

**Investigating the Role of Genetic Background and  
Niche of Origin in the Adaptation of *Pseudomonas  
aeruginosa***

by

**Nicole J. Filipow**

A thesis submitted to the Faculty of Graduate and Postdoctoral  
Affairs in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology, with a Specialization in Data Science

Carleton University  
Ottawa, Ontario

© 2017, Nicole J. Filipow

## Abstract

Prolonged respiratory infection by the opportunistic pathogen, *Pseudomonas aeruginosa*, is the major contributor to declining lung function and early mortality in individuals with cystic fibrosis (CF). Patients are believed to acquire their infections from the environment, upon which a single clone will often undergo long-term colonization in response to selection within the lung. It is unknown however, whether all genotypes are capable of causing infection, and whether all environmental strains are initially maladapted to life in the lung. To investigate these unknowns, we experimentally evolved 18 different environmental and CF-clinical genotypes within synthetic CF lung sputum (SCFM), to identify any genotypic and/or prior niche constraints on trait evolution and pathoadaptation. We found that genotype significantly constrained evolution within SCFM, which was evidenced through phylogenetic signal in the change of both traits and fitness. The traits showing the highest phylogenetic signal in this experiment included H<sub>2</sub>O<sub>2</sub> resistance, biofilm formation and twitching. We also found that environmental strains evolve differently in SCFM compared to their CF-clinical counterparts. Environmental strains significantly decreased in H<sub>2</sub>O<sub>2</sub> resistance, pigment production, and swimming whereas clinical strains slightly increased or remained unchanged. Furthermore, while the environmental strains often underwent greater fitness leaps than the clinical strains, this was not true for all strains. We hypothesize then that particular niches may better prime strains for causing infection over others. Altogether our results suggest there is widespread variation in the adaptation of *P. aeruginosa*, which is in part constrained by both genetic background, and niche of origin.

# Acknowledgements

Dr. Alex Wong first and foremost for allowing me to grow under his kind, and patient supervision over the past years. I will forever be thankful for his guidance and mentorship that made this a worthwhile journey, of which I will proudly look back on.

Dr. Rees Kassen for his scholarly wisdom, helpful discussions, and foresight that enabled this project to expand and have the meaning I had hoped for.

My committee members and all of the lab members that contributed to this project, and the overall completion of my degree. Alana Schick specifically for assisting in the outline of my lab work.

---

My parents for their never-ending pride in me, and their all-consuming support in each of my endeavors. This would not have been possible without the optimism and weekly advice from my dad, and the enthusiasm, motivational talks, and frozen food from my mum. They are to thank for every accomplishment that I have been so fortunate to have the opportunity for.

Laura and Clare for all of the laughs, chats, messages, and visits that kept me lighthearted and happy - or that at least kept “passionate Nicole” under control. Also Paul, for teaching me coding hacks and helping me with Bayes Traits.

---

But most importantly, Verity - my other half in this life. Without her, there would be no motivation in anything that I do, as she is my inspiration for everything. No one else in the world is as unlucky, yet is as spirited, as she is. May research like this allow her to live the carefree life that she *so* desperately deserves.

# Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>Table of Contents .....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>vi</b>
<b>List of Figures.....</b>	<b>vii</b>
<b>Statement of Contributions.....</b>	<b>ix</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.0 Cystic Fibrosis.....	1
2.0 <i>Pseudomonas aeruginosa</i> in CF Lung Infections .....	3
3.0 Adaptation to a CF Lung.....	4
4.0 Heterogeneity in Selection .....	6
5.0 Epistasis Influences the Availability and Fitness Effects of Beneficial Mutations.....	8
6.0 Influence of Epistasis on Adaptive Evolution within a Fitness Landscape .....	10
7.0 Historical Contingency in Evolution.....	12
8.0 Phylogenetic Signal as a Tool to Test for Effects of Genetic Background.....	14
9.0 Niche-Specific “Pre-Adaptations” .....	17
10.0 Experimental Objectives .....	18
<b>Chapter 2: Materials and Methods .....</b>	<b>21</b>
1.0 Bacterial Strains and Media .....	21
2.0 Experimental Evolution .....	25
3.0 Phenotypic Assays .....	26
3.1 Motility .....	27
3.2 Biofilm Formation .....	27
3.3 Exotoxin Production .....	28

3.4	Antibiotic and Oxidative Stress Resistance .....	28
4.0	Competitive Fitness Assay .....	29
5.0	Statistical Analyses .....	30
5.1	Phylogenetic Signal .....	30
5.2	Phylogenetic Comparative Analyses of Niche.....	31
5.3	Phylogenetic Comparative Analyses of Trait Variations.....	34
5.4	Ancestral and Evolved Population Comparisons.....	35
<b>Chapter 3:</b>	<b>Results.....</b>	<b>38</b>
1.0	Evolution of Pathoadaptive Traits and Fitness.....	38
2.0	Correlations Between $\Delta$ Traits .....	40
3.0	Phylogenetic Signal in the Propensity to Evolve and Adapt.....	43
4.0	Niche Effects in the Evolution of Traits and Fitness .....	46
5.0	Clinical Relevance of Trait Changes.....	55
6.0	Evolutionary Constraint of Ancestral Trait.....	59
7.0	Trait Predictors of Fitness .....	63
8.0	Single Isolates vs Whole Populations .....	68
9.0	Accuracy of Trait Data.....	70
<b>Chapter 4:</b>	<b>Discussions, Conclusions and Implications .....</b>	<b>72</b>
1.0	Genetic Background and Niche Influence Trait Evolution and Adaptation.....	72
2.0	The Adaptive Landscape of SCFM: A Hypothesis .....	73
3.0	A Single Strain is not the Whole Story .....	74
4.0	Implications for CF Lung Infections.....	74
<b>References</b>	<b>.....</b>	<b>76</b>
<b>Appendices</b>	<b>.....</b>	<b>91</b>
A	<i>Niche comparisons excluding En Jp119 and En Jp129</i> .....	91
B	<i>Mini literature review of Pa14 and PaO1</i> .....	94

## List of Tables

<b>Table 1.</b> Ancestral <i>P. aeruginosa</i> Strains.....	22
<b>Table 2.</b> SCFM Composition .....	24
<b>Table 3.</b> Summary of analyses that correct for the non-independence of species in a phylogenetic framework .....	36
<b>Table 4.</b> Phylogenetic signal in the change in traits and fitness, measured with three different tests.....	44
<b>Table 5.</b> Correlational analysis of niche with $\Delta$ trait.....	49
<b>Table 6.</b> Corresponding pGLS parameters for Figure 10: $\Delta$ trait as a function of niche. .	51
<b>Table 7.</b> Corresponding pGLS parameters for Figure 13: $\Delta$ trait as a function of ancestral trait. ....	61
<b>Table 8.</b> Corresponding pGLS parameters for Figure 14: fitness as a function of trait...	65
<b>Table 9.</b> Correlational analysis of fitness with trait.. ..	66

## List of Figures

<b>Figure 1.</b> Age associated lung infection by different microorganisms in Canadian cystic fibrosis patients in 2016.....	2
<b>Figure 2.</b> The possible epistatic interactions between hypothetical mutation A and hypothetical mutation B on the same genetic background. ....	9
<b>Figure 3.</b> Phylogenetic signal in the distribution of traits.....	15
<b>Figure 4.</b> Hypothetical fitness landscapes underlying the adaptation of <i>P. aeruginosa</i> to SCFM.....	20
<b>Figure 5.</b> Maximum likelihood phylogeny that depicts the shared evolutionary history of the ancestral <i>P. aeruginosa</i> strains. ....	23
<b>Figure 6.</b> Maximum likelihood phylogeny that depicts the shared evolutionary history of environmental (En) and clinical (CF) <i>P. aeruginosa</i> strains. ....	32
<b>Figure 7.</b> Heat map of $\Delta$ trait values, within a phylogenetic framework.....	39
<b>Figure 8.</b> Phylogenetic-corrected correlation matrix between $\Delta$ traits.....	41
<b>Figure 9.</b> Independent contrast analysis directly comparing environmental and clinical strains within a phylogeny, in their averaged $\Delta$ trait value across the evolved replicates of each ancestral strain. ....	47
<b>Figure 10.</b> pGLS models of $\Delta$ trait as a function of niche.. ....	50
<b>Figure 11.</b> Phylogenetic-corrected PCA of $\Delta$ traits comparing environmental and clinical niches. ....	54
<b>Figure 12.</b> Phylogenetic-corrected PCA of multivariate trait and fitness data comparing ancestral and evolved populations of environmental and clinical strains.....	56
<b>Figure 13.</b> pGLS regression models of $\Delta$ trait as a function of ancestral trait.....	60

**Figure 14.** pGLS regression models of fitness as a function of trait..... 64

**Figure 15.** Comparison of the change in fitness (s) between single isolates and the whole populations from which they were founded..... 69

## **Statement of Contributions**

In this thesis, Dr. Alex Wong, Dr. Rees Kassen, and Nicole Filipow crafted the experimental design. Nicole Filipow carried out all laboratory experiments, and all statistical analyses (detailed below). Dr. Jeremy Dettman completed the sequencing for construction of the phylogenies.

### **Wet lab experiments:**

- Experimental Evolution
- Phenotypic Assays
- Competitive Fitness Assay

### **Statistical Analyses in R:**

- Phylogenetic Signal Analysis
- Phylogenetic-Corrected Principal Component Analyses
- Phylogenetic Generalized Least Squares Analyses
- Graphing Visualizations

### **Statistical Analyses in Bayes Traits:**

- Phylogenetic-Corrected Correlational Analyses

### **Statistical Analysis in MEGA7:**

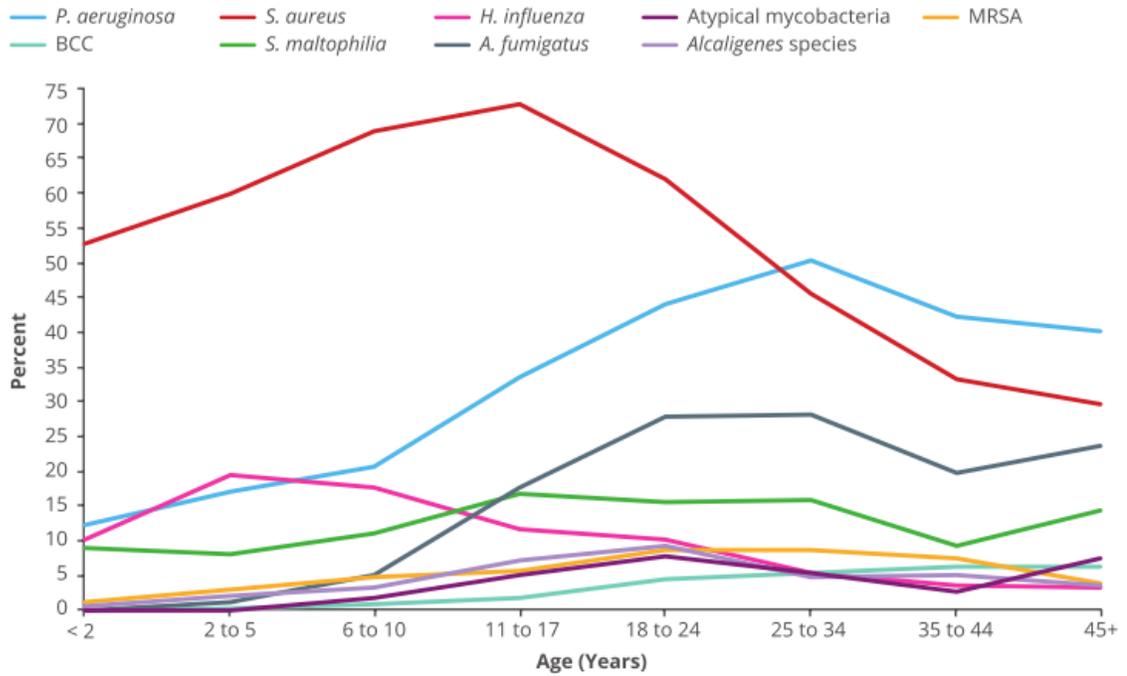
- Construction of the Phylogenies

# Chapter 1: Introduction

## 1.0 Cystic Fibrosis

Cystic fibrosis (CF) is the most common lethal recessive genetic disorder amongst Caucasians, affecting an estimated 4200 Canadians (O'Sullivan and Freedman 2009; Cystic Fibrosis Canada 2017). The underlying mechanism behind the disease is a defect in the transportation of chloride ions across cell membranes due to heritable mutations in the CF transmembrane conductance regulator (CFTR) gene (Cheng et al. 1990). While the symptoms of this defect are multifaceted and develop in many organ systems, the most serious are chronic infections of the lungs whereby concomitant respiratory complications account for ~80% of all deaths (Lyczak et al. 2002; Cystic Fibrosis Foundation 2015). The fundamental cause for frequent and chronic bacterial infection of CF lungs is still not fully understood, but it is believed that dehydration of airway surfaces due to the defective Cl<sup>-</sup> secretion results in thickened mucus and distorted epithelial cilia (Clunes and Boucher 2007). In turn, mucociliary clearance is hampered and invading bacteria are allowed to colonize and flourish within the lungs (Clunes and Boucher 2007).

Microorganisms that are frequently and repeatedly identified as causes of CF lung infection include *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Haemophilus influenzae*, *Stenotrophomonas maltophilia*, *Aspergillus fumigatus* and atypical mycobacteria (O'Sullivan and Freedman 2009). Interestingly, the balance between microbial populations fluctuates with age, and those that are less persistent in children can become predominant in adults (Figure 1) (Cystic Fibrosis Foundation 2015; Cystic Fibrosis Canada 2017). This trend is most prominently observed in the prevalence



**Figure 1.** Age associated lung infection by different microorganisms in Canadian cystic fibrosis patients in 2016. Figure obtained from The Canadian Cystic Fibrosis Registry 2016 Annual Data Report, with permission from Cystic Fibrosis Canada (Cystic Fibrosis Canada 2017).

of *P. aeruginosa*, a bacterium that has exclusively been regarded as the chief contributor to declining lung function and early mortality in individuals with CF (Lyczak et al. 2002).

## **2.0 *Pseudomonas aeruginosa* in CF Lung Infections**

Respiratory infection by *P. aeruginosa*, a Gram-negative opportunistic pathogen, has long been a significant predictor of health in individuals with CF. Children and adolescents that harbour strains are at a substantially higher risk of reduced lung function and mortality (Emerson et al. 2002; Konstan et al. 2007). In adults, infection by a single clone often develops into a chronic strain, and is associated with a rapid decline in lung function, and the onset of pulmonary exacerbations followed by their hospitalizations and associated healthcare costs (Emerson et al. 2002; Aaron et al. 2004; Flume et al. 2009; Cohen-Cymbberknoh et al. 2011).

Despite extensive therapeutic intervention, *P. aeruginosa* persists in infections as a result of high levels of resistance to host defenses and front line antibiotics (Lyczak et al. 2002). Intrinsic antibiotic resistance in *P. aeruginosa* is attributed to its exceptionally low membrane permeability, high levels of efflux pumps, and chromosomally encoded  $\beta$ -lactamase (Hancock and Speert 2000). Resistance has also been shown to readily evolve, as a result of mutations in canonical resistance genes (Smith et al. 2006; Wong and Kassen 2011; Yang et al. 2011; Wong et al. 2012; Marvig et al. 2015). Furthermore, the ability of *P. aeruginosa* to form biofilms allows it to evade both immune cells and extracellular stressors (Bjarnsholt et al. 2009). Early prevention and therapeutic measures are thereby imperative in deterring the bacterial transition to chronic pathogenicity, which has concurrently been proven to maintain patient lung function and increase life expectancy (Valerius et al. 1991; Saiman et al. 2014).

Long-term prevention of *P. aeruginosa* infection is challenging, and between 50-80% of CF adults harbor chronic strains (Cystic Fibrosis Foundation 2015; Cystic Fibrosis Canada 2017). The bacterium is ubiquitous in nature which means that patients are frequently in contact with it, despite efforts at extensive infection prevention measures, which are in place both in home and hospital settings (Saiman et al. 2014). Strains have been isolated from soils, plants, standing water, hospital and household surfaces, and most consistently from the drains and faucets of bathrooms (Regnath et al. 2004; Remold et al. 2011; Purdy-Gibson et al. 2015). This expansive habitat versatility of *P. aeruginosa* has been attributed to its complex and large genome size, which encodes numerous regulatory systems, membrane transport schemes, and virulence factors (Stover et al. 2000). While clones such as the Liverpool Epidemic Strain (LES) have been identified in spreading severe infection between patients (Salunkhe et al. 2005), CF clinical isolates are typically interspersed within a phylogeny, suggesting a non-clonal epidemic nature for the majority of infections (Dettman et al. 2013). Furthermore, the presence of unique genotypes in different patients suggests that infections originate from the environment (Burns et al. 2001). Elucidation of the evolutionary processes that drive adaptation of *P. aeruginosa* to the CF lung is thus necessary to inform better treatment and preventative measures.

### **3.0 Adaptation to a CF Lung**

During the course of infection, a single clone will undergo distinctive genetic mutations and marked phenotypic alterations that enable its long-term colonization of the lung (Struelens et al. 1993). Loss of motility, multiple virulence factors, and quorum sensing are typically seen in chronically infecting strains, as is the gain or enhancement

of biofilm formation, mucoid colonies, hypermutators, auxotrophy, and antibiotic and oxidative stress resistance (Hogardt and Heesemann 2010; Winstanley et al. 2016). As the changes in these specific traits correlate with chronic pathogenesis in a CF lung, they are accordingly termed CF pathoadaptive phenotypes.

Both longitudinal genomic studies and laboratory evolution experiments have frequently and repeatedly observed particular mutations and their associated pathoadaptive phenotypic change, where it is believed that selection within the CF lung is driving the adaptation of *P. aeruginosa* to a chronic, “clinical” phenotype of high fitness (Smith et al. 2006; Yang et al. 2011; Wong et al. 2012; Marvig et al. 2015). Profound phenotypic and genotypic diversity of a single infecting clone has also been identified. This is hypothesized to be in response to the spatial heterogeneity within a lung, which leads to the coexistence of niche-specialists (Clark et al. 2015; Winstanley et al. 2016). The transition from environment to lung is thereby a complex adaptive process that requires high bacterial evolvability in order for *P. aeruginosa* to persist and progress into a chronic strain.

It is unclear whether all strains of *P. aeruginosa* possess the same potential for adaptation to the CF lung. Although studies have shown that between 63 - 98% of children with CF become transiently infected with *P. aeruginosa* (Burns et al. 2001; Zemanick et al. 2015), and up to 80% of CF adults harbor chronic strains (Cystic Fibrosis Foundation 2015), there is profound variation in infection characteristics. For instance, the time of initial *P. aeruginosa* acquisition across individuals can differ by years (West et al. 2002; Baussano et al. 2006), there is a striking gradient of severity across infections, and some patient sputum cultures never test positive for *P. aeruginosa* (Cystic Fibrosis

Foundation 2015). Furthermore, Schelstraete et al. (2008) found that only 9 out of 50 (18%) newly infected individuals bore a genotype that matched one from their household, and Heirali et al. (2016) similarly concluded that the home environment was a rare reservoir for infection, despite the prolific isolation of *P. aeruginosa* strains from the test subject's homes. The variations in infection characteristics may be a result of evolutionary constraints on the bacterial ability to adapt to a CF lung, which are conceivably a result of the genetic diversity in *P. aeruginosa*, and the vast heterogeneity across CF patients; which may enable some individuals to better clear infections than others.

#### **4.0 Heterogeneity in Selection**

The lung environment is especially harsh and provides a unique challenge for invading bacteria. The CF lung is characterized by a dehydrated airway layer and its associated mucus buildup, chronic inflammation, low pH, a unique nutritional composition, prescribed antibiotics, oxidative stress, competing microbes, and low oxygen levels (Folkesson et al. 2012; Bhagirath et al. 2016). Moreover, this is a diverse and dynamic environment that fluctuates dramatically over time and space within the lung (Folkesson et al. 2012; Willner et al. 2012). While the specific selection pressures acting directly on certain genes and phenotypes are not wholly known, the CF lung thus presents a complex and challenging environment in which *P. aeruginosa* adapts.

Over 1500 mutations in the CFTR gene have been identified as causing CF, and different sequence variations can result in modified amounts and functionality of the transcribed CFTR protein (Bobadilla et al. 2002). This is important, because different CFTR alleles can result in altered phenotypes across patients, especially in terms of lung

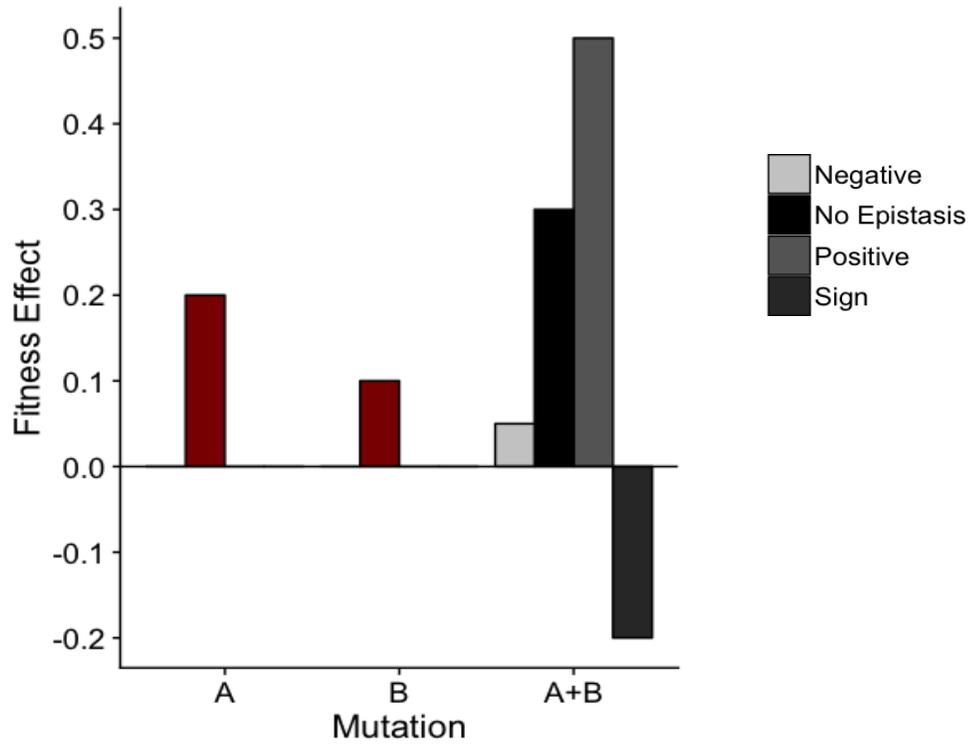
function, mucociliary clearance, immune response, and nutritional status (McKone et al. 2003; Rowntree and Harris 2013). Patients are also given varying types, doses and administrative routes of antibiotics across all age groups depending on their clinical status (Chmiel et al. 2014). In conjunction, the competing microbial constituents of the CF lung is different from patient to patient (Cystic Fibrosis Canada 2017). The selection pressures imposed on invading bacteria can therefore be quite variable between lung environments given the patient differences.

