The paediatric cystic fibrosis lung: understanding the evolving microbiome, antimicrobial resistance and novel approaches to treatment

A thesis submitted in fulfilment of the requirement of the degree of
Doctor of Medicine
Cardiff University

Dr Juliette Louise Oakley
C1849203
Supervisors: Professor David Thomas, Dr Julian Forton, Professor Tim Walsh, Dr Manon Pritchard
October 2022

<u>Acknowledgements</u>

<u>Summary</u>

CF pulmonary disease is characterised by recurrent and persistent bacterial infections, which eventually lead to premature death. Given mounting concerns regarding antimicrobial resistance, novel surveillance methodology and therapeutics are needed to enable effective treatment of CF lung infections.

Current sampling techniques for studying the paediatric CF lung microbiome are invasive and time-consuming. This thesis used paediatric induced sputum (IS) samples from the CF Sputum Induction Trial (Ronchetti et al. 2018a) to study microbial diversity and the effects of clinical treatment. The dataset represents the largest study to date using exclusively paediatric IS samples. Results demonstrated clear correlations between clinical features and diversity measures, mirroring previous literature obtained from more invasive techniques and supporting the use of IS as a routine surveillance method.

OligoG CF-5/20 is a novel anti-biofilm therapy currently undergoing clinical trials. To understand its mechanism of action and potential clinical effects, this work used Fourier transform infrared spectroscopy to analyse a small set of CF sputum samples from the CF SpIT trial and explore the interaction of OligoG CF-5/20 with respiratory mucin. Results demonstrated interaction at key mucin structures including glycan moieties and the peptide backbone, providing a potential mechanism of action to explain the modification of CF sputum.

Reflecting the intended clinical use of OligoG CF-5/20 as a prolonged treatment for patients, an evolutionary model was utilised to study the effects of OligoG CF-5/20 treatment on *P. aeruginosa*. Phenotypic and genotypic characterisation of *P. aeruginosa* exposed to 2% OligoG CF-5/20 showed a reduction in colonies with multi-drug resistant-associated phenotypes. Exposure to 2% OligoG CF-5/20 and azithromycin improved bacterial susceptibility to other classes of antibiotics.

These studies provide an insight into the role of culture-independent methods for airway sampling and the development of a novel therapy to treat multi-drug resistant bacteria in children with CF.

Abbreviations

α	Alpha
AZM	Azithromycin
β	Beta
bp	Base pair
ASL	Airway surface liquid
BA	Blood agar
BAL	Bronchoalveolar lavage
BMI	Body mass index
Ca^{2+}	Calcium cation
CF	Cystic fibrosis
CF-SpIT	CF-sputum induction trial
CFTR	Cystic fibrosis transmembrane conductance regulator
Cl^-	Chloride anion
CLSM	Confocal laser scanning microscopy
COPD	Chronic obstructive pulmonary disease
CV	Crystal violet
D	Dalton
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DNase	Dornase alfa
DTGS	Deuterated triglycerine sulfate
EPS	Extracellular polymeric substance
eDNA	Extracellular deoxyribonucleic acid
ENaC	Epithelial sodium channel
FEV1	Forced expiratory volume in 1 second
FTIR	Fourier transform infrared spectroscopy
G	G-force
GalNAc	*N*-acetylgalactosamine
GEE	Generalising estimating equations
GlcNAc	N-acetylglucosamine
H^+	Hydrogen cation
$HCO3^-$	Bicarbonate anion
Ig	Immunoglobulin
IL	Interleukin

IR	Infrared
IRT	Immunoreactive trypsin
IS	Induced sputum
KBr	Potassium bromide
LB	Luria-Bertani (agar)
LPS	Lipopolysaccharides
LRT	Lower respiratory tract
MDR	Multi-drug resistant/multi-drug resistance
MH	Mueller-Hinton (broth)
MIC	Minimum inhibitory concentration
mmol/L	Millimoles per Litre
Na^+	Sodium cation
O/N	Overnight
OTU	Operational taxonomic unit
P. aeruginosa	*Pseudomonas aeruginosa*
PCL	Periciliary layer
PCR	Polymerase chain reaction
QIC	Quasi Likelihood under Independence Model Criterion
qPCR	Quantitative polymerase chain reaction
Q-Q plot	Quantile-quantile plot
QS	Quorum sensing
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
QS	Quorum-sensing
S. aureus	*Staphylococcus aureus*
SEM	Scanning electron microscopy
SCV	Small colony variant
Spp.	Species
TNFα	Tumour necrosis factor alpha
TSB	Tryptone soy broth
UK	United Kingdom
URT	Upper respiratory tract
VNTR	Variable number of tandem repeats
v/v	Volume/volume
vWF	von Willebrand factor
WT PAO1	Wild-type PAO1 (*Pseudomonas aeruginosa*)

Preface

Presented work from thesis to date

Peer-reviewed papers published

Weiser R, **Oakley JL**, Ronchetti K, Tame JD, Hoehn S, Jurkowski TP, Mahenthiralingam E, Forton JT. (2022) The lung microbiota in children with cystic fibrosis captured by induced sputum sampling. *Journal of Cystic Fibrosis.* In Press.
DOI: https://doi.org/10.1016/j.jcf.2022.01.006
Data included in Chapter 2.

Oakley JL, Weiser R, Powell LC, Forton J, Mahenthiralingam E, Rye PD, Hill KE, Thomas DW and Pritchard MF (2021). Phenotypic and genotypic adaptations in *pseudomonas aeruginosa* biofilms following long-term exposure to an alginate oligomer therapy. *mSphere* 6(1): e01216-20.
DOI: https://doi.org/10.1128/mSphere.01216-20
Data included in Chapter 3.

Pritchard MF, **Oakley JL**, Brilliant CD, Rye PD, Forton J, Doull IJ, Ketchell I, Hill KE, Thomas DW and Lewis PD (2019). Mucin structural interactions with an alginate oligomer mucolytic in cystic fibrosis sputum. *Vibrational Spectroscopy 103:* 102932.
DOI: https://doi.org/10.1016/j.vibspec.2019.102932
Data included in Chapter 4.

Data in Chapters 3 and 4 of this thesis has been published in the aforementioned peer-reviewed articles. This includes some of the graphs, tables and illustrations included in this thesis. According to copyright, permissions have been provided by the following journals: *mSphere* and *Vibrational Spectroscopy*.

Abstracts published

Weiser R, **Oakley JL**, Mahenthiralingam E and Forton J (2020). Use of microbiota analysis to extend the validation of induced-sputum as a sampling technique for lung infections in children with cystic fibrosis. *Pediatric Pulmonology* 55: S180-S180.

Weiser R, **Oakley JL**, Mahenthiralingam E and Forton J (2019). Microbiota from paired sputum-induction and bronchoalveolar lavage samples in children with cystic fibrosis: results from The Cystic Fibrosis Sputum Induction Trial (CF-SpIT). *Journal of Cystic Fibrosis* 18(1): S35.

Weiser R, **Oakley JL**, Jones C, Mahenthiralingam E and Forton J (2019). Investigating evolution of the paediatric cystic fibrosis lung microbiota using induced sputum sampling and culture-independent techniques. *Access Microbiology 1*(1A): 822.

Oakley JL, Pritchard MF, Powell LC, Forton J, Doull I, Rye PD, Hill KE and Thomas DW (2018). Alginate oligomers as novel therapies to treat life-threatening pseudomonal multidrug resistant bacterial infections. *Journal of Cystic Fibrosis* 17(3): S19-S20.

International Presentations

Use of microbiota analysis to extend the validation of induced-sputum as a sampling technique for lung infections in children with cystic fibrosis**,** Poster presentation, **North American Cystic Fibrosis Conference, Online, 2020**

Microbiota from paired sputum-induction and bronchoalveolar lavage samples in children with CF: results from the Cystic Fibrosis Sputum Induction Trial (CF-SpIT) Oral presentation, **42nd European Cystic Fibrosis Conference, Liverpool, 2019**

Alginate oligomers as novel therapies to treat life-threatening multi-drug resistant pseudomonal infections, Oral presentation, **41st European Cystic Fibrosis Society Meeting, Belgrade, 2018**

Generating novel polymer therapies to treat life-threatening pseudomonal multi-drug resistant bacterial infections, Poster presentation, **16th International Conference on Pseudomonas, Liverpool, 2017**

National Presentations

Investigating evolution of the paediatric cystic fibrosis lung microbiota using induced sputum sampling and culture-independent techniques, Oral presentation, **Microbiology Society Annual Conference, Belfast, 2019**

Alginate oligomers as novel therapies to treat life-threatening pseudomonal multi-drug resistant infections, Poster presentation, **BPRS 20th Summer Symposium, Cambridge, 2018**

Table of contents

List of figures

Figures published in (Oakley et al. 2021).

List of tables

Chapter One

Introduction

1.1 Introduction

Cystic fibrosis (CF) is an autosomal recessive disease, affecting more than 10,400 people in the UK (Cystic Fibrosis Trust 2019, 2nd February-b). It is a life-shortening condition. Babies born today with CF have a median predicted survival of 47 years according to recent data from the UK Cystic Fibrosis Trust (Charman et al. 2018). Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene (Katkin 2019b). The effects of CFTR mutations are seen in multiple systems, including the respiratory, gastrointestinal and genitourinary systems. The disease phenotype is characterised by progressive lung disease, exocrine pancreatic insufficiency and resulting gastrointestinal malabsorption, malnutrition and growth impairment, sinus disease and CF-related diabetes (Ratjen et al. 2015). The pulmonary manifestations of the disease are the principal cause of morbidity and mortality, and their management is the focus of this work.

1.2 Cystic fibrosis

1.2.1 Genetics

Cystic fibrosis is a heterogenous disease, with a mutation in a single gene, CFTR, which is located on the long arm of Chromosome 7. The CFTR gene was first identified in 1989 and research is ongoing to fully understand its role (Riordan et al. 1989). To date, there are 2063 known mutations in the CFTR gene, though not all of these are disease causing (see 'Cystic fibrosis mutations database', and 'CFTR2'). The most common mutation is F508del, with 89.5% of the UK CF population carrying at least one of this mutation (Charman et al. 2018).

CFTR mutations are grouped in six classes according to their functional consequences. Mutations can reduce channel number, protein function or both, and the resulting phenotypic severity can vary significantly (Ratjen et al. 2015). Mutations resulting in loss of CFTR expression on the cell surface or loss of CFTR function are considered severe mutations, typically displaying both lung disease and pancreatic insufficiency (Ratjen et al. 2015). Less severe phenotypes are often caused by mutations with residual CFTR function (Ratjen et al. 2015). The patient's genotype has a significant influence on disease severity and the associated clinical manifestations, though environmental factors and non-CFTR gene modifiers also play

a role (Ratjen et al. 2015). Genotyping is important prognostically and also guides the development and implementation of existing modulatory therapies.

1.2.2 Clinical diagnosis

A newborn screening programme for CF has been in place throughout the United Kingdom since 2007. All babies are screened in the first week of life using a blood sample taken from the heel. The UK uses the immunoreactive trypsin (IRT)-DNA-IRT protocol, which identifies babies with a raised IRT and then screens these children for the most common CF mutations. Infants identified to have two CFTR disease-causing mutations, or one or no gene mutations but raised IRT, are referred to CF specialist centres for a sweat test (Edmondson et al. 2018).

Diagnosis is made following demonstration of a raised sweat chloride level (>60 mmol/L), indicating a positive sweat test, in combination with a positive newborn screen. For older children and adults not identified by screening, there should be evidence of possible CF-related disease in at least one organ plus evidence of CFTR-mutation to confirm diagnosis (Katkin 2019a).

1.2.3 Treatment for CF-related lung disease

Developments in CF management have dramatically improved patient outcomes, but treatments remain complex and onerous for patients. Despite the introduction of screening, specialist multi-disciplinary centres and expensive therapies, the disease is still associated with significant morbidity and early death (Castellani et al. 2018).

Clinical treatment predominantly focuses on preventing and treating infections (with oral, nebulised and intravenous antimicrobials), airway clearance (with mucolytic therapies and physiotherapy) and anti-inflammatory agents to reduce chronic inflammation (De Boeck and Amaral 2016). Aggressive therapy is required for proven infections, often involving hospital admissions and intravenous antibiotics.

There have been significant advances in CF therapies in recent years, particularly with regards to CFTR modulators, which work at the level of the protein dysfunction (Bell et al. 2020; Konstan and Flume 2020). However, no single treatment is yet able to prevent the sequelae that eventually lead to progressive disease, and a combination of therapies is required. Consequently, the burden of daily treatment for patients and their families remains huge (Brennan 2020).

1.3 The respiratory tract

The respiratory tract is composed of conducting airways (nose, mouth, pharynx, larynx, trachea, bronchi and conducting bronchioles) and respiratory airways (respiratory bronchioles, alveolar ducts and alveoli) (Saint-Criq and Gray 2017). Air is warmed, humidified and conducted to the respiratory airways where gaseous exchange occurs. The bronchial epithelium is predominantly composed of ciliated cells, goblet cells and basal cells, with submucosal glands including ciliated and non-ciliated serous and mucous cells residing in the cartilaginous conducting airways (Saint-Criq and Gray 2017). The alveolar epithelium consists of alveolar type I and type II cells (Saint-Criq and Gray 2017).

Airway mucus consists of 90-98% water, ions (at similar concentrations to plasma), gel-forming mucins and many other proteins, peptides and small molecules (Widdicombe and Wine 2015). It is a key component of innate immunity (Borowitz 2015). Grossly, it protects the underlying epithelium by providing a physical barrier and pH buffer, while the smaller macromolecular components provide protective antimicrobial, anti-inflammatory and antioxidant properties (Widdicombe and Wine 2015). It is continuously removed and replaced through the process of mucociliary clearance (Widdicombe and Wine 2015). Mucus is removed, along with entrapped particles, by the activity of cilia in the apical membranes of epithelial cells (Widdicombe and Wine 2015).

The conducting airway epithelium is protected from respiratory gases by a thin liquid layer called the airway surface liquid (ASL), which aids the removal of pathogens and particulate matter by mucociliary clearance (Saint-Criq and Gray 2017). This film of liquid consists of two phases: the 'periciliary layer' lies between the cilia and is free of gel-forming mucins; and a mucous gel which just touches the tips of the cilia (Widdicombe and Wine 2015). ASL hydration is carefully managed through ion/water transport across the surface epithelium, plus fluid secretion from the submucosal glands (Saint-Criq and Gray 2017).

By active secretion of chloride and bicarbonate ions, airway glands are able to facilitate secretion of water across the epithelium (Widdicombe and Wine 2015). These processes pull sodium ions into the lumen via the paracellular pathway by creating a negatively-charged environment. The resulting trans-epithelial

concentration gradient enables the osmotic movement of water into the lumen, mainly via the transcellular pathway (Widdicombe and Wine 2015). CFTR and calcium ion (Ca^{2+}) -activated chloride channels predominantly affect anion secretion at the apical membrane, though other transporters and pathways are known to be involved (Widdicombe and Wine 2015; Saint-Criq and Gray 2017). CFTR, alongside ANO1, also conducts bicarbonate ions, which contributes to the volume and pH of fluid secreted (Saint-Criq and Gray 2017).

1.4 The respiratory tract in cystic fibrosis

CFTR dysfunction has a number of key pulmonary manifestations. These include abnormal mucus adhesion and structure, delayed mucociliary clearance due to flattened cilia and airway fluid liquid (ASL) depletion, dysregulated inflammation and abnormal mucosal defences leading to infection (Ratjen et al. 2015). These combined features lead to a cascade of local airway destruction, lung injury, eventual bronchiectasis, and finally respiratory failure (Ratjen et al. 2015).

1.4.1 CFTR dysfunction

CFTR principally acts as a chloride channel, transporting ions across the apical membrane of the epithelium (Ratjen et al. 2015). It is also involved in bicarbonate secretion and inhibition of sodium transport (Ratjen et al. 2015). Chloride ion transport across the apical membrane relies on three factors: CFTR activity or open state probability; the number of CFTR channels; and the single channel conductance, which is affected by the electrochemical gradient across the membrane (Saint-Criq and Gray 2017). It is proposed that within the CF respiratory tract, impaired bicarbonate transport leads to: reduced removal of Ca^{2+} from the condensed mucins, dysfunctional mucin expansion, stasis of the mucus within ducts and increased mucus viscosity (Quinton 2008; Borowitz 2015).

Mutations in CFTR reduce electrolyte-driven fluid secretion from airway glands and surface epithelia, resulting in thickened, $HCO3^-$ -deficient secretions that are difficult to clear (Widdicombe and Wine 2015). In CF, the fluid secreted by the glands is reduced in volume whilst being increased in viscosity (Jayaraman et al. 2001; Widdicombe and Wine 2015).

CFTR has been shown to regulate other ion channels and transporters, such as the epithelial sodium channel (ENaC), as well as influencing ASL pH through reduced HCO_3^- and Cl^- secretion (Saint-Criq and Gray 2017). The effect of endogenous antimicrobials within a depleted ASL are also compromised due to the reduced bicarbonate content and increased acidity of gland secretions (Widdicombe and Wine 2015; Saint-Criq and Gray 2017). This makes mucus more difficult to clear via mucociliary or cough clearance, as well as limiting the accessibility to antimicrobials (Widdicombe and Wine 2015; Saint-Criq and Gray 2017). The resulting static, yet hyper-secreted, airway mucus provides an ideal environment for progressive bacterial colonisation and infection, stimulating the inflammatory cascade and leading to chronic airway obstruction and eventually bronchiectasis (Widdicombe and Wine 2015; Kunzelmann et al. 2017; Katkin 2019b). Mucus can also "plug" mucus glands, becoming fixed to the goblet cells or gland orifice and contributing to further airway obstruction (Ratjen et al. 2015).

1.4.2 Inflammatory lung disease

In healthy individuals, pathogens entering the airways are cleared rapidly. Following identification by epithelial cells, inflammatory pathways are activated, neutrophils move into the airway lumen and the infection is eradicated (Saint-Criq and Gray 2017). This process is regulated by CFTR and its absence results in an upregulation of proinflammatory products (Saint-Criq and Gray 2017). CFTR absence also leads to abnormal antioxidant quantities within the ASL, resulting in increased airway inflammation, epithelial damage and abnormal reconstitution of the epithelium (Saint-Criq and Gray 2017). There is also evidence that the acidic ASL pH may contribute to the patient's ability to kill bacteria by the highly pH-sensitive innate defensins within the airways (Ratjen et al. 2015). This may be related to the CFTR-dependent bicarbonate defect in the airways and the negative effect on the function of various defensins, which operate optimally at neutral or alkaline pH (Borowitz 2015).

Chronic neutrophilic inflammation is seen extensively throughout the airways in cystic fibrosis. Neutrophils act through the release and activation of enzymes, such as neutrophil elastase, as part of the defence against bacteria (Sly et al. 2013). To prevent lung damage, the lung produces α_1-antitrypsin which binds to extracellular neutrophil elastase and stops digestion of elastin (Sly et al. 2013). Active neutrophilic inflammation is associated with increased free neutrophil elastase, elastin digestion and airway damage leading to bronchiectasis (Sly et al. 2013). It is not clear whether

lung inflammation in CF only occurs as a result of infection or whether other factors, such as abnormal mucus, are proinflammatory regardless of infection (Ratjen et al. 2015). Bronchoalveolar lavage (BAL)-based studies have shown free neutrophil elastase activity in BAL fluid to be a key risk factor for early, persistent bronchiectasis in CF patients, in addition to pulmonary infection (Sly et al. 2013). In particular, lung lobes with more severe bronchiectasis are associated with more extensive inflammation (Sly et al. 2013). In one series of infants identified through newborn screening, 77.2% of patients had detectable levels of the pro-inflammatory cytokine IL-8, and 29.8% had detectable neutrophil elastase activity. The level of inflammation, regardless of infection status, was also greater than that expected in healthy infants (Sly et al. 2009). Mott et al. demonstrated neutrophilic inflammation was associated with progression of bronchiectasis and air trapping, and potentially indicative of significant future lung disease in infants with CF (Mott et al. 2012).

These studies highlight evidence of clinically significant lung disease developing in infancy, despite the absence of respiratory symptoms (Sly et al. 2009). However, the temporal relationship between infection and inflammation remains under investigation given differing outcomes from previous studies (Armstrong et al. 2005; Sly et al. 2009).

1.5 Respiratory Mucus

Mucus is produced by the healthy airway, providing a protective coating. The mucus barrier is an important component of the innate immune system, providing protection by entrapping foreign particles and pathogens and dissolving noxious gases (Thornton et al. 2008). These external insults are cleared by ciliary transport or coughing.

Contained within respiratory mucus are molecules key to host defence, including secretory IgA, collectins, defensins, cathlicidins and histatins (Thornton et al. 2008). These are likely to have direct physical interaction with mucins, as well as being simply held within solution by the biophysical properties of mucus (Thornton et al. 2008).

1.6 Respiratory mucins

1.6.1 Mucin structure

The functional and rheological properties of mucus are predominantly due to mucins. Submucosal glands contain mucous cells which secrete gel-forming mucins and fluid, and CFTR-expressing serous cells which provide the majority of the fluid from gland secretions (Widdicombe and Wine 2015; Saint-Criq and Gray 2017). Gel-forming airway mucins are also produced by goblet cells of the surface epithelium and Clara cells, with the latter forming the predominant secretory cell type within the respiratory bronchioles (Widdicombe and Wine 2015). Submucosal glands are most abundant in the nasal cavities and upper airways, as these areas are most at risk of deposition of large particles (Widdicombe and Wine 2015). They also increase at the bifurcation of airways, where there is increased turbulence and greater impact of large particles on the epithelium (Widdicombe and Wine 2015).

Mucin glycoproteins are the most significant macromolecular component of mucus gel in the healthy state (Henke et al. 2007). They are polydisperse in mass (2-40 x 10^6 Da) and length (0.5-10 µm) (Thornton et al. 1990). Mucins provide the structural framework of the mucus barrier, prevent airway dehydration, sequester pathogens and potentially clear the host-protective protein and peptides after use (Thornton et al. 2008). Mucins are classified by their MUC protein backbone, which is encoded by *MUC* genes; there are 21 known proteins within this gene family (Rose and Voynow 2006; Ma et al. 2018; Morrison et al. 2019). *MUC* genes are localised on chromosomes 1, 3, 4, 6, 7, 11, 12 and 19. There are 13 identified airway mucins (Ma et al. 2018). MUC1, MUC4, MUC5AC, MUC5B and MUC16 are the five major mucins expressed in the airways (Lillehoj et al. 2013; Taherali et al. 2018).

Respiratory mucins belong to one of three classes: membrane-tethered mucins (MUC1, MUC4, MUC16, MUC20), gel-forming mucins which are expressed by airway epithelium (MUC2, MUC5AC and MUC5B) and the secreted non-gel-forming mucin (MUC7) (Henke et al. 2004; Ma et al. 2018). The site of production depends on the specific mucin: MUC5AC and MUC5B are produced by goblet cells in the tracheobronchial surface epithelium, whereas MUC5B alone is secreted by submucosal glands (Henke et al. 2007). Typically, tethered mucins are found in the periciliary space within the respiratory tract, forming a protective barrier for the epithelium (Borowitz 2015). Gel-like mucins lie on top and trap particulate matter and microorganisms prior to removal via mucociliary clearance. MUC5B and MUC5AC

play key roles in mucociliary clearance and host defence against infection (Widdicombe and Wine 2015). Similarly, gland fluid secretion is essential to facilitate mucociliary clearance and prevent airway infections.

All gel-forming mucins, including MUC5B and MUC5AC, share a similar core protein known as the apomucin. This consists of the following structures: von Willebrand factor (vWF) domains in the N- and C-terminal regions, cysteine-rich domains throughout the protein backbone, with a cysteine knot (CK) at the C-terminal end, and a variable number of tandem repeat (VNTR) region (Morrison et al. 2019). The latter structure, the VNTR region, is the site for extensive glycosylation.

1.6.2 Mucin glycosylation

Glycosylation occurs in the endoplasmic reticulum as part of the mucin protein maturation. O-glycosylation of serine and threonine residues throughout the VNTR is initiated by linkage of *N*-acetyl-D-galactosamine (GalNAc), which can produce eight different glycan core structures. Thereafter, a backbone structure is linked and then a peripheral terminal sugar added, providing the area of mucin glycan variation. These polyanionic, hydrophilic, glycosylated blocks contain sialic and sulphate ester terminals and alternate with the cysteine-rich blocks of the core apoprotein, which are hydrophobic (Verdugo 2012). Mucins structurally resemble "bottlebrushes", with branched oligosaccharide chains arranged radially around the protein backbone (Bansil and Turner 2006). Many of the covalently attached O-glycans are sialylated or sulphated (Thornton et al. 2008). The O-linked glycans make up 80% of the mucin's molecular weight (Ma et al. 2018). With possible variation in the number of glycans per amino acid, distribution pattern and glycan size, mucins can each produce a different glycosylation profile and resulting biological properties (Morrison et al. 2019).

Gel-forming mucins are synthesised and packaged in goblet cells. The anionic sites of mucin molecules, in particular the glycan structures attached to the apomucins, are shielded by calcium and hydrogen cations in the intracellular environment. This enables these large biopolymers to be highly organised and tightly packaged inside secretory granules, prior to release (Morrison et al. 2019). It is proposed that exocytosis occurs when HCO_3^- neutralises the H^+ cations and complexes the Ca^{2+} cations, enabling the unshielded mucin anions to repel one another. The condensed mucins then expand into a mature mucus matrix and are expelled into the airway lumen (Yang et al. 2013). This process occurs by two separate signalling pathways:

a Ca^{2+}-mediated pathway which stimulates goblet cell exocytosis; and a cAMP-mediated pathway which enables HCO_3^- secretion and subsequent discharge of exocytosed mucins (Yang et al. 2013). On exocytosis, the mucins undergo volume expansion of up to 4000-fold, and this process relies on Ca^{2+}-chelation and osmotic pressure (Verdugo 2012).

1.6.3 Environmental interactions

The overall mucin structure has a significant impact on its interaction with other proteins within the lung environment. Areas of potential interaction include the formation of N- and C-terminus disulphide bonds, enabling mucin monomers to form dimers and multimers respectively. These bonds can result in linear polymers and complex multimeric networks, and will affect the mucus viscoelasticity (Morrison et al. 2019). Polymeric mucins are a subset of secreted mucins which possess cysteine-rich domains at their N and C termini. They form through end-to-end disulphide bonds and are polydisperse in mass and length, existing as a random-coil conformation in solution (Wagner et al. 2018). The gel-forming ability of these structures is affected by the polymers' degree of mucin polymerisation (Thornton et al. 2008).

In aqueous solution, mucins form reversible bonds, including chain entanglements and hydrophobic interactions, which together create a highly complex network. This is stabilised by the electrostatic repulsion provided by the negatively-charged glycan chains (Wagner et al. 2018). The strength of these interactions, and resulting elastic properties of these polymers, is affected by solution pH, and the concentration of ions (particularly Ca^{2+} and Na^+) and other small molecules within the ASL (Verdugo 2012; Wagner et al. 2018).

1.7 Mucus in health

In health, mucin production by goblet cells and submucosal glands is a carefully balanced process. The lung epithelium is able to tailor the mucin composition of mucus by altering the amounts secreted by the epithelial surface and submucosal glands; this results in mucus with different functional properties depending on the environmental challenge (Thornton et al. 2008). There is a basal level of mucin secretion at the surface epithelium, aiding constant airway protection and maintenance. Acute stimulation to degranulate can be prompted by smoke, allergens or infection, causing a rapid increase in mucin within the airways (Thornton et al.

2008; Morrison et al. 2019). This ensures that airway epithelium is primed to respond rapidly to a variety of external challenges.

Membrane-tethered mucins are attached to the airway epithelial microvilli and cilia, providing an osmotic barrier and aiding ciliary motility. Secreted mucin proteins form viscoelastic gels through the retention of water, allowing unwanted particles to be trapped and subsequently cleared. Together, these mucins interact within the ASL to enable mucociliary clearance (Ma et al. 2018). When ionic- or pH flux is altered within the ASL, or mucin regulation is altered, mucociliary transport may become ineffective (Ma et al. 2018).

Mucins appear to have a function in the management of infection by controlling virulence and suppressing disease progression, with enhanced *MUC* gene expression following exposure to bacteria, including *Pseudomonas aeruginosa* and *Staphylococcus aureus* (Wagner et al. 2018). It is speculated that MUC5AC may be an acute-response mucin and MUC5B is produced in response to more chronic airway insults, therefore increasing during infection and inflammation (Thornton et al. 2008). Other authors have reported that, in mouse models, MUC5AC has a protective role against viral infections, providing a decoy for viral receptors (Ehre et al. 2012) and MUC5B is required for mucociliary clearance and prevention of chronic infection by bacterial species (Roy et al. 2014).

Mucin can provide a nutrient source, encouraging the growth of certain commensal bacteria and promoting a stable microbial community. It may protect against epithelial adhesion and cytotoxicity of pathogens by interacting directly with them via the mucin glycan structures. These may provide a 'decoy' from host cell glycans, diverting the pathogen away from host cell glycans and potentially facilitating mucus clearance of the pathogen (Wagner et al. 2018).

1.8 Mucus in respiratory disease

Chronic muco-obstructive diseases, such as CF, chronic obstructive pulmonary disease (COPD), asthma and chronic bronchitis are all associated with upregulation of airway mucus production. This is achieved through a combination of goblet cell hyperplasia, metaplasia and gland hypertrophy, resulting in increased mucus production (Morrison et al. 2019).

The respiratory mucus of diseased airways will also contain inflammatory cells, bacteria and fungi and their resulting debris inflammatory mediators, cells and polymerized DNA from inflammatory cell necrosis (Henke et al. 2007). This is termed 'sputum' when expectorated (Ma et al. 2018). Sputum predominantly contains MUC5AC and MUC5B mucins, with only small amounts of MUC2 (Henke et al. 2007).

1.9 Mucins in cystic fibrosis

The process of exocytosis of the gel-forming mucins is adversely affected in CF. Dysfunctional or absent CFTR leads to impairment of cAMP-activated CFTR-dependent HCO_3^- secretion. This results in reduced $HCO3^-$ secretion, and therefore decreased availability of HCO_3^- to complex with the cations "shielding" mucin anions, thus limiting mucin expansion (Yang et al. 2013). Specifically, the sequestration and removal of Ca^{2+} and H^+ cations during exocytosis does not occur properly, leading to reduced mucin swelling, transportation and release into the lumen (Yang et al. 2013). Calcium-mediated mucin exocytosis occurs, but with insufficient $HCO3^-$ in the extracellular fluid due to CFTR dysfunction, the mucins are unable to expand properly. This results in aggregated mucus on the mucosal surface and an overall reduction in available mucins (Henke et al. 2004; Yang et al. 2013).

Mutations in the CFTR gene reduce sodium concentration in mucus by decreasing epithelial sodium channel (ENaC) activity, leading to reduced ASL volume (Morrison et al. 2019). Taherali et al. (2018) describe the resultant mechanism of airway dehydration, whereby mucus osmotic pressure is greater than periciliary layer (PCL) and therefore draws water from the PCL, eventually leading to its collapse. The cilia are unable to beat effectively, highly concentrated mucus adheres to the epithelium and mucociliary clearance of the highly viscoelastic mucus layer is impaired (Taherali et al. 2018). These mucins demonstrate increased polymeric entanglement with increased solute-mucin interactions occurring as a result of greater numbers of non-salt molecules in the concentrated ASL. This can also affect the viscoelasticity of the mucus gel (Morrison et al. 2019).

Goblet cell hyperplasia and upregulation of *MUC* genes both contribute to the sustained over-production of mucin reported in CF (Rose and Voynow 2006). However, studies are conflicting in their reports of mucin production in CF. Despite

goblet cell hyperplasia in the CF airway, some studies have shown that mucin production is actually reduced in CF patients not experiencing pulmonary exacerbation, with a 70% decrease in MUC5B and 93% decrease in MUC5AC in CF sputum compared to healthy subjects (Henke et al. 2004). However, during exacerbation, increases in mucin production only reach levels comparable to healthy subjects (Henke et al. 2007), showing preservation of the inflammatory and immune mediators' ability to mount an appropriate response to infection. Other studies have shown that the ratio of different mucins is varied in CF patients compared to control samples, with increases in MUC5AC and MUC5B and higher relative abundance of MUC5B in CF patients (Burgel et al. 2007; Thornton et al. 2008).

It is also well-recognised that there are practical difficulties with studying *in vivo* sputum samples and sampling mucins accurately, and therefore one should be cautious when drawing firm conclusions regarding relative quantities of specific mucins within sputum (Thornton et al. 2008; Horsley et al. 2014). There is evidence that mucins undergo proteolytic degradation prior to expectoration in CF patients with chronic infections (Ehre et al. 2014). Therefore, sputum samples may not represent the mucin composition and elasticity seen in the smaller airways (Horsley et al. 2014).

1.9.1 Structural mucin changes in CF

Airway mucins have multiple types of O-glycan chains with significant variation seen between individuals in health and disease (Lamblin et al. 2001). All mucin O-glycans share the N-acetylgalactosamine (GalNAc) residue that is linked to the central apomucin. However, thereafter there is substantial opportunity for variation, based on the subsequent core structure, the O-glycan chains attached forming the backbone, and the terminal sugar (Lamblin et al. 2001). The 'building blocks' of the carbohydrate derivative chains include galactose, N-acetylglucosamine (GlcNAc), mannose, fucose and sialic acid (Brockhausen et al. 2009). The four possible sugars present at the O-glycan chain termination are fucose, galactose, GalNAc and N-acetylneuraminic acid. There may also be the addition of tissular or histo-blood group antigens added: ABH, Secretor and Lewis; or sulphate (Lamblin et al. 2001). Lewis antigen structures are trisaccharide and tetrasaccharide capping groups which may be further sulphated and sialylated (Lewis et al. 2013a). Secreted airway mucins demonstrate increased sialylation and sulphation, with a lower O-glycosylation per protein weight (Davril et al. 1999; Venkatakrishnan et al. 2015). However, membrane-bound mucins have less sialylation and a higher degree of fucosylation on their O-

glycans (Venkatakrishnan et al. 2015). Sialic acid and sulphate both contribute anionic properties to mucins, and fucosylation provides hydrophobic properties (Rose and Voynow 2006). Modifications in these terminal groups cause changes in the viscoelastic properties of the mucus, as the addition of charged residues influences mucin aggregation (Lewis et al. 2013a).

