	<i>orfH</i>
			SNP T112P	<i>mutS</i>
			INS FS INS	<i>PA14_65940</i>
			NC T→C	<i>PA14_66350, PA14_66380</i>
			NC T→C	<i>PA14_53630, PA14_53640</i>
			NC T→C	<i>PA14_20550, amiE</i>
			SNP I141V	<i>PA14_61990</i>
			NC C→T	<i>PA14_47530, PA14_47540</i>
			SNP L320L	<i>PA14_04290</i>
			SNP S320L	<i>fliC</i>
			INS FS INS	<i>rtcR</i>
			NC A→G	<i>PA14_21300, phaJ1</i>
B2	sway	150 → 264	SNP T55A	<i>mreB</i>
			SNP R468H	<i>PA14_26640</i>
			NC 1 bp INS	<i>PA14_61390, mqoB</i>
			NC T→C	<i>pqsA, ogt</i>
			NC C→T	<i>PA14_12490, PA14_12530</i>
			NC C→T	<i>PA14_11990, proA</i>
			SNP Y233C	<i>PA14_59530</i>
			NC C→T	<i>alaS, PA14_52610</i>
			SNP R275C	<i>rpoC</i>
			SNP P135P	<i>PA14_19850</i>
			INS FS INS	<i>prmA</i>
			SNP G385D	<i>PA14_55770</i>

*'sweeps' represent time frames where mutational cohorts rapidly increased, and reached an allele frequency of 1, while 'sways' represent time frames where cohorts increased rapidly, but do not achieve fixation.

Table C.4. Mutational clusters in B3 displaying large shifts in allele frequency through time.

replicate population	adaptive event*	period of fastest increase	type	mutation	involved mutations
B3	sweep	150 → 264	DEL	12 bp DEL	<i>orfH</i>
			SNP	E84*	<i>ddl</i>
			INS	FS INS	PA14_66380
			DEL	9 bp DEL	<i>orn</i>
B3	sway	264 → 396	SNP	G1062D	<i>rpoB</i>
			SNP	D422G	PA14_13150
			NC	T→G	PA14_51840,PA14_51850
			NC	1 bp DEL	PA14_52530,PA14_52540
B3	sweep	450 → 540	INS	FS INS	PA14_65860
			SNP	P100S	PA14_25800
			SNP	D467G	<i>mutL</i>
			SNP	Y405C	PA14_13110

*'sweeps' represent time frames where mutational cohorts rapidly increased, and reached an allele frequency of 1, while 'sways' represent time frames where cohorts increased rapidly, but do not achieve fixation.