Furthermore, a specific CFTR allele can have direct clinical implications, and distinct CFTR mutations have been directly linked to consequential effects in pancreatic insufficiency (Kristidis et al. 1992). The influence of the CFTR allele on the phenotype of lung disease is less explicit, where specific mutational class rather than the mutational locus has been better attributed (de Gracia et al. 2005). Polymorphisms in modifier genes and pathways outside of the CFTR have also been associated with worsened lung disease, although it is unknown whether this is specifically a result of *P. aeruginosa* infection (Drumm et al. 2005; O’Neal et al. 2015). However, specific to *P. aeruginosa* infection, one study did find a significant association between four particular  $\Delta F508$  heterozygotes and lower rates of *P. aeruginosa* infection, than  $\Delta F508$  homozygotes (McKone et al. 2003). The existence of variable lung environments across patients from either a phenotypic or direct genotypic foundation can thereby contribute to the variations in infection statistics, in addition to the possible influence of genetic diversity in *P. aeruginosa*.

## 5.0 Epistasis Influences the Availability and Fitness Effects of Beneficial Mutations

Genetic background can profoundly shape the evolutionary trajectory of microorganisms (Tenailon et al. 2012; Wong 2017), and is a potential influence on the adaptive evolution of genetically diverse *P. aeruginosa* strains to the lungs of different CF patients. Epistasis is an important mechanism whereby genetic background influences patterns of evolution. Notably, epistasis has been found to have significant consequences for the mutational fitness effects and adaptive evolution of *P. aeruginosa* in various environments (Hall et al. 2010; Gifford et al. 2016; Vogwill et al. 2016).

Epistasis is the genomic phenomenon in which the combined fitness effect of mutations at different loci is different than what would be expected from the fitness effect of each mutation alone (Phillips 2008). Epistasis between two mutations can result in a genotype of higher, or lower fitness than expected; scenarios which are respectively termed positive and negative epistasis (Figure 2) (Phillips 2008). By contrast, sign epistasis occurs when the combined outcome of two beneficial mutations is a deleterious fitness effect, or vice versa (Figure 2) (Phillips 2008). Positive and negative epistasis have important consequences for the evolution of genetically diverse species, as a given mutation may result in varying magnitudes of fitness effects depending on the genetic background it arises in (Hietpas et al. 2011). As an example, two genetically distinct *P. aeruginosa* populations accumulated the same mutations when evolved in an identical environment, but each plateaued onto a separate fitness optimum (Gifford et al. 2016). This suggests the existence of profound epistatic interactions between mutation and genetic background. This is also true for the fitness effects of antibiotic resistance



**Figure**

2. The possible epistatic interactions between hypothetical mutation A and hypothetical mutation B on the same genetic background.

mutations in *P. aeruginosa*, which have been found to vary significantly across genetic backgrounds (Maclean and Buckling 2009; Vogwill et al. 2016).

Sign epistasis is also important in shaping evolutionary pathways, and may mean that a mutation is only beneficial on some backgrounds while detrimental on others (Weinreich et al. 2005; Wong 2017). The evolutionary trajectory between two genotypes will therefore be markedly different if a mutation is only beneficial and fixed in one background. For example, Tenaillon et al. (2012) found that populations of *Escherichia coli* evolved in high temperatures followed two distinctive and unique adaptive pathways from each other depending on an initial mutation arising in either a RNA polymerase or a termination factor. In this instance, the beneficial mutations fixed on the RNA polymerase background likely have epistatic interactions of a sign effect with the termination factor background, and vice versa.

## **6.0 Influence of Epistasis on Adaptive Evolution within a Fitness Landscape**

Evolution is often pictured as movement on a fitness landscape of adaptive peaks and maladaptive valleys within a genotypic or phenotypic space (Orr 2005). Populations closer to a given peak are more fit than those that are further away, and different peaks represent different ways of achieving high fitness (Orr 2005). As mutations arise, those that are beneficial may become fixed as they move the organism closer towards a fitness optimum. Also known as an adaptive walk, the order and identity of fixed mutations will depend on their epistatic interactions both with the genetic background they arise in, and with one another (Hietpas et al. 2011; Bank et al. 2015). The identity and number of adaptive walks available to a given genotype is ultimately limited by epistasis, which

allows only for the accumulation of particular combinations of genes and alleles to arrive at an adaptive peak (Wong 2017).

The interaction between genotype and environment will influence the topology of a fitness landscape, which may be of a rugged form with multiple peaks and valleys, or a smooth landscape of a single, global adaptive peak (Kauffman and Weinberger 1989). In a smooth landscape, maladaptive genotypes will sit further from the global optimum and require either a longer or more rapid adaptive walk to reach the fitness peak than genotypes that are well suited for the new environment. Here, diminishing returns epistasis is frequently observed, whereby mutations become less beneficial the closer the genotype is to an adaptive peak (Chou et al. 2011). This results in an inverse correlation between the rate of the adaptive walk and the number of accumulated beneficial mutations. The rate of adaptation will then differ between genotypes depending on their relative position to the adaptive peak (Chou et al. 2011). Alternatively in a rugged landscape, each genotype may be situated closest to its own local optimum where the combination of genes at each adaptive peak corresponds to variations in fitness optima (Levinthal 1997). In this case, the evolved fitness will be contingent on the ancestral genotype.

We postulate that the fitness landscape underlying the adaptation of *P. aeruginosa* specifically to a CF lung is smooth. This is suggested by the convergent evolution of CF pathoadaptive phenotypes across individual infecting clones, and the similarities between chronic strains. It seems reasonable then, that strains are climbing a single adaptive peak, where the combination of genes/phenotypes that corresponds with high fitness at the peak is the combination of pathoadaptive traits associated with chronic infection. Moreover,

the landscape underlying the adaptation to each CF lung may differ slightly given the patient differences. The reported within-patient diversity in infecting clones may also represent sub-populations climbing this peak with different levels of success, or may alternatively suggest that sub-populations are climbing different peaks in a rugged landscape.

## **7.0 Historical Contingency in Evolution**

The terms “historical contingency”, “phylogenetic inertia”, and “ancestral constraint” are often used interchangeably to describe the phenomenon by which precise evolutionary outcomes are only possible in particular starting genetic backgrounds (Blomberg and Garland 2002). In other words, evolution may be contingent on, or constrained by, a population’s history. It is theorized that history and adaptation are inversely correlated, whereby strong selection removes any effect of history and populations may reach the same adaptive peak in a fitness landscape regardless of variations in genetic background (Travisano et al. 1995). If selection is contrarily weak, then the evolutionary outcome of populations will be constrained by their history and where in a fitness landscape they start. Stochastic effects such as mutation and drift will in turn result in variation around the population mean (Travisano et al. 1995).

Historical contingency in the evolution of organisms has previously been investigated through microbial experimental evolution studies, by direct assessment of evolved populations compared to its ancestors. Often, it appears that ancestry has no effect on the final evolved fitness of different, but related genotypes, where each strain converges onto the same fitness optimum (Travisano et al. 1995; Rokyta et al. 2009; Gifford et al. 2011; Wong and Seguin 2015). In conjunction with observed adaptive

convergence, there is evidence that the rates of substitution and the tempo of adaptation vary significantly by genotype in experimentally evolved *E. coli* populations, which is strongly dependent on the ancestral fitness (Moore and Woods 2006; Wong and Seguin 2015). Maladapted genotypes thereby appear to have higher rates of adaptation, as a result of being situated further from an adaptive peak. Furthermore, Gifford et al. (2011) identified adaptive convergence as a result of large-effect mutations in the less adapted populations, and Bollback & Huelsenbeck (2009) found significant differences in the relative fitness gains of independently evolved populations. Altogether, variable rates and fitness effects of adaptive mutations, which appear to be largely influenced by ancestry, allow genetically distinct populations to reach the same fitness optimum in these experimental studies.

However, historical contingency rather than complete adaptive convergence has been implicated in other studies, and may therefore be context dependent. The constraint of ancestral genotype on endpoint fitness and trait evolution has been highlighted in genetically distinct populations of bacteriophage, bacteria and yeast (Bull et al. 2004; Hall et al. 2010; Spor et al. 2014). Melnyk & Kassen (2011) also showed that historical contingency may in fact be environment specific, where different genotypes of *Pseudomonas fluorescens* evolved in glucose appeared to converge on the same fitness optimum, while there is evidence of divergence when evolved in xylose. Strikingly, two *P. aeruginosa* populations of differing genetic backgrounds evolved in M9-serine media reached two different adaptive peaks, despite an identical starting fitness (Gifford et al. 2016). The genotype by environment interactions in these studies suggests that the

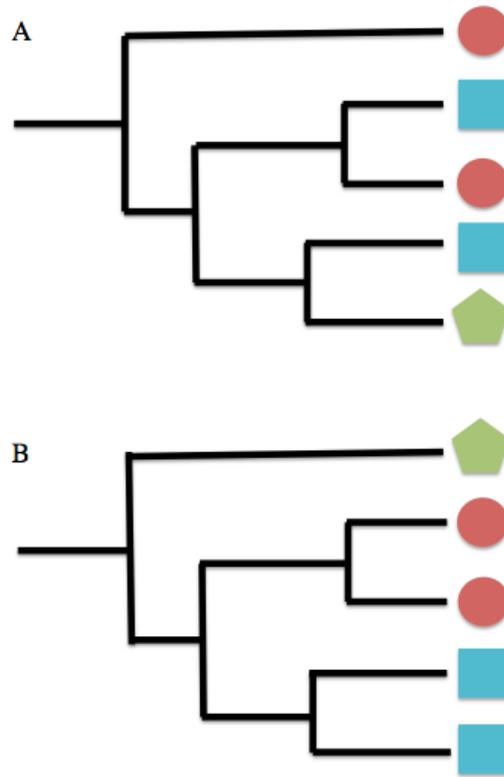
underlying adaptive landscape is rugged, and that evolved fitness is constrained by ancestry.

The described microbial evolution experiments accurately and directly ask the question “does evolved fitness depend on ancestral fitness?” These studies are limited however, in their regular use of 4 to 8 independently evolving isolates that often differ only by point mutations. They thereby discount the influence of shared phylogenetic history in evolution.

## **8.0 Phylogenetic Signal as a Tool to Test for Effects of Genetic Background**

In addition to direct fitness observations of ancestral and evolved populations, the effect of genotype on trait and fitness evolution can be measured by phylogenetic signal when relationships between the ancestral populations are known. Used largely in measuring patterns of trait distribution within the animal and plant kingdoms, phylogenetic signal is the tendency for related species to resemble each other (Blomberg et al. 2003)

In the absence of selection, traits are expected to evolve as a function of time along the branches of a given tree. Under standard phylogenetic models, traits are modeled under the assumption of Brownian motion (BM), with a small change in direction per unit time (Blomberg et al. 2003). For closely related species, a given trait will have had little time to diverge under a BM model (Figure 3B). By contrast, for species that are distantly related and share less phylogenetic history, a trait will have had much more time to diverge (Figure 3A). Variance in trait value will therefore be much greater for more divergent species. Traits whose variance between species is predicted by relatedness show a strong phylogenetic signal. Low phylogenetic signal, conversely,



**Figure 3.** Phylogenetic signal in the distribution of traits. **A)** When trait distribution does not follow a BM model of evolution and is thereby not predicted by phylogeny, phylogenetic signal is negligible. Pagel's  $\lambda$  and Blomberg's  $K$  will closely equal 0, and Abouheif's  $C_{\text{mean}}$ , and Moran's  $I$  will lie somewhere between -1 and 0. **B)** When phylogeny explains the variance in traits, phylogenetic signal is strong. Pagel's  $\lambda$ , Blomberg's  $K$ , Abouheif's  $C_{\text{mean}}$ , and Moran's  $I$  will closely equal 1.

means that variance in trait values distribution is not predicted by phylogeny (Blomberg et al. 2003). In this way, phylogenetic signal is a direct test of genetic background in the dispersal of traits.

To date, Pagel's  $\lambda$ , Blomberg's  $K$ , Abouheif's  $C_{\text{mean}}$ , and Moran's  $I$  are the four most consistently used and scrutinized methods to test for phylogenetic signal (Münkemüller et al. 2012). Pagel's  $\lambda$  is a scaling measure for the observed correlations between species relative to the correlations between species expected under BM (Pagel 1999). Blomberg's  $K$  is the ratio of the mean squared error of the tip data, over the mean squared error of the data drawn from the covariance matrix of the given phylogeny. This is then standardized against the expected ratio under BM (Blomberg et al. 2003). Abouheif's  $C_{\text{mean}}$ , and Moran's  $I$  are related autocorrelation models adapted for phylogenetic inferences, used to measure the serial independence of species given the tree topology (Moran 1950; Abouheif 1999). Specified values given by each test are described in Figure 3.

Although rare, studies in bacteria have shown weak to strong phylogenetic signal in the distribution of numerous traits including prototrophy, optimal temperature and pH growth, biofilm, and motility (Martiny et al. 2013; Krause et al. 2014; Rodriguez-Torres et al. 2017). The strength of phylogenetic signal in microbial adaptive evolution however, has not previously been investigated. We propose that testing for phylogenetic signal in the propensity of strains to evolve and adapt is a worthwhile tool, and may reveal variations in the adaptation of genetically diverse populations.

## 9.0 Niche-Specific “Pre-Adaptations”

While phylogenetic signal is an indication of the overall impact of genetic background on evolutionary outcomes, more specific predictions can be made regarding the effects of ecological niche on trait evolution. All organisms occupy some ecological space and will evolve in response to their environments. Depending on the specific selection pressures within each niche, a population may acquire different adaptations in response to those variable pressures. For example, genomic analysis of *Pseudomonas putida* from different environments revealed strong associations between strains from particular niches with genes involved in the functions required for those specific environments (Wu et al. 2011). These included functions such as heavy metal resistance, degradation of organic solvents, and high oxidative stress resistance (Wu et al. 2011).

Niche-specific adaptations of environmental and CF clinical *P. aeruginosa* strains have previously been investigated by direct assessment of the genes and phenotypes that are exclusive to each. Interestingly, marked differences between niches are often not observed and regardless of their environmental or CF lung origin *P. aeruginosa* strains display evidence of antibiotic resistance, ability to invade epithelial cells, metabolism of oil hydrocarbons, type III secretion system, and virulence genes (Alonso et al. 1999; Feltman et al. 2001; Martins et al. 2014). Furthermore, there is profound diversity across the studied strains in their capacities for these functions. This suggests that *P. aeruginosa* strains are not drastically limited by niche-specific adaptations, and may carry the genomic and phenotypic potential to readily transfer between niches. Some studies have even found that environmental isolates were more resistant to antibiotics than CF clinical isolates, and carried higher levels of exotoxins (Gad et al. 2007; Streeter et al. 2016). A

study by Head and Yu (2004) however did find particular niche differences, where CF isolates had higher biofilm formation, and much higher planktonic susceptibility to H<sub>2</sub>O<sub>2</sub>, than the environmental strains. While these studies directly assessed the phenotypes and genes specific to niches, the role of prior niche in the adaptive evolution of *P. aeruginosa* has not previously been investigated.

The influence of prior niche in the adaptation of organisms to a novel environment can also be pictured as movement on a fitness landscape. In the same way that we expect closely related organisms to cluster together in the genotypic space of a fitness landscape, we would expect organisms from the same niche to cluster together on the same level of fitness within a landscape. For example, bacterial strains from an animal infection are predicted to have a higher starting fitness in a human infection model than those strains isolated from soil, since they would have evolved prior niche-specific adaptations that would transfer over to the new environment - for example evolved responses to immune cells, or warm temperatures. We therefore expect populations from the same niche to be situated closer to an adaptive peak in novel environments that are similar to the environment the populations have originated from. The prior adaptations of CF-clinical strains may therefore facilitate the transfer of infecting strains between CF patients.

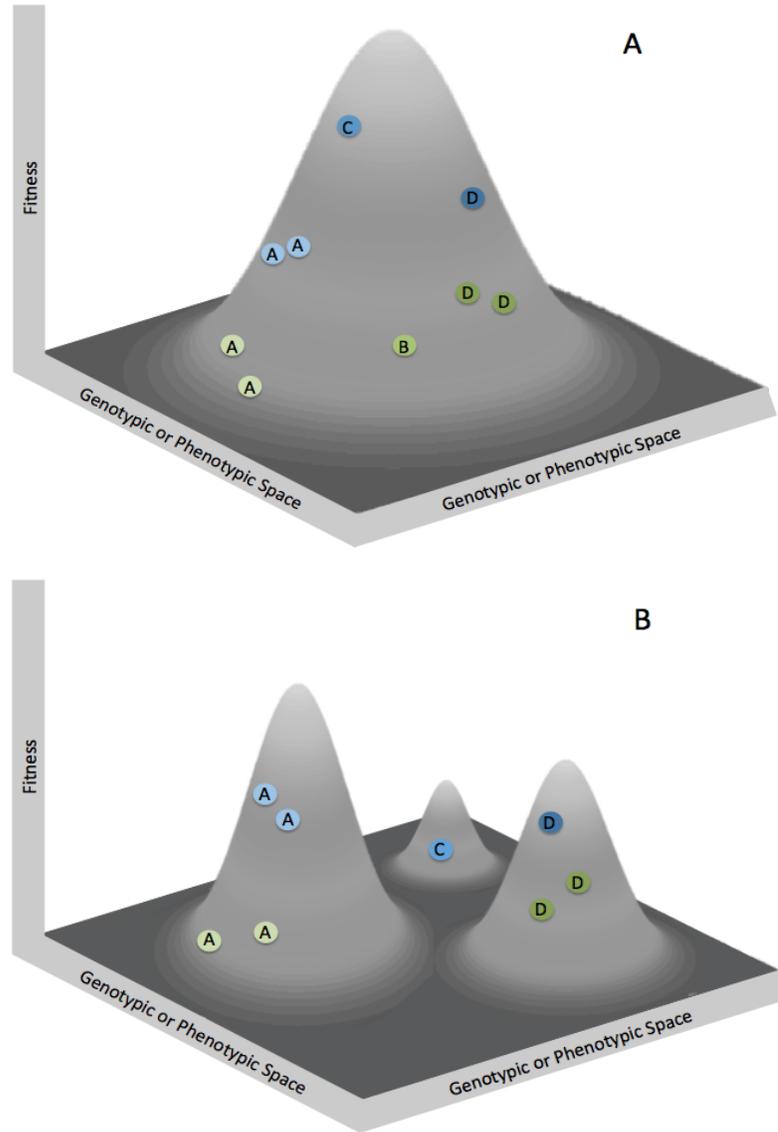
## **10.0 Experimental Objectives**

While convergent evolution has been investigated extensively in the adaptation of *P. aeruginosa* to a CF lung, the use of a single *P. aeruginosa* strain, or the sole sampling of successfully adapting isolates within a CF lung, prevents broad application of previous studies across genetically diverse environmental strains. Given this limitation, our first

goal was to determine if ancestry constrains pathoadaptation to a CF lung in *P. aeruginosa*, by inferring the contribution of phylogeny to phenotypic and fitness changes during adaptation of 18 genetically diverse strains of *P. aeruginosa* to a modified synthetic CF sputum (SCFM). We predicted that adaptation to SCFM is contingent on genotype and where an environmental strain is located in the fitness landscape, which will be evidenced by strong phylogenetic signal in the changes in traits and fitness (Figure 4). This is because closely related strains are more likely to evolve similarly to one another as a result of paralleled epistatic interactions between mutation and genotype. Furthermore, the presence of CF epidemic clones suggests that some genotypes have higher capacities for infection than others.

Our second goal was to determine if niche of origin also constrains pathoadaptation to a CF lung in *P. aeruginosa*. We predicted that all environmental strains are maladapted to SCFM and sit far from an adaptive peak compared to their CF clinical counterparts, given the evidence of selection within the lung in driving the adaptation of infecting clones (Figure 4). Environmental strains will therefore require greater fitness leaps and different phenotypic alterations than clinical strains during adaptation to SCFM, since clinical strains are presumed to have niche-specific adaptations in response to the CF lung environment that also confer high fitness in SCFM.

By investigating the drivers of early adaptation to a CF lung-like environment, we can start to elucidate any existing variation in a strains propensity to infect a CF lung, in order to hopefully bring about more efficient prevention guidelines and improved early therapeutic measures in the CF population.



**Figure 4.** Hypothetical fitness landscapes underlying the adaptation of *P. aeruginosa* to SCFM. It is predicted that closely related strains, for example of the same clades (letters), will cluster together and niche-specific pre-adaptations will mean that environmental (green) strains are overall less adapted than the clinical strains (blue). **A)** The underlying adaptive landscape is postulated to be smooth, whereby the adaptive leaps and trait evolution of each strain will be constrained by genetic background. **B)** Alternatively, the adaptive landscape could be rugged, whereby the fitness optimum available to each strain is constrained by genetic background.

## Chapter 2: Materials and Methods

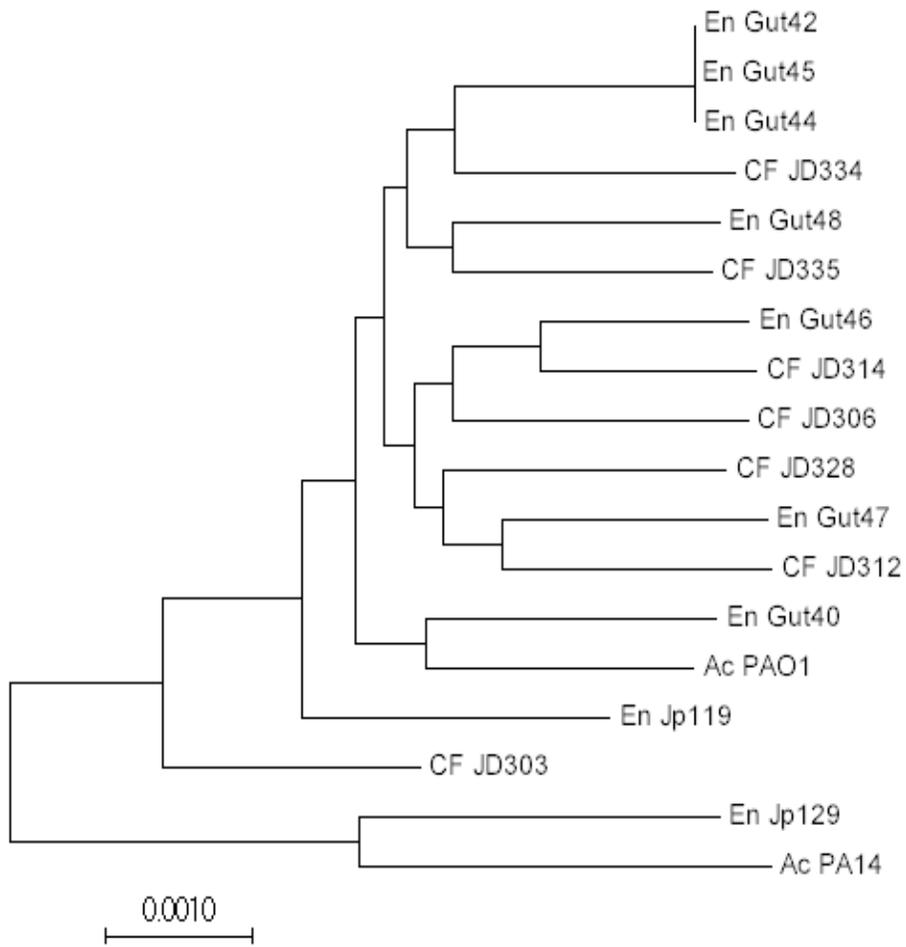
### 1.0 Bacterial Strains and Media

A diverse sample of *P. aeruginosa* strains previously collected from soil, water, and lung sputum samples from chronic CF infection was selected in order to test for the effects of genetic background and niche of origin on adaptation (Table 1). In total, we analyzed 7 clinical strains, 9 environmental strains, and the commonly used laboratory test strains Pa14 and PaO1. A phylogeny was constructed in MEGA7 (Kumar et al. 2016), using the nucleotide sequences of 380 core genes (Figure 5) (Dettman et al., personal communication). Given the high rates of horizontal gene transfer (HGT) in bacteria, inferring the proper phylogeny can be difficult. We therefore compared only the core genes that are present in all strains, and have assumed that these genes have been minimally affected by HGT. While this may not be the case, using an incorrect phylogeny for phylogenetic comparative methods is said to be better than assuming the complete independence of strains, and disregarding phylogeny altogether (Felsenstein 1985). In our construction of the tree, the phylogenetic relationships between the strains show a varied dispersion of environmental and clinical strains throughout the phylogeny (Figure 5). Glycerol (25%) freezer stocks of each strain were generated from isolated colonies grown overnight at 37°C in Lysogeny Broth (LB). These isogenic strains, referred to as the “ancestral” strains, are the parental genotypes from which each evolving population is founded in the selection experiment described below.