O-glycosylation of mucins can be altered in infection and inflammation. In CF, changes in respiratory mucins are likely predominantly related to environmental factors, rather than directly as a result of the CFTR defect (Thornton et al. 2008). CF mucins exhibit increased levels of expression of sialylated and sulphated-Lewis x determinants during inflammation and infection (Rose and Voynow 2006; Henke et al. 2007). These changes can be induced by increased inflammatory mediators, such as TNFα, as well as components of bacterial cell walls, leading to an increase in enzymatic activity and resultant alteration in mucin glycosylation (Delmotte et al. 2002; Lewis et al. 2013a). Interactions between both sialyl- and sulphated Lewis x and bacteria and viruses have been previously demonstrated (Scharfman et al. 2000).

The sulphated Lewis x antigen is the main sulphated sugar in CF respiratory mucin (Lewis et al. 2013a). This is overexpressed in the mucins of severely infected CF patients and those with evidence of inflammation, as a result of cytokine secretion (Davril et al. 1999). When compared to other respiratory diseases such as chronic bronchitis, sialyl Lewis x appears to be more prevalent in CF sputum, providing a clear correlation with specific diseases (Davril et al. 1999). It is likely that the CFTR mutation also contributes to altered terminal mucin glycosylation (Rose and Voynow 2006). Identification of sialylated and fucosylated structures in mucins in salivary secretions, where salivary glands would not typically be infected in CF, support this hypothesis (Shori et al. 2001). Similarly, the hyper-sulphation seen in a xenograft model of CF airway mucosa suggests that such alterations are due to the primary disease defect (Zhang et al. 1995).

Mucin oligosaccharides are highly heterogenous, and such diversity is likely an evolved response to the varied pathogens and other ligands entering the airway milieu (Thornton et al. 2008). Though mucus in the healthy lung has been shown to have beneficial mucin-microbe interactions, CF mucus does not seem to provide the same protection, particularly to microbes such as *P. aeruginosa* (Wagner et al. 2018). This may be related to altered mucin gene expression and production (Henke et al. 2004; Henke et al. 2007), degradation of mucin (Flynn et al. 2016), altered

glycosylation patterns affecting the signalling potential of mucin (Schulz et al. 2007) and dehydration of the mucus gel layer (Lai et al. 2009) (Wagner et al. 2018).

1.9.2 DNA in CF sputum

In health, large polymeric extracellular DNA (eDNA), predominantly derived from epithelial cell breakdown, represents about 0.02% of the total mucus mass (Lai et al. 2009). In CF, an accumulation of eDNA increases the viscosity and elasticity of mucus (Lethem et al. 1990). Increased neutrophil lysis during infective exacerbations further increases eDNA content by up to 0.5-1.5% of mucus by weight resulting in significantly altered mucus viscoelasticity (Mrsny et al. 1996; Lai et al. 2009). DNA has a significant influence on CF sputum rheology, with evidence supporting the use of DNase as a form of airway clearance (Zahm et al. 1995; Horsley et al. 2014). When analysing CF sputum, it is important to consider the role of eDNA and any contribution this may be making.

1.10 Infection in Cystic Fibrosis

Cystic fibrosis is characterised by infection of the airway early in childhood. Infection includes predominantly bacterial and fungal Infections, with eventual colonisation of the airways by dominant CF pathogens. Environmental factors within the CF lung, including hyper-secreted viscous mucus, plus the increased DNA and actin resulting from neutrophil necrosis, create a low oxygen environment in which pathogens can survive and establish drug-resistant biofilms (Moreau-Marquis et al. 2008). Mucociliary clearance of bacteria and fungi is maintained by the activity of the CFTR anion channel within the healthy lung. However, in CF, the reduction in Cl^- and $HCO3^-$ decreases the periciliary surface liquid volume and thereby effective clearance of such opportunistic bacteria (Stanton 2017). The increased quantity and thickness of the mucus produced in CF patients leads to ciliastasis and an inability to effectively clear bacteria from the airways (Matsui et al. 1998)

Children with CF will typically receive antibiotic courses following positive airway cultures and when recurrent infection or colonisation is suspected, long-term antibiotics are often commenced. Despite aggressive therapy, the effects of chronic infection will eventually lead to permanent lung damage and eventually respiratory failure and death.   This work focuses on *P. aeruginosa* infection. As one of the key CF pathogens, it represents a clear focus for CF microbiology research and development of targeted therapeutics.

1.11 *Pseudomonas aeruginosa* infection

Many children isolate *P. aeruginosa* on broncho-alveolar lavage samples within the first few years of life (Burns et al. 2001; Rosenfeld et al. 2001). Kidd et al. (2015) demonstrated successful initial eradication in 90% of cases, but 44% of children reacquired *P. aeruginosa* before the age of 5 years (Kidd et al. 2015).

P. aeruginosa secretes a variety of virulence factors which adversely affect mucociliary clearance. Pyocyanin reduces CFTR Cl^- secretion, ciliary beating and mucociliary transport; rhamnolipids promote ciliastasis; and bacterial alginate increases mucus production making immune recognition and bacterial clearance much more difficult (Ballok and O'Toole 2013).

Infections can initially be managed with acute courses of antibiotics, but most *P. aeruginosa* infections will eventually persist and become chronic. There are a number of clinical definitions for chronic pseudomonal infection, with the most recent Cochrane review (2017) requiring "the presence of *P. aeruginosa* in monthly specimens for six successive months or the development of precipitating antibodies to *P. aeruginosa* or both" (Hewer and Smyth 2017). Within the UK, 5.4% of the paediatric CF population have chronic infection, with an increase up to 44.5% of the overall adult population (Charman et al. 2018).

Prior to chronic infection, there is often intermittent isolation from respiratory tract specimens, which may represent transient colonies or simply the limitations of sampling from young patients with CF (Hewer and Smyth 2017). As infection becomes established, the quantity and type of *P. aeruginosa* alters, with an increased density of colonies present and the change from non-mucoid to mucoid phenotypes seen (Rosenfeld et al. 2001; Hewer and Smyth 2017). At this stage, pathogen elimination is considered impossible, either by host immune responses or established antimicrobial treatments (Lund-Palau et al. 2016). *P. aeruginosa* is a highly versatile opportunistic bacterium and the most prevalent pathogen in adults with CF. It is a significant cause of morbidity and mortality in CF patients, associated with accelerated disease progression (Lund-Palau et al. 2016). The resultant chronic bacterial infection, and associated airway inflammation, leads to death from respiratory failure in 80% of all patients (Bhagirath et al. 2016). Evidence suggests initial colonisation is by naturally abundant environmental strains, though patient-to-patient transmission can occur, depending on the degree and duration of close patient contact (Burns et al. 2001; Folkesson et al. 2012). CF patient susceptibility to *P. aeruginosa* is not fully understood and the proposed mechanisms relating to interactions with the CFTR protein remain unproven. It is likely that these patients have an increased risk of initial acquisition or enable chronic persistence within the lung environment, or perhaps both (Lund-Palau et al. 2016). Evidence does suggest that the hypoxic and highly viscous environment within CF mucus facilitates mucoidy phenotypes and biofilm formation, and impairs soluble host defence factors (Lund-Palau et al. 2016).

1.12 Adaptation of *P. aeruginosa* to the CF lung environment

Chronic infection occurs with expansion of the *P. aeruginosa* population within the CF airways. There is increasing diversity within the population and strains with differing traits from the primary coloniser occur, allowing them to develop and thrive (D'Argenio et al. 2007). Such traits seem to consistently occur within different CF patients, suggesting a conserved pattern of evolution is enabling adaptation (D'Argenio et al. 2007). However, studies have shown significant versatility in the pathoadaptive genes mutated within the CF lung, with resulting extensive genotypic and phenotypic diversity in these bacterial populations (Winstanley et al. 2016a; Klockgether and Tümmler 2017).

Populations are seen to differ across regions within the respiratory tract, related to variability in environmental factors such as oxygen and nutrient availability and antibiotic concentrations. Genetic compartmentalisation describes the process by which populations become geographically isolated by differing selective pressure and evolve independently (Jorth et al. 2015). This regional isolation of *P. aeruginosa* populations leads to divergent evolution of separate clonal lineages and maintenance of phenotypic diversity within different areas of the lung, with minimal mixing between geographical populations (Markussen et al. 2014; Jorth et al. 2015; Winstanley et al. 2016a).

Numerous mechanisms, including transcription factors, quorum-sensing (QS) networks, two-component systems and non-coding RNAs, enable *P. aeruginosa* to modulate virulence (Klockgether and Tümmler 2017). Evidence suggests the second messenger cyclic di-GMP has a key role in controlling complex signalling pathways and coordinating the "lifestyle transition" from motile to sessile and back again (Ha and O'Toole 2015; Valentini and Filloux 2016). Planktonic (or motile) cells change to sessile cells on attaching to a surface under certain conditions and then undergo significant physiological, metabolic and phenotypic changes (Valentini and Filloux 2016). Whereas planktonic cells are associated with acute infections, *P. aeruginosa* chronic infections occur following biofilm formation, which enables antimicrobial resistance and protection from the host immune response (Valentini and Filloux 2016).

1.12.1 Genetic adaptation

Phenotypic variability is well documented across *P. aeruginosa* strains from varying clinical and environmental sources, and not just within the CF population (Cullen et al. 2015). In response to environmental demands, *P. aeruginosa* can adapt its phenotype as necessary through reversible regulation of gene expression. In chronic infection, this reversibility is lost and mutants develop which are genetically and phenotypically different from the original strain (Bragonzi et al. 2009). Studies have shown a variety of mechanisms utilised by *P. aeruginosa* in chronic CF infections, including "loss-of-function mutations, acquisition or loss of genomic islets/islands, genome rearrangements, recombination, or point mutations" (Smith et al. 2006a; Bragonzi et al. 2009). Most pathoadaptive gene mutations are associated with loss of gene function, as demonstrated by the high frequency of frameshift mutations (Marvig et al. 2015). Horizontal gene transfer provides an additional mechanism whereby adaptation can occur (Cullen and McClean 2015).

Whole-genome sequencing technologies have enabled scientists to define the genetic basis for adaptations seen within CF lung infections (Winstanley et al. 2016a). Functional categories of mutated genes most commonly affected encode for virulence factors and regulators, small-molecule transportation (multidrug-efflux-pump genes) and antibiotic resistance (Smith et al. 2006a). However alterations in genes encoding iron acquisition, cell reproduction, quorum sensing, fatty-acid metabolism, DNA mismatch repair and anaerobic metabolism have all been demonstrated (Smith et al. 2006a). These have been termed 'pathoadaptive' traits. Mutations in multidrug efflux pumps are most commonly seen, particularly in the gene *mexZ*, which is associated with increased resistance to aminoglycosides (Smith et al. 2006a).

Infections most commonly become 'clonal', whereby a clone of cells survive and thrive through accumulation of genetic variants, enabling clonal expansion and long-term persistence/survival (Smith et al. 2006a; Bragonzi et al. 2009). The CF lung habitat appears to select for rare clones that are able to survive to become dominant members, despite regular antimicrobial exposure and the host immune system response (Klockgether and Tümmler 2017). Persistence of bacterial strains from a common lineage enables sub-populations to develop, with inevitable population diversity thereafter (Cullen and McClean 2015).

Hypermutable strains have defects in the DNA repair system or proof reading systems and therefore show an increased spontaneous mutation rate (Cullen and McClean

2015). Host environment and environmental stressors, such as antibiotics, will select for these mutator strains, as the increased mutation rates allow greater adaptation and survival in challenging conditions and acceleration of the lung adaptation process (Hoboth et al. 2009; Cullen et al. 2015). These strains are more commonly found in chronic *P. aeruginosa* infection, supporting the need for genetic and phenotypic diversity to enable long-term survival within the CF lung (Cullen and McClean 2015).

1.12.2 Virulence Factors

Although virulence factors are required for acute infection, alterations in, and downregulation of, virulence factors confer benefit to bacterial populations in chronic infection. The host's immune system will typically select those cells with recognised virulence factors, enabling survival of those with mutations. In turn, lower virulence enables populations to thrive and establish chronic infections by reducing host detection (Cullen et al. 2015). Bragonzi et al. (2009) tested clinical isolates compared to environmental and laboratory strains using murine models. They demonstrated significantly reduced virulence in CF strains compared to the laboratory strain PA14 and environmental strains, but similar to PAO1. This supports the use of PAO1 as an appropriate laboratory strain and its use in this project's laboratory model. Authors showed that strains from chronic infections caused reduced mortality in mice compared to those from early infections, supporting the loss of virulence factors in established CF infections (Bragonzi et al. 2009). However, they argue that the microevolution of these clones results in altered rather than reduced virulence following infection, leading to decreased abilities to cause acute infection but optimising persistence in the host (Bragonzi et al. 2009). Cullen et al. (2015) studied acute virulence of strains within their reference panel in the *Galleria mellonella* model. They showed significantly less virulence in those of CF origin, with a reduction in virulence seen in sequential CF strains taken over time of chronic infection (Cullen et al. 2015). This provides further evidence to support the adaptation of *P. aeruginosa* to the CF lung environment.

Pyocyanin is a major virulence factor in *P. aeruginosa* and production is thought to be downregulated with development of chronic infection (Cullen et al. 2015). This compound is a blue, redox-active phenazine, giving the typical blue/green appearance to colonies on agar plates or within solution in the laboratory, and discolouration to patient sputum samples. Pyocyanin has multiple effects within the CF lung. These include reducing cilia beat frequency and thus mucus clearance,

20

direct impact on CFTR chloride channel gates through pyocyanin-mediated ATP depletion and reduced defence against oxidative stress, and multiple actions on the host immune response (Winstanley and Fothergill 2009). Cullen et al. (2015) found variable production of pyocyanin across their reference panel, though CF strains in chronic infection typically produced lower levels than early strains.

Mowat et al. (2010) studied a widely disseminated strain of *P. aeruginosa*, the Liverpool Epidemic Strain, in sputum from ten patients with chronic CF lung infections. After characterising 15 traits, such as colony morphology, auxotrophy and pyocyanin overproduction, and defining each combination as a different haplotype, they identified 398 haplotypes from a total of 1720 isolates (Mowat et al. 2011). They demonstrated extensive diversity particularly within patients, with more composition variation seen over time than during exacerbation periods (Mowat et al. 2011).

Such diversity was also seen by Bragonzi et al. (2009) who demonstrated changes in genetic mutations between early, intermediate and late isolates when following six patients with CF over a period of 16.3 years (Bragonzi et al. 2009). Interestingly, they reported a relatively high proportion of isolates with pyocyanin overproduction during exacerbations compared to stable samples. However, this finding was not consistent throughout all samples. The authors hypothesise this could be due to either the contribution of pyocyanin to pulmonary symptoms during an exacerbation or that the CF lung environment at such times enables pyocyanin-overproducing isolates.

1.12.3 Quorum Sensing

P. aeruginosa has extremely complex quorum-sensing (QS) circuits, where "multicomponent communication and regulatory network interactions operate" (Winstanley and Fothergill 2009). QS systems are directly involved in the production and regulation of virulence factors, particularly during acute infection, but mutations in QS-related genes are commonly recognised in chronic infection (Winstanley and Fothergill 2009).

Generally, communication occurs by secretion of signalling molecules called homoserine lactones, which are secreted and prompt responses by bacteria after reaching a critical concentration within the external environment (Winstanley and Fothergill 2009). *P. aeruginosa* has a complex QS network, comprised of two independent LuxIR-type QS systems, LasIR and RhlIR, which interact with a

quinolone signal, multiple regulators and sigma factors (Winstanley and Fothergill 2009). Important QS-regulated exoproducts include pyocyanin, elastase, LasA, alkaline protease and rhamnolipids. These contribute to local tissue damage, degradation and destruction of lung structure, and inhibition of components of the immune system (Winstanley and Fothergill 2009).

The most notable virulence-related gene mutation is in *lasR*, which is a key transcriptional regulator of QS (Smith et al. 2006a). Evidence suggests that there is a strong selection pressure for loss of *lasR* function, given the high prevalence of *P. aeruginosa lasR* mutants within the CF population (Smith et al. 2006a; D'Argenio et al. 2007). Mutation is typically seen in chronic infection and leads to defective QS and loss of virulence. This suggests that the QS network may have a reducing impact on pathogenicity over time and may in fact have a negative impact on the long-term fitness. Therefore, QS mutations likely confer benefit to survival and persistence (Winstanley and Fothergill 2009).

1.12.4 Adaptations in surface structures

1.12.4.1 *Lipopolysaccharides*

Lipopolysaccharides (LPS) are major virulence factors in *P. aeruginosa* through induction of host platelet aggregation, pyrogenicity and induction of cytokines (Cullen and McClean 2015). These large molecules are embedded in the outer cell leaflet and consist of lipid A, core oligosaccharide and the highly variable long-chain O-polysaccharide (Kocincova and Lam 2011). Studies have shown loss of the O-antigen portion during CF chronic infections (Smith et al. 2006a; Winstanley and Fothergill 2009). Cullen et al. (2015) showed a variety of structural changes with regards to the A- and B- bands of the O-antigen across their reference panel, with no consistent pattern across CF vs. non-CF strains. Changes in LPS composition is associated with increased innate antibiotic resistance, and alteration of structure in chronic infection is linked to reduced host stimulation and increased survival for those cells (Cullen and McClean 2015).

1.12.4.2 *Outer membrane proteins*

Outer membrane proteins have an essential role in the within-host adaptative mechanisms displayed by Gram-negative bacteria. They are regulated by environmental factors and contribute to survival and virulence, for example by siderophore production to enable iron-uptake, providing pathways for haemoglobin

utilisation and physical protection from host complement-mediated killing (Cullen and McClean 2015).

1.12.5 Colony morphology

Airway-specific adaptations due to environmental pressures enable alteration in phenotype and genotype and will often result in changes to colony morphology and morphotype (Kirisits et al. 2005). Such changes will be visible on routine sputum culture (Clark et al. 2015). Alteration in morphology is often seen during colonisation and may be facilitated by hypermutable strains of *P. aeruginosa* (Oliver et al. 2000). Isolates can have varying morphotypes, including alterations in colony size, colour and texture, as well as visible autolysis and autoaggregative appearances (Hogardt and Heesemann 2010).

1.12.5.1 *Colony size*

Small colony variants (SCVs) are commonly seen in chronic lung infections. These variants are also described as "rough small colony variants (RSCVs), wrinkled variants, autoaggregating cells and rugose colonies" within the literature (Kirisits et al. 2005). For clarity, due to the varied phenotypic and genotypic traits of SCVs described in the literature, they have been defined as a pin point colony formed within 72 h (<1 mm in diameter) in this study (Johns et al. 2015). Slow growing colonies may be missed by standard clinical lab incubation periods (Workentine et al. 2013).

SCVs are often selected following prolonged antibiotic exposure and show a slower growth rate, smaller colony morphology, hyper-adherence, hyperpiliation, reduced motility and autoaggregative growth behaviour (Kirisits et al. 2005; Gellatly and Hancock 2013; Cullen and McClean 2015). These features can enhance biofilm formation and persistence within the CF lung (Cullen and McClean 2015). However, perhaps unsurprisingly, SCV isolates also show significant diversity in phenotype testing, and morphology cannot be used to fully predict phenotypic behaviours (von Götz et al. 2004; Kirisits et al. 2005). Literature suggests that biofilm growth results in two prevalent colony size morphologies; the wild-type (PAO1) colony morphology and the SCV colony morphology, with the former being relatively medium/large in size (Kirisits et al. 2005)

1.12.5.2 *Colony surface texture*

Poltak and Cooper (2011) demonstrated colony surface texture diversification following sub-culturing on beads. They described a predictable emergence of common heritable colony morphologies across the replicate populations. These morphotypes were described as smooth or studded (S) (at 150 generations) ruffled spreader (R) (at 300 generations) and wrinkly (W) (between 300 and 450 generations) (Poltak and Cooper 2011a). Though prolonged transfers led to a decrease in the ruffled spreader and wrinkly phenotypes, smooth phenotypes are hypothesised to persist and are likely to be essential for the overall stability of the biofilm structure (Poltak and Cooper 2011a). Boles et al (2004) hypothesised that surface texture may be related to biofilm function and, therefore, studied small smooth (termed 'mini') versus rough (termed 'wrinkly') colonies with the wild type phenotype as control. The wrinkly variant showed an accelerated biofilm formation with greater numbers of bacteria within its biofilms compared to the mini and wild-type variants (Boles et al. 2004). In addition, the wrinkly variant also demonstrated increased antibiotic and hypochlorite resistance (Boles et al. 2004).

1.12.5.3 *Colony pigmentation and opacity*

Mayer-Hamblett et al. (2014) considered colony sheen, pigmentation and autolysis as part of their phenotypic characterisation of *in vitro* samples. In cultures obtained from newly infected patients, 67% of samples showed a tan colony colour. Overall, there were low numbers of green, clear and yellow coloured colonies across all patient samples, with no obvious differences between new onset, intermittent and chronic infection groups (Mayer-Hamblett et al. 2014). Similarly, there was no significant differences in colony lysis or sheen between patient groups.

A distinct colony appearance whereby a metallic iridescent sheen is seen on the colony surface is strongly associated with *lasR* mutation (D'Argenio et al. 2007; Mayer-Hamblett et al. 2014). There is also cell autolysis, resulting in colony flattening (Hoffman et al. 2009). Inactivation of the transcriptional regulator LasR provides growth advantage within the amino acid-rich CF airways and increased β-lactamase activity, providing a degree of antibiotic resistance (D'Argenio et al. 2007; Hoffman et al. 2009).

1.12.5.4 *Colony margins*

Clark et al. (2015) described the 'halo' seen in mucoid colonies as a transparent extracellular polymeric substance (EPS) layer present around the central opaque colony. Non-mucoid colonies also had an equivalent 'halo', which they described as a 'non-concentric outer ring'. No clear patterns were identified using colony margin appearances alone (Clark et al. 2015). Sousa et al. (2013) describe the colony form as circular or irregular and the margin to be entire/irregular. This study highlighted the importance of conditions such as growth media, incubation period and number of plated colonies when assessing morphological features (Sousa et al. 2013).

1.12.6 Mucoidy

The development of the mucoid phenotype, secondary to the overexpression of alginate, is typically considered a hallmark of chronic CF infection. Mucoid isolates are rarely seen in non-CF environments and therefore are most likely related to specific CF-selective pressure (Folkesson et al. 2012; Sousa and Pereira 2014). Alginate is a major exopolysaccharide in *P. aeruginosa* biofilms. This polymer of mannuronic and guluronic acid is typically produced under microaerophilic or anaerobic conditions, as in the CF lungs (Klockgether and Tümmler 2017). Alginate biosynthesis is promoted by the protein Alg44 following binding of two c-di-GMP molecules (Whitney et al. 2015; Klockgether and Tümmler 2017). It is affected by growth conditions and mucoid strains may revert to non-mucoid phenotype on sub-culturing (Cullen et al. 2015). Alginate is secreted alongside two other exopolysaccharides, Pel and Psl, which have an important role in providing stability to the biofilm matrix, and protecting against antibiotics and host responses (Winstanley et al. 2016a). Mayer-Hamblett et al. (2014) showed the presence of mucoidy colonies had the strongest association with pulmonary exacerbation over a 1-year follow-up, suggesting this phenotypic feature may be useful as a clinical predictor.

1.12.7 Motility

P. aeruginosa is able to move by swimming, swarming and twitching, with motility determined predominantly by the environment (Overhage et al. 2008). Swimming is seen in aqueous environments, swarming on semi-solid, viscous media and twitching on solid surfaces (Overhage et al. 2008). Acute infection by *P. aeruginosa* requires pilin-mediated adherence, enabling initial cell injury and colonisation (Hogardt and Heesemann 2010). Swimming motility is enabled by the rotation of a single polar,

monotrichous flagellum (Toutain et al. 2005). A number of structures make up the flagellum structure, including the cap, filament, hook and basal body (Ha et al. 2014). The basal body contains stator and motor parts and is required to keep the flagellum in place, with the hook and filament providing both clockwise and anti-clockwise movements to enable propulsion (Ha et al. 2014). Movement is fuelled by conductance of cations (Ha et al. 2014). Functional loss of flagellum-mediated swimming motility can enable bacteria to evade the host response by providing resistance to phagocytosis (Amiel et al. 2010).

Swarming requires flagella and type IV pili for movement, and rhamnolipids to enable motion through the semi-solid media (Overhage et al. 2008). In addition to allowing bacterial movement, swarming may also contribute to early biofilm formation, and antibiotic resistance when compared to planktonic bacteria (Overhage et al. 2008). Upregulation of gene expression in swarming cells has been demonstrated, encoding for production of virulence factors such as pyoverdine, multi-drug efflux pumps which aid antibiotic resistance, and the redox-active compound phenazine, which contributes to the pulmonary tissue damage associated with *P. aeruginosa* (Overhage et al. 2008). Swarming cells exhibit upregulation of genes associated with nitrite reduction, providing fitness advantages within the nitrite-rich CF lung (Soberón-Chávez et al. 2005; Overhage et al. 2008)

Twitching motility is enabled by type 4 pili, which are the most important adhesins of *P. aeruginosa* and pull the bacterial cell along solid surfaces (Gellatly and Hancock 2013). Loss of twitching motility has been associated with infection stage and increased risk of pulmonary exacerbation when studying CF patient cultures (Mayer-Hamblett et al. 2014). The authors demonstrated persistence of both reduced twitching and swimming over time and advocated the use of defective motility as a good screening tool and potential biomarker (Mayer-Hamblett et al. 2014). Exposure to antimicrobials, including those at sub-MIC concentrations, can directly affect the surface assembly of pili and therefore inhibit the flagella-independent motility seen in *P. aeruginosa* (Wozniak and Keyser 2004). *P. aeruginosa* can switch between motile and sessile phenotypes via a variety of mechanisms. Elevation of c-di-GMP levels leads to repression of flagellum-driven swarming motility in chronic infections, with reduced levels (as seen in acute infections) upregulating the transcription of flagellar genes and promoting the motile lifestyle (Klockgether and Tümmler 2017). *P. aeruginosa* isolates from chronic CF infection lack swimming and twitching motility due to loss of flagellum and non-piliation respectively (Jain et al. 2004; Hogardt and

Heesemann 2010). Downregulation of flagella motility enables *P. aeruginosa* to avoid inducing neutrophilic extracellular traps release and subsequent phagocytosis (Floyd et al. 2016), and therefore promotes survival. Previous studies have suggested loss of motility in chronic infection, and hypothesised that non-motility may help persistence through avoidance of host immune defence systems, such as alveolar macrophages and polymorphonuclear leucocytes (Mahenthiralingam et al. 1994; Bragonzi et al. 2009; Hogardt and Heesemann 2010).

1.12.8 Biofilm Formation

A biofilm is a "structured community of bacterial cells enclosed in a self-produced polymeric matrix and adherent to inert or living surfaces" (Ha and O'Toole 2015). Biofilm formation typically occurs in chronic infections, whereby bacteria switch from the planktonic growth state to a sessile lifestyle. This is promoted by elevated levels of c-di-GMP through multiple influencers, causing upregulation of genes involved in biofilm formation, including the exopolysaccharides Psl and Pel and the adhesin CdrAB (Klockgether and Tümmler 2017). These exopolysaccharides, as well as alginate, are critical in the formation of a mature biofilm within the CF habitat (Ha and O'Toole 2015).

Biofilm communities benefit from increased horizontal gene transfer, protection from environmental stressors, such as antimicrobials, and promotion of useful metabolic interactions (Kirisits et al. 2005). However, potential disadvantages also exist, including the development of dynamic gradients of nutrient delivery and toxic metabolic product removal as a result of high cell densities (Kirisits et al. 2005). *P. aeruginosa* predominantly grows in biofilms, achieving up to 1000-fold higher tolerance to antimicrobial agents, when compared to planktonic bacteria (Häussler 2010). Biofilms require a multitude of coordinated pathways and cellular factors for development, including flagellar and twitching motility and exopolysaccharide production (Ha and O'Toole 2015).

Analysis of panel strains including CF and non-CF laboratory and environmental *P. aeruginosa* strains, Cullen et al. (2015) showed that all strains formed biofilms to some degree. However, there was diversity in terms of density of biofilm produced across the strains and time played an important role in production: strains producing rapid biofilm formation typically detached over time, whereas slower biofilm formers demonstrated increasing biofilm biomass over the duration of the assay (Cullen et al.

2015). Clark et al. (2015) used colony morphology of 235 isolates, taken from a single patient with CF over one year, to predict phenotypes such as antimicrobial susceptibilities. They found 15 distinct colony morphotypes across the sample and interestingly individual morphotypes were not useful predictors for antimicrobial susceptibility (Clark et al. 2015). The authors recommended caution with using colony morphology alone as a screening tool for guiding antibiotic susceptibility testing in CF patients (Clark et al. 2015).

1.13 Analysing biofilms

Confocal laser scanning microscopy (CLSM) and scanning electron microscopy (SEM) are imaging techniques available for biofilm analysis. In CLSM, a laser light beam is passed through the head of the microscope, down and out of the objective, and onto the microscope slide. Specimens must be prepared with fluorescent dye-staining and this enables a fluorescent image to be captured by the objective when exposed to the laser light beam. The resultant image is transferred from a module sitting in the top of the head of the microscope to a computer screen (Rowland and Nickless 2000). Software can achieve 3-dimensional images by stacking serial sections ("Z" stack), thus providing a full thickness image of the specimen (Sharp 2014). Data can be used for quantitative analysis of biofilm thickness, area and volume (Rowland and Nickless 2000). The use of LIVE/DEAD staining can enhance images by quantifying the numbers of nonviable bacteria, reflecting cell death, within the biofilm (Wimpenny et al. 2000; Khan et al. 2012c).

Images in the SEM are created by scanning a focused high-energy beam of electrons across the specimen. Interaction between the beam electrons and sample atoms produces signals that demonstrate the specimen's surface topography (Schatten 2012). The use of SEM for biofilm imaging of *P. aeruginosa* has been extensively described in the literature, with a variety of dehydration and fixation described (Jesaitis et al. 2003; Khan et al. 2012c). SEM images can provide detail regarding biofilm surface structure, including branching and porosity (Pritchard et al. 2016b), plus individual colony's structure and size (Pritchard et al. 2017a). For both imaging techniques, sample preparation is key to maintaining the structural integrity of the specimen and enabling reliable information to be obtained. Poor preparation risks causing artefacts and incorrect conclusions to be drawn. (Schatten 2012).

1.14 Experimental evolution

Experimental evolution studies the molecular basis of adaptation, enabling a greater understanding of evolutionary change within a selected population as a consequence of conditions imposed (Lang and Desai 2014; Steenackers et al. 2016). Using laboratory models, it is possible to establish multiple replicate populations with a common 'starting point' and identical conditions, and then follow the development of mutations across these populations.

By studying the parallelism in phenotypic and genotypic responses to selection, scientists can gain a much greater understanding of adaptive evolution within predetermined conditions, such as those reflecting the CF lung environment (Huse et al. 2010; Lang and Desai 2014). Such models allow for collection of samples at agreed timepoints for phenotypic and genotypic analysis and for these results to be compared to other timepoints within the evolutionary model (Lang and Desai 2014). Frozen cells taken from different time points remain viable and available for analysis at a later date (Lang and Desai 2014; Lenski 2017). Biofilm models are well suited to study the CF lung ecosystem. Conditions can be set up to enable biofilm formation, exposure to a number of antibiotics/environmental stressors, and in turn facilitate slow bacterial growth and horizontal gene transfer between multi-drug resistant (MDR) organisms (Høiby et al. 2010).

1.15 The human microbiome

The term 'microbiota' defines the microbes in or on a host, and includes bacteria, archaea, viruses, protists and fungi (Bordenstein and Theis 2015). The microbiome encompasses the genetic content of this microbiota (Bordenstein and Theis 2015). Though well defined, these terms are often used interchangeably. The human microbiome consists of the "ecological community of commensal, symbiotic, and pathogenic microorganisms that share our body" (Thomas et al. 2017). The combination of the host organism and its microbiota is known as the 'holobiont' with its resultant genome being the 'hologenome' (Bordenstein and Theis 2015). The impact of disruption of the carefully balanced holobiont can lead to alterations in hologenome expression and resultant disease (Thomas et al. 2017).

Though microbiomes include species across all major kingdoms, bacteria represent the most extensively studied phylogenetic group (Thomas et al. 2017). Prokaryotes are subdivided into two domains: Bacteria and Archaea. Many of the vast number of known genera and species are detailed in publicly-available databases, such as CORE and the Human Microbiome Project (Micah et al. 2007; Griffen et al. 2011). Databases were initially developed based on phenotypic classification of reference strains. However, with the advent of culture-independent methods, many have 16S ribosomal RNA (rRNA) gene sequences available for additional phylogenetic classification (Thomas et al. 2017).

Within the human, host cells and the microbiota present almost equal numbers within an individual (Sender et al. 2016). The gut, skin and oral cavity environments represent regions of greater microbial concentration. Humans and microbial communities have co-evolved over millions of years, enabling an effective symbiotic relationship. The immune system tolerates the microbiota well and humans benefit from the microbial contribution to host metabolism. Specifically, the microbiome enriches the metabolism of amino acids, glycans and xenobiotics, as well as synthesis of vitamins and other nutrients (Micah et al. 2007; Thomas et al. 2017). The microbiome is constantly evolving and will change in response to age, culture, environment, diet and health status (Thomas et al. 2017).

The adaptive ecosystems created by the microbiota inhabiting the human body are finely balanced and generally able to cope with constantly changing human physiology (Lloyd-Price et al. 2016). However, disturbance of this carefully-balanced

ecology can be associated with human diseases, such as type 1 and 2 diabetes, inflammatory bowel disease, asthma and cancer (Lloyd-Price et al. 2016). Identifying features to distinguish between 'healthy' and 'unhealthy' microbiomes may help with disease prevention, diagnostics and management. Though clear patterns of microbial colonisation have been associated with certain diseases when compared to healthy controls, determining features of a healthy microbiome has been somewhat more difficult (Bäckhed et al. 2012). Certainly, sufficient evidence shows that defining a healthy microbiome based on a 'core' set of microbial taxa or 'healthy microbes' is not possible (Bäckhed et al. 2012; Shafquat et al. 2014; Lloyd-Price et al. 2016).