To test for the adaptation of each strain specifically to clinically relevant conditions, a modified synthetic CF lung sputum medium (SCFM) was used (Table 2) (Palmer et al. 2007; Turner et al. 2015). This media was created primarily to mimic the nutritional

**Table 1.** Ancestral *P. aeruginosa* Strains

Niche	Strain	Isolation Location	Country	Isolation Date	Source
Environment (En)	Gut_40	King City, ON	Canada	2004	Soil
	Gut_42	King City, ON	Canada	Apr 7/10	Soil
	Gut_44	King City, ON	Canada	Apr 7/10	Soil
	Gut_45	Toronto, ON	Canada	Apr 8/10	Soil/Water
	Gut_46	Maysville, KY	USA	Apr 8/10	Soil
	Gut_47	Maysville, KY	USA	Apr 8/10	Soil
	Gut_48	Maysville, KY	USA	Apr 8/10	Soil/Water
	Jp_129	Pacific Ocean	Japan	2003	Sea Water
	Jp_119	Suruga Bay	Japan	2004	Sea Water
Clinical (CF)	JD_303	Toronto, ON	Canada	Nov/05	CF-patient
	JD_306	Toronto, ON	Canada	Nov/05	CF-patient
	JD_312	Kitchener, ON	Canada	Nov/05	CF-patient
	JD_314	Kitchener, ON	Canada	Jan/06	CF-patient
	JD_328	London, ON	Canada	May/06	CF-patient
	JD_334	Sudbury, ON	Canada	Sep/06	CF-patient
	JD_335	Toronto, ON	Canada	Oct/06	CF-patient
	PaO1				
	Pa14				



**Figure 5.** Maximum likelihood phylogeny that depicts the shared evolutionary history of the ancestral *P. aeruginosa* strains (Table 1). Branch lengths are proportional to the number of substitutions per site

Table 2. SCFM Composition

Group	Compound	Concentration (g/L)
Buffered Base	K <sub>2</sub> SO <sub>4</sub>	0.047
	KCl	1.116
	KNO <sub>3</sub>	0.035
	MOPS	2.092
	Na <sub>2</sub> HPO <sub>4</sub>	0.177
	NaCl	3.032
	NaH <sub>2</sub> PO <sub>4</sub>	0.156
	NH <sub>4</sub> Cl	0.124
Amino Acid Salts	Cysteine	0.028
	Tyrosine	0.145
	Threonine	0.128
	Histidine	0.081
	Phenylalanine	0.088
Amino Acid Solutions *	Alanine	0.159
	Arginine	0.054
	Asparagine **	0.110
	Glutamine **	0.227
	Glycine	0.090
	Isoleucine	0.147
	Leucine	0.211
	Lysine	0.311
	Methionine	0.094
	Ornithine	0.090
	Proline	0.191
	Serine	0.152
	Tryptophan **	0.002
Valine	0.130	
DNA	Double Stranded Salmon Sperm DNA	0.600
Matrix	Bovine Mucin	5.000
Supplemental	MgCl <sub>2</sub>	0.003
	FeSO <sub>4</sub> 7H <sub>2</sub> O	5.00×10 <sup>-5</sup>
	N-Acetylglucosamine	0.003
	CaCl <sub>2</sub> 2H <sub>2</sub> O	0.013
	Dextrose	0.027
	L-lactic acid	0.042

\*1.0M amino acid stock solutions were individually prepared and stored at 4°C for a maximum of 1 month. \*\*Stock solutions were prepared in NaOH

**Note:** Buffered base and amino acids were individually sterilized, combined, corrected to a pH of 6.8 with HCl, and then filter sterilized. Autoclaved biological matrix was added and the mixture was stored at 4°C for a maximum of 1 month. UV sterilized DNA was added the night prior to use, and filter sterilized supplemental ingredients were added immediately prior to use.

environment of the CF lung, which contains an abundance of amino acids, salts, and DNA. We also included bovine mucin as it provides a biological matrix to the media comparable to that of clinical sputum, which is high in mucins and notably viscous (Houdret et al. 1986). Importantly, previous studies have found that *P. aeruginosa* evolved within SCFM + mucin gain comparable mutations to those seen *in vivo*, validating its use as a study model for a cystic fibrosis lung environment (Wong et al. 2012).

## 2.0 Experimental Evolution

In order to evolve each *P. aeruginosa* strain within SCFM, a selection experiment was used in which four evolving replicate populations were founded from each ancestral strain. This was completed by initially seeding four isolated colonies from each ancestral strain individually into a separate 1.5ml aliquot of SCFM, in 24 microwell plates. Every 24 hours of stationary growth at 37°C, 15µl of each population was transferred to fresh SCFM equating to a dilution of 1:100, which results in approximately 6.64 generations per day. Daily transfers ensued for 15 days (~100 generations), where at this time point the final cultures were frozen in 25% glycerol, and designated as the “evolved whole” populations. Furthermore, a single isolate was chosen at random from each evolved whole population, and a frozen glycerol stock of each isolate was created from an overnight culture grown in LB. These strains were labeled the “evolved single” populations, and are referred to as such during the remainder of the experiment.

Multiple tests surrounding the selection experiment were in place to control for and identify any possible contamination. Throughout the transfers, one control included the growth of Pa14-*LacZ* between each test population in a checkerboard pattern within a

separate aliquot of media. After every three days of transfers, the evolving populations were streaked onto M9 minimal media agar (MMA) + X-gal (composition per liter: 6.78g Na<sub>2</sub>HPO<sub>4</sub>, 3.00g KH<sub>2</sub>PO<sub>4</sub>, 0.50g NaCl, 1.00g NH<sub>4</sub>Cl, 0.24g MgSO<sub>4</sub>, 0.01 CaCl, 8.00g glucose, 500μl X-gal, 15g agar). By the absence of PA14-*LacZ* (blue colonies) all populations were deemed free from inter-well contamination, with the exception of two replicates (34-4 and 47-4). These two populations were subsequently removed from the study. In addition, all evolved populations were streaked onto *Pseudomonas* Isolation Agar (PIA), MacConkey agar, Cetrinide agar, and M9 MMA + IPTG + X-gal, to ensure that each *P. aeruginosa* population had not been outcompeted by any possible contaminants, and that each population was free from any invading *E. coli* contaminants. In two clinical strains originally incorporated into the study, contamination was observed via these methods and the strains were therefore excluded from any further analysis.

### **3.0 Phenotypic Assays**

Each ancestral strain and its four corresponding evolved single populations were tested for multiple pathoadaptive phenotypes that have been identified by many longitudinal and genomic studies to be enhanced or lost in CF lung infections (Hogardt and Heesemann 2010; Winstanley et al. 2016). Briefly, these included forms of motility, biofilm formation, antibiotic and oxidative stress resistance, and the production of extracellular proteases and virulence factors. Specific assays used to identify each of these traits are described below. Overnight cultures were grown shaking at 150 RPM in either 4ml or 200μl of LB media at 37°C. Incubation (growth) temperatures for each assay were 37°C, unless otherwise specified. In each experiment, a trait value was found

for each evolved and ancestral population, and the change in trait value ( $\Delta Trait$ ) of each evolved strain from its ancestor was additionally calculated (Equation 1).

$$\Delta Trait \pm \sigma_M = Trait_{Evolved} - Trait_{Ancestral} \pm (\sigma_{MTraitEvolved} + \sigma_{MTraitAncestral})$$

**(Equation 1)**

### **3.1 Motility**

For swimming, overnight cultures were stab inoculated onto the centre of 0.3% agar LB plates, in replicates of 3. Plates were left upright and after 24 hours of incubation, the diameter of growth was measured. To quantify twitching motility, overnight cultures were stab inoculated onto 1.5% agar LB plates, in replicates of 4 per plate. After 48 hours of incubation, followed by 24 hours of growth at room temperature, the agar was discarded and plates were stained with 750 $\mu$ l of crystal violet dye (0.1%) for 2 minutes. The excess stain was then rinsed off with water, and the diameter of each stain was measured as an estimate of twitching motility (Rashid and Kornberg 2000).

### **3.2 Biofilm Formation**

In order to first facilitate the growth of a biofilm, overnight culture was grown in wells of 100 $\mu$ l of LB media in replicates of 6, shaking at 150 RPM for 48 hours. Planktonic cells were removed by rinsing with water, and the remaining cells aggregated to the well walls were stained with 125 $\mu$ l of crystal violet dye (0.1%) for 10 minutes. Excess stain was removed with water, and the biofilms were solubilized with 200 $\mu$ l of acetic acid (30%) for 10 minutes. The OD600 of 125 $\mu$ l of the solubilized biofilm was measured as an estimate of biofilm density (Merritt et al. 2005).

### **3.3 Exotoxin Production**

Extracellular proteases were measured by the spot inoculation of 5µl of overnight culture onto skim milk agar plates (composition per liter: 5.0g peptone, 3.0g yeast extract, 8.0g skim milk powder, 15g agar). These were completed in replicates of 4, and after 24 hours of growth the zone of hydrolysis on each plate was estimated from the diameter (Fulzele et al. 2011).

The production two of pigments involved in virulence (pyocyanin and siderophores) was measured as the optical density of culture supernatant per estimated cell count. Specifically, overnight culture was grown in 200µl LB media shaking for 24 hours. Post incubation, the cultures were diluted 1:100 in LB and the OD600 was measured as an estimated number of cells. The remaining overnight culture was centrifuged at 12000 RPM for 5 minutes, and the OD405 and OD695 of the supernatant was read. The ratio of OD405/OD600 was calculated as the measurement of siderophore production, and the ratio of OD695/OD600 was used as a measure of pyocyanin production.

### **3.4 Antibiotic and Oxidative Stress Resistance**

Minimal inhibitory concentrations (MIC's) of selected stressors were used as measures of the bacterial susceptibility to both antimicrobials and oxidative stress. The specific antibiotics tested were ciprofloxacin and tobramycin, which are commonly used to clinically treat both acute and chronic CF lung infections (Chmiel et al. 2013). Hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) and sodium nitrite (NaNO<sub>2</sub>) were furthermore included in the assay as inducers of oxidative stress, in order to measure its resistance in *P. aeruginosa* (Bollag and Henninger 1978; Bjarnsholt et al. 2005).

To test the effects of stressor on bacterial growth, overnight culture was diluted 1:100 in LB, within a 2-fold dilution of stressor, in replicates of three. The concentration ranges of each stressor were as follows: tobramycin 0.0625-32.00 $\mu$ g/ml, ciprofloxacin 0.0078-4.000 $\mu$ g/ml, H<sub>2</sub>O<sub>2</sub> 3.906-2000mM, NaNO<sub>2</sub> 0.375-48.0mg/ml. Upon incubation for 24hrs, the MIC of each compound was identified as the test concentration that inhibited the growth of cells by at least 90%, which was determined from the OD600 (Andrews 2001).

#### **4.0 Competitive Fitness Assay**

In addition to phenotyping, the ancestral and evolved fitness of each single and whole population was analyzed. Fitness was measured using a competitive fitness assay against a common competitor (Pa14-*LacZ*). Both single and whole populations were included in the assay in order to estimate the resemblance of the single isolate to its founding whole population. Furthermore, the competition experiments were carried out under the same conditions as the selection experiment, in order to identify how fitness had evolved in this environment. Specifically, whole, single and ancestral populations, as well as Pa14-*LacZ*, were grown stationary overnight, in 1.5ml SCFM. Cultures were diluted 0.5:100 in fresh SCFM in replicates of 4, with the addition of the Pa14-*LacZ* overnight culture to every replicate, at an even dilution of 0.5:100. A 50 $\mu$ l aliquot of these initial cultures was frozen in 25% glycerol and denoted as the initial sample. After stationary growth occurred for 24 hours at 37°C, another 50 $\mu$ l aliquot was frozen and labeled as the final sample. Serial dilutions of each initial and final sample were completed in 1X Minimal Salts, and the four highest dilutions were spot inoculated in

replicates of 4 onto M9 MMA + X-gal. Blue (Pa14-*LacZ*) and white (single, whole and ancestral strains) colony counts were recorded.

The selection coefficient ( $s$ ) for each replicated competition was calculated as an estimate of fitness for each population (Equation 2).

$$s = \frac{\ln\left(\frac{\textit{White final}}{\textit{White initial}}\right) - \ln\left(\frac{\textit{Blue final}}{\textit{Blue initial}}\right)}{\# \textit{generations}}$$

**(Equation 2)**

The change in fitness ( $\Delta s$ ) for each single and whole population was also calculated (Equation 1), whereby each of the four evolved strain replicates per population was compared against the fitness of its shared ancestral strain.

## **5.0 Statistical Analyses**

All statistical analyses were completed in R Software for Statistical Computing (R Core Team 2013), and Bayes Traits (Pagel 1999). The latter allows for the analysis and estimation of trait evolution in a phylogenetic framework. Graphing visualizations were completed using the ggplot2 package in R (Wickham 2009), with the exception of the  $\Delta$ trait heat map, which was visualized using the phylo.heatmap function of the phytools package in R (Revell 2012).

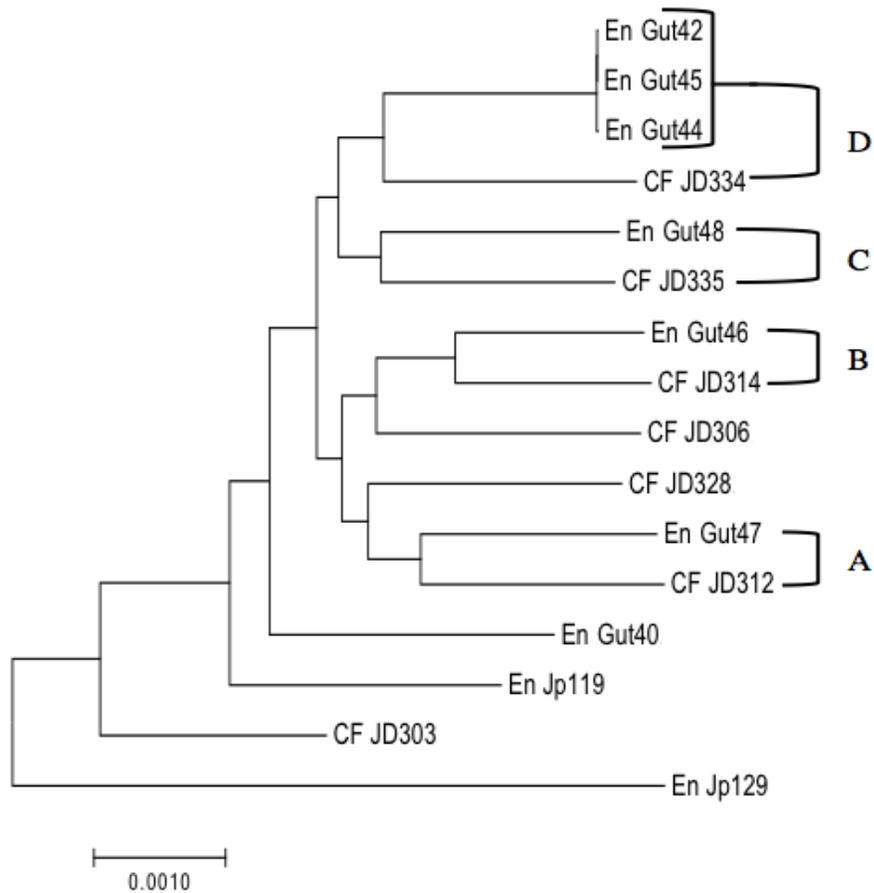
### **5.1 Phylogenetic Signal**

Phylogenetic signal of the change in a single continuous trait was assed by estimating three different parameters: Pagel's  $\lambda$ , Abouheif's  $C_{\text{mean}}$ , and Moran's  $I$ . Pagel's  $\lambda$  was estimated by using an MCMC analysis in Bayes Traits, and was confirmed via a ML analysis using the phytools package in R (Revell 2012). Abouheif's  $C_{\text{mean}}$ , and

Moran's  $I$  were both estimated using the phylosignal package in R (Keck et al. 2016). In all tests, a randomly generated dataset and a manipulated dummy dataset that mimicked a complete phylogenetic distribution of traits were included in the analysis, as a control for both maximal and minimal phylogenetic signal. The phylogeny itself that was included to analyze the distribution of  $\Delta$ traits throughout these tests was a tree with a four-way polytomy at each tip, where each tip corresponded to the evolved replicates of each ancestral strain. Since the methodology behind Blomberg's  $K$  does not allow for the inclusion of polytomies, it could not be estimated in this experiment.

## 5.2 Phylogenetic Comparative Analyses of Niche

In order to test for evolved differences between clinical and environmental strains, we had to acknowledge the non-independent sampling of strains given their shared phylogenetic history (Felsenstein 1985). We therefore carried out and compared three different analyses that correct for phylogeny when comparing individual trait changes between the two niches: a manual independent contrast analysis, a phylogenetic generalized least squares (pGLS) analysis using the nlme library of the ape package in R (Paradis et al. 2004), and a phylogenetic-corrected correlational analysis in Bayes Traits. We also did a fourth analysis: a multivariate  $\Delta$ trait comparison between niches by carrying out a phylogenetic-corrected principal component analysis (pcPCA), using the phyl.pca function of the phytools package in R (Revell 2012). All tests, which are described in detail below and summarized in Table 3, excluded Pa14 and PaO1 in the analysis and therefore corrected for non-independence of samples with the appropriate phylogeny (Figure 6).



**Figure 6.** Maximum likelihood phylogeny that depicts the shared evolutionary history of environmental (En) and clinical (CF) *P. aeruginosa* strains. Branch lengths are equivalent to the number of substitutions per site. Independent contrasts directly comparing niches are labeled A-D.

**1. Manual Independent Contrasts:** This analysis was carried out first by identifying environmental and clinical sister clades - those where the two niches share an immediate common ancestor (Felsenstein 1985). Within the *P. aeruginosa* phylogeny, four such direct contrasts between environmental and clinical niches were found, and labeled as contrasts A-D (Figure 6). Within these four contrasts, we averaged the  $\Delta$ trait values across all four replicates of each strain, and used a one-sample t-test to identify any non zero differences between environmental and clinical averaged  $\Delta$ trait values.

**2. pGLS:** Essentially a weighted regression model, sample points are given less weight in a pGLS model if there is covariance between them. In regards to the phylogenetic framework, the weighting is determined from a correlation matrix of the covariance between species based on their shared phylogenetic history (Symonds and Blomberg 2014). This covariance is altered depending on the amount of phylogenetic signal in the actual distribution of traits, where less phylogenetic signal will down-weight the covariances between species (Symonds and Blomberg 2014). This phylogenetic signal is determined through the estimation of  $\lambda$  in the distribution of the residuals, and the estimated  $\lambda$  is then applied to the correlation matrix, which is applied to the regression model (Freckleton et al. 2002). We ran a ML pGLS on all continuous  $\Delta$ traits, with niche included as the discrete predictor variable. The estimated  $\lambda$  for the correlation structure applied within the GLS models was carried out using the corPagel function of the ape package. We used the total replicated data for all environmental and clinical strains, where the incorporated phylogeny was one with a four-way polytomy at each tip. However, since the correlated matrix cannot deal with hard polytomies, we changed each zero branch length to  $10^{-8}$ .

**3. Phylogenetic-Corrected Correlations:** The correlational analysis corrects for the non-independence of samples by using a similar GLS framework to the pGLS. Here,  $\lambda$  was estimated for the correlation structure, and the analysis was carried out using an MCMC method. Bayes Traits does not allow for correlations between discrete and continuous variables if there are zero or very small branch lengths, so we averaged the  $\Delta$ trait values across replicates of each strain as was done in the independent contrasts analysis, and incorporated a tree with no polytomies (Figure 6).

**4. pcPCA:** The multivariate  $\Delta$ trait analysis was carried out on all replicated  $\Delta$ trait data, and was thereby corrected for by the environmental/clinical phylogeny (Figure 6) with four polytomies at each tip (one for each evolved replicate). Our analysis used the  $\lambda$  method and a correlation mode in calculating the principal components.

### 5.3 Phylogenetic Comparative Analyses of Trait Variations

In addition to testing for phylogenetic signal and niche effects in adaptation, we also measured different variations of trait associations to further assess any evolutionary patterns in traits and fitness. Specifically we carried out correlations between each  $\Delta$ trait, a pGLS between ancestral trait and  $\Delta$ trait, and a correlation and pGLS analysis between fitness and trait. These analyses were completed on the entire dataset of environmental, clinical, and PaO1/Pa14 strains (Figure 5). While completed in the same general manner as the niche comparison analyses, specifics to each trait comparison analysis are summarized in Table 3, and described as follows:

**$\Delta$ Trait Correlations:** Each continuous  $\Delta$ trait was correlated against all other  $\Delta$ traits. This was completed on the total replicated data while accounting for the non-independence of samples in a phylogeny that included 4 polytomies at each tip. This

analysis did not estimate  $\lambda$  for the applied correlation structure, but rather assumed a complete BM correlation structure.

**Ancestor vs  $\Delta$ Trait:** A pGLS was carried out on the ancestral value of each trait with the corresponding change in that trait. The ancestral state was the continuous predictor variable, and  $\lambda$  was estimated for the correlation structure. This was carried out on the total replicated data with a four-way polytomy tree, however this replaced the zero branch lengths with  $10^{-8}$ . A phylogenetic corrected correlational analysis was attempted to replicate the analysis, however for reasons unknown Bayes Traits could not deal with this specific dataset.

**Trait vs Fitness:** The trait values for each ancestral and evolved population were correlated with the fitness values of those same populations. This analysis thereby corrected for non-independence of samples with a 5-way polytomy at each tip - one for each ancestral strain and four for each evolved strain. In this analysis we estimated  $\lambda$  for the correlation structure. A pGLS was also carried out on the same data to replicate the analysis, where trait was the continuous predictor variable.  $\lambda$  was also estimated here for the correlation structure. The same phylogeny was incorporated where  $10^{-8}$  replaced zero branch lengths.

#### **5.4 Ancestral and Evolved Population Comparisons**

We wanted to identify if the evolved strains of each niche changed significantly from their ancestral state in a multivariate analysis. To accomplish this we carried out an additional pcPCA on each of the ancestral and evolved populations, of both environmental and clinical strains. The multivariate data included in the analysis consisted of both traits and fitness, and was corrected for by a phylogeny that included a

five-way polytomy at each tip (four for each of the evolved strains, and one for the ancestral strain). This analysis also used the  $\lambda$  method and a correlation mode in calculating the principal components.

**Table 3.** Summary of comparative analyses that correct for the non-independence of species in a phylogenetic framework

Comparison	Test	Software	Correlation Structure	Trait Data	Phylogeny	Strain Replicates	Polytomies
Niche vs $\Delta$ Trait	Independent Contrasts	Manual	Independent Contrasts	$\Delta$	Env/Clin	Averaged	None
	Correlations	Bayes Traits	$\lambda$	$\Delta$	Env/Clin	Averaged	None
	pGLS	R - nlme (phytools)	$\lambda$	$\Delta$	Env/Clin	Included	4 ( $10^{-8}$ replaces zero)
	pcPCA	R - phyl.pca (phytools)	$\lambda$	$\Delta$	Env/Clin	Included	4
$\Delta$ Traits	Correlations	Bayes Traits	Brownian	$\Delta$	Env/Clin/PA	Included	4
	pGLS	R - nlme (phytools)	$\lambda$	Ancestral / $\Delta$	Env/Clin/PA	Included	4 ( $10^{-8}$ replaces zero)
Traits vs Fitness	Correlations	Bayes Traits	$\lambda$	Evolved & Ancestral	Env/Clin/PA	Included	5
	pGLS	R - nlme (phytools)	$\lambda$	Evolved & Ancestral	Env/Clin/PA	Included	5 ( $10^{-8}$ replaces zero)
Evolved vs Ancestor	pcPCA	R - phyl.pca (phytools)	$\lambda$	Evolved & Ancestral	Env/Clin/PA	Included	5

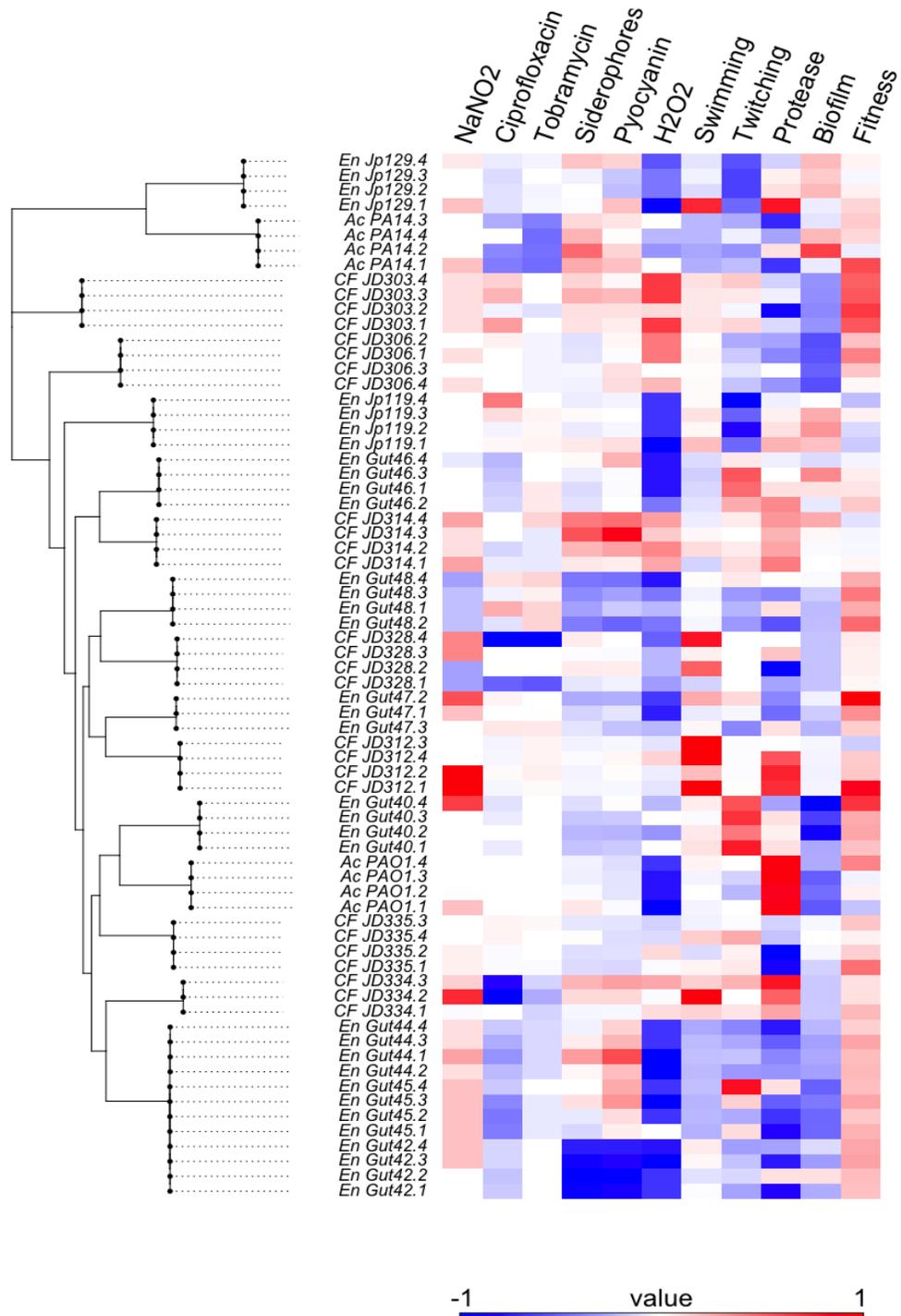
# Chapter 3: Results

## 1.0 Evolution of Pathoadaptive Traits and Fitness

We evolved 18 genetically distinct *P. aeruginosa* strains within SCFM, for ~100 generations, and measured changes in evolved trait from each ancestral state, as well as the change in fitness (Figure 7). The extent of diversity across strains in the phenotypic analysis is noteworthy. It is immediately clear that the strains are uniquely evolving within SCFM, with extensive variability in the directional changes across traits. The profound diversity in *P. aeruginosa* is evident in this study, and potential patterns of evolution explaining the observed diversity are explored in subsequent sections.

Some broad trends are also noted in the phenotypic analysis. For instance, resistance to NaNO<sub>2</sub> appears to increase or remain the same across most strains, with the exception of two clades, which decrease (En Gut48 and CF JD328) (Figure 7). By contrast, most strains show no change, or a decrease in resistance to both tobramycin and ciprofloxacin. Interestingly, however, there are also instances of slightly increased resistance, especially to ciprofloxacin, despite being evolved in the absence of drug.

As expected, almost all strains increased their fitness within SCFM (Figure 7), suggesting that the strains have indeed adapted to the environment. While it appears that some strains are decreasing in fitness, a one-sample t-test reveals that these decreases are not significantly different from zero ( $p > 0.05$ ). We have assumed that strains that do not show significantly increased fitness gains are initially situated closer to an adaptive peak within SCFM, than those strains with larger gains. While quite variable across strains and niches (which are explored in subsequent sections), two epidemic strains (CF\_JD334<sub>LESA</sub> and CF\_JD328<sub>LESB</sub>) included in this study show relatively low changes in fitness



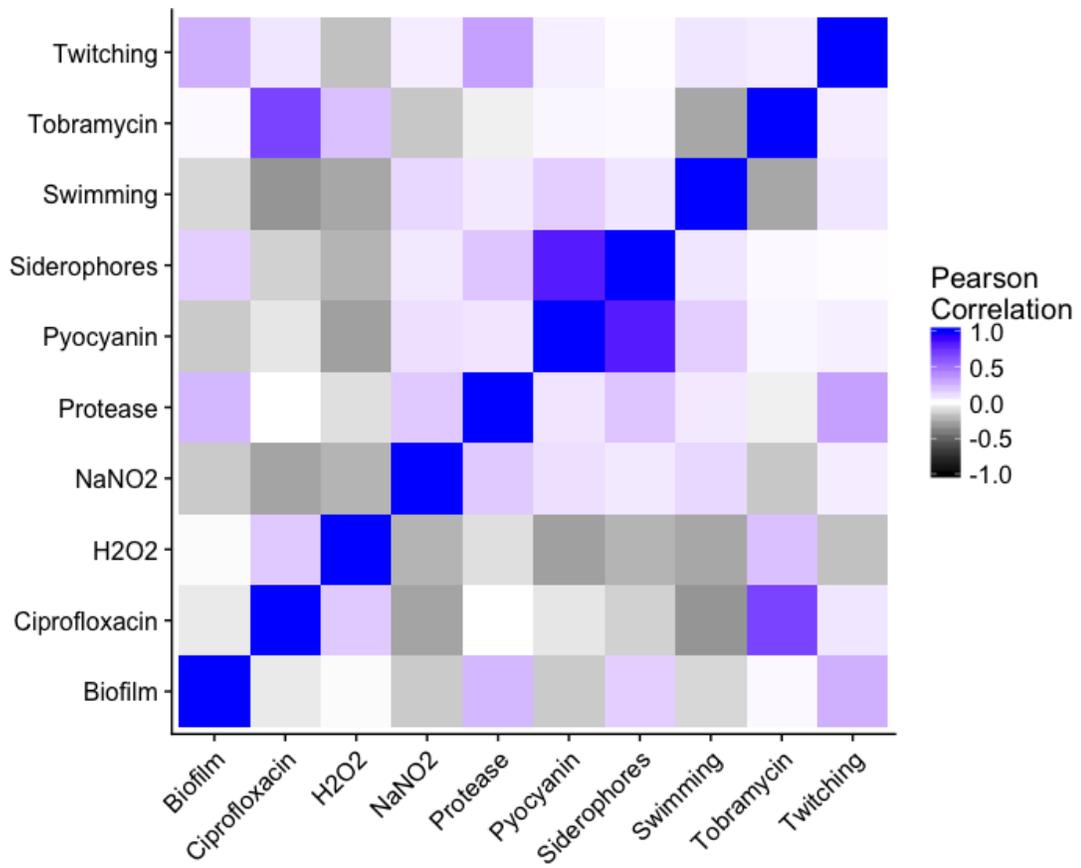
**Figure 7.** Heat map of  $\Delta$ trait values, within a phylogenetic framework. Each trait is scaled between -1 and 1, where negative blue boxes indicate a decrease in trait value, and positive red boxes indicate an increase in trait value. White signifies no change.

( $\Delta_{SCF\_JD335} = 0.25$ ,  $\Delta_{SCF\_JD328} = 0.05$ ) (Figure 7). This suggests that these epidemic strains are well adapted to a lung environment, which is why they have a high potential of transfer between patients.

## 2.0 Correlations Between $\Delta$ Traits

We carried out an analysis of correlations between trait changes, in order to identify which traits tend to change in a similar manner (Figure 8). This is important because we cannot treat the  $\Delta$ traits as independent samples if there are strong correlations between them, given that there may be a single genetic basis underlying the correlated trait changes. Because we also cannot assume the independence of strains given their shared phylogenetic history, we tested for  $\Delta$ trait correlations while correcting for phylogeny. We found that pyocyanin and siderophore production were very strongly associated ( $r = 0.82$ ), as were ciprofloxacin and tobramycin resistance ( $r = 0.68$ ) (Figure 8). No other trait combinations showed moderate or strong correlations, with an overall observance of weak or no associations overall ( $r < 0.3$ ).

Correlated change in pyocyanin and siderophore production is not unexpected, as both are pigmented molecules in *P. aeruginosa* that contribute to its overall virulence (Meyer 2000; Lau et al. 2004). Siderophores are essential molecules in the uptake of iron - something that is a scarcity in most biological systems and requires competition for (Meyer 2000). In *P. aeruginosa*, pyoverdine is the major siderophore. Pyoverdine also has direct toxic effects, where it has been shown to mediate the killing of leukocytes, and regulate virulence factor production (Becerra et al. 2001; Lamont et al. 2002). Pyocyanin itself is a zwitterion that can easily cross membranes, and has a wide range of effects in



**Figure 8.** Phylogenetic-corrected correlation matrix between  $\Delta$ traits. The correlation structure was estimated from a complete BM assumption, and the analysis was completed using a MCMC method.

causing death of host cells, and competing microbes (Lau et al. 2004). Where evolution in SCFM does not involve the competition from other microbes or a host immune system, pyocyanin and pyoverdine are concurrently not as useful as when in the presence of microbes and host cells. Therefore, they are somewhat functionally linked and it is no surprise that they evolve with one another in SCFM, which is evidenced in their correlational values. Alternatively, the two traits could share a genetic basis, although a brief review of the literature did not find anything to suggest this.

The observed correlation may otherwise be an experimental artifact. The pyoverdine and pyocyanin assay was completed on the same day and on the same cultures. The optical density at each corresponding wavelength may have thereby included detection of both pigments, which could have resulted in the observed correlations. The correlated result is consistent with another study however (Clark et al. 2015), so we are optimistic that experimental flaw is not the case. However, to be sure of this the assay should be repeated on select cultures on different days.

The correlated changes in antibiotic resistance profiles between tobramycin and ciprofloxacin appear to be the result of overall decreases in resistance observed across most strains. When broadly looking at trends in resistance, it generally appears that strains are increasing in susceptibilities to both antibiotics (Figure 8), which is likely due to the cost of resistance in the absence of antibiotic (Andersson and Levin 1999). Clark et al. (2015) also observed correlated resistance profiles between ciprofloxacin, ceftazidime, and aztreonam, suggesting overall that antibiotic resistance and susceptibility can evolve together. There are two notable reasons that this may occur in *P. aeruginosa*. Firstly, biofilms in particular have been shown to increase combined resistances to antibiotics,

and are a major reason for antibiotic failure in treating CF infections (Bjarnsholt et al. 2009). Secondly, efflux pumps can play a role in generating multidrug resistance depending on their substrate specificity. Relevant here, the MexXY multidrug efflux system has been shown to result in cross-resistance to both fluoroquinolones (ciprofloxacin) and aminoglycosides (tobramycin) (Morita et al. 2012). A point mutation in this efflux pump is a plausible reason for the observed correlation between the tobramycin and ciprofloxacin susceptibility profiles in this experiment. It also appears that En Gut48 increases its resistance to both antibiotics throughout the ~100 generations in SCFM (Figure 8), which may have arisen via either of these mechanisms. This is especially concerning, as drug resistance may be capable of evolving in the absence of a drug selection pressure.

Regardless of the underlying genomic or functional cause, we generally refer to pyocyanin and pyoverdine production as pigment production, and resistances to tobramycin and ciprofloxacin as antibiotic resistance throughout the remainder of the results. This is because we cannot assume that the trait changes of each pair are independent samples. We do however attempt to correct for strong correlations between traits by relevant pcPCA's described in subsequent sections.

### **3.0 Phylogenetic Signal in the Propensity to Evolve and Adapt**

We tested for the strength of phylogenetic signal in the change of each measured trait and fitness, by estimating Pagel's  $\lambda$ , Abouheif's  $C_{\text{mean}}$  and Moran's  $I$  (Table 4). The distribution of  $\Delta\text{fitness}$  and all  $\Delta\text{traits}$  revealed non-zero values of phylogenetic signal, which suggests that phylogeny is influential to some degree for the evolution of all CF pathoadaptive traits (Table 4). Notably, there was consistently high and significant

**Table 3.** Phylogenetic signal in the change in traits and fitness, measured with three different tests. Bolded values are significant for the  $\lambda$  test ( $p < 0.05$ ).

	Measured Variable	$C_{\text{mean}}^*$	$I^*$	$\lambda$	$p$ -value ( $\lambda$ )
$\Delta$ Trait	Biofilm	0.574	0.568	0.774	< <b>0.05</b>
	Proteases	0.484	0.475	0.697	< <b>0.05</b>
	Swimming	0.348	0.342	0.664	< <b>0.05</b>
	Twitching	0.565	0.561	0.742	< <b>0.05</b>
	H <sub>2</sub> O <sub>2</sub>	0.614	0.605	0.863	< <b>0.05</b>
	NaNO <sub>2</sub>	0.300	0.298	0.230	> 0.05
	Siderophores	0.671	0.670	0.347	< <b>0.05</b>
	Pyocyanin	0.569	0.567	0.091	> 0.05
	Ciprofloxacin	0.362	0.354	0.543	< <b>0.05</b>
	Tobramycin	0.317	0.312	0.619	< <b>0.05</b>
Fitness	$\Delta$	0.277	0.267	0.577	< <b>0.05</b>
	Evolved	0.251	0.244	0.515	< <b>0.05</b>

\*All measured traits and fitness showed significant signal ( $p < 0.05$ )

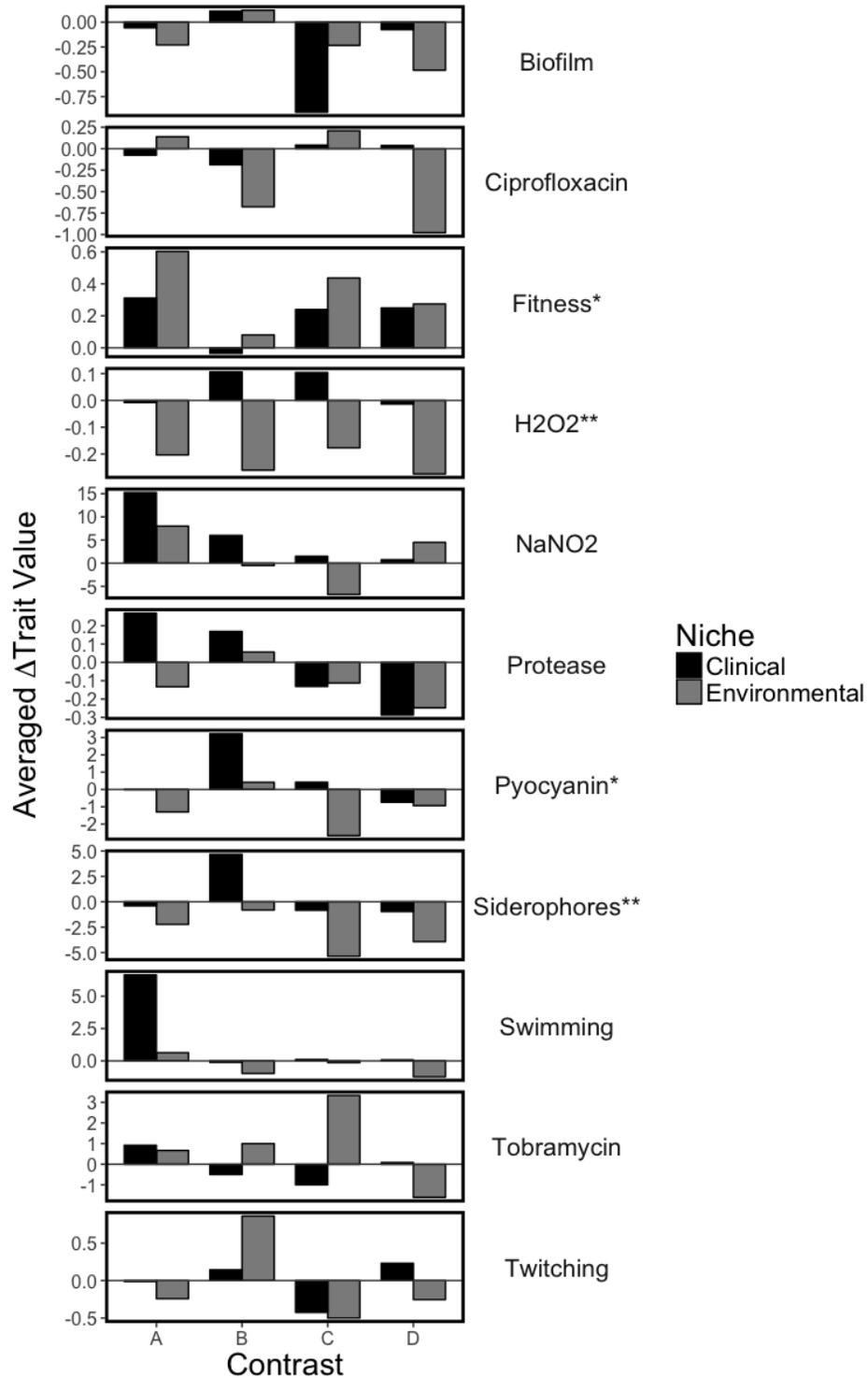
phylogenetic signal in changes in H<sub>2</sub>O<sub>2</sub> resistance ( $\lambda = 0.863, I = 0.605, C_{\text{mean}} = 0.614$ ), biofilm ( $\lambda = 0.774, I = 0.568, C_{\text{mean}} = 0.574$ ), and twitching ( $\lambda = 0.742, I = 0.561, C_{\text{mean}} = 0.565$ ). The repeatable strength of phylogenetic signal across all three tests is robust evidence for the strong influence of phylogeny in the evolution of these pathoadaptive traits, within SCFM. While there are slight discrepancies between tests in other traits, there is still moderate phylogenetic signal observed repeatedly in changes in proteases ( $\lambda = 0.697, I = 0.475, C_{\text{mean}} = 0.484$ ), antibiotic resistance ( $\lambda \sim 0.600, I \sim 0.330, C_{\text{mean}} \sim 0.340$ ), and swimming ( $\lambda = 0.664, I = 0.342, C_{\text{mean}} = 0.348$ ). Signal in the change in pigment production and NaNO<sub>2</sub> resistance however are quite variable between tests, and not always significant ( $p < 0.05$ ), so it is inconclusive whether there is reliable signal in the evolution of these traits.

There is significant phylogenetic signal in  $\Delta$ fitness across all tests of intermediate strength ( $\lambda = 0.577, I = 0.267, C_{\text{mean}} = 0.277$ ) (Table 4). This is almost mirrored in the phylogenetic signal of evolved fitness, which is slightly weaker than that of  $\Delta$ fitness ( $\lambda = 0.515, I = 0.244, C_{\text{mean}} = 0.251$ ). The moderate strength of phylogenetic signal in  $\Delta$ fitness is very noteworthy. This suggests that the fitness changes required for early adaptation of *P. aeruginosa* to SCFM is dependent on genetic background. While it is difficult to make strong inferences concerning the underlying fitness landscape within SCFM, it is clear that not all strains are responding the exact same way to the environment. There is genetic variation in a strain's propensity to adapt, suggesting that some strains are more maladapted than others to life in the lung and require greater fitness changes, as indicated by a moderate phylogenetic signal in  $\Delta$ fitness.