Features of a healthy microbiome may include prevalent organisms or molecular pathways and certain levels of ecological diversity or stability (Lloyd-Price et al. 2016). Stability refers to the microbial community's resistance (ability to resist change following ecologic stress) and resilience (ability to return to equilibrium state following a stressor) (Bäckhed et al. 2012). Though stability is generally maintained in the absence of perturbations, extrinsic stressors such as diet and antibiotic exposure can disrupt the microbiota, particularly within the gut (Bäckhed et al. 2012; Consortium 2012). Lack of diversity or evenness appears to be linked with reduced ability to tolerate perturbation, resulting in greater susceptibility to developing disease (Virgin and Todd 2011; Bäckhed et al. 2012).

It is important to consider the structure and function of the microbial community when assessing microbiome health. Structure describes the number and types of microbes present, and function refers to the metabolic activity and resultant end-products from microbial activity (Bäckhed et al. 2012). Evidence suggests that despite often varying compositions between subjects, microbial function is fairly consistent in similar ecosystems (Consortium 2012). Community composition is most comparable within habitats, for example similarities between oral communities are greater between individuals than multiple habitats within the same subject (Lloyd-Price et al. 2016). However, intra-individual variability over time is less than inter-individual changes within the same habitats (Lloyd-Price et al. 2016). Microbial function can be fully interrogated by applying metatranscriptomics and metaproteomics, as the microbiome is also transcriptionally regulated (Bäckhed et al. 2012).

Microbiome exposure to antibiotics can be both helpful and deleterious. Though typically the impact of antibiotics is short-lived, there is evidence that some impacts can be longer-lasting (Blaser and Falkow 2009). Antibiotics cause ecological

disturbances to the microbiota, namely as a result of effects on susceptible bacteria and the potential persistence of resistant strains (Bäckhed et al. 2012). It is proposed that these imbalances can contribute to the individual's risk of infection and disease, and may explain the rising trends of obesity, allergies and asthma (Blaser and Falkow 2009). Antibiotics, however, can also help to manage dysbiosis and improve clinical outcome, such as in the treatment of *Clostridium difficile*-associated diarrhoea (Bäckhed et al. 2012).

Recent advances in next-generation sequencing have significantly increased our understanding of the human microbiome and the possible functions and contributions of our resident microbiota (Man et al. 2017). Although the majority of studies have focused on gut microbiome and microbiota-derived metabolites, interest has expanded to other surfaces of the human body, including the respiratory tract mucosa (Man et al. 2017).

1.16 Experimental tools for studying the human microbiome

Culture-independent analytical techniques have revolutionised our understanding of the human microbiome and its interaction with the host. They have confirmed that cultured microorganisms represent only a small proportion of the true microbial communities present on earth, and previous estimates of microbial diversity, species richness and species abundance have been significantly underestimated (v. Wintzingerode et al. 1997).

Ribosomal genes are present in all organisms and are required for protein synthesis. They are ideal for the study of microbial taxonomy, as they have been present since the beginning of evolution, demonstrate variability between different species, and enable consistent taxonomic differentiation (Peix et al. 2009). However, though some authors suggest differentiation can be as precise as genus and species (Peix et al. 2009), others argue that species-level resolution may not be feasible (Kuczynski et al. 2012). An example of taxonomic classification of the bacterial species *P. aeruginosa* is shown below (**Table 1.1**) (Todar 2006).

Table 1.1. Taxonomic classification of *Pseudomonas aeruginosa*

Taxonomic Rank	Classification
Kingdom	Bacteria
Phylum	Proteobacteria
Class	Gamma Proteobacteria
Order	Pseudomonadales
Family	Pseudomonadaceae
Genus	*Pseudomonas*
Species	*Pseudomonas aeruginosa*

The ribosome is 70S and comprises of a large 50S subunit and small 30S subunit. The 'S' stands for the Svedberg unit. This is a non-SI unit measuring sedimentation rate and measures "particle size based on its rate of travel in a tube subjected to high *g* force" (Marchesi and Ravel 2015). Ribosomes typically contain 50-60% RNA in the form of three ribosomal RNA (rRNA) molecules: 5S, 16S and 23S rRNA (Noller 1984). 16S rRNA is part of the structure of the small subunit of bacterial ribosomes, containing approximately 1500 nucleotides, bound to 21 proteins (Marchesi and Ravel 2015). The 16s rRNA gene is comprised of nine variable regions (V1 – V9) interspersed throughout the highly conserved 16S sequence (Johnson et al. 2019). The 16S rRNA gene was selected as the most useful for the study of prokaryote classification, as the 5S rRNA gene is too small and the 23S rRNA gene too conserved to allow for useful differentiation between genus and species (Peix et al. 2009).

Analysis of the 16S rRNA genes from biological samples is possible using an ever-increasing range of molecular methods (McGinn and Gut 2013; Muzzey et al. 2015). Amplification and sequencing of the 16S rRNA genes within a sample, followed by taxonomic assignment to each sequence, allows differentiation of the microbiota to the taxonomic level ranging from phylum to species (Marchesi and Ravel 2015). Though the entire gene can be sequenced, originally done using Sanger sequencing (McGinn and Gut 2013), this process is highly intensive and expensive. Therefore, most studies focus on a specific region of the gene to save cost and effort and enable higher throughput (Johnson et al. 2019). This has been achieved through the development of a range of methodologies known as 'next generation sequencing', which share many features of Sanger sequencing. However, they incorporate a few fundamental differences to technique which have enabled a massive upscale of data output (Muzzey et al. 2015).

Using technologies such as the Illumina sequencing platform, short sequences (≤ 300 bases) can be interrogated. Sub-regions of the gene either use a single variable region of 16S rRNA, such as V4, or expand up to three variable regions, such as V1-V3 or V3-V5 (Johnson et al. 2019). The Illumina platform compensates for the potential limitations associated with shorter read lengths by using a technique called 'paired-end sequencing' on fragments of 100 to 300 bases. This uses specific, short read lengths and effectively overlaps the reads, thus providing greater total fragment length, sequence accuracy and overall data quality (Kuczynski et al. 2012).

1.17 Analytical tools for studying the microbiota

Metataxonomics uses high-throughput techniques to create a metataxonomic tree showing the relationships between all sequences obtained from the entire microbiota (Marchesi and Ravel 2015). It relies on the amplification and sequencing of taxonomic marker genes (16S rRNA amplicons) as a representation of the microbiota's phylogenetic diversity (Kunin et al. 2010; Marchesi and Ravel 2015).

Following DNA extraction from the required sample, the 16S rRNA gene is amplified using appropriately-selected amplicon primers. After PCR amplification and sequencing, the sequences must be processed to discard erroneous reads and maximise data quality. Sequences are then clustered based on similarity to generate operational taxonomic units (OTUs). For example, the software package 'mothur' is a bioinformatics pipeline that can be used to neaten sequences by trimming, screening and aligning, and then assigning sequences to OTUs (Schloss et al. 2009). OTUs are then compared to existing reference databases to infer likely taxonomic classification and identification of the bacterial species (Kim et al. 2017; Johnson et al. 2019).

OTUs are based on sequence identity (% ID) and historically thresholds of % ID were used to represent taxonomic classification levels; with 97% for species and 95% for genera (Goodrich et al. 2014). However, the use of thresholds to guarantee correct identification of 'bacterial species' using sequencing of the 16S rRNA to represent the entire gene is probably unrealistic (Schloss 2010).

Management of the large data output from high-throughput sequencing has led to the development of physical and computational infrastructures. These enable microbiology bioinformaticians to produce, analyse and share data from large datasets (Connor et al. 2016). The Cloud Infrastructure for Microbial Bioinformatics (CLIMB) facility is an example of such a computing resource, allowing training, expertise and data processing within a single computational resource (Connor et al. 2016).

1.17.1 Diversity measures

Output from sequencing, in the form of OTUs, requires further statistical analysis to provide meaningful data and enable assessment of the bacterial communities within

the sample. The study of bacterial community diversity uses measures including community richness and evenness. Community richness describes the number of different species present within a defined niche; evenness demonstrates the relative abundance of these species within the community (Caverly et al. 2015; Kim et al. 2017). As species richness and evenness increase, so does the diversity (Kim et al. 2017). Ecosystems with increased richness and evenness are typically longer established and more stable, with greater ability to resist environmental perturbations such as antibiotic use (Relman 2012).

Shannon-Weaver and Simpson diversity indices are commonly-used methods for measuring bacterial diversity based on OTUs. These provide mathematical measures for species diversity within a community, inferring community composition (Kim et al. 2017). Though both measure species richness and evenness, the Shannon diversity index places more weight on species richness (Lemos et al. 2011), whilst the Simpson diversity index more heavily biases towards species evenness (Simpson 1949). Shannon diversity values increase "as the number of species increases and as distribution of individuals among species becomes more even" (Ludwig et al. 1988; Lemos et al. 2011). Simpson's diversity index indicates species dominance and is expressed from 0 to 1, with an increasing index with decreasing diversity (Simpson 1949; Lemos et al. 2011). Haegeman et al. (2013) used a set of *in silico* communities and demonstrated robust estimation of community diversity using Shannon and Simpson diversity indices. To optimise accuracy, it is necessary to ensure standardised sample size and sequence reads, normalisation and overall quality control.

It is possible to use diversity measures within a single sample (alpha diversity) and between different samples (beta diversity). Beta diversity compares microbiome communities across populations using pairwise distance measures (Maziarz et al. 2018). A number of tools exist to assess the dissimilarity between two community populations, including the Bray-Curtis dissimilarity index (Bray and Curtis 1957). Bray-Curtis is a quantitative tool to measure beta diversity and uses sequence abundance (Goodrich et al. 2014). Bray-Curtis dissimilarity uses the species-wise differences between samples relative to the total abundance of species between two communities (Ricotta and Podani 2017). The index describes the "proportion of the total species abundances in which the two plots differ" (Ricotta and Podani 2017). An index score of zero would indicate completely identical communities, with a score of 1 suggesting no commonality.

1.17.2 Technical considerations

Deep sequencing has a number of clear technical benefits. It requires only small sample volumes for analysis, which can be extremely useful when studying areas with lower microbial biomass such as the lungs. Also, the ability for PCR to amplify small amounts of DNA within a sample means that organisms occurring in very small numbers can still be detected and identified, allowing for a greater understanding of true species diversity (v. Wintzingerode et al. 1997). Limitations to sampling the lung microbiome include the aforementioned difficulties with low biomass, accessibility and potential contamination from multiple sources. Sampling the lung microbiome is challenging due to physical difficulties accessing the lower respiratory tract. Primarily, sampling occurs by bronchoscopy, which is an invasive procedure requiring general anaesthesia. The process requires passage of the bronchoscope through the upper respiratory tract, with a high potential risk for contamination by residing microbiota (Charlson et al. 2011). There is also a risk of contamination from sampling fluid, dust, cleaning reagents and instruments that may confound the data, given the relatively low biomass from lower respiratory tract samples (Charlson et al. 2011; Salter et al. 2014).

Technical limitations of deep sequence surveying may alter the reported taxonomic distributions and frequencies within a dataset (Salter et al. 2014). Potential issues relate to collection and appropriate storage of the sample, DNA extraction, choice of appropriate amplifying primers, read length and depth, and sequencing and bioinformatics analysis technology (v. Wintzingerode et al. 1997; Kunin et al. 2010; Salter et al. 2014). In addition, there is a risk of contamination by introducing microbial DNA during sample preparation, by 'sterile' water, PCR reagents and the DNA extraction kits. This is a particular problem when sampling from low biomass environments such as the lung, where contaminating DNA can generate misleading results (Salter et al. 2014). These risks can be controlled for, but not completely ruled out, by ensuring negative control samples are checked at each stage of the process (Salter et al. 2014).

When reporting results, it is important to quantify DNA on the initial clinical samples, ensure concomitant sequencing of negative controls and to describe identification and removal of contaminants (Salter et al. 2014). This should allow accurate and clinically appropriate conclusions to be drawn from the data thereafter. Another significant limitation of culture-independent analysis is that it cannot distinguish

between recently killed bacteria and live bacteria, making it impossible to confirm active infection or colonisation (Bassis et al. 2015).

1.18 The lung microbiome

1.18.1 The lung microbiome in health

Throughout the respiratory tract, specialised bacterial communities exist within specific sites. These have key roles in maintaining and protecting human health (Man et al. 2017). Adult human airways have a surface area 40 times larger than that of skin, with niche-specific bacterial communities present throughout the airway surface (Man et al. 2017). The upper respiratory tract (URT) represents the most densely populated site. The URT is also the site of primary colonisation by potential respiratory bacterial pathogens before they progress to causing infection within the upper or lower respiratory tract (Bogaert et al. 2004). The established respiratory microbiota plays an important role in preventing colonisation by 'colonisation resistance'. If colonisation is successful, the resident microbiota can then prevent "pathogen overgrowth, inflammation and subsequent local or systemic spread" by repressing growth and also enhancing interactions with the host immune system (Man et al. 2017).

It is likely that the respiratory microbiota is required for the structural maturation of the respiratory tract, as well as the development of effective local immunity (Gollwitzer et al. 2014; Yun et al. 2014). Murine models have demonstrated smaller lungs and a reduced number of mature alveoli in germ-free rodents (Wostmann 1981; Yun et al. 2014; Man et al. 2017).

In healthy individuals, the lung has a particularly low microbial biomass, which makes the microbiome more difficult to study than other environments such as the gut (Lloyd-Price et al. 2016). Traditionally, the lower airways of healthy individuals were considered sterile, but more recently both murine and human samples have challenged this concept and demonstrated clear presence of microbial communities through 16S rRNA sequencing techniques (Goulding et al. 2007; Charlson et al. 2011). Charlson et al. demonstrated that the healthy lung had the same microbial composition as the upper respiratory tract, but with 2 to 4 logs lower biomass (Charlson et al. 2011). The authors concluded that the lower amounts of bacterial sequences within the lower respiratory tract were mostly due to micro-aspiration of URT microbiota and bronchoscopy carryover from the procedure.

In healthy individuals, the lung microbiota forms as a result of bacteria entering the lungs by direct mucosal dispersion and micro-aspiration from the URT. It is proposed that the lung microbiome composition is determined by three key processes: immigration, elimination and the relative growth rates of the microbial community members (Bassis et al. 2015). Host defences must clear or prevent significant growth of microbes to prevent infection (Dickson et al. 2014; Marsh et al. 2016). Bassis et al. (2015) demonstrated that though adult lungs shared bacterial communities with those identified in the mouth, the similarity between the two communities' composition was variable. The lung did not share common features with the nasal microbiome (Bassis et al. 2015). In contrast, Marsh et al. (2016) showed that the combination of nasal and oral microbiota samples more effectively represented the microbiota seen on paediatric BAL lavage samples than oral samples alone. They found that paired BAL and upper airway (oropharyngeal (OP) and nasopharyngeal (NP) combined) were >50% similar in 69% of their paediatric samples (Marsh et al. 2016). Though the authors claim this is an "imperfect, but reliable, representation" of the BAL microbiota, it could be argued that even collecting OP and NP samples would still miss substantial numbers of microbiota compared to BAL. Contributions from gastric microbiota via gastro-oesophageal reflux or microbiota in inspired ambient air have been identified as other potential bacteria sources for the lungs, but evidence is negligible (Man et al. 2017).

Evidence suggests that the healthy lung microbiota is being formed *in utero*. This is predominantly through transfer of maternal antibodies and microbial molecules via the placenta and will contribute to the development of the immune system postnatally (de Agüero et al. 2016; Man et al. 2017). After birth, neonates are rapidly colonised, with niche differentiation in the URT within the first week of life and population composition predicting microbiota stability during the first 2 years of life (Bosch et al. 2016). However, it is unclear when a stable respiratory microbiota is fully established and it likely occurs during the first few years (Man et al. 2017).

Early colonisation is altered by the mode of delivery, with babies born by Caesarean section showing a delay in development of 'healthy' respiratory microbiota profiles containing protective commensals (Bosch et al. 2016). This may impact the infant's respiratory health in the long term. Similarly, breastfed infants also transition towards a 'healthy' microbiota more quickly, with potential benefits including reduced respiratory infections and wheezing in the first year of life (Biesbroek et al. 2014).

It is hypothesised that community composition in healthy lungs represents transiently present microorganisms derived from the URT, in direct comparison to the resident communities seen in individuals with chronic respiratory diseases (Man et al. 2017). Limited studies in preterm neonates have shown that the lower respiratory tract (LRT) microbiota is fairly simple and dominated by *Staphylococcus* spp., *Ureaplasma* spp. and *Acinetobacter* spp. (Man et al. 2017). In older infants, children and adults, the LRT microbiota was typically dominated by species seen in the URT, including "*Moraxella* spp., *Haemophilus* spp., *Staphylococcus* spp. and *Streptococcus* spp., but lacked other typical URT species, such as *Corynebacterium* spp. and *Dolosigranulum* spp." (Marsh et al. 2016; Man et al. 2017)

Once established, a number of factors can act as perturbations to the respiratory microbiota equilibrium. These include antibiotic treatment (at any time) and smoking, with the latter affecting the URT rather than LRT microbiota (Jakobsson et al. 2010; Lim et al. 2016). In contrast, exposure to beneficial bacteria can have a positive effect on development of a healthy microbiome composition, though the order and timing of exposure is critical (Man et al. 2017). A number of keystone species may have particularly significant benefits on the function and well-being of the URT ecosystem; for example, *Dolosigranulum* spp. and *Corynebacterium* spp. are associated with exclusion of potential respiratory pathogens such as *Streptococcus pneumoniae*, and maintenance of respiratory health (Pettigrew et al. 2012; Biesbroek et al. 2014; Man et al. 2017).

1.18.2 The lung microbiome in cystic fibrosis

Long-established culture-dependent sampling methods have provided clear evidence of pathogens, such as *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Haemophilus influenzae* and *Burkholderia cepacia* complex (BCC), within the CF airways (Acosta et al. 2017). These were often considered 'mono-species' infections, but data obtained by culture-independent microbial detection methods now suggests this is over-simplistic (Coburn et al. 2015; Cuthbertson et al. 2020). Methods such as 16S rRNA gene sequencing have demonstrated that the CF microbiome is highly diverse, and evolves over time, most likely in relation to patient demographics, clinical status and treatment exposure (Coburn et al. 2015; Acosta et al. 2017; Cuthbertson et al. 2020). However, when compared to healthy controls, it is clear that the CF microbiome is less diverse. There is less variation between the microbiomes of

patients with CF (inter-CF sample distance within clusters) than between healthy controls (inter-control sample distance) (Blainey et al. 2012).

Studies have shown that it is not simply the presence/absence of pathogens such as *Pseudomonas* that determines lung disease, but perhaps more importantly the pathogens' relative abundance, dominance over other species and stability (despite perturbation) that is more important to patient outcomes (Carmody et al. 2013; Coburn et al. 2015).

The conceptual framework of the island model has been used to describe the origin of infective pathogens within the lower airway. The upper airways, represented as the mainland, are microbiota-rich. These represent the source of 'migrants' to the microbiota-poor lower airways, with the latter described as the 'islands'. There is likely variable penetration of the lower airways, represented as different regions or 'islands' (Boutin and Dalpke 2017). Healthy individuals are able to clear pathogens and maintain neutrality within the lower airways. However, for patients with CF, such migration may enable the entry of opportunistic CF pathogens such as *S. aureus* as early as the first year of life. As CF mucus appears to facilitate colonisation by *S. aureus*, this may explain such early-life infections (Boutin and Dalpke 2017).

Evidence is clear that the nose, throat, oral cavity and lower airways are distinct ecological niches, with differing patterns of evolution. There are common patterns of species diversity, abundance and interaction, with changes seen with increasing age in both healthy individuals and people with CF (Boutin and Dalpke 2017). It is therefore of paramount importance to ensure that lower airway samples are obtained to study the lung microbiome, as community composition cannot be inferred by looking at upper airway samples exclusively.

The lung microbiome in less severe disease appears to be dominated by the same genera in the oropharynx and the lower airways (Boutin and Dalpke 2017). Spatial heterogeneity is apparent in CF lungs whereas healthy patients appear to have homogeneity across the lobes (Brown et al. 2014; Jorth et al. 2015; Winstanley et al. 2016a). The neutral model, whereby the lung microbiota resembles those of the throat, is seen in younger children with CF. However, aging leads to an imbalance in migration and elimination of bacteria within the lungs. The CF lung environment favours establishment of pathogens and independent, regional growth within the lower airways (Boutin and Dalpke 2017).

1.19 Analysing sputum

1.19.1 Culture-dependent and culture-independent methods

A key mainstay of CF management focuses on early identification and treatment of respiratory infections. Traditionally, this has been achieved through airway sampling and use of microscopy, culture and antibiotic sensitivity testing. Using this approach has enabled clinicians to provide targeted antibiotic therapy whenever possible (Caverly et al. 2015).

However, it is likely that culture-dependent methods are not sufficient to fully understand the CF lung environment. Data using culture-independent methods provides far greater detail about the CF lung microbiome (Blainey et al. 2012). To optimise study design and data output, obtaining airway samples needs to be sustainable, tolerable and repeatable.

1.19.2 Airway sampling

For older children and adults with CF, respiratory samples are easily obtained as expectorated sputum. However, most young children are unable to expectorate sputum and will typically swallow secretions (Forton 2015). Therefore, alternative airway sampling methods are commonly used, including cough swabs and oropharyngeal swabs. However, these are likely to capture upper airway organisms as well as the sought-after lower airway organisms. To optimise lower airway sampling, bronchoscopy with BAL should be used. Though considered the gold standard sampling method, this procedure requires an anaesthetic and hospital admission.

1.19.3 The CF-Sputum Induction Trial (CF-SpIT)

More recently, studies have explored the use of induced sputum as a non-invasive lower airway sampling method. CF-SpIT is a prospective, internally-controlled interventional trial. Patients were recruited from those attending the CF service, Children's Hospital for Wales (Cardiff, UK). The trial explored lower airway pathogen detection in sputum induction vs. cough swab, and sputum induction vs. single-lobe, two-lobe and six-lobe bronchoalveolar lavage (Ronchetti et al. 2018b). This work supported the use of induced sputum (IS) as a standard of care in combination with BAL to sample the lower airway using conventional microbiology. To further develop the work of the CF-Sputum Induction Trial, a study was undertaken using thirty within-patient, time-matched samples from four lower airway niches (IS and 3 BAL samples)

using molecular detection techniques (Weiser et al. 2022). Comparing BAL samples from different lobes, authors showed that lower airway microbiota were non-uniform across multiple compartments within the lung in a significant proportion of children.

The study then compared patterns of bacterial diversity between IS and BAL samples. 80% of IS samples captured a meaningful representation of the lower airway microbiota identified amongst the BAL samples (Weiser et al. 2022). The degree of concordance varied between full capture of the pattern seen in the BAL profiles, typically amongst those patients with no evidence of compartmentalisation, to identifying the dominant taxon and overall taxa picture but at different relative abundances. 20% of matched samples demonstrated dissimilarity between IS and all-matched BAL samples (Weiser et al. 2022). To reflect the greater diversity within IS samples, pathogens were considered present at a relative abundance of >0%, with a pragmatic cut off of >5% relative abundance for BAL pathogens. At these defined detection thresholds, IS identified one or more pathogens in 96% of samples. Specifically, IS detected the CF pathogens *Haemophilus, Pseudomonas* and *Staphylococcus* on BAL in 100%, 64% and 43% of samples respectively (Weiser et al. 2022). Overall, this study validated the use of IS to study lower airway microbiota using molecular detection methods.

1.20 Novel techniques for analysing sputum

1.20.1 Fourier transform infrared spectroscopy

Fourier transform infrared spectroscopy (FTIR) is a method of determining the structure of small molecules, including their chemical composition and even potentially their architectural formation (Barth 2007). It can be used on a wide range of materials and offers a non-invasive, rapid, high-throughput and low-cost method, requiring very small samples (Barth 2007). Though previously utilised for chemical compounds, there is increasing use of FTIR as a tool for analysis of biological specimens (Su and Lee 2020). Infrared spectroscopy studies the interaction of a sample's chemical bonds with the radiation of a light source. Each molecule produces a unique spectrum based on the wavelength and quantity of infrared radiation absorbed, with a resultant signature spectral fingerprint of absorbance peaks seen (Balan et al. 2019; Su and Lee 2020). This represents the stretching and bending vibrations between bonds within the molecule. Infrared spectroscopy uses infrared light to cause a molecule to enter a higher vibrational state at the point of the light-matter interaction. A transfer of energy occurs at certain wavelengths, depending on

the composition of the material being studied, and these energy level transitions result in a spectrum (Baker et al. 2016). The spectrum is made up of peaks and bands that can be interpreted both qualitatively (peak position) and quantitively (peak intensity and relative intensity) (Baker et al. 2016).

1.20.2 FTIR and analysis of biological materials

Biological materials absorb energy in the mid infrared region (4000 to 400 cm^{-1}) of the electromagnetic spectrum. This region is the most frequently analysed. The FTIR spectrum can be divided into five regions according to the main macromolecules (**Table 1.2**) (Baker et al. 2016).

Studies typically focus on the following spectral regions: the fingerprint region (1450-600 cm^{-1}), and the Amide I and II regions (1700 to 1500 cm^{-1}) (Su and Lee 2020). Development of spectral libraries has provided a reference for comparison of spectrum curves and identification of functional groups present and even the overall chemical composition (Balan et al. 2019). This technology is often used to detect a change of functional group in molecules from tissues or cells, enabling comparison between 'healthy and disease' or 'treated and control' (Lewis et al. 2010). Biochemical or morphological changes at the molecular level can be identified, compared to established biomarkers and known morphological alterations seen in disease, and used as a screening and diagnostic tool (Su and Lee 2020).

Table 1.2. Key macromolecules and associated wavenumbers

Macromolecule and proposed vibrational mode	Wavenumber cm^{-1}
-CH2 and -CH3 groups of fatty acids and proteins	3050–2800
C=O stretching vibrations from lipid esters	1800–1700
C=O, N–H, and C–N modes from Amide I and II protein bands	1700–1500
Phosphate vibrations from nucleic acids	1225 and 1080
Carbohydrate absorption	1200–900

1.20.3 FTIR and analysis of sputum

FTIR can be performed on a variety of biological materials, including blood, tissues, urine, extracellular vesicles, bile and sputum (Su and Lee 2020). FTIR is able to detect biochemical compositions within biological material, including nucleic acids, proteins, lipids and carbohydrates. It identifies the molecular conformation of the structure, including functional groups, bonding types and intermolecular interactions (Su and Lee 2020). Other potential uses of FTIR analysis of sputum include the identification of bacterial infection. Bosch et al. (2008) formulated a protocol for identification of gram-negative rod species, such as *Burkholderia cenocepacia* species, which they reported was reliable and enabled rapid testing. Based on the changes to sputum seen with exposure to bacteria, it is possible that alterations in sputum related to therapeutics may also be identifiable using FTIR.

1.21 Antibiotic use within the CF population

Children with CF typically commence prophylactic antibiotics from the point of diagnosis, though the CF START study is currently underway to consider whether this is the best approach to prevent infections (Southern 2020). Exacerbations are characterised by signs of pulmonary infection, such as increased cough, increased sputum production or discolouration and temperature, and associated with reduction in lung function from baseline. Exacerbations are typically managed with antibiotics. Most patients are commenced on oral antibiotics initially, with intravenous antibiotics used for exacerbations unresponsive to oral antibiotics, or for more unwell patients. It is standard treatment to use two broad-spectrum antibiotics intravenously, for example an aminoglycoside and a beta-lactam antibiotic in combination (Daniels et al. 2013).

During acute respiratory exacerbations, there is evidence of oxidative damage in the CF lung, which increases *P. aeruginosa* diversity and promotes the emergence of antibiotic-resistant bacteria (Boles and Singh 2008). Hypermutable isolates are also associated with accelerated resistance to antimicrobials, particularly with antibiotic exposure (Oliver 2004). Prevalence ranges between populations, but is thought to increase in chronic infections compared to acute infections (Fothergill et al. 2010). Fothergill et al. (2010) studied 40 isolates from CF patients with the *P. aeruginosa* Liverpool Epidemic Strain (LES) before, during and after antibiotics. They demonstrated a shift in colony morphology from smooth, pigmented morphotype to

an increase in white mucoid isolates, with overall general increase in morphology diversity (Fothergill et al. 2010). There was no increase in hypermutable isolates, but variation in antibiotic susceptibility profiles and virulence factor production was demonstrated (Fothergill et al. 2010).

With use of broad spectrum antibiotics, it is also important to consider the effects of antibiotic treatment of CF lung infections on the lungs' microflora and co-colonising microorganisms (Cullen and McClean 2015). Daniels et al. (2013) demonstrated significantly increased abundance of *P. aeruginosa* compared to other bacterial species during antibiotic therapy, whilst non-pseudomonal species were shown to fall. There is evidence that these wider species may alter the virulence of *P. aeruginosa*, thus contributing to treatment outcomes (Daniels et al. 2013). The authors also demonstrated a reduction in the mean bacterial taxa richness after antibiotic initiation. The environmental stress caused by antibiotic treatment also drives adaptation of bacterial populations, with facilitated survival of those with innate or acquired antibiotic resistance mechanisms (Wozniak and Keyser 2004; Cullen and McClean 2015).

1.22 Antimicrobial resistance

Antimicrobial resistance (AMR) occurs when the infection-causing micro-organism is able to survive exposure to a medicine that would have previously stopped growth and/or caused death (O'Neill 2016). This enables resistant strains to increase and has led to the emergence of 'superbugs' such as Methicillin-resistant *Staphylococcus aureus* (MRSA) and multi-drug resistant (MDR) tuberculosis. Such infections have become so widespread that there are now implications for prevention and treatment of infections ranging from 'simple' pneumonias to those infections associated with chemotherapy and life-saving surgical operations (O'Neill 2016).

1.22.1 Implication of antimicrobial resistance

The cost of AMR to patients, clinicians and economies is an increasing problem worldwide. It is estimated that 25,000 patients die from a serious resistant bacterial infection acquired in hospitals every year in the European Union alone (Organisation 2014). Predictions suggest that by 2050, 10 million lives a year will be at risk from drug-resistant infections (O'Neill 2016). The economic burden of AMR for the world's population is difficult to accurately measure, though it has been estimated to cost the

US healthcare system $21 to $34 billion, accompanied by more than 8 million additional days in hospital (Organisation 2014). This is anticipated to rise to a cumulative 100 trillion US dollars of economic output by 2050 (O'Neill 2016).

Despite providing life-saving treatments for severe infections, antibiotics are commonly inappropriately prescribed for even minor illnesses (Fleming-Dutra et al. 2016), in addition to widespread use by the farming industry for food animal therapy, disease prevention and growth promotion (Van Boeckel et al. 2015; WHO 2017). Though AMR is a naturally-occurring adaptive mechanism, the misuse and overuse of antimicrobials in humans, animals and the environment means this process is facilitated, with transmission of resistance occurring between humans, animals and the environment (WHO 2017). As a result, previously treatable infections are becoming impossible to control.

1.22.2 AMR and cystic fibrosis

For patients with CF, AMR has far-reaching consequences for the successful treatment of both acute and chronic infections. Antibiotic insensitivity and frank resistance of *P. aeruginosa* within CF lungs is high and has increased over the past twenty years. Pitt et al. (2003) demonstrated resistance to at least two antibiotics in approximately 40% of isolated strains (Pitt et al. 2003). This was supported by a nationwide survey in 2001, which indicated 98 of the most resistant isolates being *P. aeruginosa* from CF patients (Henwood et al. 2001). More recently, the prevalence of multi-drug resistant or extensively-drug resistant strains are rising worldwide, compromising appropriate available treatments and resulting in significant morbidity and mortality (Oliver et al. 2015). *P. aeruginosa* has remarkable intrinsic antimicrobial resistance, but its ability to further develop resistance also stands it apart from most other bacteria. In response to the environmental stress induced by antibiotics, *P. aeruginosa* undergoes a significant change in gene expression; the *algU* (also known as *algT*) regulon leads to downregulation of genes affecting bacterial motility, virulence and overall metabolism, whilst upregulating those related to membrane permeability and drug efflux (Folkesson et al. 2012).

Current intensive use of antibiotics in CF patients is thought to have enabled *P. aeruginosa* colonisation within the lung as, prior to such treatments, children would have succumbed to severe *S. aureus* infections in early childhood (Folkesson et al. 2012). According to the Cystic Fibrosis Foundation's most recent report, of those

patients positive for *P. aeruginosa* in 2017, 17.9% were reported to carry MDR strains (Registry 2018). Therapeutic options are limited by the ongoing emergence of antimicrobial-resistant strains, making the need for alternative treatments an urgent priority (Gellatly and Hancock 2013). Antibiotic resistance is seen to affect all classes of antibiotics, most commonly through chromosomally-encoded mechanisms. Resistance to beta-lactams and aminoglycosides in particular is linked to imported genes encoding for drug-inactivating enzymes through horizontal gene transfer (Folkesson et al. 2012; Oliver et al. 2015). Genetic mutations in antibiotic-resistance associated genes are amongst the most common seen in the adaptation process enabling *P. aeruginosa* survival. These typically reflect the individual's exposure to particular antibiotic regimes and often occur as *de novo* mutations with clonal expansion thereafter.

The fitness cost of antibiotic resistance should be considered, as this may affect bacterial virulence (positive or negative effect) and virulence traits (Oliver et al. 2015). Fitness cost can be defined as a "detrimental impact on the ability of an organism to survive and propagate in a particular environment" (Clark et al. 2015). Rapid changes in antimicrobial susceptibility profiles have been reported within 7 days after antibiotic administration (Clark et al. 2015). However, the fitness cost of maintaining such resistance after removal of the specific antimicrobial agent means selection of this phenotypic trait may not continue.

Chronic infection with *P. aeruginosa* provides an additional challenge, as it often appears that the infection is susceptible to antibiotics *in vitro* but cannot be eradicated within the lungs. It has been proposed that this phenomenon is created by persister cells. These are dormant phenotypic variants, which make up a small proportion of the population (Mulcahy et al. 2010). These specialised survivors are naturally present, rather than appearing through mutations. They have intrinsically increased tolerance of antimicrobial treatment due to their innate inactivity, and may represent a key cause of persistent infection despite use of treatment that appeared effective *in vitro* (Mulcahy et al. 2010).