The slight decrease in phylogenetic signal in evolved fitness is possible evidence for selection erasing some effect of history, and driving the adaptive convergence of these strains. This would suggest that strains are starting to converge onto some optimum as a result of selection, whereas their fitness gains are somewhat dictated by phylogeny and where in a fitness landscape they begin. However, our measure of fitness is assumed to be transitive - that is that high fitness against Pa14-*LacZ* is consistent with high fitness in SCFM. As this may not be the case, a more robust measure of fitness, and a longer selection experiment is needed in order to properly investigate the tradeoff effect between adaptation and history.

#### **4.0 Niche Effects in the Evolution of Traits and Fitness**

We compared the individual  $\Delta$ traits between niches using three different analyses to correct for phylogeny, as well as compared the two niches with a multivariate analysis. When comparing niches using the manual independent contrasts analysis, we found that siderophore production and resistance to H<sub>2</sub>O<sub>2</sub> were significantly decreased in environmental strains, compared to clinical strains ( $p < 0.05$ ) (Figure 9). Furthermore, change in pyocyanin production was decreased in environmental strains of near significance ( $p < 0.07$ ) when compared to the clinical strains, and change in fitness was slightly higher amongst environmental strains ( $p < 0.07$ ). No significant niche differences were observed for the rest of the traits, which include biofilm, antibiotic resistance, resistance to NaNO<sub>2</sub>, proteases, swimming and twitching ( $p > 0.07$ ) (Figure 9).



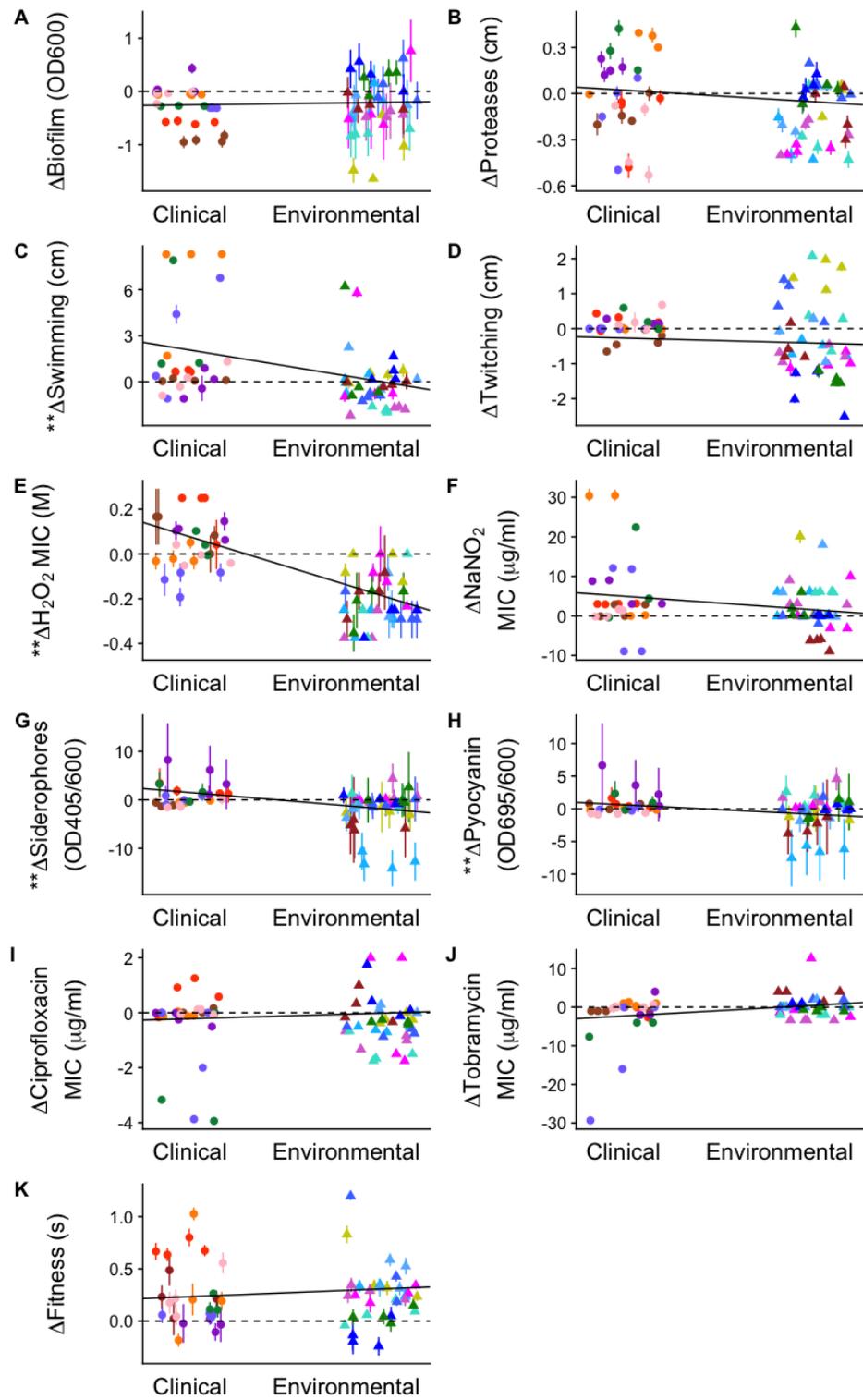
**Figure 9.** Independent contrast analysis directly comparing environmental and clinical strains within a phylogeny, in their averaged  $\Delta$ trait value across the evolved replicates of each ancestral strain. Stars equate to significance in niche differences (\*\* =  $p < 0.05$ ; \* =  $p < 0.07$ )

The phylogenetic-corrected correlational analysis for niche comparisons revealed similar trends for changes in siderophore production, and resistance to H<sub>2</sub>O<sub>2</sub>. There is evidence that these phenotypes decreased more, or increased less, in environmental strains; this effect was extremely pronounced for  $\Delta$ H<sub>2</sub>O<sub>2</sub> (siderophores:  $r = -0.45$ ,  $\log\text{BF} > 2$ ; H<sub>2</sub>O<sub>2</sub>:  $r = -0.84$ ,  $\log\text{BF} > 10$ ) (Table 5). Changes in fitness are also significant in this analysis, but a low correlation coefficient indicates that environmental strains are increasing in fitness only slightly more than clinical strains ( $r = 0.09$ ,  $\log\text{BF} > 2$ ). There are discrepancies with the independent contrasts analysis for changes in swimming and tobramycin resistance however, for which there is some evidence of differences in this correlational analysis (swimming:  $r = -0.5$ ;  $\log\text{BF} > 2$ ; tobramycin:  $r = 0.41$ ,  $\log\text{BF} > 2$ ). Furthermore,  $\Delta$ pyocyanin did not show differences between niches in this analysis (Table 5).

In conjunction with the independent contrasts and correlational analyses, the pGLS revealed significant differences between niches in the changes in H<sub>2</sub>O<sub>2</sub> resistance ( $n = 62$ ,  $\lambda = 0.62$ ,  $p = 0.00$ ) and siderophores ( $n = 62$ ,  $\lambda = 0.19$ ,  $p = 0.00$ ) (Figure 10, Table 6). Similar to the correlational analysis, swimming was also significantly different between the niches, showing much lower increases in the environmental strains ( $n = 62$ ,  $\lambda = 0.56$ ,  $p = 0.03$ ) (Figure 10, Table 6). Pyocyanin production was also significant in this analysis as it was in the independent contrasts and shows more decreases in the environmental strains ( $n = 62$ ,  $\lambda = -0.03$ ,  $p = 0.01$ ) (Figure 10, Table 6).

**Table 4.** Correlational analysis of niche with  $\Delta$ trait. R is the correlation for environmental strains compared to clinical strains. When positive, environmental strains are changing less negatively (or more positively) than clinical strains. Vice versa for negative. The stronger R is, the more different the  $\Delta$ trait values between the two niches. Stars equate to strength of evidence (\*= $\log$ BF > 2; \*\*= $\log$ BF>10).

Trait	R	95% CI	$\lambda$ Estimate	95% CI	LogBF
Biofilm	0.08	(-0.08 - 0.22)	0.57	(0.04 - 0.93)	0.77
Proteases	-0.19	(-0.36 - -0.04)	0.99	(0.40 - 1.00)	-0.35
Swimming	-0.50	(-0.61 - -0.36)	0.60	(0.04 - 0.95)	*4.81
Twitching	-0.16	(-0.30 - 0.00)	0.49	(0.02 - 0.90)	-0.29
H <sub>2</sub> O <sub>2</sub>	-0.84	(-0.88 - -0.77)	0.72	(0.08 - 0.95)	**18.08
NaNO <sub>2</sub>	-0.35	(-0.48 - -0.20)	0.81	(0.09 - 0.98)	1.50
Siderophores	-0.45	(-0.58 - -0.31)	0.16	(0.01 - 0.54)	*3.83
Pyocyanin	-0.34	(-0.49 - -0.17)	0.14	(0.01 - 0.52)	1.42
Ciprofloxacin	0.12	(-0.07 - 0.25)	0.75	(0.13 - 0.97)	-0.19
Tobramycin	0.41	(0.25 - 0.53)	0.56	(0.03 - 0.96)	*3.15
Fitness	0.09	(-0.07 - 0.21)	1.00	(0.75 - 1.00)	*3.32



**Figure 10.** pGLS models of  $\Delta$ trait as a function of niche. Raw data is plotted with the pGLS regression line included (solid). Colours represent the replicates of each strain. Stars equate to significance (\*\* =  $p < 0.05$ ).

**Table 5.** Corresponding pGLS parameters for Figure 10:  $\Delta$ trait as a function of niche. Bolded values are significant ( $p < 0.05$ ).

$\Delta$	Intercept	Slope (SE)	$\lambda$ Estimate	$p$ -value
Biofilm	-0.25	0.05 (0.21)	0.82	0.83
Protease	0.02	-0.08 (0.09)	0.61	0.42
Swimming	2.03	-2.03 (0.93)	0.56	<b>0.03</b>
Twitching	-0.27	-0.14 (0.39)	0.78	0.72
H2O2	0.07	-0.26 (0.04)	0.62	<b>0.00</b>
NaNO2	4.95	-3.40 (2.51)	0.28	0.18
Siderophores	1.48	-3.26 (1.08)	0.19	<b>0.00</b>
Pyocyanin	0.60	-1.45 (0.56)	-0.03	<b>0.01</b>
Ciprofloxacin	-0.22	0.20 (0.40)	0.60	0.62
Tobramycin	-2.43	2.85 (1.69)	0.46	0.10
Fitness	0.24	0.07 (0.11)	0.61	0.53

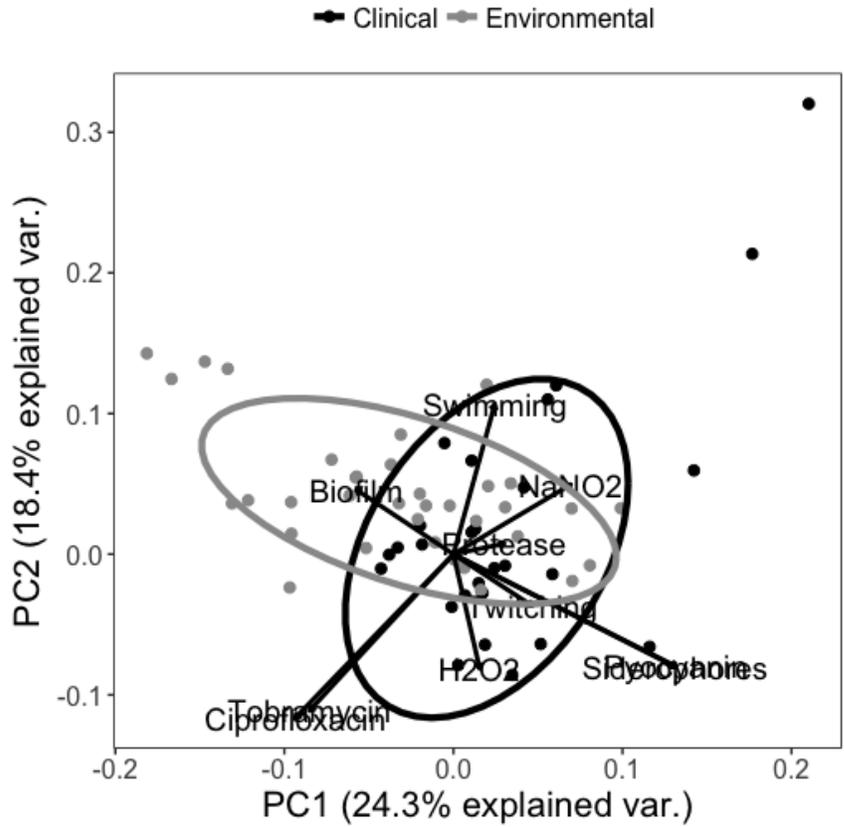
The consistency across all three analyses is strong evidence that environmental strains decrease significantly in H<sub>2</sub>O<sub>2</sub> resistance and siderophore production compared to clinical strains. The discrepancies amongst the other tests are more difficult to interpret, as each test has its own limitations. The independent contrasts analysis excludes some strains and averages values across replicates, the correlational analysis also averages values across replicates, and the pGLS analysis modifies the zero branch lengths of the phylogeny. Since the pGLS analysis alters the data in the most minimal way, it is likely the most robust measure. We are therefore confident that changes in pyocyanin are also significantly different between the two niches, which is corroborated with our  $\Delta$ trait correlational analysis that shows pyocyanin is strongly correlated with siderophores ( $r = 0.82$ ) (Figure 8). Given the pGLS and correlational analysis results, we are also confident that swimming is significantly different between the two niches.

The large decrease in H<sub>2</sub>O<sub>2</sub> resistance in environmental strains compared to the clinical strains is consistent with the study by Head and Yu (2004), which showed planktonic clinical cells had much higher susceptibility. This is believed to be a result of biofilm protection against oxidative stress. It is reasonable then, that the environmental strains, having started with high resistances, increased in susceptibility in becoming more clinical-like. The virulence pigments also decrease more in environmental strains, which is expected, since a loss of virulent factors are often observed in chronic lung infections. Again, the environmental strains possibly lose the pigments they start with, while clinical strains have already lost them.

Since changes in fitness were higher in the environmental strains in the correlational analysis, and almost significant in the independent contrasts analysis but not

at all significant in the pGLS, we investigated the results in more depth. As the independent contrasts exclude certain strains from its analysis, further investigation revealed that two of the excluded environmental strains, En Jp119 and En Jp129, undergo little change in fitness. These strains are also unique, as they are the only test strains isolated from outside of North America, and the only two from coastal sea water sources (Table 1). When excluded from the correlation and independent contrasts analyses, the change in fitness is significantly different and greater in the environmental strains than clinical strains (independent contrasts:  $p \ll 0.05$ ; correlations:  $r = 0.38$ ,  $\log\text{BF} = 3.06$ ) (See Appendix A). When these strains are excluded from the pGLS, while still not significant, the slope of the pGLS is much steeper between the two niches ( $n = 54$ ,  $p = 0.22$ ,  $\lambda = 0.44$ , slope = 0.14) (See Appendix A). Given the observed variations between strains, we cannot conclude that all environmental strains are maladapted to a CF lung, however they appear to be generally less adapted than the clinical strains. This raises the question as to whether strains from particular niches are better primed for infection, and can more readily transition to life in the lung.

We also did a multivariate analysis by carrying out a pcPCA of the  $\Delta$ traits, to further identify any differences between environmental and clinical strains (Figure 11). A PCA is a multivariate analysis that identifies the different combinations of traits that best explain the variance in the data. This takes possibly correlated traits and converts them into a series of linear principal components (PC), which explain the variance in the data, in an ordinal structure. For instance, our analysis found that PC1 explained 24.3% of the variance in the data, and PC2 explained 18.4% of the variance (Figure 11). Furthermore



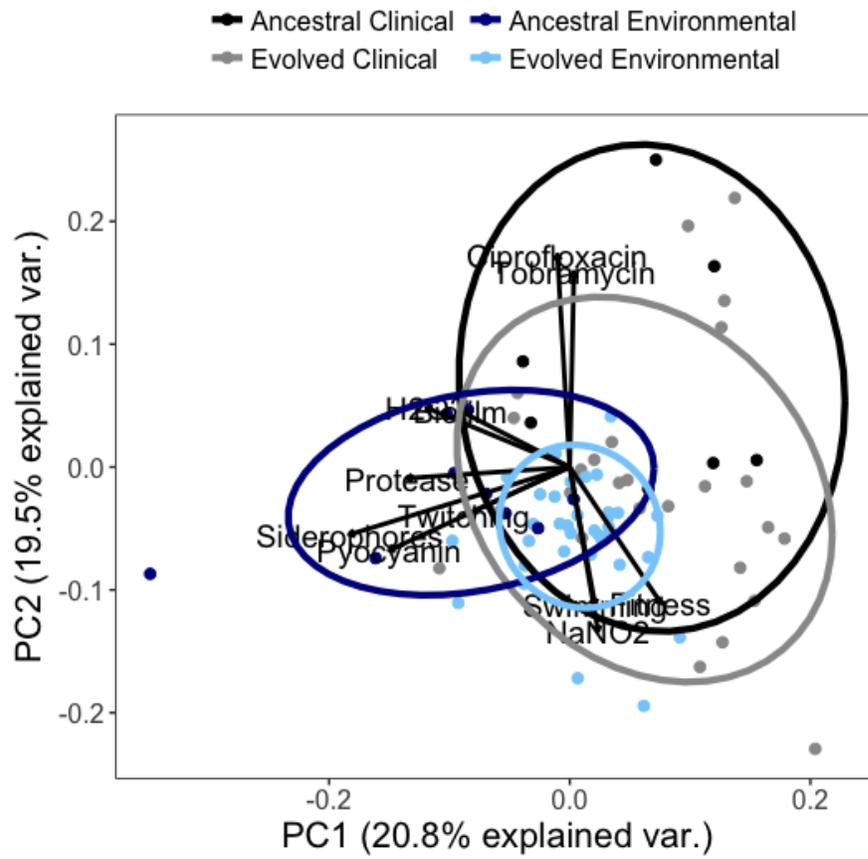
**Figure 11.** Phylogenetic-corrected PCA of  $\Delta$ traits comparing environmental and clinical niches. Ellipses equate to coverage of 85% of data.

in this analysis, the samples are corrected for non-independence due to their shared phylogenetic history, resulting in a pcPCA of  $\Delta$ traits. When comparing the niches in the multivariate  $\Delta$ trait analysis, the pcPCA shows a difference between environmental and clinical strains along PC1 (Figure 11). This niche difference was confirmed as significant through a pGLS of PC1 ( $n = 62$ ,  $\lambda = 0.26$ ,  $p = 0.00$ ). Differences between niches in PC1 mostly consist of changes in pigment production (loading  $> 0.8$ ), which is consistent with both the independent contrasts and correlations. Antibiotic resistance is also a major contributor to PC1 (loading  $> 0.5$ ), which was not significantly observed in the prior analyses (Figure 11).

The combined analyses investigating the niche differences in trait evolution and fitness gains throughout evolution in SCFM are compelling evidence that environmental and clinical strains are changing differently from one another both in terms of individual traits, and some combination of traits.

## **5.0 Clinical Relevance of Trait Changes**

Given the niche differences in the various trait changes, we sought to determine what the relevance of these trait changes are - are environmental strains transitioning to a clinical phenotype? To investigate this, we performed a pcPCA of each ancestral and evolved population's trait and fitness values, of both niches, to identify how the evolved populations are changing from their ancestors (Figure 12). An analysis of niche overlap revealed the highest overlap between the ancestral and evolved clinical strains, with a Schoener's D of 51%, compared to 44% for the ancestral and evolved environmental strains. This firstly suggests that clinical strains are overall changing less within SCFM, than the environmental strains.



**Figure 12.** Phylogenetic-corrected PCA of multivariate trait and fitness data comparing ancestral and evolved populations of environmental and clinical strains. Ellipses equate to coverage of 85% of data.

We carried out paired, two sample t-tests and an analysis of Cohen's  $d$  to determine the significance and effect size of the evolved transitions in SCFM, along the first two principal components. We found that along PC1, environmental and clinical strains are both significantly changing throughout evolution in SCFM (Figure 12). Environmental strains are increasing along PC1 with a large effect size ( $p = 8.3 \times 10^{-5}$ ,  $d = -0.76$ ), and clinical strains are slightly decreasing along PC1 with a small effect size ( $p = 0.04$ ,  $d = 0.45$ ). Along PC2, the transition is also significant from ancestral to evolved state, in both environmental and clinical strains. Environmental strains are greatly shifting downwards ( $p = 5.8 \times 10^{-9}$ ,  $d = -1.30$ ), whereas in the same direction clinical strains are moderately shifting ( $p = 0.01$ ,  $d = -0.56$ ).

The transition along PC2 in both niches suggests an evolved response to the unique factors within SCFM that are not mirrored in the environment or a CF lung. This is likely attributed to the absence of antibiotics and competition from immune cells and microbes. In support of this, antibiotic susceptibility largely influences PC2 (loading > 0.7). Additionally, a gain in fitness largely contributes to PC2, suggesting that SCFM is not a perfect proxy for the CF lung, and both strains require adaptations that are specific to this environment. However, environmental strains are transitioning along PC1 towards a more "clinical-like" state, evidenced by the significant transition of the majority of strains in the rightward direction of a large effect, whereas the clinical strain's shift along PC1 is only marginally significant, and of a small effect (Figure 12). The right shift in environmental strains along PC1 also appears to be directed by the loss of proteases, twitching, H<sub>2</sub>O<sub>2</sub> resistance, pigment production, and biofilm formation (loadings > 0.4), which are often observed in clinical infections, with the exception of biofilm reduction. A

gain in fitness along PC1 additionally appears to shift environmental strains to a more clinical phenotype (loading = 0.34).

The more static nature of the clinical strains compared to environmental strains, especially along PC1, suggests overall that they are better adapted to SCFM. Even though SCFM is only a nutritional media that lacks many aspects of the cystic fibrosis lung including competing microbes, the immune system, epithelium, and antibiotics, there appear to be similar cues within the media that are present within a CF lung.

Furthermore, the loss of traits that direct the environmental strains to a more clinical phenotype include those that are known to be lost during adaptation to the CF lung. The only inconsistency is the general loss of biofilm formation in our experiment, which is typically enriched in CF infections (Bjarnsholt et al. 2009). The nature of our experimental design may be a factor in this, in which daily transfers during the selection experiment occurred from the middle of the medium without mixing of the populations, and could have selected only for planktonic cells. However, Pa14 in our experiment did increase in biofilm formation, which is consistent with other studies. It may be that biofilm is not always necessary for increased fitness in SCFM as it is in the lung, which could be a result of the lack of stressors present within the media.