Antimicrobial susceptibility testing is notoriously difficult with such significant phenotypic diversity occurring within the same host. Limitations occur as testing will be undertaken on single isolates deemed to be representative of the overall population, whereas in fact there exists a variety of isolates with varying degrees of antibiotic resistance (Kidd et al. 2018). This limits the use of antimicrobial tests on

predicting clinical outcome and optimising patient therapy (Winstanley et al. 2016a). Understanding adaptive mechanisms for survival within the CF lung is essential to developing effective antimicrobial treatments.

1.22.3 Future challenges

Recommended interventions to tackle AMR have included a global public awareness campaign, improvements in clinical diagnostic technology to avoid incorrect usage, reduction of antibiotic use within the agricultural sector and, finally, improving the antibiotic development pipeline (O'Neill 2016). With rapidly emerging drug resistance and few novel antibiotics in the development pipeline, there is an urgent need for alternative antimicrobial therapies. Development of novel therapies is required, particularly directed to combat biofilm growth and tackle multi-drug resistant strains, such as those seen in *P. aeruginosa* CF lung infections (Lopez-Causape et al. 2015). This will provide the opportunity to optimise treatments for people with CF, in whom the emergence of AMR is an increasing concern.

1.23 Novel therapies

Despite significant advances in the treatment of CF, including CFTR modulator therapies, there is still a need for new therapeutics for chronic infection and inflammation and airway clearance (van Koningsbruggen-Rietschel et al. 2020). The low-molecular weight alginate oligosaccharide OligoG CF-5/20 is a novel therapeutic currently being developed. It offers a potential alternative means to improving mucociliary clearance in CF, with additional abilities to reduce bacterial lung infections.

1.23.1 Structure of alginates

Alginate is a natural polysaccharide which is widely used in the pharmaceutical, food, printing and textile industries (Qin 2008). The commercial process of alginates from farmed brown seaweed has been in place since the early 20th Century (Hay et al. 2013a). A number of brown seaweeds contain alginate, with the most commercially-relevant being *laminaria*, *macrocystis* and *ascophyllum* (Qin 2008). *Laminaria hyperborean*, for example, contains 17-33% and 25-30% alginate content on a dry weight basis for the fronds and stems respectively (Qin 2008). The chemical structure of alginate is a linear polymeric acid composed of 1,4-linked β-D-mannuronic acid (M) and α-L-guluronic acid (G) residues (Qin 2008). These residues alter stereo-

chemically due to their differing structures at C-5. Alginate is a block copolymer of β-D-mannuronic acid and α-L-guluronic acid, with three different blocks possible: GG blocks contain only units derived from L-guluronic acid; MM blocks contain only units of D-mannuronic acid; and MG blocks consist of alternating units from D-mannuronic acid and L-guluronic acid (Qin 2008). MM block segments produce a linear, ribbon-like, flexible polymer as a result of the β (1→4) linkages formed by mannuronic acid (Qin 2008; Yang et al. 2011a). GG block segments form α (1→4) linkages by guluronic acid, resulting in a much stiffer molecular chain and a folded, rigid structural conformation (Qin 2008; Yang et al. 2011a). The physical properties of different alginates will vary according to both the G and M content, as well as the relative ratio of the three blocks within the structure (Qin 2008). It is important to quantify the relative uronic acid content within the alginate to fully predict its physical properties and therefore its value within industry.

1.23.2 Bacterial biosynthesis of alginates

Alginates are produced by two genera of bacteria: *Pseudomonas* and *Azotobacter* (Hay et al. 2013a). Alginate biosynthesis can be divided into a four-step process, including precursor synthesis, polymerisation, periplasmic modification/transit and secretion, and is controlled by 13 core genes (Hay et al. 2013a; Hay et al. 2014). The mucoid strain of *P. aeruginosa*, typically seen in respiratory samples from older CF patients with chronic infection, demonstrates hypersecretion of alginate to aid thick, highly-structured biofilm formation (Nivens et al. 2001; Hay et al. 2009a; Hay et al. 2009b; Hay et al. 2013b). The alginate produced is an O-acetylated linear polymer of D-mannuronate and L-guluronate residues (DeVries 1994). The complex mechanism of alginate overexpression within CF isolates is predominantly controlled by the master regulator alternate sigma factor AlgU (previously called AlgT or σ22) (Nivens et al. 2001; Hay et al. 2014). AlgU is encoded in an operon containing four other genes: *mucA, mucB, mucC and mucD*, all of which modulate its activity. Mutations in *mucA* are most commonly seen in *P. aeruginosa* CF strains, and have been directly linked to the overproduction of alginate within these strains (Hay et al. 2014). This occurs as a result of mutations inactivating the anti-sigma factor *mucA* and switching the regulatory system into a permanent 'on' state (Hay et al. 2014). Interestingly, the action of suppressor mutations at the algT locus cause mucoid CF strains to rapidly switch to the non-mucoid phenotype when grown in laboratory medium rather than within CF respiratory epithelium (Nivens et al. 2001).

Both the respiratory condition of the CF lung and the host's inflammatory response to the presence of *Pseudomonas* increase alginate synthesis and lead to conversion to the mucoid phenotype (Kipnis et al. 2006). Conditions of environmental stress, such as the presence of antimicrobials, oxidising agents, elevated temperatures and osmotic imbalances, many of which are commonly seen in CF patients, will all activate AlgU with subsequent loss of the regulatory cascade and overproduction of alginate (Hay et al. 2014). This process aids anchorage of *P. aeruginosa* to the colonised respiratory epithelium, protects it against phagocytosis and antibiotics, and can also attenuate the host response (Kipnis et al. 2006). Though it is not crucial to biofilm development, the presence of alginate has been shown to aid maturation of the biofilm and enable the formation of thick and highly structured biofilms with differentiated microcolonies (Nivens et al. 2001; Kipnis et al. 2006; Hay et al. 2014). Alginate also maintains hydration of the cells in desiccated conditions, which may be particularly helpful within the CF lung where there is recognised ASL dehydration (Hay et al. 2014).

1.23.3 Industrial production of alginate

Alginate has a number of significant properties which have enabled its widespread use in a variety of industries. It is considered non-toxic, non-immunogenic, biodegradable and biocompatible (Qin 2008; Rehm 2010; Yang et al. 2011a; Khan et al. 2012b). Sodium alginate is considered a safe product by the U.S Food and Drug Administration. As a natural polymer, alginate also represents a renewable resource with a potentially unlimited supply within the natural environment (Qin 2008; Rehm 2010). Alginates with a G content of >50% do not create an immune response within the body, whereas a high M content will stimulate the immune system. This is of relevance when producing a commercially-viable product which needs to be non-immunogenic (Christensen 2011).

As naturally-occurring alginate has a variable G/M composition and molecular weight with associated variation in material properties, it has been necessary to develop methods of generating polymers of defined weight and composition (Rehm 2010; Khan et al. 2012b). This has enabled products that can be used as stabilisers, viscosifiers and gelling agents in the food, drink, pharmaceutical and printing industries (Hay et al. 2013b). Within medicine, alginates have been developed for use in advanced wound management (Qin 2008). Production of alginate requires treatment of raw seaweed by an aqueous alkali solution, such as sodium hydroxide,

enabling conversion of the various salt forms into a water-soluble sodium alginate (Qin 2008). Following filtration, the sodium alginate can be precipitated by the addition of calcium chloride and a water-soluble sodium alginate is produced following purification and conversion (Qin 2008).

1.23.4 OligoG CF-5/20

OligoG CF-5/20 is a novel pharmaceutical preparation derived from linear alginate polysaccharides extracted from the brown seaweed *Laminaria hyperborea*. The initial polysaccharides produced are arranged in blocks of homopolymeric guluronic (G) or mannuronic (M) acid and heteropolymeric blocks of alternating G and M (Ermund et al. 2017). Through a process of selective hydrolysis and fractionation, a low-molecular-weight polydisperse oligosaccharide (mean Mn 3200 $gmol^{-1}$) enriched from sodium alginate polysaccharide is produced (Hengzhuang et al. 2016). This contains a mixture of G-rich polysaccharides, composed of more than 85% guluronic acid blocks with 1-30 guluronic units (G1-G30) (Ermund et al. 2017). Through further fractionation to remove the short polymers, the pharmaceutical preparation of Oligo-G CF-5/20 is produced. It has a mean degree of polymerisation (DPn) of 16. The process is completed by purification through charcoal filters and spray-drying (Khan et al. 2012b). The majority of material contains molecules between 5 and 20 G units, with most polysaccharides containing 12 to 14 G units (Ermund et al. 2017). OligoG CF-5/20 also contains M blocks (1,4-linked β-D-mannuronic acid), though these represent <15% of the oligomer composition (Hengzhuang et al. 2016). OligoG CF-5/20 is produced by AlgiPharma AS (Sandvika, Norway), who supplied the product for use in this research project.

1.23.5 Established effects of OligoG CF-5/20

Prior studies have demonstrated that OligoG CF-5/20 alters the visco-elastic properties of mucin/alginate gels and mucin/DNA gels (Nordgård and Draget 2011; Pritchard et al. 2016a). In addition to direct effects on mucin, OligoG CF-5/20 has also been shown to chelate calcium, enabling it to detach CF mucus and facilitate the process of normal mucin unfolding (Ermund et al. 2017). A CF mouse model demonstrated that OligoG CF-5/20 treatment reduced the accumulation of mucin, normalising the mucus phenotype and improving long-term survival (Vitko et al. 2016). *In vitro* studies have shown that OligoG CF-5/20 disrupts bacterial biofilm formation and growth (Pritchard et al. 2017a; Powell et al. 2018). It also reduces pseudomonal growth and microcolony formation (Pritchard et al. 2017a). This is

thought to explain its ability to potentiate the effect of antibiotics (Khan et al. 2012a; Pritchard et al. 2017a). Its antimicrobial effects are not related to interactions with lipopolysaccharide (LPS) or cell permeability (Pritchard et al. 2017b). However, there is evidence that it reduces the expression of the *las* and *rhl* components of *P. aeruginosa* quorum sensing, with likely influences on virulence and biofilm development (Powell et al. 2014b; Jack et al. 2018).

Clinical trials are underway, with initial results showing repeated inhalation of OligoG CF-5/20 dry powder (DPI) was safe in adults. However, the study was unable to show significant treatment benefit with OligoG CF-5/20 compared to placebo (van Koningsbruggen-Rietschel et al. 2020). Recommendations were made to reduce the dose of OligoG CF-5/20 DPI in future phase 2B clinical studies, which are currently being performed under the framework of HORIZON2020.

1.24 Aims and objectives

1.24.1 Aims

Use of culture-independent methods have demonstrated the complexity of the CF lung microbiome. Current methods for studying the paediatric CF lung microbiome are invasive and time-consuming, limiting their role in regular surveillance. The value of induced sputum for culture-dependent sampling has already been demonstrated. IS has also been shown to be representative of broncho-alveolar lavage in a small dataset. This thesis will use a large sample set of paediatric IS samples to establish microbial diversity and consider the effects of clinical treatment on diversity measures.

OligoG CF-5/20 is a novel therapy currently undergoing clinical trials. To understand its mechanism of action and potential clinical effects, further data is required. This thesis will use a FTIR to analyse CF sputum that has been treated with OligoG CF-5/20 to explore potential antimicrobial interactions with *ex vivo* samples.

Clinical use of OligoG CF-5/20 will likely entail repeated treatment courses or prolonged patient use over weeks to months. It is paramount to understand the effects of OligoG CF-5/20 treatment on relevant CF bacteria such as *P. aeruginosa*, particularly with regards to pathogenicity. This thesis will use a novel culture-dependent method to study the effects of prolonged OligoG CF-5/20 by assessing the phenotypic and genotypic characterisation of *P. aeruginosa*.

1.24.2 Objectives

1. Analyse paediatric IS samples from the CF SpIT trial using culture-independent methods to establish microbial diversity within the CF lung using a new sampling method. A clinical database will be constructed from patients' records and correlations between microbiota diversity and clinical parameters will be evaluated and assessed.

2. Identify possible structural interactions between OligoG CF-5/20 and *ex-vivo* CF sputum using FTIR. Samples from the CF SpIT trial, provided by patients with CF, will be used to analyse this novel therapy's mechanism of action on sputum.

3. Explore the effects of prolonged exposure to OligoG CF-5/20 on *P. aeruginosa* using a novel culture-dependent model to represent the CF lung environment. A range of assays will be utilised to characterise any phenotypic or genotypic alterations that may arise in *P. aeruginosa* colonies.

Chapter Two

Induced sputum samples can be used to investigate microbial diversity in children with cystic fibrosis

2.1 Introduction

Development of cystic fibrosis lung disease starts in the first year of life. Disease stems from early infection, inflammation and resultant functional and structural lung disease. Identification and timely treatment of infection is critical to reduce lung damage and disease progression.

2.1.1 Traditional airway sampling methods

Traditional methods of culturing viable organisms from respiratory samples have aided decision-making regarding the commencement of appropriate antibiotics during pulmonary exacerbations (Caverly et al. 2015). However, paradox responses to antibiotic therapy, whereby there may be clinical improvement despite evidence of antibiotic-resistant organisms on laboratory testing, suggest that culture-dependent methods may not sufficiently explain airway microbial ecology within the CF lung (Blainey et al. 2012).

Frequent and timely microbial surveillance represents one of the key cornerstones of effective CF care. Lower airway sampling within the paediatric population is notoriously difficult as young children are generally well, cough-free and do not expectorate sputum, even during a pulmonary exacerbation (Forton 2015). Therefore, sampling has been traditionally limited to non-invasive cough swabs or oropharyngeal swabs, in which diagnostic accuracy is questionable and many true positives are likely to be missed (Rosenfeld et al. 1999). Rosenfeld et al. (1999) demonstrated a sensitivity of 44% for identification of *Pseudomonas aeruginosa* on oropharyngeal cultures compared to BAL culture.

The alternative method of BAL represents an invasive procedure commonly requiring a general anaesthetic, skilled personnel and increased costs (Zampoli et al. 2016). However, this remains the gold-standard method for airway sampling (De Blic et al. 2000) and is considered the best sampling technique for infants and young children with CF (Brennan et al. 2008).

2.1.2 Induced sputum

Numerous studies have explored and demonstrated the benefits of using Is samples as a non-invasive lower airway sampling method. Authors have described the procedure as safe, well-tolerated, feasible and repeatable, making it a useful tool for both one-off and longitudinal monitoring (Blau et al. 2014; Forton 2015; Ronchetti et al. 2018b).

Sensitivity and specificity of IS when compared directly to other sampling methods have been variable. There is clear evidence that IS roll superior to upper airway sampling for children with CF (Zampoli et al. 2016; Ronchetti et al. 2018b). Comparisons have also been made between culture-based bacterial and pathogen yields of IS with the gold standard invasive technique of BAL (Blau et al. 2014; D'Sylva et al. 2017; Ronchetti et al. 2018b). Microbiological sensitivity for pathogen detection using IS compared to BAL ranges from 27.3% (D'Sylva et al. 2017) to 69% (Ronchetti et al. 2018b). In part, such differing results reflect variations in study design. However, it is also likely that IS, similar to expectorated sputum in older children, may be representing a different compartment of the CF lung from BAL, and therefore the samples will not necessarily completely align (Forton 2015).

2.1.3 Defining the airway microbiota

Through significant technological developments, it is now possible to study the lung microbiome in much greater depth by analysis of 16S rRNA gene sequences obtained via high-throughput, next-generation sequencing approaches (Blainey et al. 2012). These technologies have facilitated the study of evolutional ecology of the CF lung microbiota, which has the potential to be used as a biomarker for host phenotype, predicting outcomes such as disease prognosis and response to treatment (Kuczynski et al. 2012).

Whilst most studies of lower airway microbiota in children with CF have been performed on BAL or expectorated sputum samples, there have been some small studies using IS. Weiser et al. (2022) used matched sets of thirty IS and BAL samples for within-patient comparison of lower airway microbiota. In addition to demonstrating clear evidence of compartmentalisation in children with CF, they also validated the use of IS as a key tool in lower airway sampling. The authors reported that 50% of IS

samples closely resembled at least one of the matched BAL compartments, and a further 30% were related in composition (Weiser et al. 2022). IS detected 86.2% of the top 5 genera identified on matched BAL samples based on relative abundance. When identifying specific CF pathogens, IS demonstrated variable performance depending on the pathogen, with 43-100% sensitivity and 73-100% negative predictive values (Weiser et al. 2022).

The current study will use the complete IS sample set (136 samples) available from the CF-SpIT cohort to explore the clinical application of microbiota profiling in the paediatric CF population.

2.2 Aims

The hypothesis of this study was that culture-independent methods could be used on paediatric IS samples to identify correlations between patients' clinical features and patterns of microbial diversity. Specific aims were:

- To perform culture-independent microbiota analysis on IS samples from children with CF.
- To describe the respiratory microbiota in the CF lung within the paediatric population using non-invasive IS sampling.
- To compose a clinical database for all patients within the dataset and to explore clinical correlations with diversity measures.

2.3 Materials and methods

2.3.1 Study design and participants

CF-SpIT is a prospective, internally-controlled interventional trial performed at the Noah's Ark Children's Hospital for Wales, Cardiff, United Kingdom. Eligible study participants were children with CF aged between 6 months and 18 years. All children attending the Children's Hospital for annual review in the outpatient clinic, treatment of an acute chest exacerbation, clinically-indicated bronchoscopy and/or any routine surgery under general anaesthetic were eligible for inclusion (Ronchetti et al. 2018b).

Though the full CF-SpIT trial included samples obtained through cough swab, IS and BAL, this study focuses on IS samples only. The procedure for sputum induction has been described previously (Ronchetti et al. 2018b). Briefly, this requires nebulised delivery of 7% hypertonic saline (sodium chloride), physiotherapy performed by a specialist physiotherapist and collection of a sputum sample, either by expectoration or using oropharyngeal suction in children unable to spontaneously expectorate.

The CF-SpIT study has been subject to Institutional Review by the Cardiff and Vale Research Review Service (CaRRS; Project-ID-11-RPM-5216). Ethical approval was obtained from the South Wales Research Ethics Committee (11/WA/0334). The study has been registered with the UK Clinical Research Network (14615) and the International Standard Randomised Controlled Trial Network Registry (12473810).

Results from CF-SpIT using conventional microbiology (Ronchetti et al. 2018b) and 30 paired IS and six-lobe BAL samples (Weiser et al. 2022) have been published previously. The focus of this project was to analyse the complete set of IS samples using culture-independent methods.

2.3.2 Clinical database data collection

Patient data was obtained from two clinical digital workspaces available for routine patient care within the Cardiff and Vale University Health Board. Clinical data focused on clinical status at the time of sampling, clinically-relevant parameters, such as growth and lung function, and current and historical treatment regimens. When calculating duration of treatments, including mucolytics and antibiotics, patient letters

were used. The database only holds electronic letters from 2007 onwards. If duration of treatment was less than a year, this was labelled as 0.5 years. If it was not possible to measure the exact start date due to lack of documentation or presumed documentation prior to electronic notes availability, the following age of initiation was presumed:

1. Hypertonic saline - 6 years old
2. DNase - 6 years old
3. Colomycin - 6 years old
4. Flucloxacillin – 0 years old (following positive screening result in neonatal period).

These ages were chosen based on the CF pulmonary medication guidelines (Mogayzel Jr et al. 2013). Duration of treatment also reflects the child's age when commenced pre-2007 records. Historically, flucloxacillin was commenced following a positive screening test, typically during the first 1-2 months of life. However, for many patients, treatment has been intermittent during childhood. For children in whom flucloxacillin use was unclear, or they had previously received treatment but were not on it at the time of testing, it was recorded that they were not on treatment.

Predicted lung function for age, ethnicity and height, plus Z-scores, were calculated using the European Respiratory Society Global Lung Initiative (GLI) Spirometry Task Force online calculator (Quanjer et al. 2012). The World Health Organisation (WHO) AnthroPlus anthropometric calculator (Blössner et al. 2009) was used for weight, height and BMI Z-score calculations. Lung function was performed on all children aged 5 years and above.

Any patients on continuous oral antibiotic treatment for any period of time, for example children receiving treatment for non-tuberculous mycobacteria, had 'number of courses' calculated as total days of treatment divided by 14 days, as a 14-day course is considered 'standard' oral treatment.

2.3.3 Sample collection and storage

IS samples were collected between 3rd April 2012 and 25th April 2018. Each sample was divided to allow for analysis within different arms of the CF-SpIT trial, including the culture-dependent and culture-independent analyses. One aliquot was frozen at -80°C within 30 minutes of collection to be used subsequently for DNA extraction. The DNA extraction protocol was adapted from previous CF studies (Ronchetti et al. 2018b; Weiser et al. 2021), allowing for variable sample volume and viscosity.

As described in (Weiser et al. 2022), the protocol set up varied by volume as follows: (a) if the sample volume was >800 µl, 800 µl was used; (b) if the sample volume was between 400-800 µl, all the sample was extracted; (c) if the sample volume was <400 µl, the volume was increased to 400 µl using sterile water and all the sample used. Sample mixtures were added into sterile 2 ml non-stick tubes and centrifuged at 21100 x g for 5 mins. Supernatant was removed to reduce total volume to 400 µl, and pelleted material was resuspended by pipetting. The 400 µl resuspension was added directly into Maxwell® 16 Tissue DNA purification kit cartridges, and extracted using the automated Maxwell® 16 instrument (Promega, Southampton, UK). Blank controls for the DNA extraction kits were included. Eluted DNA was stored at -20°C. DNA extraction was also completed by Dr Julian Forton, Dr Cerith Jones and Dr Rebecca Weiser.

2.3.4 16S rRNA gene sequencing and bacterial diversity analysis

2.3.4.1 16S rRNA gene sequencing

Sample library preparation and 16S rRNA gene sequencing was performed at the Cardiff University Genomics Research Hub. The V4-region of the 16S rRNA gene was amplified and sequenced as previously described (Kozich et al. 2013), generating 250 bp paired-end reads on the Illumina MiSeq platform.

Bioinformatic analysis was carried out using a virtual machine hosted by the Cloud Infrastructure for Microbial Bioinformatics (CLIMB) consortium (Connor et al. 2016). Quality control and Illumina adapter trimming of the raw sequencing reads was performed using FastQC (Andrews 2010) v0.11.5 and Trim Galore! (Krueger) v0.4.3 for paired end reads. Trimmed reads were analysed with mothur (Schloss et al. 2009) 139 v1.39.5 using the MiSeq SOP pipeline. OTUs were defined using a cut off value of 97% and sequence reads were subsampled to 1000 reads per sample. After

subsampling, sequence coverage was calculated in mothur as >97.4%. OTUs with <10 reads across the entire dataset were excluded.

Taxonomic classification was obtained using SILVA reference files (release 132). OTUs were consolidated to genus level for downstream analysis and OTUs that could not be classified to genus level were left at family level. Raw sequence data have been submitted to the European Nucleotide Archive under project number PRJEB34389. Thanks to Dr Rebecca Weiser who performed the bioinformatic and bacterial diversity analysis.

2.3.4.2 Sequencing controls

Samples were sequenced across two MiSeq runs, each containing a mock community control (HM-783D Microbial Community B; BEI resources). Analysis of the mock communities confirmed consistent and accurate microbiota profiling (Weiser et al. 2022).

Blank DNA extraction controls were included for three different batches of Maxwell® 16 Tissue DNA purification kits. One DNA extraction control was included in the first sequencing run, and two were included in the second. After sequencing of the 16S rRNA gene reads within these blanks, the taxa present and their abundance was evaluated using R statistical software. A consistent contamination signature was identified, with sequence reads from the family *Enterobacteriaceae* being the dominant contaminants in all three controls (>84% relative abundance). No other contaminants were found >1% relative abundance in more than one blank. Weiser et al. (2022) developed a decontamination procedure based on sequence read numbers and 16S rRNA gene copies (determined via qPCR) to selectively remove *Enterobacteriaceae* (and other contaminant) reads from low biomass BAL (n=48/90) and IS (n=6/30) samples most at risk of contamination (<6000 sequencing reads and/or gene copies). Full details of the decontamination procedure are given in Weiser et al. (2022).

As qPCR data were not available for the full set of IS samples (n=136), determination of 'low biomass' status and decontamination using the method of Weiser et al. (2022) was not possible. Removing *Enterobacteriaceae* completely was not an option as IS samples have previously been identified as culture-positive for *Enterobacteriaceae*. Therefore, to avoid introducing bias in the datasets, the samples were analysed

without modification in relation to the controls, and the limitation of this approach is acknowledged.

2.3.5 Statistical analysis

Statistical analysis was carried out in Microsoft Excel, SPSS and R statistical software (R-Core-Team 2013).

Alpha diversity indices for IS samples were calculated using the vegan package in R. This data analysis was performed by Dr Rebecca Weiser. Shannon index measures were utilised as the principal diversity measure for clinical correlation analysis.

The clinical database was completed on Microsoft Excel and analysed using IBM SPSS Statistics Version 27 and Excel. Statistical correlations were tested using clinical parameters and diversity measures. The diversity data were normally distributed, calculated by the distribution of the sum of the residuals using the Q-Q plot in SPSS. Transformation of the diversity data was not required. Clinical data, such as patient age or lung function, was assessed to establish whether it was normally distributed using the Q-Q plot in SPSS. Correlation or comparison of mean values was performed for IS diversity measures and clinical data using Pearson correlation and Spearman rank or Independent samples t-test (all 2-tailed) respectively.

Where patients had provided multiple samples, their contribution to the overall data output was controlled using Generalised Estimating Equations (GEE) in SPSS. Multiple samples were controlled for using the patient's case number. The working correlation matrix structure was either 'Independent' or 'Unstructured'. The choice of which structure to use was based on the Goodness of Fit 'Quasi Likelihood under Independence Model Criterion (QIC)'. Each clinical correlation test was considered separately. The structure giving the lowest QIC value was considered the better correlation structure. A linear model was used, with Wald Chi-square statistics reported. Statistical significance was determined by the conventional p-value threshold of <0.05 for all tests. To decrease the false discovery rate when testing multiple samples, the Benjamini-Hochberg procedure was applied to p-values generated from GEE calculations.

2.4 Results

2.4.1 Clinical data

There were 136 induced sputum samples from 86 individual patients. Patients provided between one and six samples over the six-year study period. Patient age at time of collection ranged between 6 months and 17.7 years, with a mean age of 8.2 years. 97.8% were Caucasian, with the remaining population (2.2%) of South East Asian ethnicity. 63.2% were defined as unwell or had a wet cough at the time of sampling.

The CF SpIT trial originally performed 200 paired cough swab and sputum-induction procedures in 124 patients, with 33 sputum-induction procedures being unsuccessful. Further samples, including bronchoalveolar lavage and throat swabs, were obtained from some patients. These samples were not included within this work. There were 176 patients who were eligible for inclusion, as patients within Cardiff and Vale University Health Board's paediatric CF service. 136 of 200 sputum samples had sufficient quantity available for diversity analysis **(Figure 2.1)**.

Table 2.1. describes the patients' clinical demographics, including current clinical status, traditional markers of disease state such as lung function and growth, and treatments prescribed.

<u>Figure 2.1 Participant Flow Diagram</u>

Samples provided by study participants, including numbers of induced sputum (IS) samples provided by individual patients and available for specific analysis (lung function/body mass index (BMI) and treatment exposure).

131 patients enrolled from the Paediatric Cystic Fibrosis service, Cardiff and Vale University Health Board

Samples provided:

136 induced sputum (86 patients)

176 bronchoalveolar lavage samples (70 patients; up to 4 samples per procedure)

22 cough swabs (20 patients)

17 throat swabs (17 patients)

Data from cough swabs, throat swabs and BAL excluded from this analysis.

Induced sputum samples (86 patients, 136 IS samples)

1 sample (56 patients)

2 samples (17 patients)

3 samples (8 patients)

4 samples (4 patients)

6 samples (1 patient)

Lung function (58 patients, 89 IS samples)

1 sample (39 patients)

2 samples (12 patients)

3 samples (4 patients)

4 samples (2 patients)

6 samples (1 patient)

Patients <5 years excluded

Treatment exposure (including mucolytic and antimicrobial therapies)

Contributing patients and sample numbers detailed within individual results sections.

BMI (75 patients, (115 IS samples

1 sample (51 patients)

2 samples (14 patients)

3 samples (6 patients)

4 samples (3 patients)

6 samples (1 patient)

Patients <2 years excluded

67

Table 2.1. Clinical demographics.

Clinical demographics	
Age, years (min, max)	8.3 (0.47, 17.7)
Male	55.9 %
Genotype	Homozygote delta F508 47.8 % Heterozygote delta F508 41.9 % Other 10.3 %
Mean BMI for age (Z-score)	0.31
Mean FEV1 for age (Z-score), at time of sampling	-1.71
Mean FEV1 % predicted, at time of sampling (min, max)	79% (32, 116)
Mean FEV % predicted, best of year (min, max)	95.4% (69, 127)
Hypertonic saline use (% of study participants)	73.5 %
DNase use (% of study participants)	75.7 %
Flucloxacillin prophylaxis (% of study participants)	69.1 %
Azithromycin prophylaxis (% of study participants)	55.1 %
Nebulised colomycin (% of study participants)	44.1 %
Nebulised tobramycin (% of study participants)	14.7 %
Previous *Pseudomonas aeruginosa* isolation (% of study participants)	72.1 %
Intravenous antibiotics received in last 2 years (% of study participants)	63.2 %
Mean number of intravenous antibiotic courses within last 2 years (min, max)	1.4 (0, 7)
Mean number of days of intravenous antibiotic use in last 2 years (min, max)	17.7 (0, 98)
Acute oral antibiotics in last 2 years (% of study participants)	99.3 %
Mean number of oral antibiotic courses within last 2 years (min, max)	8 (0, 21)
Mean number of oral antibiotic courses within last 2 years (min, max)	8.8 (0, 52)
Mean number of days of oral antibiotic use in last 2 years (min, max)	143 (0, 730)
Unwell or wet cough at time of sample (% of study participants)	63.2 %
Inpatient at time of sample (% of study participants)	41.2 %

2.4.2 Markers of disease severity

Across the patient population, there was a statistically significant negative correlation between lung function, using FEV1 Z-score, and increasing age (Pearson correlation: -2.85, 2-tailed significance: 0.007; GEE (n=89) <0.001) (**Figure 2.2.**).

Figure 2.3. also demonstrates a statistically significant negative correlation between BMI Z-score for over 2-year-olds and increasing age (Pearson correlation: -0.288, 2-tailed significance: 0.002; GEE (n=115) <0.001). BMI Z-scores are not included for children below two years, as this measure is not validated for younger children.

Mucociliary clearance nebulised therapies, DNase and hypertonic saline, were used by 75.7% and 73.5% of the population respectively (**Table 2.1.**). Samples were provided by patients receiving prophylactic oral antibiotics (flucloxacillin and azithromycin) and nebulised antibiotics (colomycin and tobramycin). Nebulised therapy is generally commenced as an anti-pseudomonal treatment in response to repeated positive respiratory cultures, recurrent exacerbations, and potentially declining clinical status with positive *P. aeruginosa* cultures. This is likely reflected by the number of patients in whom *P. aeruginosa* has previously been isolated on culture (72.1%).

Use of acute courses of oral antibiotics was seen in almost all patients (99.3%), with a mean of eight courses over two years (range 0-21) (**Table 2.1.**). In a small number of patients, continuous antibiotics were required for the full two years prior to sampling as treatment for non-tuberculous mycobacteria. Intravenous antibiotics were received by 63.2% of the patient group, with a mean of 1.4 courses (range 0-7) delivered in the past two years.
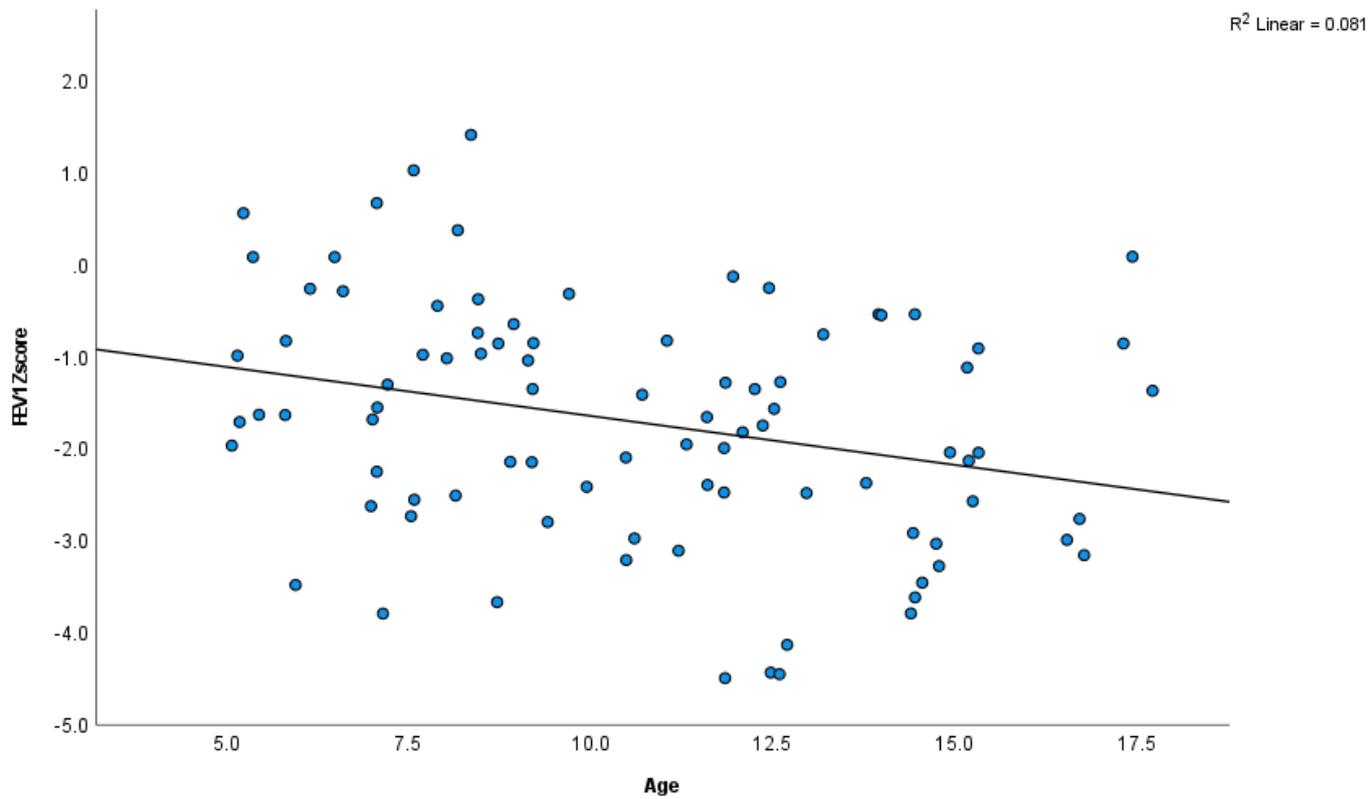
Figure 2.2. Scatterplot of FEV1 Z-score by age.

Negative correlation between lung function and age, with declining FEV1 Z-score with increasing patient age. Pearson correlation: -2.85, 2-tailed significance: 0.007; GEE (n=89) <0.001.
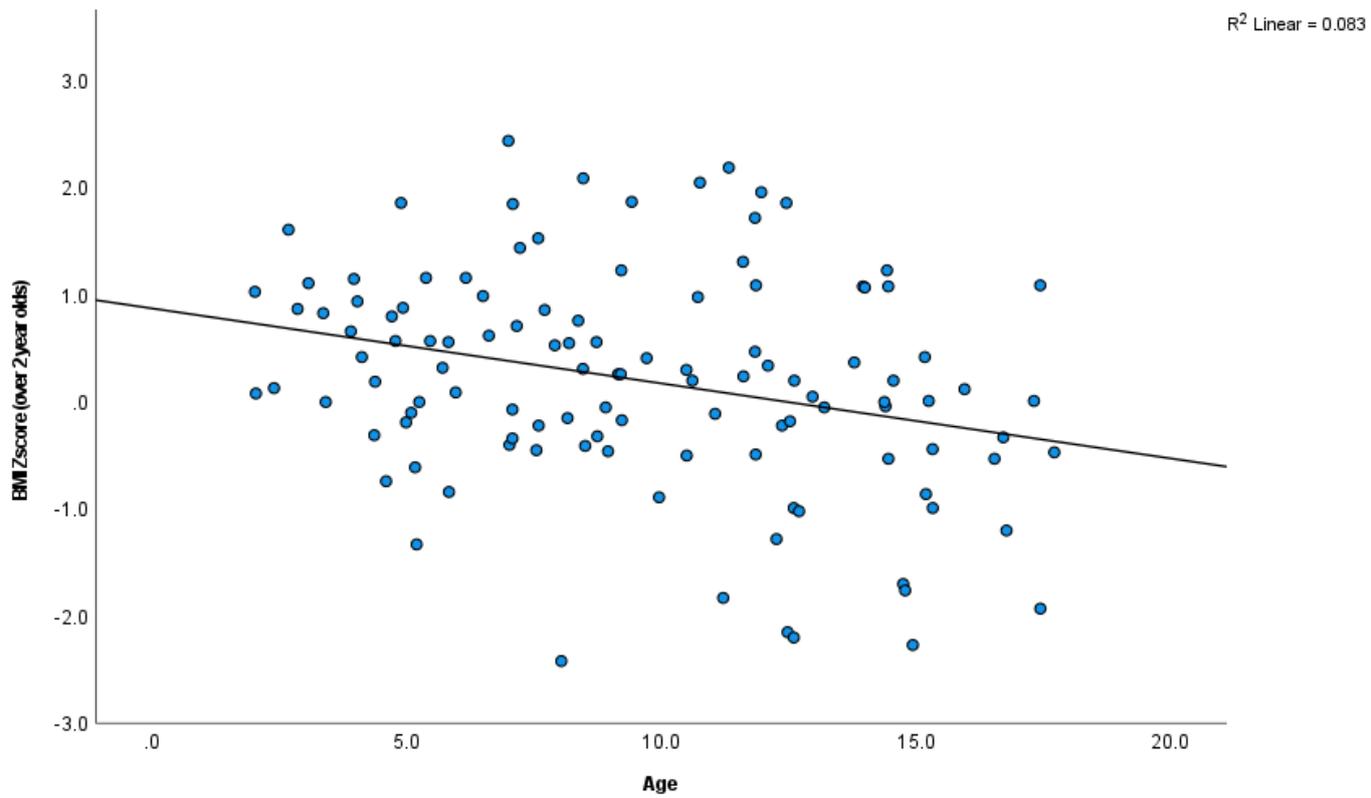
Figure 2.3. Scatterplot of BMI Z score for over 2-year-olds by age.

Negative correlation between body mass index (BMI) Z-scores and age, with fall in BMI with increasing patient age. Pearson correlation: -0.288, 2-tailed significance: 0.002; GEE (n=115) <0.001.

2.4.3 16S rRNA gene sequencing statistics and data analysis

Successful sequencing and data analysis were completed on 136 induced sputum samples out of 141 samples processed. Initial sequencing failed for the five excluded samples and further attempts to re-sequence were not possible due to insufficient sample volumes. A total of 2649426 sequence reads were obtained across the dataset of 136 IS samples; the average number of sequence reads per sample was 19481 (range: 1159-188190).

After normalising to 1000 reads, and excluding any OTUs below 10 reads across the whole group, each sample had a mean of 997 reads (range 955-1000). From 136 IS samples, 272 bacterial operational taxonomic units (OTUs) were identified, which represented 120 genus/family groups.

The top 25 genera were calculated based on the number of reads across the whole dataset for each genus (**Table 2.2.**). Of note, *Prevotella_6*, *Prevotella_7* and *Selenomonas_3* are not subgroups of *Prevotella* and *Selenomonas* respectively. They have been identified as separate genera by the bioinformatics pipeline and databases used.

Samples contained an average (mean) of 14.5 of the top 25 genera (range 5-21). The top five genera identified were *Veillonella*, *Prevotella_7*, *Haemophilus*, *Streptococcus* and *Neisseria*. All five were found in two-thirds of the induced sputum samples, with 97.7% of samples containing at least three of these genera **(Table 2.3)**. The proportion of the top ten genera identified is also shown in **Table 2.3**, with 23% of samples containing all ten genera at varying relative abundance.

Table 2.2. Top 25 genera from the induced sputum samples.

Genera with the highest frequency of reads across the induced sputum samples, with % of samples containing each genus. *Enterobacteriaceae: it was not possible to further classify this family.

Number of reads	Genera	%Samples containing genera
41514	*Veillonella*	98.53
19302	*Prevotella_7*	96.32
16061	*Haemophilus*	78.68
12437	*Streptococcus*	99.26
7775	*Neisseria*	74.26
6328	*Pseudomonas*	59.56
5935	*Enterobacteriaceae**	68.38
5137	*Prevotella*	80.15
2079	*Granulicatella*	83.09
1474	*Actinomyces*	77.94
1386	*Leptotrichia*	51.47
1175	*Burkholderia-Caballeronia-Paraburkholderia*	5.88
1156	*Porphyromonas*	66.18
1137	*Moraxella*	14.71
1048	*Capnocytophaga*	58.09
1028	*Alloprevotella*	61.76
1014	*Staphylococcus*	24.26
917	*Gemella*	66.18
877	*Fusobacterium*	61.03
865	*Selenomonas_3*	46.32
854	*Prevotella_6*	53.68
547	*Megasphaera*	36.03
524	*Rothia*	50
416	*Enterococcus*	2.21
392	*Lautropia*	36.76

Table 2.3. Presence of top 5 genera and top 10 genera across induced sputum samples.

Percentage of IS samples containing the top 5 and top 10 genera, based on those listed in **Table 4.2**.

Number of top 5 genera present in sample	% Samples containing top 5 genera
5 of 5	66.91
4 of 5	15.44
3 of 5	15.44
2 of 5	2.21
1 of 5	0
Number of top 10 genera present in sample	**% Samples containing top 10 genera**
10 of 10	22.79
9 of 10	20.59
8 of 10	25.74
7 of 10	16.17
6 of 10	4.41
5 of 10	4.41
4 of 10	3.68
3 of 10	0
2 of 10	0
1 of 10	0

2.4.4 Cystic fibrosis pathogens

Reflecting previous literature, the following organisms were considered cystic fibrosis airway pathogens: *Haemophilus influenzae*, *Staphylococcus aureus*, methicillin-resistant *Staphylococcus aureus* (MRSA), *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex species, non-tuberculous Mycobacterium species, *Achromobacter xylosoxidans*, *Stenotrophomonas maltophilia*, and *Klebsiella pneumoniae* (Ronchetti et al. 2018b).

Species-level classification is challenging using 16S rRNA gene sequencing analysis, but genera containing the aforementioned CF pathogens were identified. Certain bacterial groups, such as *Enterobacteriaceae*, which includes the pathogen *Klebsiella pneumoniae*, are difficult to distinguish below family level using this approach (Jovel et al. 2016). These genera and the family *Enterobacteriaceae* were considered pathogens for this study.

Pathogen detection threshold was set at >0% relative abundance, and considered clinically relevant at this level. Samples containing pathogens at threshold >0% relative abundance were labelled 'pathogen positive' as previously described in (Weiser et al. 2022). Using the presence of a pathogen at any relative abundance reflected the greater diversity seen in IS samples compared to BAL samples and ensured that no pathogens were missed due to the high diversity.

Table 2.4. shows the average relative abundance of each pathogen across all samples. At least one of the eight pathogens was detected at >0% relative abundance in 129/136 (94.9%) samples. No pathogen was present in all samples.

Haemophilus is the most commonly identified genera, with a mean relative abundance of 11.8%. *Haemophilus*, *Pseudomonas*, *Burkholderia* and the family *Enterobacteriaceae* had the highest relative abundance within individual samples at 86.4%, 97.6%, 96.2% and 71.4% respectively. This suggests species dominance within these patient samples.

This sample set includes data for 13 infants, with children aged 6 months to 1 year. Amongst this group, 15% had evidence of *Haemophilus* dominance (2 children) and 23% had *Veillonella* dominance (3 children), with the remaining 62% having a diverse

community containing a varying relative abundance of pathogens and other relevant genera (**Table 2.5**).

Table 2.4. Percentage (%) relative abundance for seven key cystic fibrosis pathogens and their representative genera.

Enterobacteriaceae: it was not possible to further classify this family.

Relative abundance (%)	*Haemophilus*	*Stenotrophomonas*	*Pseudomonas*	*Staphylococcus*
Mean	11.84	0.04	4.67	0.75
Mode	0	0	0	0
Minimum	0	0	0	0
Maximum	86.40	2.00	97.60	39.10

Relative abundance (%)	*Burkholderia*	*Achromobacter*	*Mycobacteria*	*Enterobacteriaceae**
Mean	0.87	0.001	0.01	0.44
Mode	0	0	0	0
Minimum	0	0	0	0
Maximum	96.20	0.10	1.13	71.37

Table 2.5. Exploration of pathogen dominance in infants with cystic fibrosis

Data from 13 infants (aged 6-12 months). Five children demonstrated pathogen dominance with either *Haemophilus* or *Veillonella*, with the remaining infants having a diverse pathogen community.

Sample code	Pathogen dominant/diverse	% Relative abundance of pathogens
CF48	Diverse	48.6% *Veillonella*, 10.6% *Haemophilus*, 33% *Streptococcus*
CF130	*Haemophilus* dominant	85.9% *Haemophilus*
CF132	Diverse	34.7% *Haemophilus*, 22.6% *Streptococcus*
CF74	*Veillonella* dominant	57% *Veillonella*, 19.3% *Streptococcus*
CF108	Diverse	41% *Veillonella*, 30.8% *Haemophilus*
CF140	*Haemophilus* dominant	33.5% *Haemophilus*, 15.4% *Veillonella*
CF200	Diverse	40% *Veillonella*, 26.4% *Haemophilus*
C201	Diverse	35.5% *Veillonella*, 26.7% *Haemophilus*
CF135	*Veillonella* dominant	60.3% *Veillonella*
CF172	*Veillonella* dominant	56.3% *Veillonella*, 13.2% *Streptococcus*
CF202	Diverse	47% *Moraxella*, 19.2% *Veillonella*, 9% *Haemophilus*
CF121	Diverse	41.4% *Veillonella*, 26.6% *Haemophilus*, 9% *Streptococcus*
CF127	Diverse	22.6% *Veillonella*, 14.1% *Haemophilus*, 13% *Streptococcus*

Percent relative abundance for each cystic fibrosis pathogen is shown in **Figure 2.4**. Haemophilus is present at varying levels of relative abundance across 78.7% of samples. *Staphylococcus* was similarly spread across levels of relative abundance, though it was only identified in 24.2% of samples. *Pseudomonas* and *Enterobacteriaceae* are similarly present across 59.6% and 68.4% of samples respectively. However, in contrast to *Haemophilus*, these genera appear most commonly present at lower levels of percent relative abundance (>0-1 and >1-2.5%), with a small number of samples with high relative abundance indicating pathogen dominance in these patients.

Achromobacter and mycobacteria were present in low levels, being identified in only 1.5 and 1.4% of samples respectively, both at low levels of relative abundance (<2.5%). Similarly, *Stenotrophomonas* was present in 7.3% of samples, again at low relative abundance of 2.5%. *Burkholderia* was identified in 5.8% of samples, but likely presented the dominant pathogen in two of these samples, being present at >10-25% and >50-100% respectively.
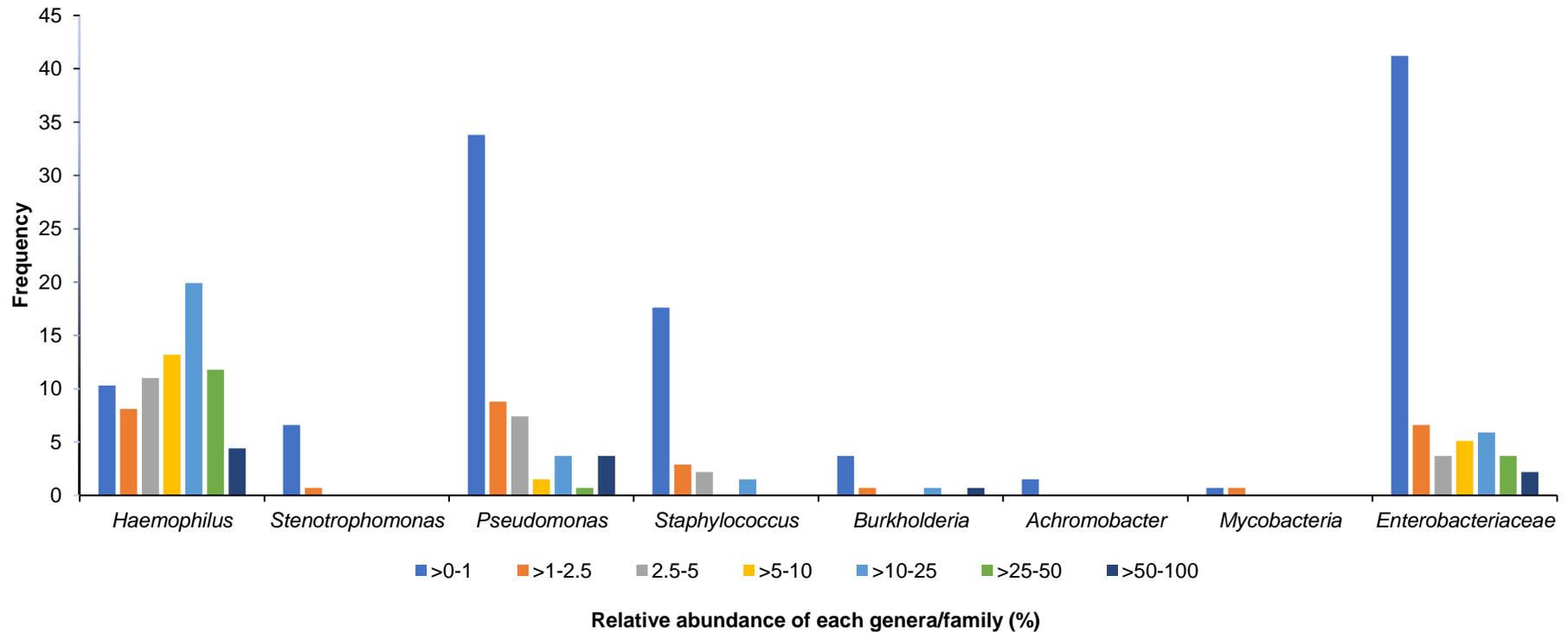
Figure 2.4. Relative abundance (%) of each cystic fibrosis pathogen by genera/family.

Presence of all CF pathogens at >0% relative abundance, with frequency of occurrence across all samples.

2.4.5 Diversity analysis

The Shannon diversity index was calculated for all samples. This provides a measure for species richness and evenness, weighting towards species richness. The Q-Q plot, calculated by the distribution of the sum of the residuals, demonstrates normal distribution of diversity data (**Figure 2.5.**).

The mean diversity index value for induced sputum samples was 1.85 (min. 0.13, max 2.88, SD 0.59), with higher values indicating increasing numbers of species and greater evenness of species among individuals.
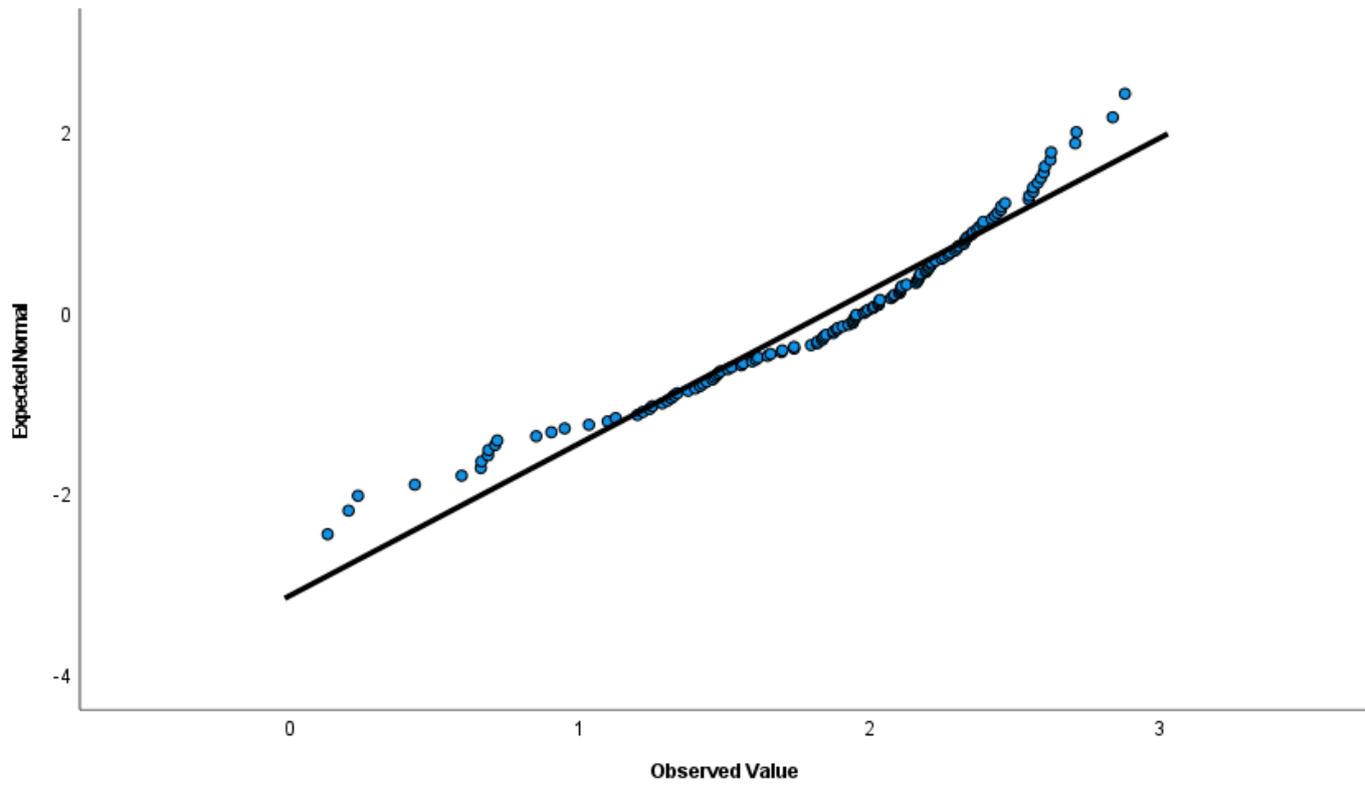
Figure 2.5. Q-Q plot of Shannon diversity demonstrating normal distribution of diversity data.

2.4.6 Effects of clinical features on diversity

Correlations between key clinical features and Shannon diversity measures were explored. Pearson correlations, both 1- and 2-tailed, were undertaken for clinical data which was normally distributed. For clinical data with non-normal distribution, Spearman rank was used. Independent-samples t-test was used for data with binary outcomes, such as treatments received; 'yes' or 'no'. **Table 2.6a** shows the statistical test used for each clinical correlation.

Table 2.6b details the correlation values using Pearson correlation and Spearman rank., with statistically significant values (p-value <0.05) highlighted. All samples were included without control for multiple samples from the same patient at this stage.

Results demonstrate significant negative Pearson correlation between Shannon diversity index and the following clinical measures: age; duration of use of DNase and azithromycin; number of *P. aeruginosa* isolates over the past 3 years and acute intravenous antibiotic therapy measures. The Spearman rank test showed significance negative correlation between Shannon diversity and duration of use of tobramycin. This suggests that diversity reduces as these measures increase; for example, increasing age is associated with reducing Shannon diversity.

Table 2.6c shows comparison of clinical factors with categorical groups ('yes' or 'no') using the Independent samples t-test. This demonstrates significant differences in mean values for those patients with DNase, tobramycin and azithromycin use, as well as those who had received acute intravenous antibiotics within the last 2 years. There were also significant differences for those patients with a history of *P. aeruginosa* isolation. This suggests that exposure to these treatments and *P. aeruginosa* infection was associated with significantly lower mean diversity values than those without exposure.

Table 2.6a. Statistical tests used to establish relationships between clinical features and Shannon Diversity.

Data was assessed on SPSS using visualisation of the Q-Q plot to establish normal distribution. Relationships between the clinical feature and Shannon diversity were tested using either Pearson Correlation, Independent samples T-test or Spearman rank.

	Normal distribution on Q-Q plot	Statistical testing method		
		Pearson Correlation	Independent Samples T-test	Spearman rank
Age	Yes	Yes		
BMI Z score	Yes	Yes		
BMI Z score for over 2 years	Yes	Yes		
FEV1 for age Z score	Yes	Yes		
FEV1 % predicted	Yes	Yes		
FVC Z score	Yes	Yes		
FEV1/FVC Z score	Yes	Yes		
Hypertonic Saline use (Y/N)	No		Yes	
Hypertonic saline duration of use	Yes	Yes		
DNase use	No		Yes	
DNase duration of use	Yes	Yes		
Colomycin use	No		Yes	
Colomycin duration of use	Yes	Yes		
Tobramycin use	No		Yes	
Tobramycin duration of use	No			Yes
Azithromycin use	No		Yes	
Azithromycin duration of use	Yes	Yes		
Flucloxacillin use	No		Yes	
Flucloxacillin duration of use	Yes	Yes		
P. aeruginosa ever isolated	No		Yes	
P. aeruginosa No. of isolates in 3 years	No			
P. aeruginosa No. of years since first isolated	Yes	Yes		

Age at first isolation of *P. aeruginosa*	No			Yes
Acute IV antibiotics in last 2 years	No		Yes	
Acute IV antibiotics, no. of courses	Yes	Yes		
Acute IV antibiotics, no. of days	Yes	Yes		
Acute oral antibiotics in last 2 years	No		Yes	
Acute oral antibiotics, no. of courses	Yes	Yes		
Acute oral antibiotics, no. of courses (including continuous)	Yes	Yes		
Acute oral antibiotics, no. of days	Yes	Yes		
Unwell/wet cough	No		Yes	
Inpatient at time of sample	No		Yes	

Table 2.6b. Correlations between clinical features and Shannon Diversity
using Pearson correlation and Spearman rank.

Statistically significant values (p-value <0.05) highlighted.

	Number of samples	Pearson Correlation	Spearman rank	Significance (2 tailed)
Age	136	-0.207		**0.015**
BMI Z score	136	-0.036		0.675
BMI Z score for over 2 years	115	-0.131		0.163
FEV1 for age Z score	89	0.146		0.172
FEV1 % predicted	89	0.166		0.120
FVC Z score	89	0.132		0.217
FEV1/FVC Z score	89	0.028		0.794
Hypertonic saline duration of use	100	-0.082		0.419
DNase duration of use	103	-0.251		**0.011**
Colomycin duration of use	60	-0.207		0.112
Tobramycin duration of use	20	-0.358	-0.452	**0.045**
Azithromycin duration of use	74	-0.229		**0.050**
Flucloxacillin duration of use	136	-0.44		0.614
P. aeruginosa No. of isolates in 3 years	136	-0.352		**0.000**
P. aeruginosa No. of years since first isolated	96	-0.165		0.109
Age at first isolation of *P. aeruginosa*	98		-0.136	0.182
Acute IV antibiotics, no. of courses	136	-0.210		**0.014**
Acute IV antibiotics, no. of days	136	-0.242		**0.005**
Acute oral antibiotics, no. of courses	133	0.080		0.363
Acute oral antibiotics, no. of courses (including continuous)		0.014		0.870
Acute oral antibiotics, no. of days	136	0.054		0.529

Table 2.6c. Comparison of mean values for relevant clinical features and Shannon Diversity using the Independent samples t-test.

Statistically significant values (p-value <0.05) highlighted.

	Number of samples	Mean (Yes)	Mean (No)	Independent sample t-test	Significance (2 tailed)
Hypertonic Saline use (Y/N)	136	1.863	1.805	-0.503	0.616
DNase use	136	1.786	2.041	2.193	**0.03**
Colomycin use	136	1.758	1.919	1.586	0.115
Tobramycin use	136	1.428	1.920	3.587	**<0.001**
Azithromycin use	136	1.731	1.991	2.604	**0.010**
Flucloxacillin use	136	1.910	1.709	-1.852	0.066
P. aeruginosa ever isolated	136	1.780	2.024	2.197	**0.03**
Acute IV antibiotics in last 2 years	136	1.730	2.053	3.190	**0.002**
Acute oral antibiotics in last 2 years	136	1.844	2.350	0.854	0.395
Unwell/wet cough	136	1.760	1.972	1.893	0.061
Inpatient at time of sample	136	1.871	1.832	-0.376	0.707

2.4.6.1 Controlling for multiple samples from individual patients

To control for multiple samples from individual patients, Generalised Estimating Equations (GEE) were calculated for each clinical correlation. **Table 2.7.** demonstrates the number of subjects and samples provided per subject based on the patient's case number. Selection of the working correlation matrix (WCM) structure with the lowest QIC value, either 'Independent' or 'Unstructured', is detailed for each clinical correlation. Using the lowest QIC value to inform the selected WCM structure within the GEE model aims to appropriately select the best working covariance structure and optimise estimations made about the relationship between covariates and response (diversity measures).

The correlation between clinical data and Shannon diversity index was studied using GEE to control for the effects of multiple samples from the same patient (**Table 2.8.**). Significant results using the Wald Chi Square test are highlighted. A p-value <0.05 was considered significant.

There was a negative correlation between Shannon diversity and the following: age; use of DNase, colomycin, tobramycin and azithromycin; duration of use of DNase, colomycin and tobramycin; *P. aeruginosa* ever isolated and number of isolates within 3 years; acute intravenous and oral antibiotic use in the last two years; number of days treated with acute intravenous antibiotics; being unwell or presence of wet cough at the time of sampling. This suggests reduced diversity index as each of these clinical features is present or increases. There was a positive correlation between Shannon diversity and the following: FEV1 Z-score and FEV1 percent (%) predicted. This suggests increasing diversity with higher FEV1 scores, representative of better lung function.

2.4.6.2 Benjamini-Hochberg procedure for multiple tests

Using the Benjamini-Hochberg procedure for multiple tests, a corrected significance level was calculated at 0.02031. P-values obtained from GEE performed on clinical correlations were adjusted accordingly (*q*-values). **Table 2.8** highlights all p-values which reached the corrected significance level after the Benjamini and Hochberg procedure. Age, DNase duration of use and number of days treated with acute intravenous antibiotics were no longer statistically significant when the Benjamini-Hochberg correction was applied. For completion, the more conservative Bonferroni corrected significance level was 0.00156.

Table 2.7. Selection of Working Correlation Matrix structure.

Calculation of the 'Goodness of Fit' QIC value for each clinical parameter is listed, with the most appropriate WCM structure selected on the basis of the lowest QIC value. (QIC: Quasi-likelihood under Independence Model Criterion; WCM: Working Correlation Matrix)

	Number of samples	Number of subjects	Number of measurements per subject Minimum, Maximum	Correlation matrix dimension	Working correlation matrix		Selected WCM Structure
					Independent Structure Goodness of Fit QIC value	Unstructured Structure Goodness of Fit QIC value	
Age	136	86	1, 6	6	50.012	50.668	Independent
BMI Z score	136	86	1, 6	6	51.368	51.991	Independent
BMI Z score for over 2 years	115	75	1, 6	6	44.653	46.450	Independent
FEV1 Z score	89	58	1, 6	6	37.789	36.616	Unstructured
FEV1 % predicted	89	58	1, 6	6	37.757	36.932	Unstructured
FVC Z score	89	58	1, 6	6	38.279	38.449	Independent
FEV1/FVC Z score	89	58	1, 6	6	38.331	38.869	Independent
Hypertonic Saline use (Y/N)	136	86	1, 6	6	51.907	53.408	Independent
Hypertonic saline duration of use	100	60	1, 5	5	39.052	53.486	Independent
DNase use	136	86	1, 6	6	49.534	49.838	Independent

DNase duration of use (Y/N)	103	64	1, 6	6	40.091	39.687	Unstructured
Colomycin use(Y/N)	136	86	1, 6	6	50.666	49.217	Unstructured
Colomycin duration of use	60	38	1, 6	6	29.382	27.046	Unstructured
Tobramycin use (Y/N)	136	86	1, 6	6	48.226	47.041	Unstructured
Tobramycin duration of use	20	12	1, 4	4	10.099	15.976	Independent
Azithromycin use (Y/N)	136	86	1, 6	6	49.689	47.818	Unstructured
Azithromycin duration of use	74	50	1, 6	6	32.084	30.129	Unstructured
Flucloxacillin use (Y/N)	136	86	1, 6	6	51.496	54.353	Independent
Flucloxacillin duration of use	131	86	1, 6	6	52.440	53.817	Independent
P. aeruginosa ever isolated (Y/N)	136	86	1, 6	6	49.699	49.926	Independent
P. aeruginosa No. of isolates in 3 years	136	86	1, 6	6	43.529	55.610	Independent
P. aeruginosa No. of yrs. since first isolated	96	59	1, 6	6	40.680	40.072	Unstructured
Age at first isolation of *P. aeruginosa*	98	59	1, 6	6	40.457	39.709	Unstructured
Acute IV antibiotics in last 2 years (Y/N)	136	136	1, 6	6	48.224	47.335	Unstructured
Acute IV antibiotics, no. of courses	136	86	1, 6	6	50.863	55.482	Independent

Acute IV antibiotics, no. of days	136	86	1, 6	6	50.159	55.065	Independent
Acute oral antibiotics in last 2 years	136	86	1, 6	6	49.333	51.452	Independent
Acute oral antibiotics, no. of courses	133	84	1, 6	6	50.359	50.454	Independent
Acute oral antibiotics, no. of courses (incl. continuous)	136	86	1, 6	6	50.861	52.593	Independent
Acute oral antibiotics, no. of days	136	86	1, 6	6	51.095	52.514	Independent
Unwell/wet cough	136	86	1, 6	6	49.944	49.551	Unstructured
Inpatient at time of sample	136	86	1, 6	6	51.352	53.470	Independent

Table 2.8. Clinical correlations with Shannon diversity, controlling for multiple samples from individual patients with generalising estimating equations.

* For binary responses where comparison made to 'yes/no treatment' (1) versus 'no/not treated' (0).