Altogether, these results are consistent with previous studies that have found that SCFM is a comparable growth medium to natural CF lung sputum, and presents some of the same selection pressures that are present *in vivo* (Fung et al. 2010; Wong et al. 2012; Turner et al. 2015). More consistency in future studies should include antibiotics in SCFM at concentrations that are typically present in clinical CF sputum, as the lack of

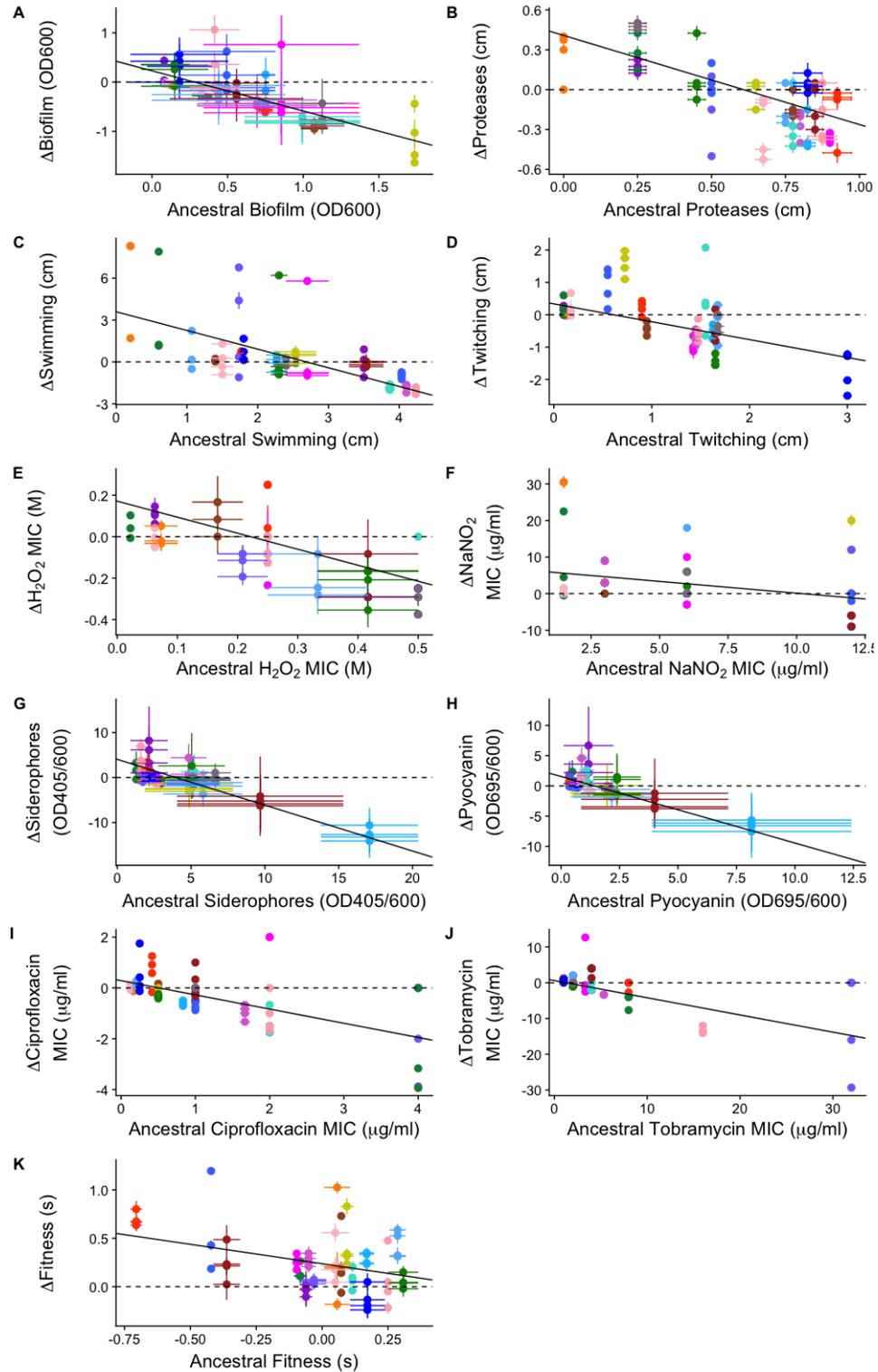
antibiotic in our study appears to drive the paralleled evolution of clinical and environmental strains in SCFM, along PC2.

## **6.0 Evolutionary Constraint of Ancestral Trait**

Given the evidence of phylogenetic signal of varying strengths in the evolution of traits and fitness, we attempted to identify if the observed constraint of genetic background is influenced by the ancestral state of the genotype. We investigated the change in trait values when compared to the ancestral trait value through a pGLS analysis to see if the directionality in the evolution of traits is contingent on how high or low that trait is to begin with (Figure 13, Table 7). Furthermore this analysis allows us to identify if the changes across traits are consistent with stabilizing or directional selection acting on them. Stabilizing selection results in the population mean converging onto a single non-extreme value, and is evidenced by a decreasing trend in evolved populations that begin at a high state, and an increasing trend in evolved populations that begin at a low state (Hansen 1997). Directional selection is alternatively the convergence of the population mean onto an extreme value (either high or low) (Lenski and Travisano 1994).

We found across all traits, a significant and negative association between ancestral state and change in trait value (slope  $< 0$ ,  $p < 0.05$ ) (Figure 13, Table 7). This means that overall, strains that start with high values tend to decrease more (or increase less) than strains that start with low values. This is strong evidence that the constraint of genetic background in the propensity to evolve is somewhat predicted by the ancestral value.

In particular in proteases, swimming, and H<sub>2</sub>O<sub>2</sub> resistance, there is some evidence that the evolution of traits is of a stabilizing nature (Figure 13B,C,E). In these traits, low



**Figure 13.** pGLS regression models of  $\Delta$ trait as a function of ancestral trait. Raw data is plotted with the pGLS regression line included (solid). Colours represent the replicates of each strain. All models are significant ( $p < 0.05$ ).

**Table 6.** Corresponding pGLS parameters for Figure 13:  $\Delta$ trait as a function of ancestral trait. Bolded values are significant ( $p < 0.05$ ).

$\Delta$	Intercept	Slope (SE)	$\lambda$ Estimate	$p$ -value
Biofilm	0.23	-0.82 (0.11)	0.36	<b>0.00</b>
Protease	0.41	-0.67 (0.12)	0.38	<b>0.00</b>
Swimming	3.60	-1.34 (0.24)	0.32	<b>0.00</b>
Twitching	0.34	-0.55 (0.17)	0.64	<b>0.00</b>
H2O2	0.17	-0.77 (0.11)	0.56	<b>0.00</b>
NaNO2	6.56	-0.64 (0.25)	0.12	<b>0.01</b>
Siderophores	4.07	-1.02 (0.07)	0.70	<b>0.00</b>
Pyocyanin	1.55	-1.10 (0.08)	0.57	<b>0.00</b>
Ciprofloxacin	0.29	-0.56 (0.08)	0.11	<b>0.00</b>
Tobramycin	0.63	-0.48 (0.07)	0.24	<b>0.00</b>
Fitness	0.24	-0.40 (0.18)	0.40	<b>0.03</b>

ancestral values tend to increase, and high ancestral values tend to decrease, evidenced by the negative slope of the pGLS models crossing zero at some intermediate value. This also appears to be slightly true for biofilm and twitching (Figure 13A,D). This suggests that evolution in SCFM is favoured by the convergence of these traits onto some optimum value, which is consistent with stabilizing selection. However since it is unknown if the directionality of these trait changes to an optimum value is concurrent with high fitness, it is inconclusive whether selection is actually acting on these traits.

In regards to changes in pigments and antibiotic resistance, those with low ancestral values do not especially increase when evolved in SCFM ( $\Delta$ trait for min ancestral trait  $\sim 0$ ) (Figure 13G, H, I, J). High ancestral values conversely precede large decreases in trait value, suggesting that SCFM favors the complete loss of the trait. As previously described, a widespread loss of pigments and antibiotic resistance could be a result of the high carrying costs without much conferred benefit within SCFM, since there are no antibiotics, competing microbes, or immune cells present in the media.

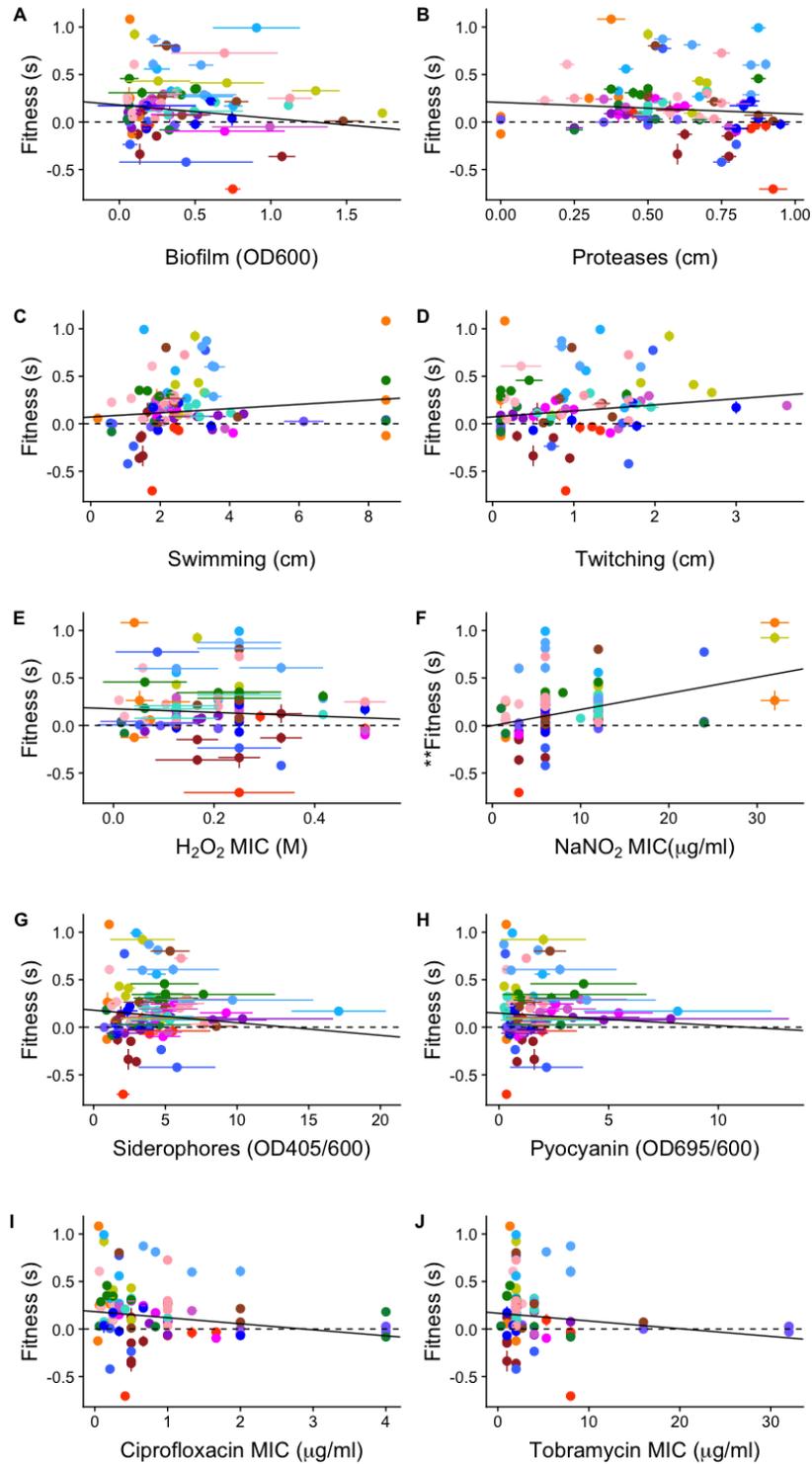
Resistance to  $\text{NaNO}_2$  is contrarily gained across most strains, where selection appears to be driving the evolution of resistance regardless of the ancestral state. This is evidenced by the shallow slope and apparent increase in resistance of the majority of the strains (Figure 13F, Table 7). A review of the literature reveals a cascade of regulatory genes involved in the respiration of nitrate during anaerobic growth (Schreiber et al. 2007). A byproduct of this reaction is nitrite, which *P. aeruginosa* is also capable of using as an electron acceptor (Kerschen et al. 2001). *P. aeruginosa* also possesses an excretory system for nitrite so as to avoid toxic levels building within the cell (Kerschen et al. 2001). While the selection experiment was carried out aerobically, the high

viscosity of mucin is believed to drive the formation of microcolonies (Fung et al. 2010). Microcolonies are a form of unattached biofilm, in which growth is thought to occur anaerobically (Sriramulu et al. 2005). As our biofilm assay tested for attached biofilms, we do not know what the prevalence of microcolonies is within our populations. It could be an explanation however, as to why *P. aeruginosa* in SCFM + mucin may switch to utilizing nitrate as an electron acceptor during anaerobic growth. Furthermore, since  $\text{KNO}_3$  is an ingredient in SCFM, it is possible that the evolved utilization of nitrate and nitrite as electron receptors, and the subsequent excretion of nitrite from the cell in this environment allow *P. aeruginosa* to withstand high amounts of nitric stress.

The changes in fitness compared to the ancestral value also show a negative slope, where genotypes that start with high fitness values increase less than genotypes that start with low fitness in SCFM (Figure 13K). This strongly suggests that there is widespread variation in a strain's starting fitness within a lung, and that more maladapted genotypes require greater fitness gains during adaptive evolution. Furthermore, not all strains are ending up on the same fitness values, or trait values, after the 100 generation selection experiment. This also suggests that the early adaptation to SCFM is variable amongst strains, and that the phenotypic and fitness changes required to reach a fitness peak are met with varied levels of success across strains.

## **7.0 Trait Predictors of Fitness**

In order to identify which traits may specifically be under directional selection, we carried out a pGLS and correlational analysis comparing trait values with fitness (Figure 14, Table 7; Table 8). Those traits whose absolute values correlate with fitness are likely to be under directional selection. While we acknowledge that competitive



**Figure 14.** pGLS regression models of fitness as a function of trait. Raw data is plotted with the pGLS regression line included (solid). Colours represent the replicates of each strain. Stars equate to significance (\*\* =  $p > 0.05$ )

**Table 7.** Corresponding pGLS parameters for Figure 14: fitness as a function of trait.  
 Bolded values are significant ( $p < 0.05$ ).

Trait	Intercept	Slope (SE)	$\lambda$ Estimate	$p$ -value
Biofilm	0.18	-0.14 (0.09)	0.47	0.12
Protease	0.21	-0.12 (0.15)	0.42	0.43
Swimming	0.07	0.02 (0.02)	0.39	0.19
Twitching	0.07	0.06 (0.05)	0.38	0.17
H2O2	0.18	-0.19 (0.25)	0.42	0.44
NaNO2	0.00	0.02 (0.00)	0.48	<b>0.00</b>
Siderophores	0.18	-0.01 (0.01)	0.45	0.31
Pyocyanin	0.15	-0.01 (0.02)	0.42	0.55
Ciprofloxacin	0.18	-0.06 (0.04)	0.40	0.07
Tobramycin	0.17	-0.01 (0.01)	0.40	0.20

**Table 8.** Correlational analysis of fitness with trait. Stars equate to strength of evidence (\*= $\log\text{BF} > 2$ ; \*\*= $\log\text{BF}>10$ ).

Trait	R	95% CI	$\lambda$ Estimate	95% CI	LogBF
Biofilm	-0.28	(-0.32 - -0.22)	0.46	(0.26 - 0.65)	*6.45
Proteases	-0.19	(-0.21 - -0.16)	0.52	(0.34 - 0.68)	*2.63
Swimming	0.19	(0.17 - 0.22)	0.35	(0.17 - 0.54)	*3.10
Twitching	0.14	(0.13 - 0.14)	0.55	(0.63 - 0.39)	0.59
H <sub>2</sub> O <sub>2</sub>	-0.04	(-0.07 - -0.02)	0.47	(0.30 - 0.64)	-1.76
NaNO <sub>2</sub>	0.42	(0.40 - 0.45)	0.40	(0.22 - 0.61)	**17.63
Siderophores	-0.08	(-0.11 - -0.03)	0.47	(0.28 - 0.65)	1.06
Pyocyanin	-0.03	(-0.06 - -0.01)	0.42	(0.23 - 0.63)	0.46
Ciprofloxacin	-0.12	(-0.15 - -0.09)	0.51	(0.33 - 0.67)	1.72
Tobramycin	-0.11	(-0.14 - -0.09)	0.65	(0.50 - 0.79)	0.18

fitness against Pa14-*LacZ* may be an imperfect proxy for fitness, it allows for an accurate comparison of fitness values between the competing genotypes.

We found in both analyses, that resistance to NaNO<sub>2</sub> was positively correlated with fitness, of intermediate strength ( $n = 88$ ,  $\lambda = 0.48$ ,  $p = 0.00$ ;  $r = 0.42$ ,  $\log\text{BF} = 17.63$ ) (Figure 14, Table 8; Table 9). Furthermore in the correlational analysis, biofilm formation is negatively correlated with fitness, but of a modest strength ( $r = -0.28$ ,  $\log\text{BF} = 6.45$ ). Additionally, proteases and swimming show weak correlations with fitness, in which proteases are negatively correlated ( $r = -0.19$ ,  $\log\text{BF} = 2.63$ ) and swimming is positively correlated ( $r = 0.19$ ,  $\log\text{BF} = 3.10$ ) (Table 9). However, the significance in these results are not mirrored in the pGLS, which shows a high degree of scatter around the regression lines (Figure 14), so it is difficult to draw conclusions regarding directional selection on these particular traits.

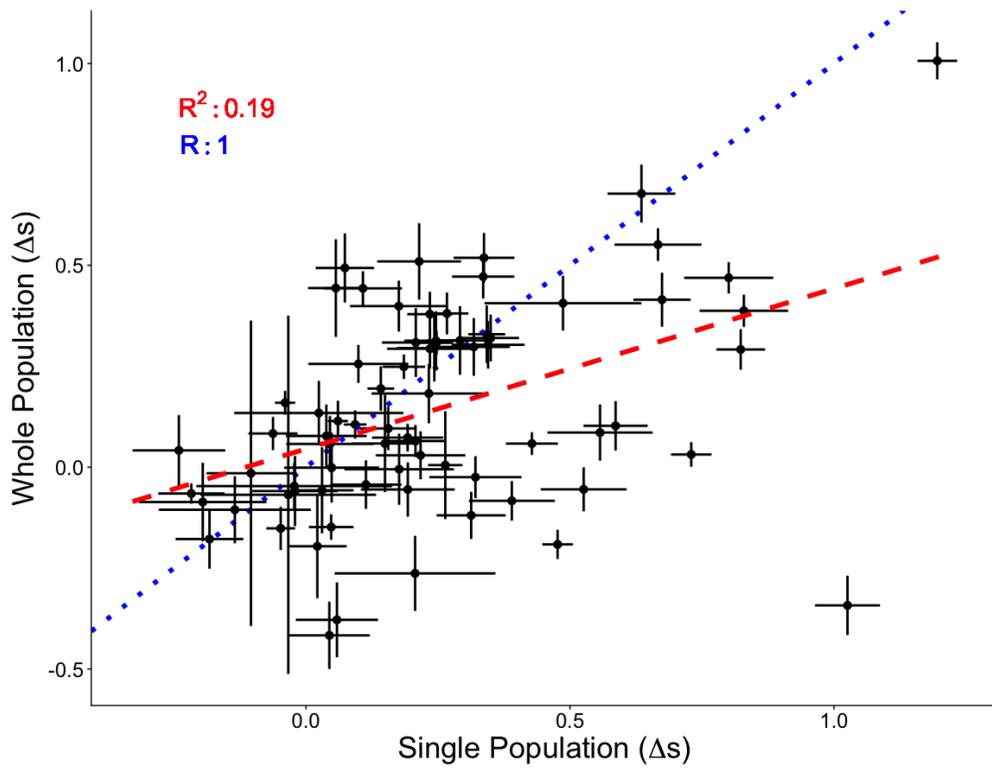
The moderate correlation of fitness with NaNO<sub>2</sub> resistance is intriguing, and corroborated in our analysis of ancestor vs  $\Delta$ trait, which showed evidence for directional selection on NaNO<sub>2</sub> resistance. Although evolved competitive fitness may not be truly representative of evolved fitness in SCFM, further investigation reveals that the majority of strain replicates are increasing in NaNO<sub>2</sub> resistance (87%). Additionally in support of the significant biofilm correlation, there is widespread reduction across replicates (81%). This suggests that biofilm and NaNO<sub>2</sub> resistance are indeed under directional selection. Furthermore in support of selection acting directly on NaNO<sub>2</sub> resistance, there was very low, insignificant, and inconsistent phylogenetic signal observed in the changes in NaNO<sub>2</sub> resistance ( $\lambda = 0.230$ ,  $I = 0.298$ ,  $C_{\text{mean}} = 0.300$ ) (Table 4). This strongly suggests that selection is driving the evolution of this trait, and erasing an effect of phylogeny.

The benefits of NaNO<sub>2</sub> resistance have already been discussed as a possible result of nitrate respiration during anaerobic growth in microcolonies. The possible advantageousness of low biofilm is contrary to what was expected. As previously stated however, the selection experiment may have selected for planktonic cells, and against biofilm. Furthermore, slow growth typically occurs in biofilms, which may have been a disadvantage when competed against Pa14-*LacZ*.

## 8.0 Single Isolates vs Whole Populations

In a typical experimental evolution study, fitness assays and/or phenotyping are carried out on a single isolate (or, more rarely, multiple isolates) from a population. However, the extent of within-population diversity is unknown, and it is unclear how representative of the whole population the single isolate may be. To account for this, and reassure us of our phylogenetic analysis, we measured the competitive fitness of all ancestral, single, and whole populations to see if the change in fitness of the single isolates is a good estimation of the change in fitness of the whole populations. We found an increasing, linear trend between both single and whole  $\Delta s$  values for each population, which was expected ( $r = 0.44$ ) (Figure 15).

The absence of a 1:1 correlation and the presence of a flatter slope in the fitness trend suggest that on average the single isolates are more fit than the whole populations that they originate from. Combined with consistent and strong scatter around the mean ( $r^2 = 0.19$ ), this highlights the probable within-population diversity and existence of what may be a multitude of competing factors. Clonal interference, and frequency dependent selection may especially be contributing to the maintenance of diverse genotypes (Levin 1988; Gerrish and Lenski 1998).



**Figure 15.** Comparison of the change in fitness ( $s$ ) between single isolates and the whole populations from which they were founded. Blue dotted line is an expected 1:1 correlation, red dashed line is the actual regression line.

It is difficult to capture the diversity and dynamics of any population in biology. However based on our results of a linear trend between fitness of isolates and their founding populations, we do believe that the testing of randomly chosen single colonies isolated from a whole population is an acceptable standard. A better future practice would be to measure a handful of within-population isolates for stronger replication.

## **9.0 Accuracy of Trait Data**

Pa14 and PaO1 were included in the study in part as a control for some of the phenotypic assays, since many studies include them in their analysis. We found what was expected for biofilm formation, which increases in Pa14. PaO1 is known to produce relatively high amounts of biofilm, which was also observed here. Upon evolution in SCFM this biofilm production decreased in PaO1. Motility was decreased in Pa14, while slightly decreased or remained unchanged in PaO1 replicate populations. This is reassuring, since motility is typically lost in CF infections (Marvig et al. 2015).