Statistically significant values (p-value <0.05) highlighted. ** p-value remains significant with corrected significance level following Benjamini-Hochberg procedure for multiple samples.

| | No. of samples | No. of subjects | Beta coefficient | Std. Error | 95% Wald Confidence Interval | | Hypothesis Test | | |
					Lower	Upper	Wald Chi-Square	df	Significance
Age	136	86	-0.025	0.0112	-0.047	-0.003	4.932	1	**0.026**
BMI Z score	136	86	-0.019	0.0441	-0.106	0.067	0.193	1	0.66
BMI Z score for over 2 years	115	75	-0.078	0.0456	-0.167	0.012	2.895	1	0.089
FEV1 Z score	89	58	0.137	0.0254	0.087	0.187	29.154	1	**<0.001 **
FEV1 % predicted	89	58	0.013	0.0023	0.008	0.017	32.207	1	**<0.001 **
FVC Z score	89	58	0.061	0.0508	-0.038	0.161	1.463	1	0.226

FEV1/FVC Z score	89	58	0.016	0.0565	-0.095	0.127	0.08	1	0.777
Hypertonic Saline use*	136	86	0.058	0.1252	-0.187	0.303	0.214	1	0.644
Hypertonic saline duration of use	100	60	-0.021	0.0244	-0.069	0.027	0.763	1	0.382
DNase use*	136	86	-0.256	0.1066	-0.465	-0.047	5.757	1	**0.016 ****
DNase duration of use	103	64	-0.042	0.0207	-0.082	-0.001	4.094	1	**0.043**
Colomycin use*	136	86	-0.269	0.077	-0.42	-0.118	12.215	1	**<0.001 ****
Colomycin duration of use	60	38	-0.034	0.0088	-0.051	-0.017	15.107	1	**<0.001 ****
Tobramycin use*	136	86	-0.723	0.0924	-0.904	-0.542	61.145	1	**<0.001 ****
Tobramycin duration of use	20	12	-0.113	0.0315	-0.175	-0.052	12.967	1	**<0.001 ****
Azithromycin use*	136	86	-0.328	0.0798	-0.485	-0.172	16.905	1	**<0.001 ****
Azithromycin duration of use	74	50	-0.026	0.0203	-0.066	0.013	1.707	1	0.191
Flucloxacillin use*	136	86	0.201	0.134	-0.061	0.464	2.258	1	0.133
Flucloxacillin duration of use	131	86	-0.006	0.0148	-0.035	0.023	0.176	1	0.675
P. aeruginosa ever isolated*	136	86	-0.245	0.105	-0.45	-0.039	5.43	1	**0.02 ****

P. aeruginosa No. of isolates in 3 years	136	86	-0.048	0.0046	-0.057	-0.039	109.498	1	**<0.001 **
P. aeruginosa No. of yrs. since first isolated	96	59	-0.018	0.0155	-0.048	0.012	1.334	1	0.248
Age at first isolation of *P. aeruginosa*	98	59	0.011	0.0197	-0.028	0.049	0.3	1	0.584
Acute IV antibiotics in last 2 years*	136	86	-0.423	0.0906	-0.6	-0.245	21.783	1	**<0.001 **
Acute IV antibiotics, no. of courses	136	86	-0.074	0.0396	-0.152	0.003	3.525	1	0.06
Acute IV antibiotics, no. of days	136	86	-0.007	0.003	-0.012	-0.001	4.751	1	**0.029**
Acute oral antibiotics in last 2 years*	136	136	-0.507	0.0559	-0.616	-0.397	82.174	1	**<0.001 **
Acute oral antibiotics, no. of courses	133	84	0.012	0.0119	-0.011	0.035	1.007	1	0.316
Acute oral antibiotics, no. of courses including continuous	136	86	0.001	0.0063	-0.011	0.014	0.042	1	0.838

Acute oral antibiotics, no. of days	136	86	0	0.0005	-0.001	0.001	0.485	1	0.486
Unwell/wet cough	136	86	-0.264	0.0819	-0.425	-0.104	10.401	1	**0.001** **
Inpatient at time of sample	136	86	0.039	0.0974	-0.152	0.23	0.159	1	0.69

2.4.6.3 Age, lung function and clinical status at time of sampling

The following graphs show the relationship between Shannon diversity and the clinical features for those factors deemed to have significant correlation through GEE.

Figure 2.6a. shows significant negative correlation between age and diversity, with reduction in diversity index as patient age increases (Pearson correlation (2-tailed) 0.015; GEE (n=86) 0.026).

Figures 2.6b. and **c.** both show significant positive correlation between FEV1 parameters, with increasing diversity with better lung function (FEV1 Z-score: Pearson correlation (2-tailed) 0.172; GEE (n=58) <0.001; FEV1 % predicted: Pearson correlation (2-tailed) 0.120; GEE (n=58) <0.001). However, **Figure 2.6c.** demonstrates that patients with the lowest Shannon diversity (<0.5), suggestive of species dominance, can have variable lung function, from low (32.1%) to within normal range (86.8%). Therefore, lung function as a single clinical marker may be insufficient to identify those children with species dominance.

Lung function parameters have an age-independent association with Shannon diversity. Using the co-variates of age and FEV1 % predicted, FEV1 % predicted continues to show a positive correlation with diversity (GEE (n=58) Age 0.584, FEV1 % predicted <0.001). Similarly, using co-variates of age and FEV1 Z-score, the positive correlation between FEV1 Z-score and diversity is seen independent of age (GEE (n=58) Age 0.599, FEV1 % predicted <0.001). These results show an age-independent association between lung function and Shannon diversity.
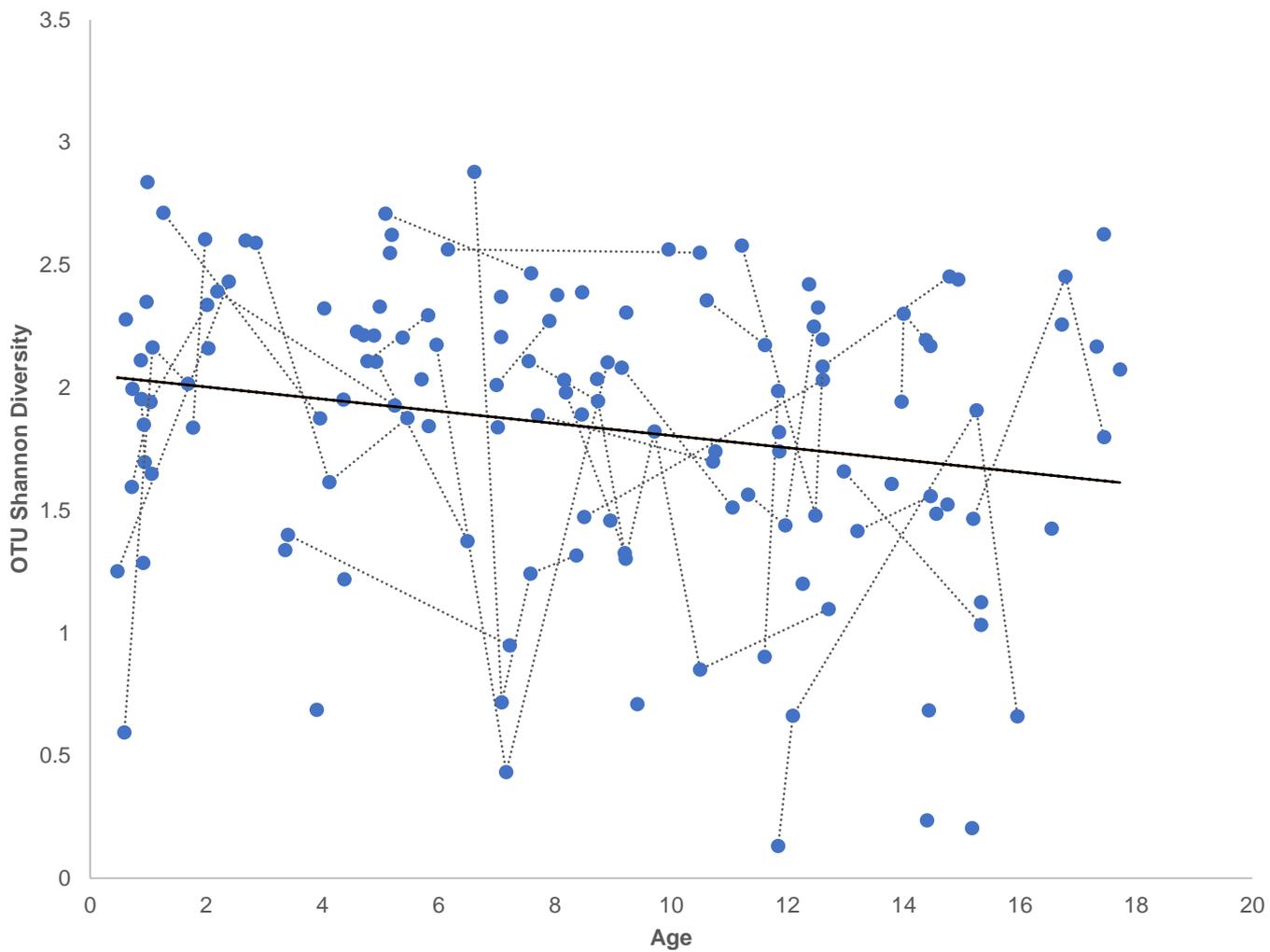
<u>Figure 2.6a. Shannon diversity by age, with best line fit.</u>

136 samples from 86 patients; samples from the same patient are joined by a grey line. Line of best fit for all samples in black: R^2 linear = 0.043. Significant negative correlation between age and diversity, with reduction in diversity index as patient age increases. Pearson correlation (2-tailed) 0.015; GEE (n=86) 0.026.
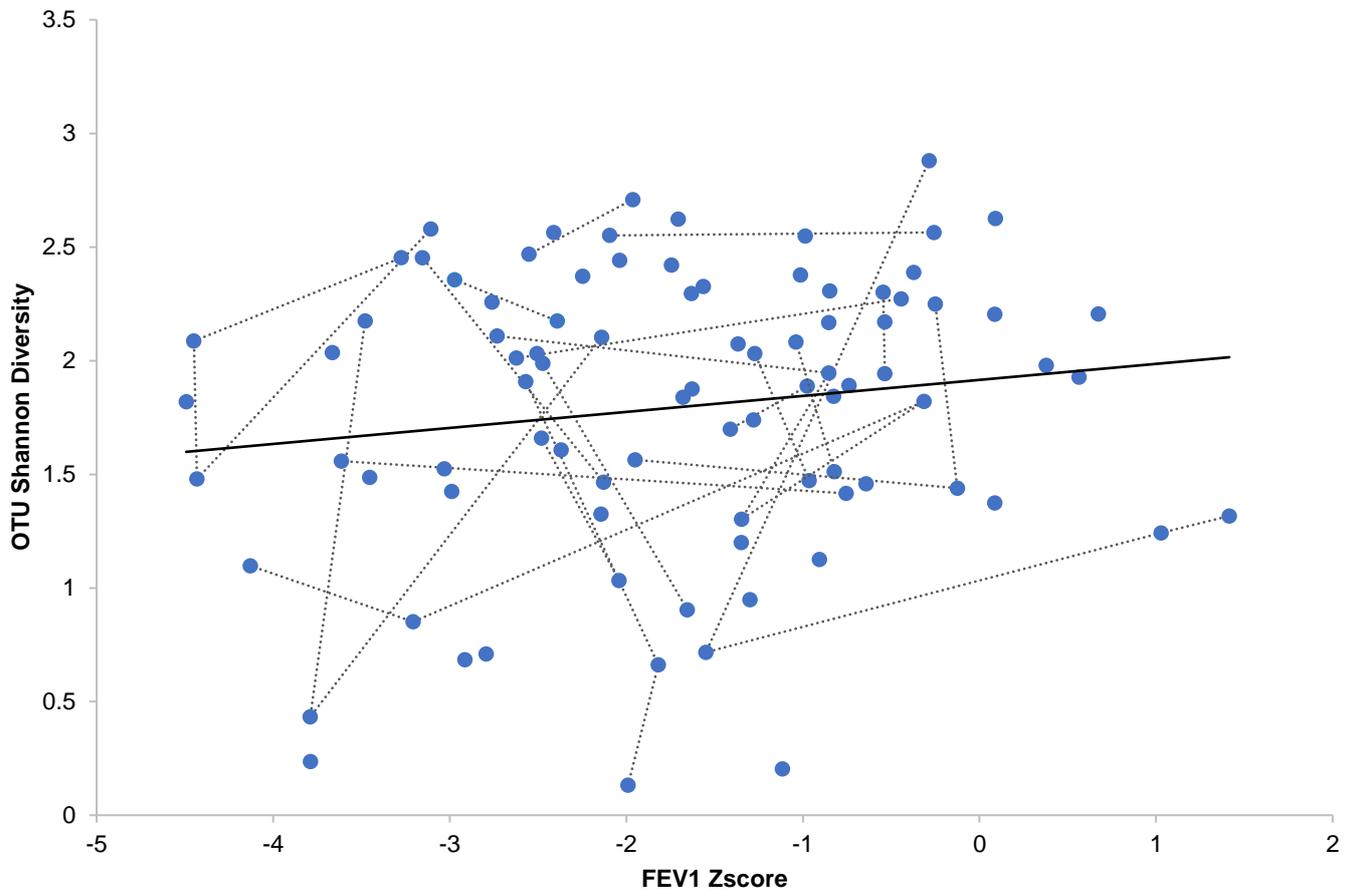
Figure 2.6b. Shannon diversity by FEV1 Z-score, with best line fit.

89 samples from 58 patients; samples from the same patient are joined by a grey line. Line of best fit for all samples in black: R^2 linear = 0.021. Significant positive correlation between FEV1 Zscore, with increasing diversity with better lung function. Pearson correlation (2-tailed) 0.172; GEE (n=58) <0.001.
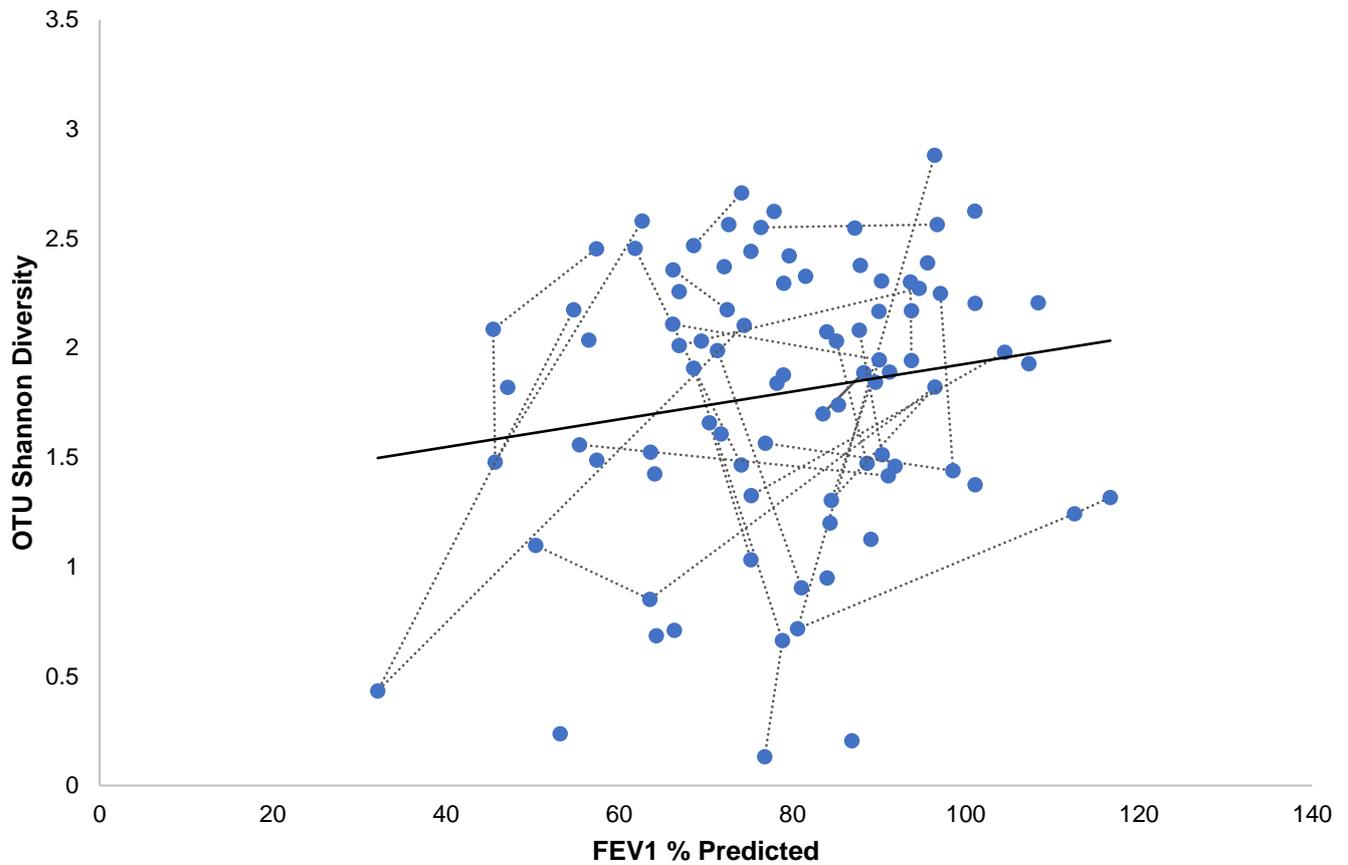
Figure 2.6c. Shannon diversity by FEV1 percent (%) predicted, with best line fit.

89 samples from 58 patients; samples from the same patient are joined by a grey line. Line of best fit for all samples in black: R^2 linear = 0.027. Significant positive correlation between FEV1 % predicted, with increasing diversity with better lung function. Pearson correlation (2-tailed) 0.120; GEE (n=58) <0.001.

2.4.6.4 *Pseudomonas* status

For patients with a positive isolation of *P. aeruginosa* at least once before (72.1% of samples (98 samples; 59 patients), Shannon diversity is significantly lower than for those patients without positive culture (Independent samples t-test (2-tailed) 0.03; GEE (n=86) 0.02).

Figure 2.7a. shows that with increasing numbers of positive isolates, diversity reduces significantly (Pearson correlation (2-tailed) <0.001; GEE (n=86) <0.001). Twenty-two of 86 patients (25.5%) have had three of more isolates of *P. aeruginosa* within the last three years. Amongst those children that have ever isolated *P. aeruginosa*, three or more positive PSA cultures within the last 3 years was associated with significantly lower Shannon diversity (Independent samples t-test (2-tailed) 0.005; GEE (n=59) <0.001) (**Figure 2.7b**).

Patients were considered to have chronic pseudomonas infection if they had isolated mucoid *P. aeruginosa* or had three or more samples within a year showing *P. aeruginosa* positivity. Patients with chronic infection have significantly lower Shannon diversity (Pearson correlation (2-tailed) <0.001; GEE (n=86) <0.001) compared to those who with no history of PSA positivity or acute/intermittent PSA infection only. Chronic PSA infection in those children that had ever isolated *P. aeruginosa*, was also associated with significantly lower Shannon diversity compared to those with acute/intermittent infection only (Independent samples t-test (2-tailed) <0.001; GEE (n=59) <0.001).
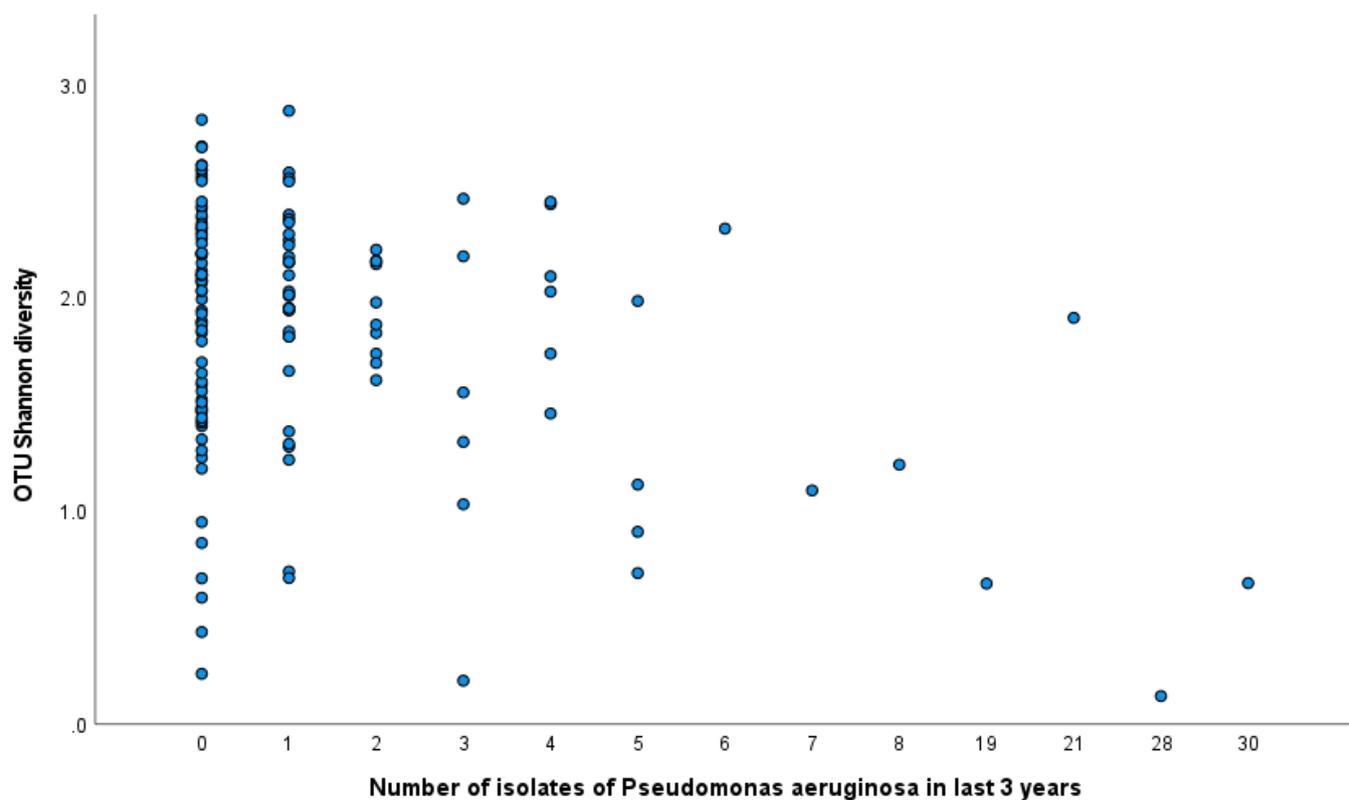
Figure 2.7a. Scatterplot of Shannon Diversity by number of *Pseudomonas aeruginosa* isolates in the last three years.

136 samples from 86 patients; Pearson correlation (2-tailed) <0.001; GEE (n=86) <0.001. Increasing numbers of positive isolates of *P. aeruginosa* is associated with significant reduction in Shannon diversity.
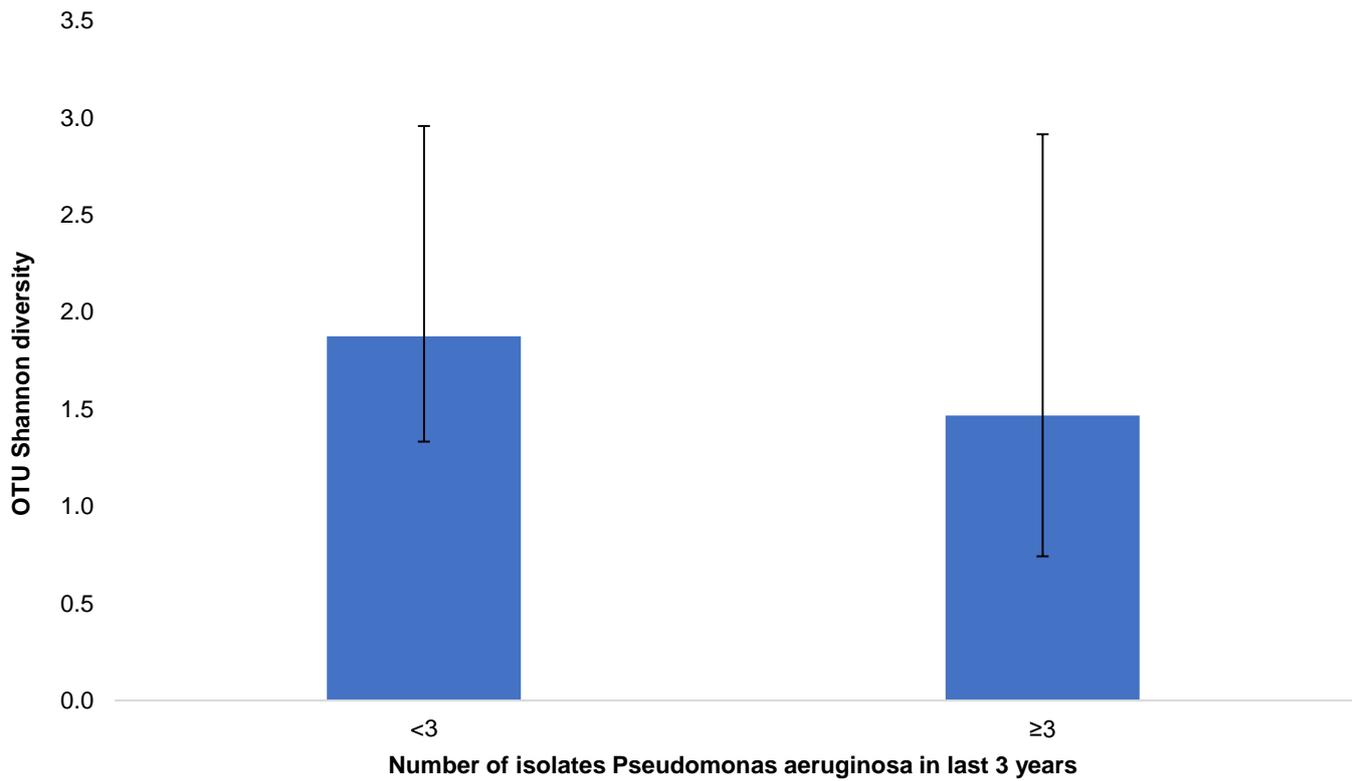
Figure 2.7b. Mean Shannon Diversity by number of isolates in the last three years, amongst children that had isolated *P. aeruginosa* at least once before (+/- 2 SD).

98 samples from 59 patients; Independent samples t-test (2 tailed) 0.005; GEE (n=59) 0.016. Significant reduction in mean diversity seen in those with >3 isolates of *P. aeruginosa* in the last 3 years compared to those with <3 isolates.

2.4.6.5 Oral therapies

All patients received oral antibiotics, either as acute oral courses or as prophylactic therapy. There were statistically significant correlations between oral antimicrobial therapies and Shannon diversity. Diversity was significantly lower in the 55.1% of patients receiving azithromycin prophylaxis (Independent samples t-test (2-tailed) 0.01; GEE (n=86) <0.001).

Diversity is significantly lower in those receiving oral antibiotics within the last two years. However, this result is of limited value as there was only patient that did not receive oral antibiotics (0.7% of the sample group).

2.4.6.6 Nebulised therapies

Correlation between Shannon diversity and the use of DNase, colomycin and tobramycin respectively showed significantly lower diversity index measures for patients on these nebulised therapies following analysis with GEE (**Table 2.8**).

Figures 2.8a, b and **c.** are scatterplots demonstrating the significant negative correlation between duration of nebulised therapy use for DNase, colomycin and tobramycin respectively. For all therapies, increased duration of use is associated with a reduction in the Shannon diversity index.
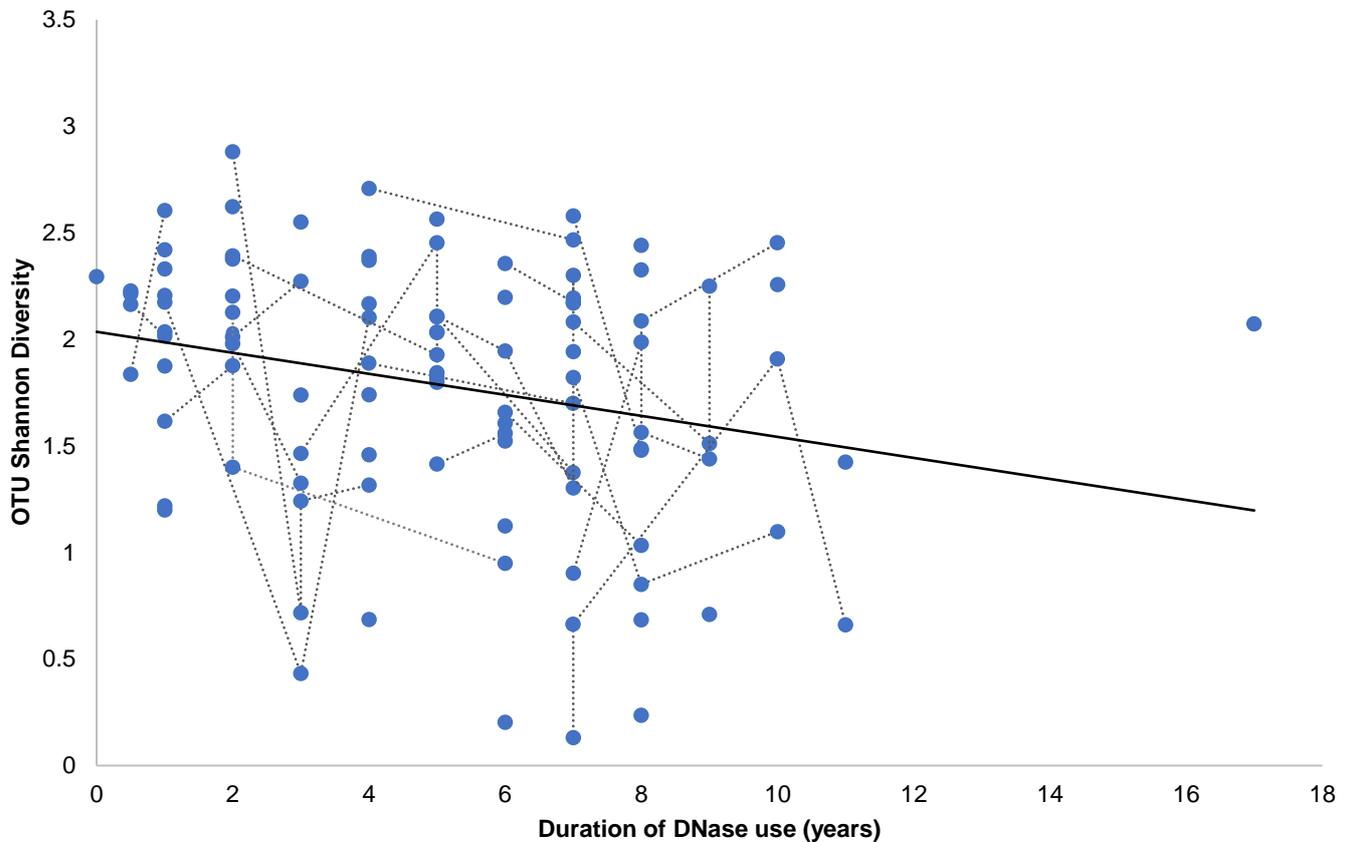
103 samples from 64 patients; samples from the same patient are joined by a grey line. Line of best fit for all samples in black: R^2 linear = 0.063. Pearson correlation (2-tailed) 0.011; GEE (n=86) 0.043. The box shows the number of samples provided by individual patients. Significant negative correlation between duration of DNase use and Shannon diversity.
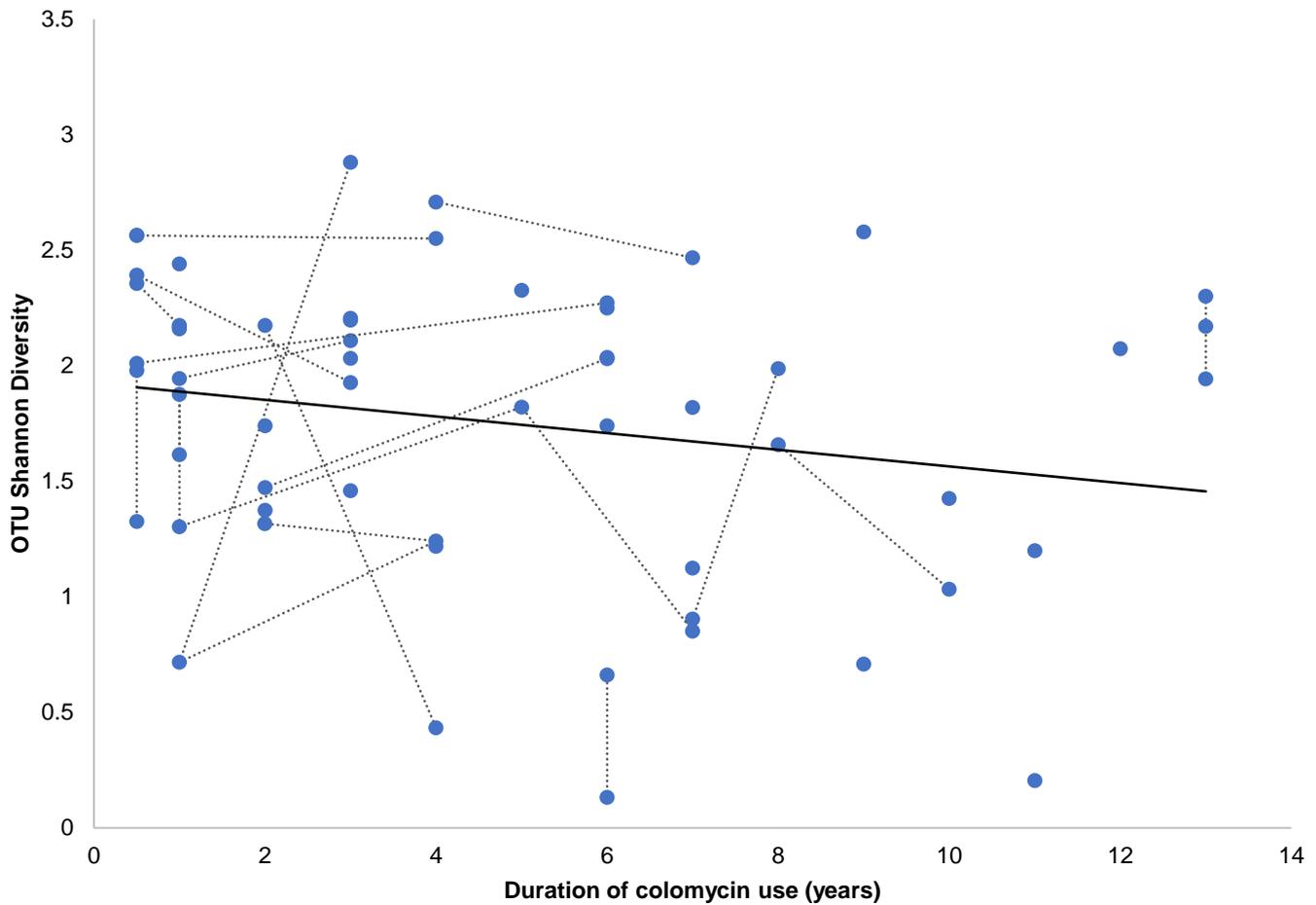
‘Duration of Colomycin’ samples (38 patients, 60 IS samples)
1 sample (23 patients)
2 samples (12 patients)
3 samples (1 patient)
4 samples (1 patient)
6 samples (1 patient)

Figure 2.8b. Scatterplot of Shannon Diversity by duration of use of the nebulised antibiotic therapy, Colomycin in years, with best line fit.

60 samples from 38 patients; samples from the same patient are joined by a grey line. Line of best fit for all samples in black: R^2 linear = 0.043. Pearson correlation (2-tailed) 0.112; GEE (n=86) <0.001. The box shows the number of samples provided by individual patients. Significant negative correlation between duration of colomycin use and Shannon diversity.
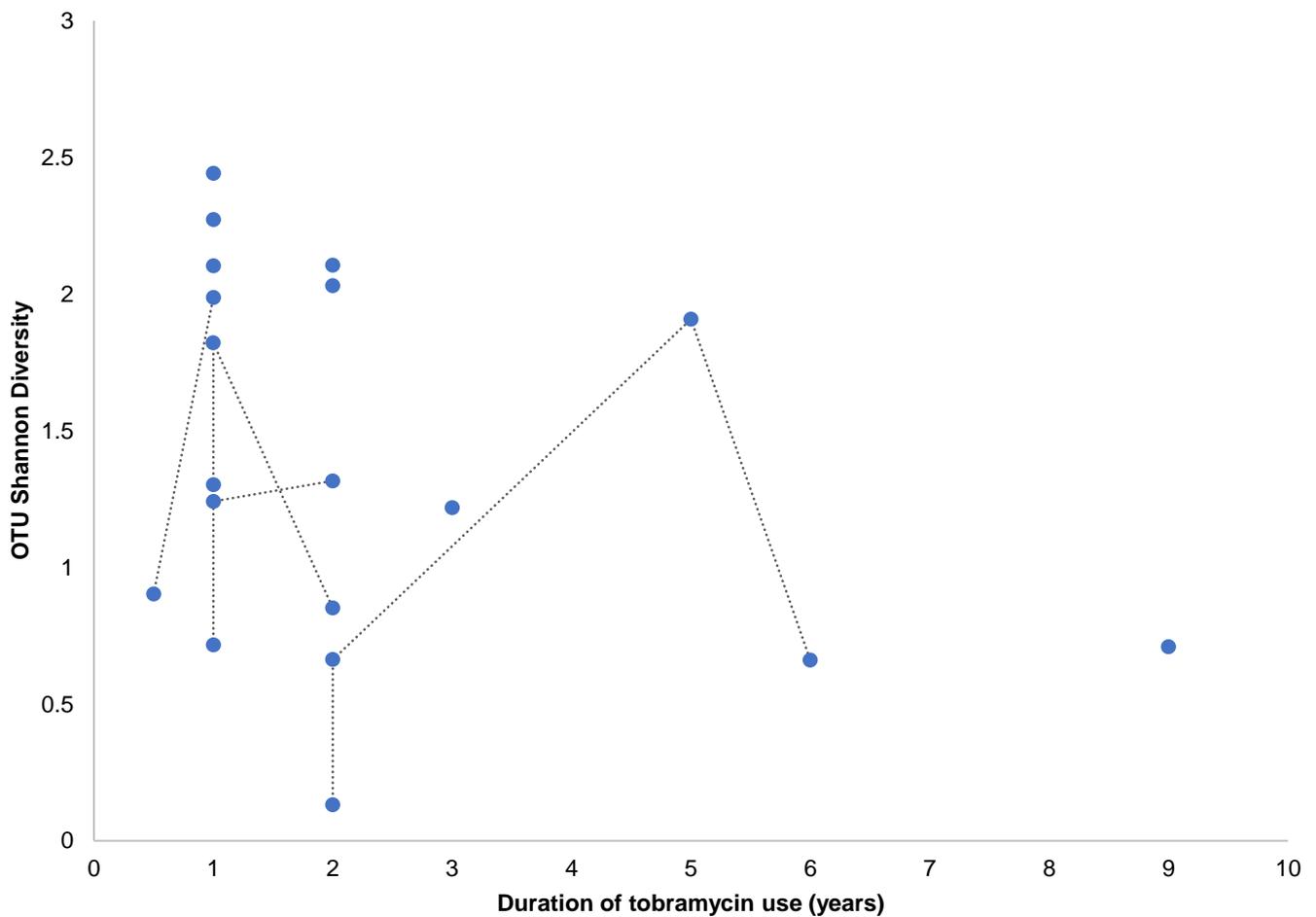
'Duration of Tobramycin' samples (12 patients, 20 IS samples)

1 sample (8 patients)

2 samples (1 patient)

3 samples (2 patients)

4 samples (1 patient)

Figure 2.8c. Scatterplot of Shannon Diversity by duration of use of the nebulised antibiotic therapy, Tobramycin, in years.

20 samples from 12 patients; samples from the same patient are joined by a grey line. Spearman rank (2-tailed) 0.045; GEE (n=86) <0.001. The box shows the number of samples provided by individual patients. Significant negative correlation between duration of tobramycin use and Shannon diversity.

2.4.6.7 Intravenous therapies

There was a statistically significant correlation between use of intravenous antibiotics and Shannon diversity. There was significantly lower diversity seen in the 63.2% of patients that received intravenous antibiotics in the two years prior to sampling. The scatterplot, **Figure 2.9**, demonstrates the reducing diversity index with increasing days of intravenous therapy over the past two years.
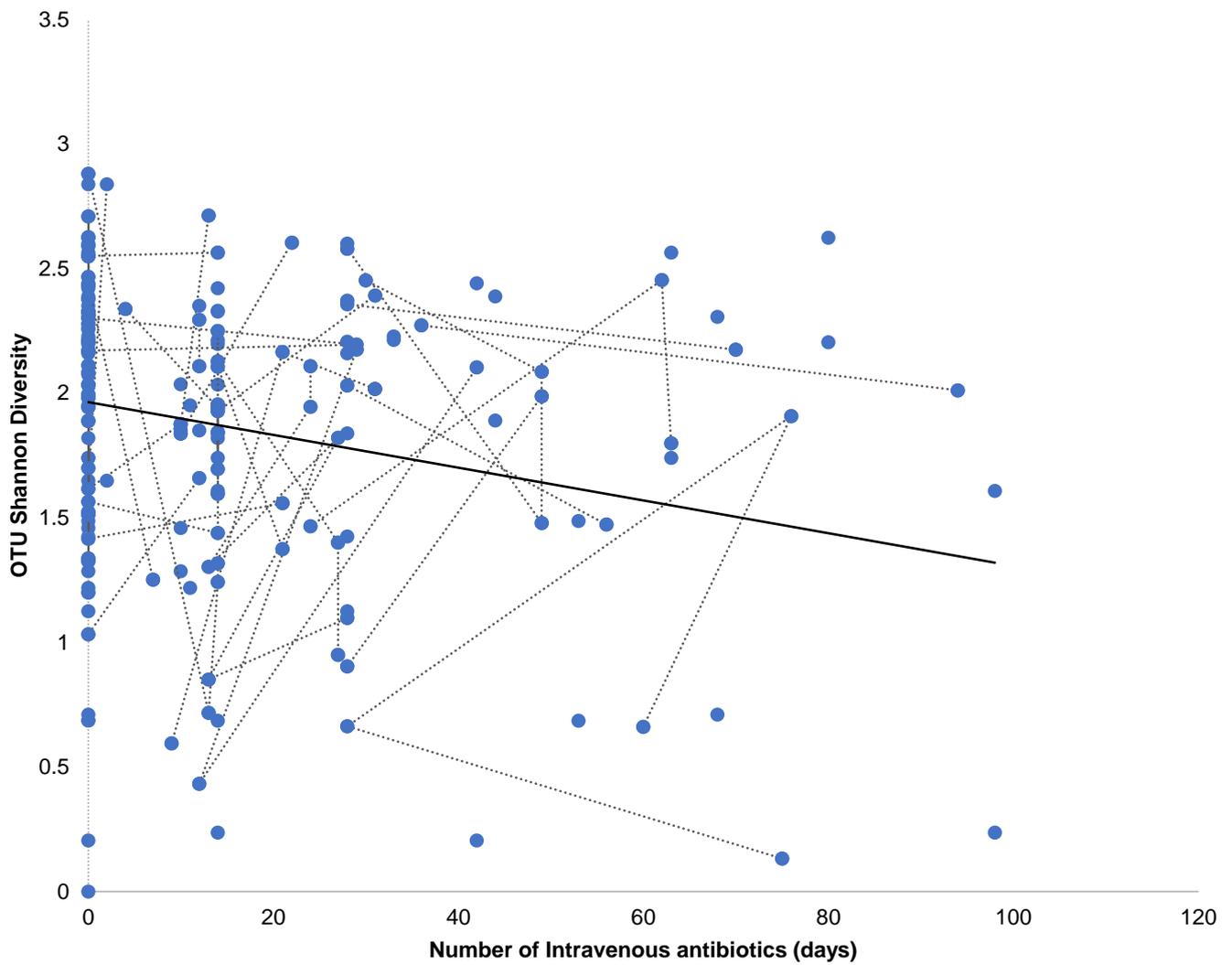
Figure 2.9. Scatterplot of Shannon Diversity by number of intravenous antibiotic days in the last 2 years, with best line fit.

136 samples from 86 patients; samples from the same patient are joined by a grey line. Pearson correlation (2-tailed) 0.005; GEE (n=86) 0.029. Significant negative correlation between duration of intravenous antibiotics received and Shannon diversity.

2.5 Discussion

Chronic pulmonary infection and associated inflammation are the primary cause of morbidity and premature death in people with cystic fibrosis. Culture-based approaches have historically been the only option available for infection surveillance and management, with a focus placed on the testing and identification of a relatively small number of bacterial species associated with CF lung infections (Rogers et al. 2004; Cuthbertson et al. 2020).

More recently, studies using molecular techniques have provided a greater understanding of the complexity of the CF lung microbiome (Rogers et al. 2004). Culture-independent data has enabled exploration of the polymicrobial bacterial community composition within the CF lung. Use of diversity measures, including species richness and evenness, and the division into core and satellite species has expanded our understanding of the lung metacommunity (Van Der Gast et al. 2011). The value of culture-independent methods is clear, but the implementation of this technique is yet to be incorporated into routine clinical practice. In the paediatric population, this is in part due to the difficulty of obtaining valid airway samples. (Weiser et al. 2022) showed IS adequately reflected the microbiota trends seen in BAL samples which are currently considered the gold standard investigation. These results validated the use of IS as a non-invasive technique for children of all ages.

This study uses the full dataset of IS samples obtained through the CF-SpIT trial to explore trends in microbiota diversity according to the patient's clinical status. It represents the largest study to date using exclusively paediatric IS samples. The success of the IS procedure was described by Ronchetti et al. (2018) as 84%, with the procedures well-tolerated by almost all children. The current study demonstrates the ease of undertaking further in-depth analysis using sequencing, with a success rate of 96% of samples.

A clinical database was successfully compiled from hospital records and enabled correlations between clinical measures and diversity data. The patient population is representative of many other CF centres, with similar patterns of lung function decline and falling BMI Z-scores with age (Cystic Fibrosis Trust 2019, 2nd February-a).

Trends in community composition amongst adult patients have been linked to clinical status measures, such as lung function (Cuthbertson et al. 2020) and pulmonary exacerbations (Zemanick et al. 2013). These studies used expectorated sputum and BAL samples respectively to demonstrate these relationships. Microbial diversity trends using CF SpIT IS samples correlated well with existing evidence from more invasive BAL samples in paediatrics and expectorated sputum in adult studies (Harris et al. 2007; Cox et al. 2010; Coburn et al. 2015). This suggests that IS represents a valid alternative for microbiome surveillance for children in whom sputum expectoration is generally not possible and repeated BALs would be unrealistic and otherwise unwarranted.

Results from this study show a clear reduction in diversity with increasing age, reflecting previous trends seen in non-induced sputum datasets. Authors have previously shown that, as with 'healthy' infants, children with CF demonstrate respiratory microbiota diversification initially, with diversity evolving to pathogen-centred communities with increasing age (Cox et al. 2010; de Koff et al. 2016).

The current study reported presence/absence of genera at >0% relative abundance, based on the higher diversity seen in IS samples compared to BAL samples (Weiser et al. 2022). Commonly seen genera included *Veillonella*, *Prevotella*, *Streptococcus*, *Haemophilus*, *Granulicatella*, *Actinomyces and Neisseria*, all of which were present in >70% of samples. These results from IS samples reflect similar findings reported using expectorated sputum samples (Coburn et al. 2015).

Altered patterns of microbiota development have been reported in CF infants as early as the first month of life (Prevaes et al. 2016). The decrease in diversity from adolescence through to adulthood occurs alongside the long-recognised decline in lung function (Caverly et al. 2015; Linnane et al. 2021). Differences between microbial communities in CF infants and healthy controls even before the introduction of antibiotics or pulmonary exacerbations might suggest that these are inherent to lung disease in CF (de Koff et al. 2016). The current study provides data for 13 children with CF aged 6 months to 1 year, and 62% have a diverse community. In the

remaining infants, there was evidence of *Haemophilus* and *Veillonella* dominance, similar to that seen in previous studies (Coburn et al. 2015; Prevaes et al. 2016). With increasing age, studies have reported clear evidence of increasing relative abundance of 'typical CF pathogens' such as *Pseudomonas*, *Staphylococcus* and *Haemophilus* in older children (Linnane et al. 2021). Within this study, there was a high detection of CF pathogens, with 94.9% of IS samples having at least one pathogen present. Similar to the paediatric samples from Coburn et al.'s (2015) study, *Haemophilus* was the most commonly identified genera. Other CF pathogens were present at variable degrees of relative abundance, but there were also a number of patients with clear pathogen dominance.

Interestingly, this study has shown the high numbers of children with at least one previous positive airway culture of *P. aeruginosa*, with significantly lower diversity in these patients compared to those with negative cultures for *P. aeruginosa*. In particular, of those that had isolated *P. aeruginosa* previously, the presence of chronic pseudomonas infection was associated with lower diversity. This suggests that the microbiota might alter according to acute versus chronic infection, implying that early effective treatment and eradication of *Pseudomonas* infection may help to preserve diversity within the CF lung.

This study showed statistically significant correlations between decreasing FEV1 Z-score and FEV1 % predicted and a reduction in Shannon diversity index. Given this was an exclusively paediatric cohort with most children having normal lung function (mean FEV1 of 79% predicted), it is interesting to note that trends in decreasing diversity with declining lung function are significant enough to show correlation. Numerous studies have found correlations between reducing microbial diversity, increasing age and declining lung function (Cox et al. 2010; Acosta et al. 2017). Results have demonstrated a reduction in the bacterial community's taxonomic richness, evenness and overall diversity with increasing age (Webb et al. 2008; Cox et al. 2010). Coburn et al. (2015) showed similar trends in adult patients, with lower microbiota diversity in advancing lung disease. However, unlike the current study, they did not demonstrate this trend in their paediatric cohort and proposed this was most likely due to the less severe lung disease typically seen in children (Coburn et al. 2015).

The well-described relationship between lung function and diversity has led to the proposal that diversity may be a useful biomarker for disease prognostication. Acosta

et al. (2018) reported that CF microbial diversity in 104 young adults was predictive of disproportionate lung function decline and progression to end-stage lung disease (defined as death/transplantation <25 years) over the ensuing 5 years. The current data showed an age-independent association between lung function and diversity using IS samples. Decline in lung function is associated with a fall in diversity measures, regardless of age. This was an exciting result as it suggests that, even in a paediatric population where lung function is generally well-maintained, there is sufficient association between diversity and this well-established marker of disease severity demonstrated using this sampling technique.

There have been significant advances in CF management and patient outcomes, predominantly due to pulmonary and nutritional therapies (West and Flume 2018). However, disease-modifying therapies may not only improve disease prognosis, but may also support the evolution of an altered microbiome. The precise impact of these therapies on the microbiota in still under investigation and many aspects remain unanswered.

Antibiotics are key treatments used to prevent and treat recurrent pulmonary exacerbations. It is thought that many exacerbations occur following disturbance of a previously stable chronic bacterial infection, though it is likely that viruses and newly-acquired bacterial infections also play a role (Stokell et al. 2015). Within our patient group, almost two thirds of children were unwell or had a wet cough at the time of sampling, suggestive of a pulmonary exacerbation. There was a significant reduction in diversity in the 'unwell' group compared to those that were well or cough-free. Almost all patients within the current study had received acute courses of oral antibiotics within the preceding two years and therefore it is difficult to draw conclusions regarding the direct effect of these courses on diversity. However, considering the effect of acute intravenous antibiotic courses, there was clear evidence of a reduction in diversity for children that had received at least one course in the last 2 years. Shannon diversity was also negatively correlated with increasing numbers of IV antibiotics days, suggesting that perhaps increasing numbers of antibiotic courses have a cumulative effect on diversity. However, this may simply reflect the severity of the child's overall clinical status in terms of increasing age, reducing lung function and/or increasing pulmonary exacerbations.

Repeated antibiotic courses may alter the healthy commensals within the respiratory microbiome, with evidence showing clear alteration in the abundance of upper

respiratory tract commensals following perturbation by antibiotics. Prevaes et al. (2016) demonstrated the effect of first antibiotic treatment in their infant group causing increased potential pathogens and reduced 'healthy commensals' for a number of months after therapy. Linnane et al. (2021) showed no significant correlations between recent oral or intravenous antibiotic use and alpha diversity measures. Acosta et al. (2017) performed a retrospective analysis to explore the impact of chronic therapies on the CF microbiota when compared to those samples without exposure (Acosta et al. 2017). They showed significant differences for patients treated with inhaled colistin and tobramycin and oral ciprofloxacin, but no significant changes in community composition for those exposed to inhaled DNase, inhaled aztreonam and oral azithromycin (Acosta et al. 2017; Acosta et al. 2018). Linnane et al. (2021) showed no association between prophylactic flucloxacillin or long-term azithromycin treatment and alpha diversity measures.

The current study shows clear negative correlations between nebulised DNase, colomycin and tobramycin and diversity, but no evidence of correlation with hypertonic saline or oral fluxcloxacillin use. In contrast to Acosta et al.'s (2017) study, there was also a negative correlation between azithromycin use and diversity measures. Considering the use of nebulised antibiotics is typically considered for children with recurrent positive cultures, most commonly for *P. aeruginosa*, it is possible that the lower diversity within these groups is a reflection of chronic *P. aeruginosa* infection and evolving pathogen dominance. Alternatively, it may be that these treatments are also affecting other members of the airway community. It is difficult to infer causality from this data, and further longitudinal studies would help address these questions.

Though many of the findings are similar to previously published studies, direct comparison with this dataset may not be appropriate, as these used alternative airway sampling methods such as BAL and nasopharyngeal swabs. However, data comparing matched IS and BAL samples from the CF SpIT trial showed a capture rate of 50-80% by IS samples, supporting the alignment of these sampling methods (Weiser et al. 2022). The technique employed for IS Spit clearly different from other lower airway sampling methods and it is likely that sputum induction captures a different niche of the lower airway microbiome from BAL (Forton 2015). Blau et al. (2014) recognised the higher prevalence and bacterial load of upper respiratory tract flora in IS compared to BAL. The authors suggest this is because sputum induction is more likely to be contaminated by upper airway secretions during the procedure. It

is certainly possible that this may be reflected in the diverse genera seen in many of the current study's samples, but the similarities in trends seen in previous studies using alternative airway sampling methods make this potentially less relevant.

As with all microbiota studies, it is important to recognise that culture-independent technologies can detect bacterial DNA at extremely low concentrations and DNA from known CF pathogens may be present when the culture is negative (ven Belkum et al. 2000; Flight et al. 2015). This study had a high prevalence of CF pathogens, with at least one pathogen identified at >0% relative abundance in 94.9% of samples. Though detection of low concentrations of DNA may represent early infection or a low-level bystander, this concept raises the question of how we should best define a clinically-relevant infection using such sensitive technologies (Weiser et al. 2022). Given the relatively low biomass obtained from respiratory samples and the associated difficulties with reliable molecular detection, it is important to consider the risk of low biomass contamination leading to inaccurate conclusions regarding the presence/absence of pathogens (Davis et al. 2018). Mock communities were utilised for both 16S rRNA gene bacterial diversity sequencing runs and demonstrated good correlation with the relative abundances of taxa within the test samples. This provided reassurance that the study results offered accurate microbiota profiling (Weiser et al. 2022). Comprehensive determination of 'low biomass' status was not possible for the 136 IS samples analysed in this study, as qPCR data were not available. However, data comparing IS and BAL samples, incorporating 30 IS samples included within this study, showed that patterns were replicated in multiple samples from the same individual, supporting a true biological signal (Weiser et al. 2022). Further studies are required to understand the relevance of low-level pathogen detection and provide guidance to clinicians regarding the level at which treatment should be considered.

This study included a mixture of stable and potentially exacerbating patients. This reflects the original CF-SpIT study design, which included BAL sampling. From an ethical perspective, performing a bronchoscopy would only be appropriate if there was a clinical need for this invasive procedure. Therefore, for most children in whom a bronchoscopy (and the concurrent IS sample) is performed, there will have been prior clinical concerns, such as recurrent exacerbations or current significant illness not responding to standard therapies. It is important to recognise this as previous studies have shown differing results in terms of correlation between antibiotic use and diversity depending on the clinical status of their patient population at the time of sampling (Zemanick et al. 2017; Linnane et al. 2021).

There were a number of patients that provided multiple IS samples during the study period. Statistical methods (GEE) were used to control for their individual contributions to the study. Linnane et al. (2021) had a small sub-set of sequential annual BAL samples. Interestingly, they demonstrated variable diversity results at different time points for a single patient. Therefore, it would have been useful to further explore the data for our patients providing multiple samples and consider whether similar patterns could also be seen within this study. Further longitudinal studies would be helpful (Goodrich et al. 2014) to expand this dataset further.

Controlling for multiple testing using the Benjamini and Hochberg procedure was undertaken as part of the statistical analysis. This was important in recognition that performing multiple correlation tests may result in false positive correlations being identified. However, most of the existing significant correlations remained statistically significant using this method, despite reducing the corrected significance level to a p-value of 0.02031. The Benjamini and Hochberg procedure decreases the false discovery rate and therefore reduces type I errors. It is considered less conservative than the Bonferroni correction, which was also provided, but preferable as it is less likely to increase type II errors.

Despite the limitations reported above, this study also has many clear strengths. It provides a large dataset of IS samples for microbiota analysis. The sequencing success rate was excellent and significantly better than many previous studies (Zemanick et al. 2017; An et al. 2018). The clinical database was collated using multiple patient sources to ensure complete data collection. The approach to sampling, DNA extraction and microbiota analysis was in line with previously published, robust microbiome studies (Zemanick et al. 2017; Linnane et al. 2021).

The CF-SpIT study validated sputum induction as a non-invasive, well-tolerated reliable tool for assessing the conventional microbiology of the lower airway (Ronchetti et al. 2018b). As a result, many centres in the UK are now performing sputum induction routinely, with centres in Europe, the USA and Australia also incorporating this technique into routine practice. The benefit for children has been well-described, but there is also increasing interest amongst adult physicians whose patients are no longer spontaneously expectorating due to effective modulator therapy. A follow-on study by (Weiser et al. 2022) demonstrated that IS could also provide adequate sampling for culture-independent microbiota analysis. The current work has built on these studies, providing the largest dataset of induced sputum

samples in children with CF. Data has clearly demonstrated that diversity output from IS samples aligns well with published literature, with similar trends in clinical correlations seen in BAL and expectorated sputum samples.

As a non-invasive and effective method, IS provides the potential for frequent serial sampling from children of all ages. Monitoring changes in diversity over time may represent a new tool for disease monitoring, like FEV1 in current practice, and eventually could be considered part of routine assessment for all children. This technique is ideally placed for use in longitudinal studies of the lower airway microbiota. Its use could be wide-reaching, including studying microbiota evolution in large paediatric datasets, introduction of microbiota profiling and prognosticating for children with CF, and potentially incorporation into new clinical trial protocols as a mechanism for monitoring responses to novel therapies.

This study has demonstrated clear correlations between patient's clinical features and Shannon diversity. The existing dataset could be used to explore other diversity measures, including inter-patient variability. In particular, it would be beneficial to consider the genera present using hierarchical clustering and whether patients can be allocated to clearly-defined clades. Using these clades, it would be interesting to compare diversity measures and clinical features to see if these are affected. This could be undertaken using simply pathogen-diverse versus pathogen-dominant clades, or potentially using CF pathogen-specific clades for comparison. Furthermore, previous studies have explored the relationship between inflammatory markers and diversity in BAL samples (Linnane et al. 2021) and it would be useful to review this using the current IS samples.

Longitudinal studies using this patient group, plus ideally additional patients from other CF centres, would be useful to expand on trends seen in this study. Key questions include what level of relative abundance should be considered indicative of infection and is a pathogen presence/absence precise enough to inform clinical decisions? Should we have different thresholds for samples obtained from BAL versus IS? Considering sputum induction likely represents a different lower airway niche from BAL, should it be used as an additional investigation rather than a replacement of BAL sampling? What effects are patient therapies, such as CF modulators and novel treatments like OligoG CF-5/20, having on the lung microbiome and do these effects matter? Given the differences between the paediatric and adult CF microbiome, do we have a window of opportunity to intervene (de Koff et al. 2016)

and could we use the microbiota to guide us? Answering these questions might also improve our understanding of patterns of infection, enable earlier disease intervention, and improve outcomes for children in later life.

2.6 Conclusion

It is likely that establishing the lung microbiota is just one piece of a complicated puzzle in understanding cystic fibrosis. This research demonstrates the use of a non-invasive airway sampling method, with excellent sequencing success and trends in diversity and clinical correlations which match those of previously published studies performed using invasive techniques. The opportunities that IS microbiota analysis may provide for improving the knowledge-base, treatment options and overall health for children with CF are exciting and warrant further research.

If, as a scientific and medical community, we agree that microbiota analysis has a significant part to play in the clinical care of our patients, we need a less invasive method to obtain this data. Induced sputum sampling should be considered a strong contender.

Chapter Three

Mucin structural interactions with OligoG CF-5/20 in cystic fibrosis sputum

3.1 Introduction

Cystic fibrosis is associated with increased airway secretions and poor mucus clearance. There appear to be multiple factors within CF sputum and the CF airways which contribute to the overall disease phenotype.

The physical properties of CF sputum are affected by the polymeric mucins, as well as the cell surface-bound mucins, host DNA and actin, inflammatory cells, bacteria and viruses (Voynow and Rubin 2009; Ehre et al. 2014). Alterations in mucin terminal carbohydrate moieties are associated with respiratory disease. CF mucins have been shown to possess increased levels of fucose, galactose, GlcNAc (Bhat et al. 1996), sialic acid and sulphate (Davril et al. 1999; King and Rubin 2002). Increases in glycosylation and branching in CF mucins result in a higher tendency to 'gel' and obstruct mucociliary transport *in vivo* (Bhat et al. 1996).

3.1.1 Mucins and therapeutics

Mucins play an important role in the innate immune system, providing a physical barrier to chemical, enzymatic and mechanical insults (Linden et al. 2008). Though key in defence, mucins may actually hinder drug delivery by binding to inhaled agents and removing them through mucociliary clearance (Widdicombe 1997; Taherali et al. 2018). Methods to disrupt these muco-adhesive interactions, allowing crossing of the mucin-protective layer of the lung, would greatly improve drug delivery (Bansil and Turner 2006). Muco-adhesive systems aim to prolong contact of the required drug at the target site. Mechanisms to enable this require interaction with the mucin glycoprotein through disulphide bonding, hydrogen bonding, and electrostatic and/or hydrophobic interactions (Taherali et al. 2018).

The choice of the mucin experimental system used to test therapeutics is important. Numerous authors have demonstrated the limitations of commercially-produced mucins (Wagner et al. 2018), which are typically reconstituted mucin gels. Studies have shown that such gels have significant structural differences compared to native mucin, and though they often have less heterogeneity, the pattern of diffusion and interactions with key ions such as calcium do not reflect naturally-occurring mucus (Wagner et al. 2018). Due to these limitations, this study incorporated the use of

human sputum, which it is hoped will more closely represent the mucus seen *in vitro* within the CF lung environment.

3.1.2 Alginate oligomers

Previous *in vitro* studies of OligoG CF-5/20 have demonstrated its ability to alter the viscoelastic properties of mucin/alginate gels, mucin/DNA gels and CF sputum (Nordgård and Draget 2011; Pritchard et al. 2016b). OligoG CF-5/20 interacts with respiratory mucins, inducing alterations in mucin surface charge and porosity (Pritchard et al. 2016b). High MW alginates, as secreted by virulent CF respiratory bacteria such as a mucoid *P. aeruginosa*, increase the elasticity and viscosity of mucus, whilst conversely, low Mw alginates have been shown to reduce the bulk elasticity and viscosity of mucus (Sletmoen et al. 2012). Further studies are indicated to better understand how OligoG CF5/20 is able to alter CF sputum viscoelasticity.

3.1.3 Fourier transform infrared spectroscopy (FTIR)

3.1.3.1 Using FTIR for sputum analysis

Sputum is a complex biological material containing numerous biochemical substances. FTIR should in theory be able to create a distinct infrared 'fingerprint' and enable visualisation of the chemical bonds present within sputum (Whiteman et al. 2008). Mucus, and its associated glycans and glycosylated structures, predominate the sputum spectrum, facilitating effective analysis using FTIR. By accessing spectral libraries, the associated biological components can then be identified and comparisons made between different sputum samples.

Results suggest FTIR has high sensitivity and specificity in diagnosing disease using bronchoalveolar lavage and induced sputum samples (Lewis et al. 2010). Whiteman et al. (2008) used FTIR to compare the spectral profiles of induced sputum samples between patients with chronic obstructive pulmonary disease (COPD) and healthy volunteers. They demonstrated clear differences, with shifts in peak position and altered intensity, particularly in the amide and glycogen rich regions (Whiteman et al. 2008). Such differences were hypothesised to be related to airway inflammation, bacterial presence and pharmacological treatments for COPD (Whiteman et al. 2008).

Lewis et al. (2010) demonstrated significant changes in wavenumber in sputum samples which enabled discrimination between lung cancer sputum and healthy control sputum. In particular, increased absorbance was noted again in the glycogen rich region, amide I and II regions and with bonds associated with nucleic acids (Lewis et al. 2010). The authors proposed these changes may be directly related to structural changes caused by lung cancer, and could enable use of FTIR as a diagnostic tool for lung cancer.

Lewis et al. (2013) also explored the secondary structure of mucin using Attenuated total reflectance (ATR)-FTIR analysis of bovine salivary mucin. Results showed the combination of secondary structures in the protein backbone of mucins, including β-sheet, β-turns and random coil, and their associated wavenumbers (Lewis et al. 2013b). The spectral library created described a range of mucin sugars, with different glycosylation patterns, as well as Lewis antigens (Lewis et al. 2013a). These studies support the use of FTIR as a useful tool for characterising alterations in mucin structures following exposure to respiratory therapeutics.

3.2 Aims

This study uses Attenuated Total Reflectance Fourier-transform infrared spectroscopy (ATR-FTIR) to examine direct physicochemical interactions of OligoG CF-5/20 with the mucin component of CF sputum, in order to better understand how it is able to alter CF sputum viscoelasticity.

Using mucin infrared (IR) profiles from a published spectral library (Lewis et al. 2013a), this study aims to identify key IR wavelengths and use these to describe the structural changes and electrostatic interactions that occur within mucin glycan moieties and peptide backbone following *ex vivo* OligoG CF-5/20 treatment of CF sputum samples.

3.3 Materials and methods

3.3.1 Patient samples

Induced sputum samples were collected from paediatric patients with CF taking part in the CF SpIT trial. As previously described, informed written consent was obtained (CF Sputum Induction Trial [CF-SpIT] UKCRN: 14615) and study approval from the Cardiff and Vale Research Ethics Committee was established. Samples were collected by expectoration, after which they were stored by freezing at -80°C. Prior to subsequent use, sputum samples were defrosted overnight at 4°C.

3.3.2 Alginate oligosaccharide (OligoG CF-5/20)

OligoG CF-5/20 was produced from the stem of brown seaweed *Laminaria hyperborea* and provided by AlgiPharma AS (Sandvika, Norway). OligoG CF-5/20 has a high guluronate content (>85%) and a mean degree of polymerisation (Dpn) of 16. The synthesis of OligoG CF-5/20 is described in detail elsewhere (Khan et al. 2012c), but in short requires a process of purification and fractionation, further purification using charcoal filters then spray drying. Phase 1 studies have confirmed safe working concentrations of 2% to 10% (Khan et al. 2012c). The concentration of 0.2% OligoG CF-5/20 (v/v) was utilised for FTIR analysis reflecting the previous pilot study (Pritchard et al. 2016b).

3.3.3 FTIR of CF sputum

FTIR analysis was performed on CF patient sputum samples (n=13) treated with 0.2% OligoG CF-5/20 (v/v) or dH_2O (v/v). CF sputum samples were divided in half and treated with 0.2% (v/v) OligoG CF-5/20 or dH_2O (control) prior to centrifugation at 4000 rpm for 5 mins prior to incubating statically at 37°C for 60 mins (Pritchard et al. 2016a). Reference spectra of 0.2% OligoG CF-5/20 in dH_2O were also generated.

Following incubation, 3 µl of each sample was pipetted in triplicate onto a 96-well silicon plate (Bruker Optics Inc., Coventry, UK). Following drying at room temperature (1 hour), high-throughput FTIR analysis was performed using a Bruker Vertex 70 in transmission mode with a KBr beamsplitter, DTGS detector and HTS-XT attachment.

Spectra were acquired within the range of 4000-400cm^{-1}, at a resolution of 4cm^{-1} with each spectrum being the averaged result of 32 scans. Each silicon plate was scanned in triplicate to ensure technical reproducibility of individual sample spectra. With thanks to Dr Manon Pritchard and Dr Charles Bright who helped with sample processing.

3.3.4 FTIR data processing and analysis

All data processing, analysis and visualisations were performed using Bruker OPUS 7.5 software with data subsequently analysed in Microsoft Excel. The whole infrared spectra were pre-processed prior to further analysis by baseline correction using the automatic 'rubberband' correction and vector normalisation. Absorbance spectra were smoothed using a nine-point Savitzky-Golay algorithm and second derivative spectra were calculated using a 9-point window (Lewis et al. 2010; Baker et al. 2016).

Based on the previously-published mucin glycosylation IR spectral library (Lewis et al. 2013a), the 'fingerprint' spectral region (absorbance values between wavenumbers 1800 and 900 cm^{-1}) was assessed. The peak picking algorithm provided by the OPUS 7.5 was used to identify predominant peaks within samples. Statistically significant differences between peak positions or absorbance intensity at key wavenumbers in treated and control CF sputum samples were determined using the Mann-Whitney *U* Test. The R-statistical computing environment was utilised for this statistical analysis, using in-built algorithms and code developed previously. All data processed using the R-statistical package was undertaken by Dr Charles Brilliant.