We additionally did a miniature review of the literature, where studies that include the trait analysis of both Pa14 and PaO1 wildtype strains were highlighted (see Appendix B). We then compared the relative ratios of our ancestral results for Pa14 and PaO1 with the ratios between the trait values in these studies, as a type of control. Challenging here was the limited amount of studies that include both wildtype strains in the analysis, and report the results of both wildtype strains. Secondly, the ratios of traits between Pa14 and PaO1 in the highlighted studies vary greatly between each other, so it is difficult to know which result to base ours off of. The actual MIC results for PaO1 and Pa14 also varied greatly between studies, so it was problematic to compare our MIC values with studies that included either strain in their experiment.

It is likely that the specific Pa14 and PaO1 strains used across studies differ slightly from one another, which is concerning. In fact, a study that analyzed the genomes of PaO1 strains from different laboratories identified an alarming amount of differences in deletions, inversions and single nucleotide substitutions between strains (Klockgether et al. 2010). This calls in to questions the repeatability of some studies that use these strains, and is a plausible reason for the vast differences in phenotyping that are observed across studies. While this makes it difficult to control for accuracy in our study, all strains were equally treated in each assay and replicates and controls were included within each assay, which allows us to be confident of our results when comparing the trait values between different strains.

# Chapter 4: Discussions, Conclusions and Implications

## 1.0 Genetic Background and Niche Influence Trait Evolution and Adaptation

Charles Darwin in the Origin of Species first described phylogenetic inertia, when he explained the “Law of Conditions of Existence”. Darwin suggested that the traits of any species are built upon the preexisting adaptations (Darwin 1859). In our experiment, we attempted to identify how much the existing genetic background within *P. aeruginosa* constrains not only trait evolution, but also adaptation. We found significant evidence for the contribution of phylogeny in the propensity to evolve and adapt, specifically to a CF lung-like environment. Phylogenetic signal was detected across all traits and fitness, and was moderate to strong across many including H<sub>2</sub>O<sub>2</sub> resistance, biofilm formation and twitching motility. Our results suggest that the early evolution of fitness gains and pathoadaptive traits is contingent, at least in part, on genetic background and where in a phylogeny a strain is located. Furthermore, the constraint of genetic background was mirrored in the constraint of the ancestral state, and genotypes that have a high trait value consistently decreased more (or increased less) than genotypes that started with a low trait value. While not specifically investigated, epistasis is likely the underlying mechanism for the constraint of genetic background in the propensity of *P. aeruginosa* to evolve and adapt to SCFM.

Since niche specific adaptations can occur in either the accessory genome, or from mutations in the core genome, the phylogeny of strains does not take into question the prior influence of niche. Our comparison of strains between two broad niches enabled us to understand how the evolved differences of niche-specific selection contribute to

adaptation of a novel environment. We found that in SCFM environmental and clinical strains differ significantly in the evolution of certain pathoadaptive traits including H<sub>2</sub>O<sub>2</sub> resistance, pigment production, and swimming, and there is slight evidence that they differ overall in their fitness gains. On average, environmental strains were less adapted to SCFM than their clinical counterparts, and required different phenotypic alterations. Noteworthy however was that this is not true for all strains. The results from two environmental isolates suggest that prior niche-specialized adaptations allow some strains to more readily transition between environments.

## **2.0 The Adaptive Landscape of SCFM: A Hypothesis**

Given our results implicating both niche and genetic background in the fitness gains of *P. aeruginosa* to SCFM, we can start to hypothesize the underlying fitness landscape of adaptation to SCFM. In regards to the phylogenetic inference in the evolution of traits and fitness, it appears that closely related genotypes have similar capacities to adapt than distantly related genotypes. Variations in fitness leaps also suggest that strains are initially situated at variable distances from an adaptive peak, and that most environmental strains lie further from a peak than the clinical strains they are closely related to. Important to note here, is that not all environmental strains sit further from a peak than all clinical strains, but within those closely related clades, environmental strains tend to be more maladapted.

Whether the landscape is smooth or rough can only be determined from the endpoint state of the evolved isolates, which was not the focus of this study. However, our PCA analysis suggests that all strains, regardless of environmental or clinical origin

are converging on a similar region of multivariate trait space. This finding is consistent with a single adaptive peak in SCFM.

### **3.0 A Single Strain is not the Whole Story**

A notable outcome of this study is the wide range of diversity of both ancestral traits, and patterns of evolution in the 18 investigated *P. aeruginosa* strains. This incredible diversity indicates the two standard laboratory strains of *P. aeruginosa* may not be representative of the species as a whole. The investigation of a multitude of strains may be required to draw broad conclusions regarding its involvement in infectious disease. We hope that this study will bring about the inclusion of multiple strains from a variety of niches into further studies that assess the adaptation of *P. aeruginosa*.

### **4.0 Implications for CF Lung Infections**

The outcomes of this study have major implications for the lung infections of people with CF. Importantly the findings are a confirmation that patients with CF should never be in contact: our results suggest that clinical strains are generally better adapted to life in the lung and can readily switch between lung-like environments without requiring extensive adaptive leaps and pathoadaptive changes. Secondly, it is hope that infection may not be inevitable. The variation in adaptive leaps necessary for different strains is evidence of this. Particular genotypes may be readily eradicated before undergoing those major phenotypic alterations required for persistence. Therefore our findings are a verification that newborn screening for CF should be maintained and supported, so that early antibiotic therapy can help prevent the adaptive leaps required by less pre-adapted genotypes. Lastly, the results of the study are evidence that patient treatment should be

specific, and infecting strains should be genotyped. The profound diversity in *P. aeruginosa*, and patient heterogeneity in CF, is support that treatments should be tailored to the specified bacterial genotype/environment lung combination of each patient.

In addition to validating current practices and suggestions for direct clinical implications, outcomes from this study can allow us to begin to explore the answers to questions such as “which genotypes are primed for infection?” and “do these high infectivity genotypes cluster to particular niches?”. We can start to discover targets of drug therapy in genotypes of high infection potential, to prevent their early colonization of the lung. We can also start to identify niches that harbour high infectivity genotypes, to inform better prevention measures from these environments. With *P. aeruginosa* the number one threat to the life of CF individuals, any measures taken to inform better infection prevention guidelines and early treatment options are significant, and positive steps.

## References

- Aaron, S. D., K. Ramotar, W. Ferris, K. Vandemheen, R. Saginur, E. Tullis, D. Haase, D. Kottachchi, M. St Denis, and F. Chan. 2004. Adult Cystic Fibrosis Exacerbations and New Strains of *Pseudomonas aeruginosa*. *Am. J. Respir. Crit. Care Med.* 169:811–815.
- Abouheif, E. 1999. A Method for Testing the Assumption of Phylogenetic Independence in Comparative Data. *Evol. Ecol. Res.* 1:895–909.
- Alonso, A., F. Rojo, and J. L. Martínez. 1999. Environmental and Clinical Isolates of *Pseudomonas aeruginosa* Show Pathogenic and Biodegradative Properties Irrespective of their Origin. *Environ. Microbiol.* 1:421–430.
- Andersson, D. I., and B. R. Levin. 1999. The Biological Cost of Antibiotic Resistance. *Curr. Opin. Microbiol.* 2:489–493.
- Andrews, J. M. 2001. Determination of Minimum Inhibitory Concentrations. *J. Antimicrob. Chemother.* 48, Suppl.:5–16.
- Bank, C., R. T. Hietpas, J. D. Jensen, and D. N. A. Bolon. 2015. A Systematic Survey of an Intragenic Epistatic Landscape. *Mol. Biol. Evol.* 32:229–238.
- Baussano, I., I. Tardivo, R. Bellezza-Fontana, M. P. Forneris, A. Lezo, L. Anfossi, M. Castello, V. Aleksandar, and E. Bignamini. 2006. Neonatal Screening for Cystic Fibrosis does not Affect Time to First Infection with *Pseudomonas aeruginosa*. *Pediatrics* 118:888–895.
- Becerra, C., I. Albesa, and A. J. Eraso. 2001. Leukotoxicity of Pyoverdinin, Production of Reactive Oxygen Species, and Effect of UV Radiation. *Biochem. Biophys. Res. Commun.* 285:414–418.
- Bhagirath, A. Y., Y. Li, D. Somayajula, M. Dadashi, S. Badr, and K. Duan. 2016. Cystic Fibrosis Lung Environment and *Pseudomonas aeruginosa* Infection. *BMC Pulmoary*

Med. 16:1–22.

- Bjarnsholt, T., P. Ø. Jensen, M. Burmølle, M. Hentzer, J. A. J. Haagensen, H. P. Hougen, H. Calum, K. G. Madsen, C. Moser, S. Molin, N. Høiby, and M. Givskov. 2005. *Pseudomonas aeruginosa* Tolerance to Tobramycin, Hydrogen Peroxide and Polymorphonuclear Leukocytes is Quorum-Sensing Dependent. *Micrbiology* 151:373–383.
- Bjarnsholt, T., P. Ø. Jensen, M. J. Fiandaca, J. Pedersen, C. R. Hansen, C. B. Andersen, T. Pressler, M. Givskov, and N. Høiby. 2009. *Pseudomonas aeruginosa* Biofilms in the Respiratory Tract of Cystic Fibrosis Patients. *Pediatr. Pulmonol.* 44:547–558.
- Blomberg, S. P., and T. Garland. 2002. Tempo and Mode in Evolution: Phylogenetic Inertia, Adaptation and Comparative methods. *J. Evol. Biol.* 15:899–910.
- Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution* (N. Y). 57:717–745.
- Bobadilla, J. L., M. Macek Jr., J. P. Fine, and P. M. Farrell. 2002. Cystic Fibrosis: A Worldwide Analysis of CFTR Mutations–Correlation With Incidence Data and Application to Screening. *Hum. Mutat.* 19:575–606.
- Bollag, J.-M., and N. M. Henninger. 1978. Effects of Nitrite Toxicity on Soil Bacterial under Aerobic and Anaerobic Conditions. *Soil Biol. Biochem.* 10:377–381.
- Bollback, J. P., and J. P. Huelsenbeck. 2009. Parallel Genetic Evolution Within and Between Bacteriophage Species of Varying Degrees of Divergence. *Genetics* 181:225–234.
- Bull, J. J., M. R. Badgett, R. Springman, and I. J. Molineux. 2004. Genome Properties and the Limits of Adaptation in Bacteriophages. *Evolution* (N. Y). 58:692–701.
- Burns, J. L., R. L. Gibson, S. McNamara, D. Yim, J. Emerson, M. Rosenfeld, P. Hiatt, K. McCoy, R. Castile, A. L. Smith, and B. W. Ramsey. 2001. Longitudinal Assessment

- of *Pseudomonas aeruginosa* in Young Children with Cystic Fibrosis. *J. Infect. Dis.* 183:444–452.
- Cheng, S. H., R. J. Gregory, J. Marshall, S. Paul, D. W. Souza, G. A. White, C. R. O’Riordan, and A. E. Smith. 1990. Defective Intracellular Transport and Processing of CFTR Is the Molecular Basis of Most Cystic Fibrosis. *Cell* 63:827–834.
- Chmiel, J. F., T. R. Aksamit, S. H. Chotirmall, E. C. Dasenbrook, J. S. Elborn, J. J. LiPuma, S. C. Ranganathan, V. J. Waters, and F. A. Ratjen. 2014. Antibiotic Management of Lung Infections in Cystic Fibrosis: I. The Microbiome, Methicillin-Resistant *Staphylococcus aureus*, Gram-Negative Bacteria, and Multiple Infections. *Ann. Am. Thorac. Soc.* 11:1120–1129.
- Chmiel, J. F., M. W. Konstan, and J. S. Elborn. 2013. Antibiotic and Anti-Inflammatory Therapies for Cystic Fibrosis. *Cold Spring Harb. Perspect. Med.* 3:a009779.
- Chou, H.-H., H.-C. Chiu, N. F. Delaney, D. Segre, and C. J. Marx. 2011. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science* (80- .). 332:1190–1192.
- Clark, S. T., J. Diaz Caballero, M. Cheang, B. Coburn, P. W. Wang, S. L. Donaldson, Y. Zhang, M. Liu, S. Keshavjee, Y. C. W. Yau, V. J. Waters, D. E. Tullis, D. S. Guttman, and D. M. Hwang. 2015. Phenotypic Diversity within a *Pseudomonas aeruginosa* Population Infecting an Adult with Cystic Fibrosis. *Nat. Sci. Reports* 5:10932.
- Clunes, M. T., and R. C. Boucher. 2007. Cystic Fibrosis: The Mechanisms of Pathogenesis of an Inherited Lung Disorder. *Drug Discov. Today Dis. Mech.* 4:63–72.
- Cohen-Cymbarknoh, M., D. Shoseyov, and E. Kerem. 2011. Managing Cystic Fibrosis: Strategies that Increase Life Expectancy and Improve Quality of Life. *Am. J. Respir. Crit. Care Med.* 183:1463–1471.
- Cystic Fibrosis Canada. 2017. The Canadian CF Registry 2016 Annual Data Report.

Retrieved from: <http://www.cysticfibrosis.ca/blog/2016-registry-annual-data-report/>.

- Cystic Fibrosis Foundation. 2015. Cystic Fibrosis Foundation Patient Registry 2015 Annual Data Report. Bethesda, Maryland.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, Or, the Preservation of Favoured Races in the Struggle for Life*. Murray, J, London.
- de Gracia, J., F. Mata, A. Alvarez, T. Casals, S. Gatner, M. Vendrell, D. de la Rosa, L. Guarner, and E. Hermosilla. 2005. Genotype-Phenotype Correlation for Pulmonary Function in Cystic Fibrosis. *Thorax* 60:558–563.
- Dettman, J. R., N. Rodrigue, S. D. Aaron, and R. Kassen. 2013. Evolutionary Genomics of Epidemic and Nonepidemic Strains of *Pseudomonas aeruginosa*. *PNAS* 110:21065–21070.
- Dietrich, L. E. P., A. Price-Whelan, A. Petersen, M. Whiteley, and D. K. Newman. 2006. The Phenazine Pyocyanin is a Terminal Signalling Factor in the Quorum Sensing Network of *Pseudomonas aeruginosa*. *Mol. Microbiol.* 61:1308–1321. Blackwell Publishing Ltd.
- Drumm, M. L., M. W. Konstan, M. D. Schluchter, A. Handler, R. Pace, F. Zou, M. Zariwala, D. Fargo, A. Xu, J. M. Dunn, R. J. Darrah, R. Dorfman, A. J. Sandford, M. Corey, J. Zielenski, P. Durie, K. Goddard, J. R. Yankaskas, F. A. Wright, and M. R. Knowles. 2005. Genetic Modifiers of Lung Disease in Cystic Fibrosis. *N. Engl. J. Med.* 353:1443–1453.
- Emerson, J., M. Rosenfeld, S. McNamara, B. Ramsey, and R. L. Gibson. 2002. *Pseudomonas Aeruginosa* and Other Predictors of Mortality and Morbidity in Young Children With Cystic Fibrosis. *Pediatr. Pulmonol.* 34:91–100.
- Felsenstein, J. 1985. Phylogenies and the Comparative Method. *Am. Nat.* 125:1–15.
- Feltman, H., G. Schulert, S. Khan, M. Jain, L. Peterson, and A. R. Hauser. 2001. Prevalence of Type III Secretion Genes in Clinical and Environmental Isolates of

- Pseudomonas aeruginosa*. *Microbiology* 147:2659–2669.
- Fernández, L., E. B. M. Breidenstein, D. Song, and R. E. W. Hancock. 2012. Role of Intracellular Proteases in the Antibiotic Resistance, Motility, and Biofilm Formation of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 56:1128–1132.
- Flume, P. A., P. J. Mogayzel, K. A. Robinson, C. H. Goss, R. L. Rosenblatt, R. J. Kuhn, and B. C. Marshall. 2009. Cystic Fibrosis Pulmonary Guidelines: Treatment of Pulmonary Exacerbations. *Am. J. Respir. Crit. Care Med.* 180:802–808.
- Folkesson, A., L. Jelsbak, L. Yang, H. Krogh Johansen, O. Ciofu, N. Høiby, and S. Molin. 2012. Adaptation of *Pseudomonas aeruginosa* to the Cystic Fibrosis Airway: an Evolutionary Perspective. *Nat. Rev. Microbiol.* 10:841–851.
- Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic Analysis and Comparative Data: A Test and Review of Evidence. *Am. Nat.* 160:712–726.
- Fulzele, R., E. DeSa, A. Yadav, Y. Shouche, and R. Bhadekar. 2011. Characterization of Novel Extracellular Protease Produced by Marine Bacterial Isolate from the Indian Ocean. *Brazilian J. Microbiol.* 42:1364–1373. Elsevier.
- Fung, C., S. Naughton, L. Turnbull, P. Tingpej, B. Rose, J. Arthur, H. Hu, C. Harmer, C. Harbour, D. J. Hassett, C. B. Whitchurch, and J. Manos. 2010. Gene Expression of *Pseudomonas aeruginosa* in a Mucin-Containing Synthetic Growth Medium Mimicking Cystic Fibrosis Lung Sputum. *J. Medican Microbiol.* 59:1089–1100.
- Gad, G. F., R. A. El-Domany, S. Zaki, and H. M. Ashour. 2007. Characterization of *Pseudomonas aeruginosa* Isolated from Clinical and Environmental Samples in Minia, Egypt: Prevalence, Antibigram and Resistance Mechanisms. *J. Antimicrob. Chemother.* 60:1010–1017.
- Gerrish, P. J., and R. E. Lenski. 1998. The Fate of Competing Beneficial Mutations in an Asexual Population. *Genetica* 102/103:127–144.
- Gifford, D. R., S. E. Schoustra, and R. Kassen. 2011. The Length of Adaptive Walks Is

- Insensitive To Starting Fitness in *Aspergillus nidulans*. *Evolution* (N. Y). 65:3070–3078.
- Gifford, D. R., M. Toll-Riera, and R. C. MacLean. 2016. Epistatic Interactions Between Ancestral Genotype and Beneficial Mutations shape Evolvability in *Pseudomonas aeruginosa*. *Evolution* (N. Y). 70:1659–1666.
- Hall, A. R., V. F. Griffiths, R. C. MacLean, and N. Colegrave. 2010. Mutational Neighbourhood and Mutation Supply Rate Constrain Adaptation in *Pseudomonas aeruginosa*. *Proc. R. Soc. B* 277:643–650.
- Hancock, R. E. W., and D. P. Speert. 2000. *P. aeruginosa* Antibiotic Resistance in *Pseudomonas aeruginosa*: Mechanisms and Impact on Treatment. *Drug Resist. Updat.* 3:247–255.
- Hansen, T. F. 1997. Stabilizing Selection and the Comparative Analysis of Adaptation. *Evolution* (N. Y). 51:1341–1351.
- Head, N. E., and H. Yu. 2004. Cross-Sectional Analysis of Clinical and Environmental Isolates of *Pseudomonas aeruginosa* : Biofilm Formation , Virulence , and Genome Diversity. *Infect. Immun.* 72:133–144.
- Heirali, A., S. McKeon, S. Purighalla, D. G. Storey, L. Rossi, G. Costilhes, S. J. Drews, H. R. Rabin, M. G. Surette, and M. D. Parkins. 2016. Assessment of the Microbial Constituents of the Home Environment of Individuals with Cystic Fibrosis (CF) and their Association with Lower Airways Infections. *PLoS One* 11:e0148534.
- Hietpas, R. T., J. D. Jensen, and D. N. A. Bolon. 2011. Experimental Illumination of a Fitness Landscape. *PNAS* 108:7896–7901.
- Hogardt, M., and J. Heesemann. 2010. Adaptation of *Pseudomonas aeruginosa* During Persistence in the Cystic Fibrosis Lung. *Int. J. Med. Microbiol.* 300:557–562.
- Houdret, N., J.-M. Perini, C. Galabert, A. Scharfman, P. Humbert, G. Lamblin, and P. Roussel. 1986. The High Lipid Content of Respiratory Mucins in Cystic Fibrosis is

- Related to Infection. *Biochim. Biophys. Acta* 880:54–61. Elsevier.
- Kauffman, S. A., and E. D. Weinberger. 1989. The NK Model of Rugged Fitness Landscapes And Its Application to Maturation of the Immune Response. *J. Theor. Biol.* 141:211–245.
- Keck, F., F. Rimet, A. Bouchez, and A. Franc. 2016. phylosignal: an R Package to Measure, Test, and Explore the Phylogenetic Signal. *Ecol. Evol.* 6:2774–2780.
- Kerschen, E. J., V. R. Irani, D. J. Hassett, and J. J. Rowe. 2001. *snr-1* Gene Is Required for Nitrate Reduction in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 183:2125–2131.
- Klockgether, J., A. Munder, J. Neugebauer, C. F. Davenport, F. Stanke, K. D. Larbig, S. Heeb, U. Schöck, T. M. Pohl, L. Wiehlmann, and B. Tümmler. 2010. Genome Diversity of *Pseudomonas aeruginosa* PAO1 Laboratory Strains. *J. Bacteriol.* 192:1113–1121.
- Konstan, M. W., W. J. Morgan, S. M. Butler, D. J. Pasta, M. L. Craib, S. J. Silva, D. C. Stokes, M. E. B. Wohl, J. S. Wagener, W. E. Regelman, and C. A. Johnson. 2007. Risk Factors For Rate of Decline in Forced Expiratory Volume in One Second in Children and Adolescents with Cystic Fibrosis. *J. Pediatrics* 151:134–139.
- Krause, S., P. M. Van Bodegom, W. K. Cornwell, and P. L. E. Bodelier. 2014. Weak phylogenetic signal in physiological traits of methane-oxidizing bacteria. *J. Evol. Biol.* 27:1240–1247.
- Kristidis, P., D. Bozon, M. Corey, D. Markiewicz, J. Rommens, L.-C. Tsui, and P. Durie. 1992. Genetic Determination of Exocrine Pancreatic Function in Cystic Fibrosis. *Am. J. Hum. Genet.* 50:1178–1184.
- Kukavica-Ibrulj, I., A. Bragonzi, M. Paroni, C. Winstanley, F. Sanschagrin, G. A. O’Toole, and R. C. Levesque. 2008. In Vivo Growth of *Pseudomonas aeruginosa* Strains PAO1 and PA14 and the Hypervirulent Strain LESB58 in a Rat Model of Chronic Lung Infection. *J. Bacteriol.* 190:2804–2813.

- Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.
- Lamont, I. L., P. A. Beare, U. Ochsner, A. I. Vasil, and M. L. Vasil. 2002. Siderophore-Mediated Signaling Regulates Virulence Factor Production in *Pseudomonas aeruginosa*. *PNAS* 99:7072–7077.
- Lau, G. W., D. J. Hassett, H. Ran, and F. Kong. 2004. The Role of Pyocyanin in *Pseudomonas aeruginosa* Infection. *Trends Mol. Med.* 10:599–606.
- Lenski, R. E., and M. Travisano. 1994. Dynamics of Adaptation and Diversification: A 10,000-Generation Experiment with Bacterial Populations. *PNAS* 91:6808–6814.
- Levin, B. R. 1988. Frequency-Dependent Selection in Bacterial Populations. *Philos. Trans. R. Soc. B* 319:459–472.
- Levinthal, D. A. 1997. Adaptation on Rugged Landscapes. *Manage. Sci.* 43:934–950.
- Lyczak, J. B., C. L. Cannon, and G. B. Pier. 2002. Lung Infections Associated with Cystic Fibrosis. *Clin. Microbiol. Rev.* 15:194–222.
- Maclean, R. C., and A. Buckling. 2009. The Distribution of Fitness Effects of Beneficial Mutations in *Pseudomonas aeruginosa*. *PLoS Genet* 5:e1000406.
- Martins, V. V., A. Pitondo-Silva, L. de Melo Manço, J. P. Falcão, S. dos S. Freitas, W. D. da Silveira, and E. G. Stehling. 2014. Pathogenic Potential and Genetic Diversity of Environmental and Clinical Isolates of *Pseudomonas aeruginosa*. *Apmis* 122:92–100.
- Martiny, A. C., K. Treseder, and G. Pusch. 2013. Phylogenetic Conservatism of Functional Traits in Microorganisms. *ISME J.* 7:830–838.
- Marvig, R. L., L. M. Sommer, S. Molin, and H. K. Johansen. 2015. Convergent Evolution and Adaptation of *Pseudomonas aeruginosa* within Patients with Cystic Fibrosis. *Nat. Genet.* 47:57–64.

- McKone, E. F., S. S. Emerson, K. L. Edwards, and M. L. Aitekne. 2003. Effect of Genotype on Phenotype and Mortality in Cystic Fibrosis: a Retrospective Cohort Study. *Lancet* 361:1671–1676.
- Melnyk, A. H., and R. Kassen. 2011. Adaptive Landscapes in Evolving Populations of *Pseudomonas fluorescens*. *Evolution* (N. Y). 65:3048–3059.
- Merritt, J. H., D. E. Kadouri, and G. A. O’Toole. 2005. Growing and Analyzing Static Biofilms. *Curr. Protoc. Microbiol.* (Supplement B), 1B.1.1–1B.1.18.
- Meyer, J. M. 2000. Proverdines: Pigments, Siderophores and Potential Taxonomic Markers of Fluorescent *Pseudomonas* Species. *Arch. Microbiol.* 174:135–142.
- Moore, F. B. G., and R. Woods. 2006. Tempo and Constraint of Adaptive Evolution in *Escherichia coli* (Enterobacteriaceae, Enterobacteriales). *Biol. J. Linn. Soc.* 88:403–411.
- Moran, P. A. P. 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37:17–23.
- Morita, Y., J. Tomida, and Y. Kawamura. 2012. MexXY Multidrug Efflux System of *Pseudomonas aeruginosa*. *Front. Microbiol.* 3:1–13.
- Münkemüller, T., S. Lavergne, B. Bzeznik, S. Dray, T. Jombart, K. Schiffers, and W. Thuiller. 2012. How to Measure and Test Phylogenetic Signal. *Methods Ecol. Evol.* 3:743–756.
- O’Neal, W. K., P. Gallins, R. G. Pace, H. Dang, W. E. Wolf, L. C. Jones, X. Guo, Y.-H. Zhou, V. Madar, J. Huang, L. Liang, M. F. Moffatt, G. R. Cutting, M. L. Drumm, J. M. Rommens, L. J. Strug, W. Sun, J. R. Stonebraker, F. A. Wright, and M. R. Knowles. 2015. Gene Expression in Transformed Lymphocytes Reveals Variation in Endomembrane and HLA Pathways Modifying Cystic Fibrosis Pulmonary Phenotypes. *Am. J. Hum. Genet.* 96:318–328.
- O’Sullivan, B. P., and S. D. Freedman. 2009. Cystic Fibrosis. *Lancet* 373:1891–1904.

- Orr, H. A. 2005. The Genetic Theory of Adaptation: a Brief History. *Nat. Rev. Genet.* 6:119–127.
- Pagel, M. 1999. Inferring the Historical Patterns of Biological Evolution. *Nature* 401:877–884.
- Palmer, K. L., L. M. Aye, and M. Whiteley. 2007. Nutritional Cues Control *Pseudomonas aeruginosa* Multicellular Behavior in Cystic Fibrosis Sputum. *J. Bacteriol.* 189:8079–8087.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of Phylogenetics and Evolution in R Language. *Bioinformatics* 20:289–290.
- Phillips, P. C. 2008. Epistasis—the Essential Role of Gene Interactions in the Structure and Evolution of Genetic Systems. *Nat. Rev. Genet.* 9:855–867.
- Purdy-Gibson, M. E., M. France, T. C. Hundley, N. Eid, and S. K. Remold. 2015. *Pseudomonas aeruginosa* in CF and Non-CF Homes is Found Predominantly in Drains. *J. Cyst. Fibros.* 14:341–346. European Cystic Fibrosis Society.
- R Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria.
- Rashid, M. H., and A. Kornberg. 2000. Inorganic Polyphosphate is Needed for Swimming, Swarming, and Twitching Motilities of *Pseudomonas aeruginosa*. *PNAS* 97:4885–4890.
- Regnath, T., M. Kreuzberger, S. Illing, R. Oehme, and O. Liesenfeld. 2004. Prevalence of *Pseudomonas aeruginosa* in Households of Patients with Cystic Fibrosis. *Int. J. Hyg. Environ. Health* 207:585–588.
- Remold, S. K., C. K. Brown, J. E. Farris, T. C. Hundley, J. A. Perpich, and M. E. Purdy. 2011. Differential Habitat Use and Niche Partitioning by *Pseudomonas* Species in Human Homes. *Microb. Ecol.* 62:505–517.

- Revell, L. J. 2012. phytools: An R Package for Phylogenetic Comparative Biology (and Other Things). *Methods Ecol. Evol.* 3:217–223.
- Rodriguez-Torres, M. D., Á. Islas-Robles, Z. Gómez-Lunar, L. Delaye, I. Hernández-González, V. Souza, M. Travisano, and G. Olmedo-álvarez. 2017. Phenotypic Microdiversity and Phylogenetic Signal Analysis of Traits Related to Social Interaction in *Bacillus* spp. from Sediment Communities. *Front. Microbiol.* 8:1–13.
- Rokyta, D. R., Z. Abdo, and H. A. Wichman. 2009. The Genetics of Adaptation for Eight Microvirid Bacteriophages. *J. Mol. Evol.* 69:229–39. Springer.
- Rowntree, R. K., and A. Harris. 2013. The Phenotypic Consequences of CFTR Mutations. *Ann. Hum. Genet.* 67:471–485.
- Saiman, L., J. D. Siegel, J. J. Lipuma, R. F. Brown, E. A. Bryson, M. J. Chambers, V. S. Downer, J. Fliege, L. A. Hazle, M. Jain, B. C. Marshall, C. O'Malley, S. R. Pattee, G. Potter-Bynoe, S. Reid, K. A. Robinson, K. A. Sabadosa, H. J. Schmidt, E. Tullis, J. Webber, and D. J. Weber. 2014. Infection Prevention and Control Guideline for Cystic Fibrosis: 2013 Update. *Infect. Control Hosp. Epidemiol.* 35:s1–s67.
- Salunkhe, P., C. H. M. Smart, J. A. W. Morhan, S. Panagea, M. J. Walshaw, C. A. Hart, R. Geffers, B. Tümmler, and C. Winstanley. 2005. A Cystic Fibrosis Epidemic Strain of *Pseudomonas aeruginosa* Displays Enhanced Virulence and Antimicrobial Resistance. *J. Bacteriol.* 187:4908–4920.
- Schelstraete, P., S. Van Daele, K. De Boeck, M. Proesmans, P. Lebecque, J. Leclercq-Foucart, A. Malfroot, M. Vaneechoutte, and F. De Baets. 2008. *Pseudomonas aeruginosa* in the Home Environment of Newly Infected Cystic Fibrosis Patients. *Eur. Respir. J.* 31:822–829.
- Schreiber, K., R. Krieger, B. Benkert, M. Eschbach, H. Arai, M. Schobert, and D. Jahn. 2007. The Anaerobic Regulatory Network Required for *Pseudomonas aeruginosa* Nitrate Respiration. *J. Bacteriol.* 189:4310–4314.
- Smith, E. E., D. G. Buckley, Z. Wu, C. Saenphimmachak, L. R. Hoffman, D. A.

- D'Argenio, S. I. Miller, B. W. Ramsey, D. P. Speert, S. M. Moskowitz, J. L. Burns, R. Kaul, and M. V Olson. 2006. Genetic Adaptation by *Pseudomonas aeruginosa* to the Airways of Cystic Fibrosis Patients. *PNAS* 103:8487–8492.
- Spor, A., D. J. Kvitek, T. Nidelet, J. Martin, J. Legrand, C. Dillmann, A. Bourgeois, D. de Vienne, G. Sherlock, and D. Sicard. 2014. Phenotypic and Genotypic Convergences are Influenced by Historical Contingency and Environment in Yeast. *Evolution* (N. Y). 68:772–790.
- Sriramulu, D. D., H. Lünsdorf, J. S. Lam, and U. Römling. 2005. Microcolony Formation: a Novel Biofilm Model of *Pseudomonas aeruginosa* for the Cystic Fibrosis Lung. *J. Med. Microbiol.* 54:667–676.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. L. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K-S Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. W Hancock, S. Lory, and M. V Olson. 2000. Complete Genome Sequence of *Pseudomonas aeruginosa* PAO1, an Opportunistic Pathogen. *Nature* 406:959–964.
- Streeter, K., C. Neuman, J. Thompson, E. Hatje, and M. Katouli. 2016. The Characteristics of Genetically Related *Pseudomonas aeruginosa* from Diverse Sources and their Interaction with Human Cell Lines. *Can. J. Microbiol.* 62:233–240.
- Struelens, M. J., V. Schwam, A. Deplano, and D. Baran. 1993. Genome Macrorestriction Analysis of Diversity and Variability of *Pseudomonas aeruginosa* Strains Infecting Cystic Fibrosis Patients. *J. Clin. Microbiol.* 31:2320–2326.
- Symonds, M. R. E., and S. P. Blomberg. 2014. A Primer on Phylogenetic Generalised Least Squares. Pp. 105–130 in L. Garamszegi, ed. *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*. Springer, Berlin, Heidelberg.

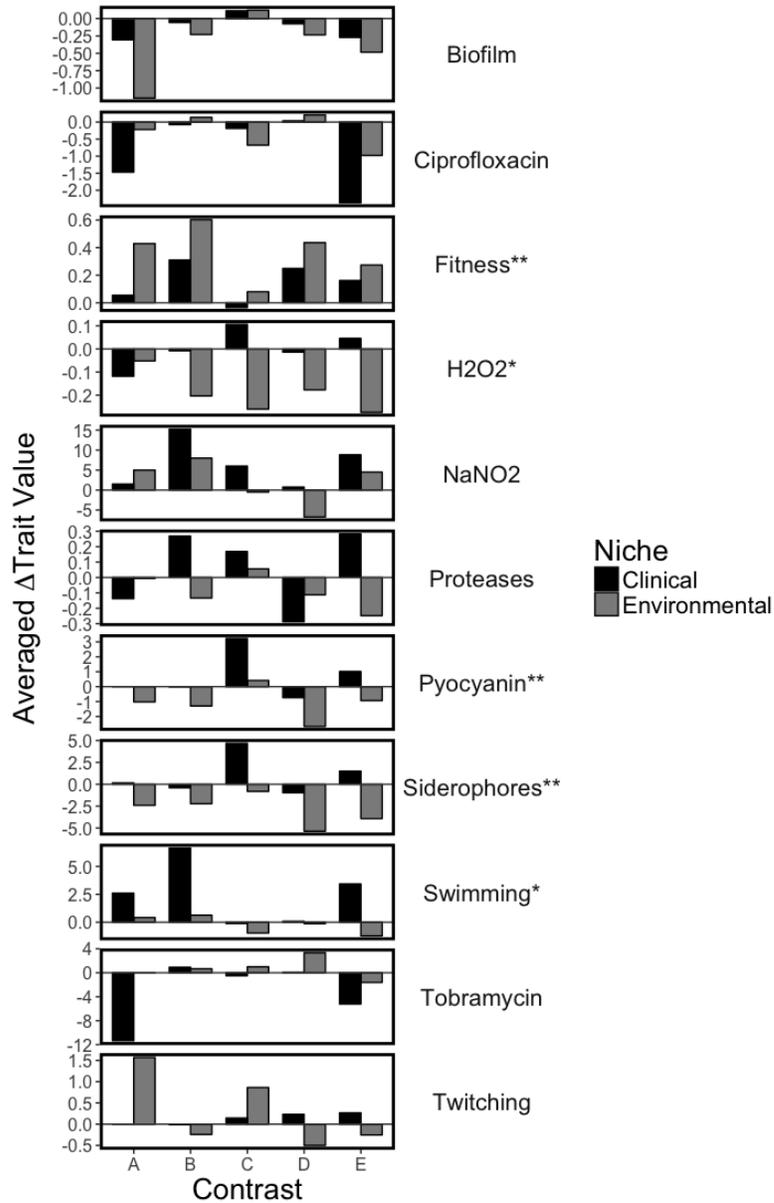
- Tenaillon, O., A. Rodríguez-verdugo, R. L. Gaut, P. McDonald, A. F. Bennett, A. D. Long, and B. S. Gaut. 2012. The Molecular Diversity of Adaptive Convergence. *Science* (80-. ). 335:457–461.
- Travisano, M., J. A. Mongold, A. F. Bennett, and R. E. Lenski. 1995. Experimental Tests of the Roles of Adaptation, Chance and History in Evolution. *Science* (80-. ). 267:87–90.
- Turner, K. H., A. K. Wessel, G. C. Palmer, J. L. Murray, and M. Whiteley. 2015. Essential Genome of *Pseudomonas aeruginosa* in Cystic Fibrosis Sputum. *PNAS* 112:4110–4115.
- Valerius, N. ., C. Koch, and N. Hoiby. 1991. Prevention of Chronic *Pseudomonas aeruginosa* Colonisation in Cystic Fibrosis by Early Treatment. *Lancet* 338:725–726.
- Vogwill, T., M. Kojadinovic, and R. C. Maclean. 2016. Epistasis Between Antibiotic Resistance Mutations and Genetic Background Shape the Fitness Effect of Resistance across Species of *Pseudomonas*. *Proc. R. Soc. B* 283:20160151.
- Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: Sign Epistasis and Genetic Constraint on Evolutionary Trajectories. *Evolution* (N. Y). 59:1165–1174.
- West, S. E. H., L. Zeng, B. L. Lee, M. R. Kosorok, A. Laxova, M. J. Rock, M. J. Splaingard, and P. M. Farrell. 2002. Respiratory Infections With *Pseudomonas aeruginosa* in Children With Cystic Fibrosis. *Jama* 287:2958–2966.
- Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.
- Willner, D., M. R. Haynes, M. Furlan, R. Schmieder, Y. W. Lim, P. B. Rainey, F. Rohwer, and D. Conrad. 2012. Spatial Distribution of Microbial Communities in the Cystic Fibrosis Lung. *ISME J.* 6:471–474.
- Winstanley, C., S. B. Kaye, T. J. Neal, H. J. Chilton, S. Miksch, and C. A. Hart. 2005.

- Genotypic and Phenotypic Characteristics of *Pseudomonas aeruginosa* Isolates Associated with Ulcerative Keratitis. *J. Med. Microbiol.* 54:519–526.
- Winstanley, C., S. O'Brien, and M. A. Brockhurst. 2016. *Pseudomonas aeruginosa* Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends Microbiol.* 24:327–337.
- Wong, A. 2017. Epistasis and the Evolution of Antimicrobial Resistance. *Front. Microbiol.* 8:246.
- Wong, A., and R. Kassen. 2011. Parallel Evolution and Local Differentiation in Quinolone Resistance in *Pseudomonas aeruginosa*. *Microbiology* 157:937–944.
- Wong, A., N. Rodrigue, and R. Kassen. 2012. Genomics of Adaptation during Experimental Evolution of the Opportunistic Pathogen *Pseudomonas aeruginosa*. *PLoS Genet.* 8:e1002928.
- Wong, A., and K. Seguin. 2015. Effects of genotype on rates of substitution during experimental evolution. *Evolution (N. Y.)*. 69:1772–1785.
- Wozniak, D. J., T. J. O. Wyckoff, M. Starkey, R. Keyser, P. Azadi, G. A. O'Toole, and M. R. Parsek. 2003. Alginate is Not a Significant Component of the Extracellular Polysaccharide Matrix of PA14 and PAO1 *Pseudomonas aeruginosa* Biofilms. *PNAS* 100:7907–7912.
- Wu, X., S. Monchy, S. Taghavi, W. Zhu, J. Ramos, and D. van der Lelie. 2011. Comparative Genomics and Functional Analysis of Niche-Specific Adaptation in *Pseudomonas putida*. *Microbiol. Rev.* 35:299–323.
- Yang, L., L. Jelsbak, R. L. Marvig, S. Damkiaer, C. T. Workman, M. H. Rau, S. K. Hansen, A. Folkesson, H. K. Johansen, O. Ciofu, N. Hoiby, M. O. A. Sommer, and S. Molin. 2011. Evolutionary Dynamics of Bacteria in a Human Host Environment. *PNAS* 108:7481–7486.
- Zemanick, E. T., B. D. Wagner, C. E. Robertson, M. J. Stevens, S. J. Szeffler, F. J.

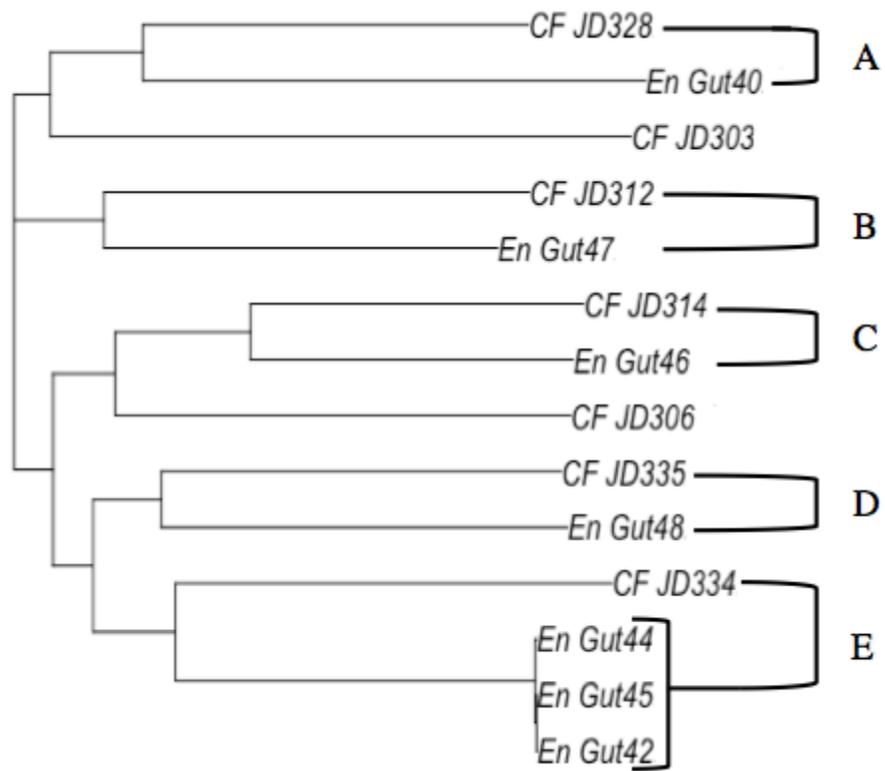
Accurso, S. D. Sagel, and J. K. Harris. 2015. Assessment of Airway Microbiota and Inflammation in Cystic Fibrosis using Multiple Sampling Methods. *Ann. Am. Thorac. Soc.* 12:221–229.

# Appendices

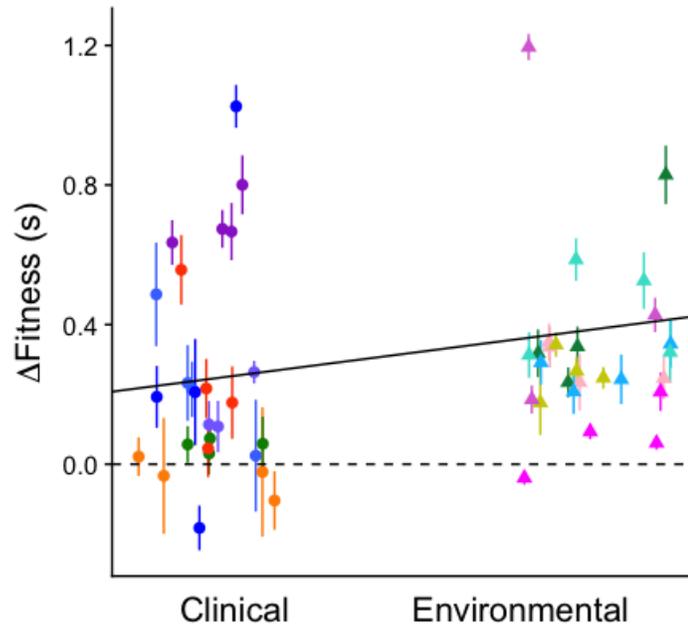
## A Niche comparisons excluding *En Jp119* and *En Jp129*



**Figure A1.** Independent contrasts analysis directly comparing environmental and clinical strains, based on a phylogeny that excludes environmental strains *En Jp119* and *En Jp129*. A correlational analysis of  $\Delta$ fitness and niche also revealed a significant difference between the two niches ( $r = 0.28$ ,  $\log\text{BF} > 3.06$ )



**Figure A2.** Phylogenetic tree excluding test strains PaO1, Pa14; and environmental strains En Jp119, En Jp129. Independent contrasts directly comparing environmental and clinical strains are labeled A-E.



**Figure A3.** pGLS analysis comparing fitness changes between niches, that exclude environmental strains En Jp119, En Jp129.  $p = 0.22$ ,  $\lambda = 0.44$ , intercept = 0.25 slope = 0.14 (0.11),  $n=54$ .

**B Mini literature review of Pa14 and PaO1**

**Table B1.** The ratios of Pa14 and PaO1 in their ancestral trait values across studies that have previously investigated both strains in conjunction.

Trait	Finding	Our Result	Literature		
Pyocyanin	Author	Filipow	Dietrich		
	Ratio	1:2.4	10:1		
	Relativity	Pa14<PaO1	Pa14>PaO1		
Twitching	Author	Filipow	Winstanley	Head	Kukavica-Ibrulj
	Ratio	1:3	1:1.86	26:1	-
	Relativity	Pa14<PaO1	Pa14<PaO1	Pa14>PaO1	Pa14=PaO1
Swimming	Author	Filipow	Head	Kukavica-Ibrulj	
	Ratio	1.74:1	0.8:1.0	-	
	Relativity	Pa14>PaO1	Pa14<PAO1	Pa14=PaO1	
Biofilm	Author	Filipow	Head	Kukavica-Ibrulj	
	Ratio	1:2.7	-	1:0.91	
	Relativity	Pa14<PaO1	Pa14<PaO1	Pa14>PaO1	
Tobra	Author	Filipow	Fernández	Wozniak	
	Ratio	3:1	1:1	1:1	
	Relativity	Pa14>PaO1	Pa14:PaO1	Pa14:PaO1	
Cipro	Author	Filipow	Fernández	Wozniak	
	Ratio	2:1	1:1	1:5	
	Relativity	Pa14>PaO1	Pa14:PaO1	Pa14<PaO1	