3.4 Results

3.4.1 Analysis of clinical samples

FTIR samples were collected from 13 patients (mean age 13.1 years; range 7-17 years; male to female ratio, 9:4 respectively). Mean FEV1 % predicted was 71.9% (range 47 to 99%) and FEV1 Z-score for age ranged from -0.1 to -4.5. All patients had previously isolated *Pseudomonas aeruginosa*, with 5/13 positive with the current sample (**Table 3.1**). All patients were using the nebulised therapy DNase to aid mucociliary clearance.

All patients (except one) who provided induced sputum samples were taking antibiotics (**Table 3.1**). No correlation was found between the wavenumber changes reported here and antibiotic use, sex, age, use of DNase or FEV1.

Table 3.1. Patient data for sputum samples at time of sampling

Patient	Age (Years)	Sex	Antibiotics	Cultured Pathological Bacteria (Isolated in past 12 months)	FEV1 % predicted (Z score)	BMI centile (Z score)
1	16	M	AZM, FLC	*H. influenzae, S. aureus, S. maltophilia,*	62 (-3.2)	25 (-1.2)
2	17	M	AZM, CST^{IV}, FLC, DOX, MEM^{IV}, TOB^{neb}	*H. influenza, S. aureus, S. maltophilia,*	*	9 (-1.93)
3	14	F	AZM, CLR, ETB, RIF	*S. maltophilia*	64 (-3.0)	9 (-1.7)
4	15	F	AZM, CLM^{neb} FLC	*P. aeruginosa*	87 (-1.1)	50 (0.42)
5	15	M	AZM, CLM^{neb} FLC	*P. aeruginosa, S. aureus*	75 (-2.0)	25 (-0.99)
6	9	M	AZM, FLC, CLM^{neb}/TOB^{neb}	*S. marcescens, S. aureus*	84 (-1.3)	91 (1.23)
7	16	F	AZM, CLM^{neb}	*H. parainfluenzae*	64 (-2.9)	25 (-0.53)
8	14	M	CLR, MIN, MEM^{neb}, MXF	*M. abscessus*	57 (-3.5)	50 (0.2)
9	14	M	AZM	*H. influenzae, MRSA, S. aureus*	64 (-2.9)	91 (1.23)
10	11	M	AZM, CLM^{neb}	*P. aeruginosa, MRSA*	47 (-4.5)	50 (-0.49)
11	11	F		*B. cenocepacia*	99 (-0.1)	98 (1.96)
12	12	M	AZM	*P. aeruginosa* (mucoid)	78 (-1.8)	50 (0.34)
13	7	M	AZM, FLC, CLM^{neb}/TOB^{neb}	*B. cenocepacia, P. aeruginosa, S. maltophilia, S. aureus*	81 (-1.5)	91 (1.85)

<u>Table 3.1</u>

Clinical data from 13 patient samples. Details include age, gender, antibiotic use, culture positivity over the last 12 months, lung function and growth. There were no correlations identified between wavenumber and clinical features listed.

<u>Abbreviations:</u>

M, male; *F*, Female; BMI centile, body mass index (kg/m^2) centile; *FEV1*, forced expiratory volume in one second; *IV*, intravenous antibiotic; *Neb,* nebulised antibiotic.

Antibiotics: *AZM*, azithromycin; *CLR*, clarithromycin; *CST*, colistin; *CLM*, colomycin; *DOX*, doxycycline; *ETB*, ethambutol; *FLC*, flucloxacillin; *MEM*, meropenem; *MIN*, minocycline; *MXF*, moxifloxacin; *RIF*, rifampicin; *TOB*, tobramycin.

Bacteria: *B. cenocepacia, Burkholderia cenocepacia; H. influenzae, Haemophilus influenzae; H. parainfluenzae, Haemophilus parainfluenzae; M. abscessus, Mycobacterium abscessus; M. avium-intracellulare, Mycobacterium avium-intracellulare, P. aeruginosa, Pseudomonas aeruginosa; S. aureus, Staphylococcus aureus; S. maltophilia, Stenotrophomonas maltophilia; S. marcescens, Serratia marcescens;* MRSA*, methicillin resistant S. aureus.*

* Missing data

3.4.2 Sputum and OligoG CF-5/20 spectra

FTIR was used to generate absorbance spectra between wavenumbers 4000 to 400 cm^{-1} to establish biochemical changes following 0.2% (v/v) OligoG CF-5/20 incubation with CF sputum samples. Analysis was then performed on infrared (IR) wavenumbers within the 1800 to 900 cm^{-1} 'fingerprint' region to focus on key structural changes in mucin proteins and associated glycosylation.

Figure 3.1a shows the mean absorbance spectra for both untreated and treated CF sputum samples, highlighting the wavenumber regions where IR absorbance patterns were altered post-treatment, as indicated by a shift up or down in wavenumber. The spectra were aligned with a representative OligoG CF-5/20 IR spectrum to easily identify regions of absorbance change that occurred due to OligoG CF-5/20 being present in the samples.

Raw absorbance IR spectra are a series of many peaks merged into broad peaks and used to determine the exact position of peak changes between untreated and OligoG CF-5/20 treated sputum spectra. Meanwhile, second derivative spectra are used to enhance the separation of any overlapping peaks and study these areas in closer detail. Comparisons were made of treated and untreated CF sputum second derivative spectra with a second derivative spectrum of 0.2% (v/v) OligoG CF-5/20 alone (**Figure 3.1b**). This allowed determination of potential structural changes or interactions at key mucin peaks following OligoG CF-5/20 incubation of CF-patient sputum.

<u>Figures 3.1a and b. Raw absorbance and second derivative infrared spectra for untreated and OligoG CF-5/20 treated sputum.</u>

<u>Figure 3.1a. Mean infrared (IR) spectra from the 1800 cm^{-1} to 900 cm^{-1} IR wavenumber region of: untreated CF sputa; treated CF sputa; and OligoG CF-5/20 in water.</u>

Baseline-corrected, vector-normalised absorbance spectra. Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line. This shows the mean absorbance spectra for both untreated and treated CF sputum samples, highlighting the wavenumber regions where IR absorbance patterns were altered post-treatment. The spectra were aligned with a representative OligoG CF-5/20 IR spectrum.
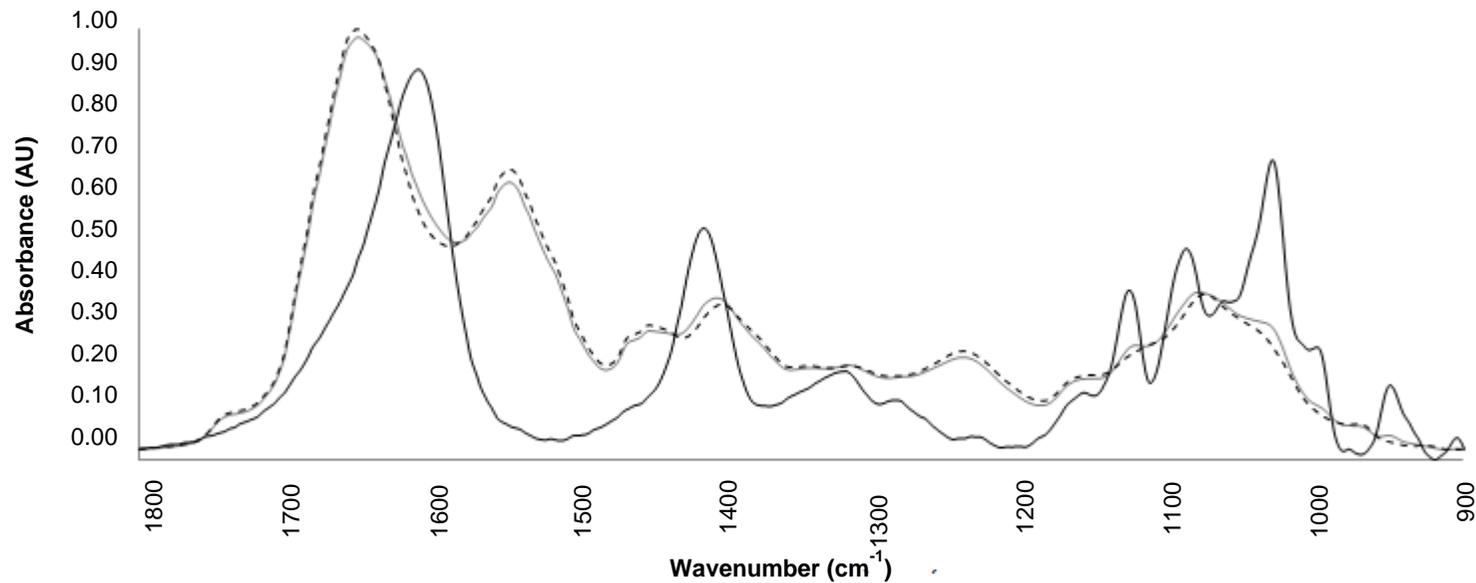
Figure 3.1b. Mean second derivative spectra for untreated and OligoG CF-5/20- treated sputum.

Baseline-corrected, vector-normalised absorbance spectra. The peaks are negative and point down in second derivative IR spectra.

Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line. Comparisons were made of treated and untreated CF sputum second derivative spectra with a second derivative spectrum of 0.2% (v/v) OligoG CF-5/20 alone. This allowed determination of potential structural changes or interactions at key mucin peaks following OligoG CF-5/20 incubation of CF-patient sputum.
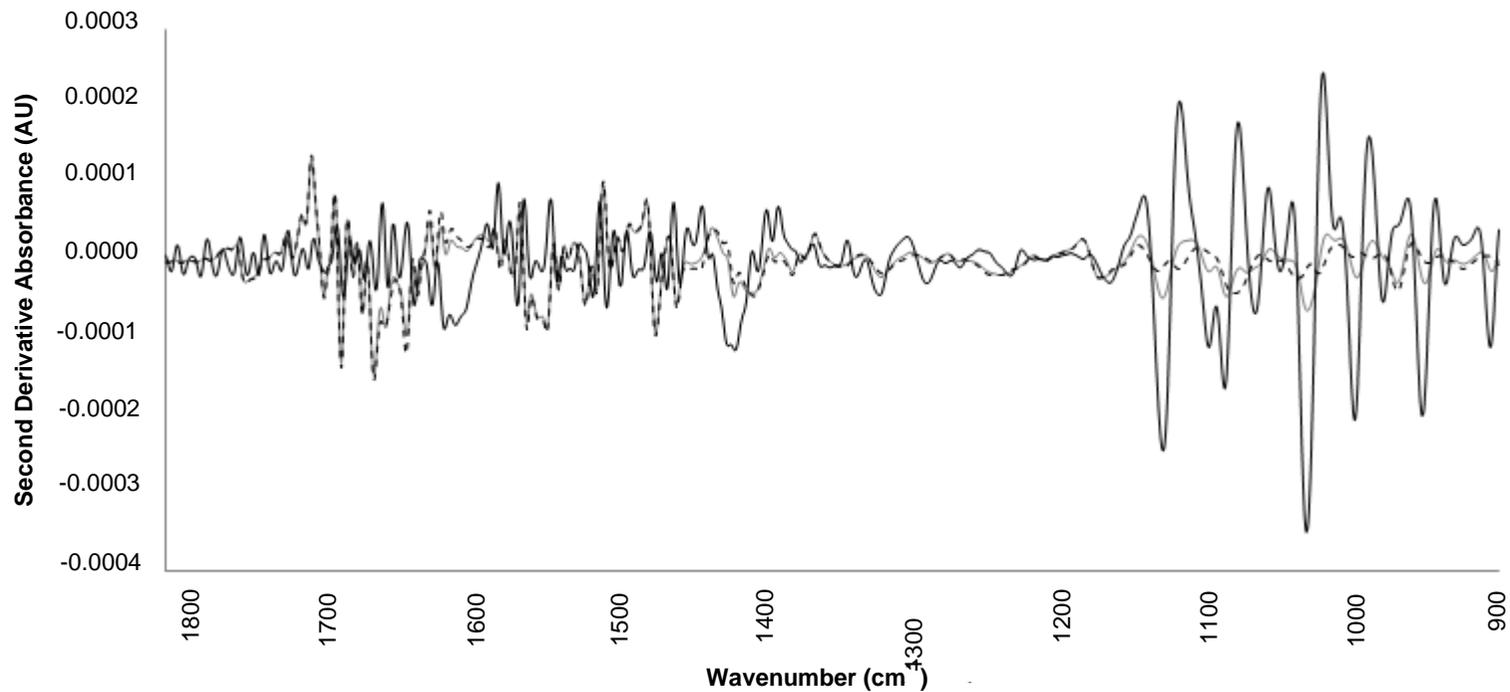
3.4.3 Mucin glycan changes in the presence of OligoG CF-5/20

The untreated sputum spectrum demonstrated a major glycan-associated peak at 1078 cm^{-1} (**Figure 3.2a**). This peak shifted to 1070 cm^{-1} in the OligoG CF-5/20 treated sputum spectrum, and no peak was evident in the OligoG CF-5/20 only spectrum at the same wavenumbers. **Figure 3.2b** shows that the distributions of major glycan-associated peak apex positions for all samples in the OligoG CF-5/20 treated sputum spectrum were significantly different from the untreated control samples (**Table 3.2**, p-value <0.05).

A peak at 1053 cm^{-1} was completely lost following sputum treatment with OligoG CF-5/20, although comparison with the OligoG CF-5/20 only reference spectrum shows the clear presence of a peak maxima. This suggests that the second derivative peak loss at 1053cm^{-1} may not be indicative of interaction between OligoG CF-5/20 and the sputum mucins, rather it could purely be indicative of OligoG CF-5/20 being present in the sample masking the presence of 1053cm^{-1}.

Figure 3.2a. Mean second derivative spectra for OligoG CF-5/20 treated and untreated sputum within the 1040 cm^{-1} to 1090 cm^{-1} wavenumber range indicating IR absorbance from mucin glycans and OligoG CF-5/20.

Baseline-corrected, vector-normalised absorbance spectra. The peaks are negative and point down in second derivative IR spectra. Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line. The OligoG CF-5/20 spectrum has been scaled by a factor of 5 to facilitate visualisation of all spectra in the plot. Spectra show IR peak at 1078 cm^{-1} shifted to 1070 cm^{-1} following OligoG CF-5/20 incubation and the peak at 1053 cm^{-1} in the untreated spectrum was completely lost.
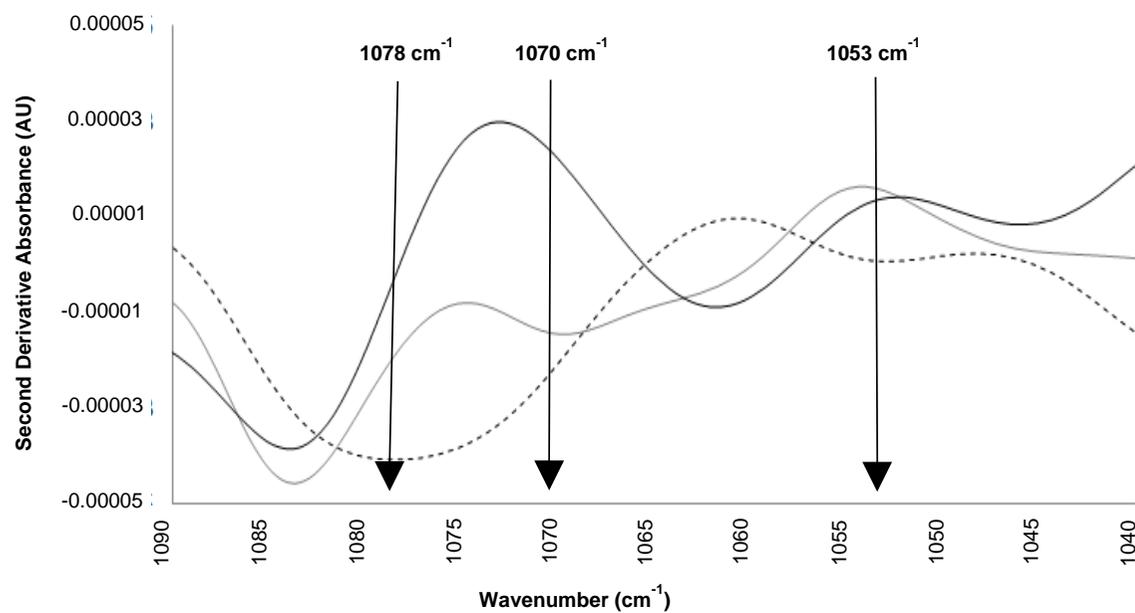
Figure 3.2b. Boxplots of second derivative peak apex positions between 1080 cm^{-1} and 1060 cm^{-1}.

Baseline-corrected, vector-normalised absorbance spectra. Boxplots show the distributions of second derivative peak apex positions centred around 1078 cm^{-1} and 1070 cm^{-1} in untreated (light grey box) and treated (dark grey box) samples respectively. *P <0.008, mean of n=13 sputum samples for treated and untreated groups; 3 replicates.

Statistically significant differences between peak positions in treated and untreated CF sputum samples were determined using the Mann-Whitney *U* Test.
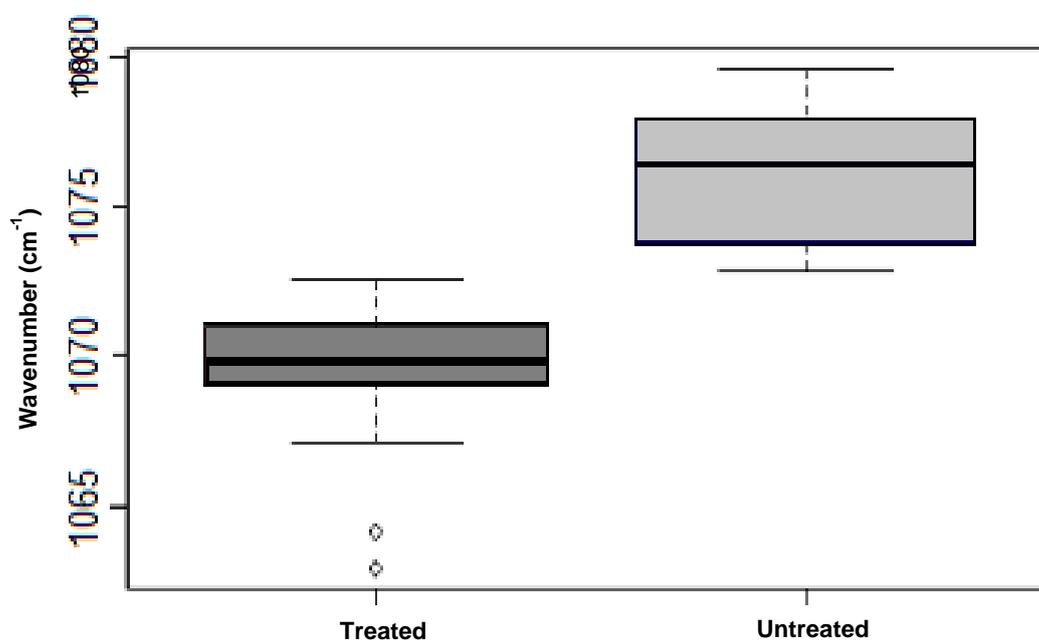
Table 3.2. Results of Paired Mann-Whitney *U* tests for differences of means in absorbance intensity at the specified wavenumbers and peak position around the wavenumbers. Mean of n=13 treated and untreated groups; 3 replicates, * p-value =<0.05.

Structural Interaction	Wavenumbers (cm^{-1})	Peak Shift or Absorbance Change	*p-value*
Random coil	1652	Second Derivative Peak Shift	0.3591
		Second Derivative Absorbance Change	0.1793
β-sheet	1637	Second Derivative Peak Shift	0.1070
		Second Derivative Absorbance Change	0.01876*
Sulphated-Lewis x	1240	Peak Shift	0.4973
		Absorbance Decrease in Treated Group	0.0078*
Sulphated-Lewis x	1116	Peak Loss in Treated CF Sputum	3.436e-11*
		Second Derivative Absorbance Change	7.105e-15*
Mucin glycans	1080	Peak Shift	0.0002*
		Absorbance Change	0.1055

3.4.4 OligoG CF-5/20 interactions with the Lewis x antigen

The absorbance at 1116 cm^{-1} was lost following OligoG CF-5/20 treatment (**Figure 3.3a**), suggesting an interaction of OligoG CF-5/20 with the sulphated-Lewis x antigen (Lewis et al. 2013a). This trend was seen across all samples (**Figure 3.3b**), with a statistically significant change in absorbance (**Table 3.2**).

The absorbance at, and peak positions around, 1240 cm^{-1}, was also examined as this wavenumber is also indicative of absorbance by the sulphated-Lewis x antigen (Lewis et al. 2013a) **(Figure 3.4a)**. A small shift in peak position towards lower wavenumbers in the OligoG CF-5/20 incubated sputum spectra, relative to untreated control sputum spectra was observed (**Figure 3.4b**). Additionally, a statistically significant decrease in absorbance was observed at 1240cm^{-1} in the incubated sputum spectra (**Figure 3.4c**, **Table 3.2**).

Figure 3.3a. Mean second derivative spectra for OligoG CF-5/20 treated and untreated sputum surrounding 1116 cm^{-1} indicating IR absorbance from sulphated-Lewis antigen.

Baseline-corrected, vector-normalised absorbance spectra. The peaks are negative and point down in second derivative IR spectra. Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line. The IR peak at 1116 cm^{-1} was lost following OligoG CF-5/20 incubation. The new peak occurring at 1096 cm^{-1} in the treated spectrum is suggestive of interaction between OligoG CF-5/20 and sulphated-Lewis antigen.
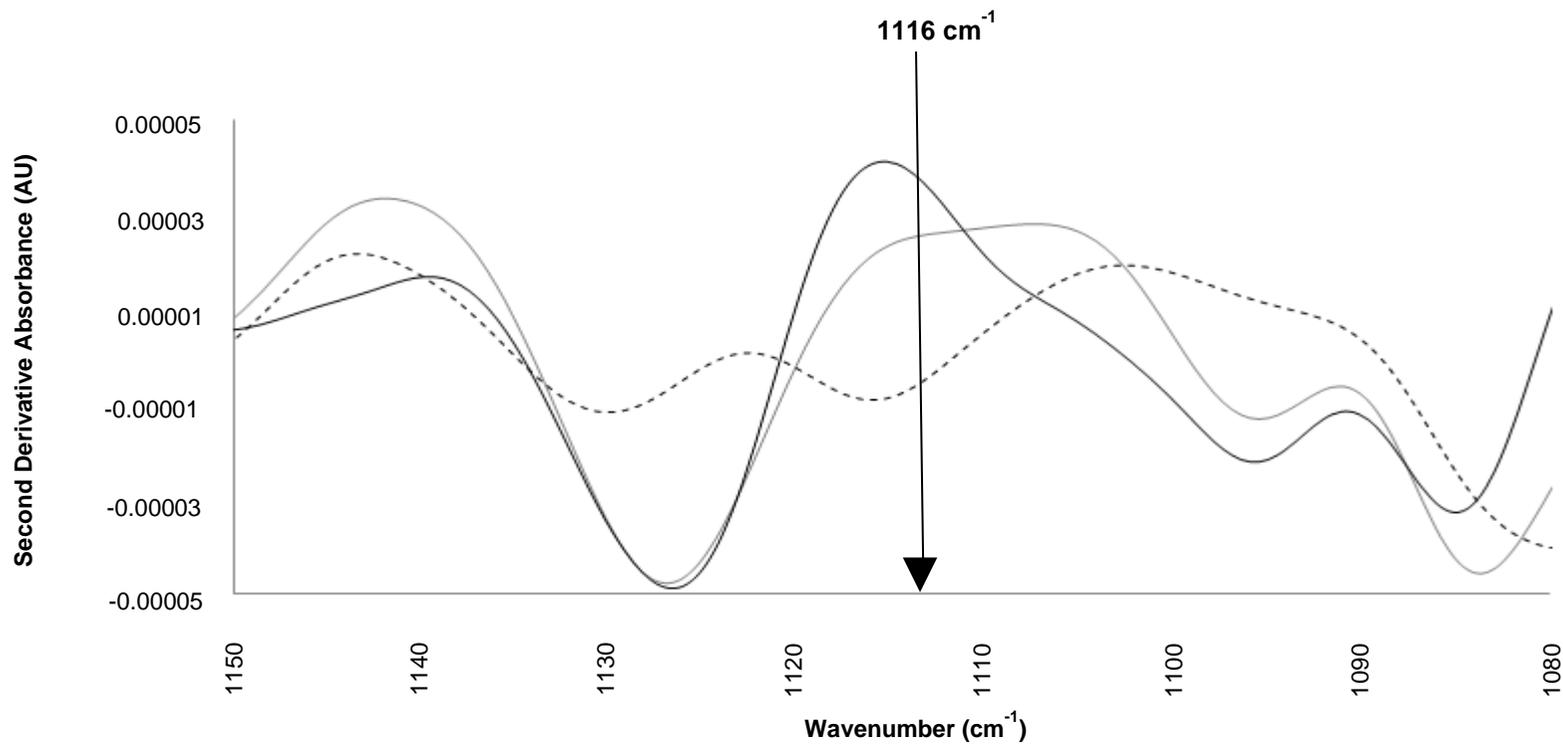
Figure 3.3b. Boxplots of second derivative negative peak height distributions centred around 1116 cm^{-1} in untreated and treated samples respectively.

* p-value <0.05, mean of n=3 replicates

Baseline-corrected, vector-normalised absorbance spectra. Statistically significant differences between peak height distributions in treated (dark grey box) and untreated (light grey box) CF sputum samples were determined using the Mann-Whitney *U* Test. P <0.008, mean of n=13 sputum samples for each group; 3 replicates.
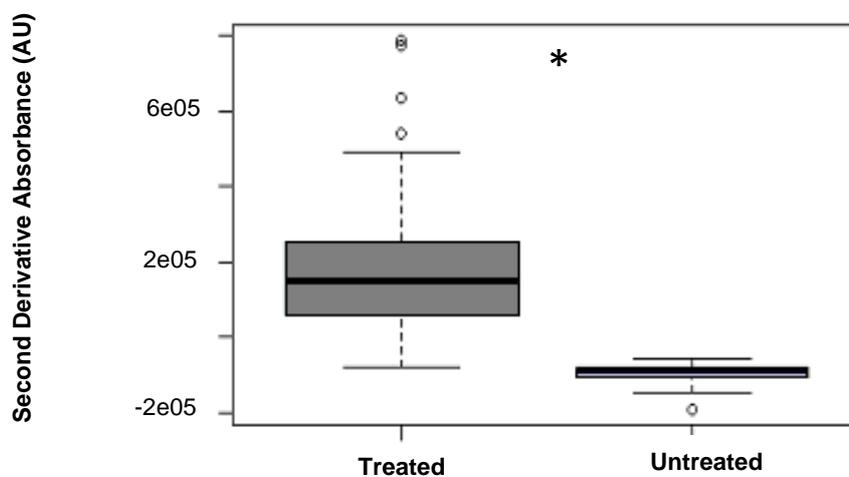
Figure 3.4a. Mean raw absorbance at 1240 cm^{-1} of untreated and treated samples showing lower absorbance in the treated sputum spectrum.

Baseline-corrected, vector-normalised absorbance spectra. Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line.
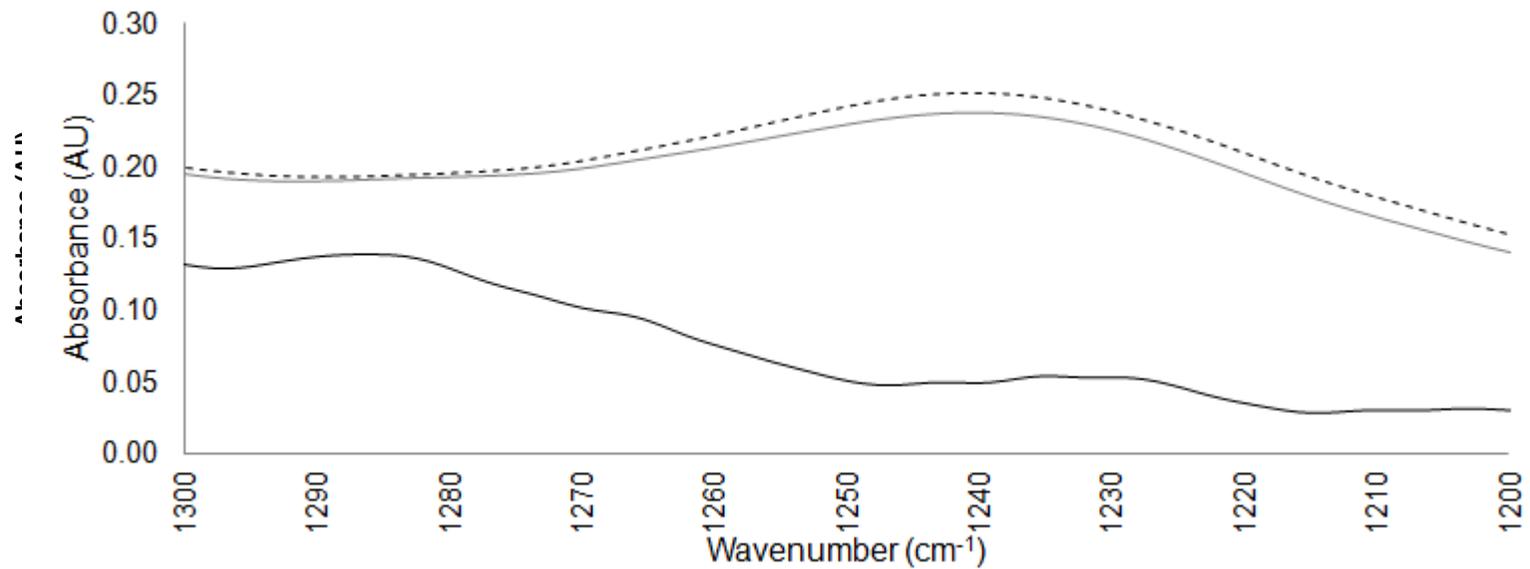
Figure 3.4b. Boxplots showing the distribution of wavenumbers in untreated and treated samples centred around 1240 cm[-1].

Baseline-corrected, vector-normalised absorbance spectra. Boxplots show the distributions of wavenumber in untreated (light grey box) and treated (dark grey box) sputum samples respectively. Mean of n=13 sputum samples for treated and untreated groups; 3 replicates. Statistically significant differences between wavenumber distribution were determined using the Mann-Whitney *U* Test.

A small shift in peak position towards lower wavenumbers in the OligoG CF-5/20 incubated sputum spectra, relative to untreated control sputum spectra was observed
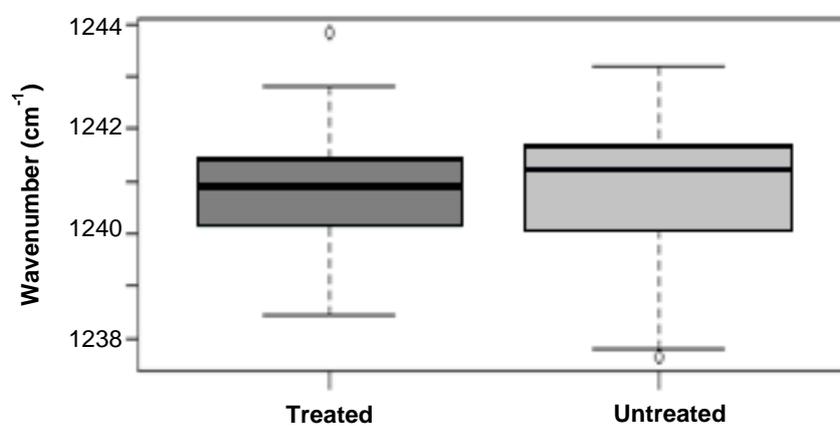
Figure 3.4c. Boxplots showing the relative absorbance in untreated and treated samples centred around 1240 cm^{-1}.

Baseline-corrected, vector-normalised absorbance spectra. Boxplots show the distributions of wavenumber in untreated (light grey box) and treated (dark grey box) sputum samples respectively. Mean of n=13 sputum samples for treated and untreated groups; 3 replicates. Statistically significant differences between wavenumber distribution were determined using the Mann-Whitney U Test; * p-value <0.008. A statistically significant decrease in absorbance was observed at 1240cm^{-1} in the incubated sputum spectra
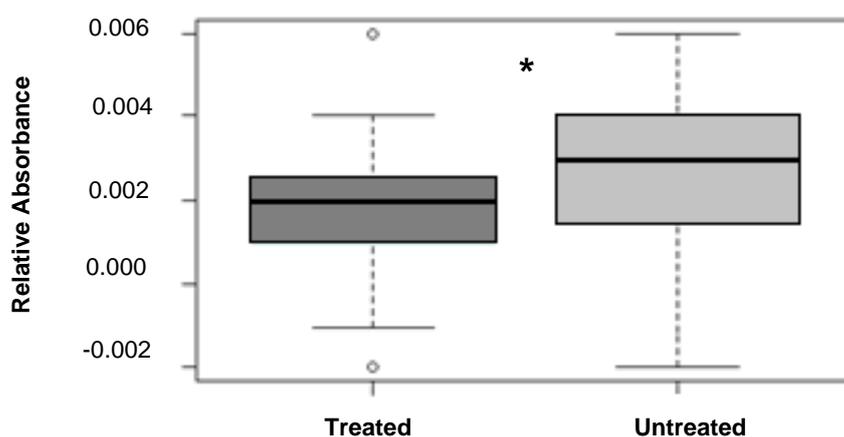
3.4.5 Mucin protein backbone changes in the presence of OligoG CF-5/20

The Amide I region between wavenumbers 1628 cm^{-1} and 1664 cm^{-1} was examined for potential interactions between OligoG CF-5/20 and the mucin peptide backbone, including areas representative of the protein secondary structures. This range was selected to include those peaks related to the random coil (1652 cm^{-1}) and β-sheet (1637cm^{-1}) secondary structures that predominate in sputum mucins and to avoid amino sugar absorbance at 1626 cm^{-1} (**Figure 3.5a**) (Lewis et al. 2013b).

No statistically significant (p-value >0.05) shifts in peak position around 1652cm^{-1} were observed (**Figure 3.5b i**)**.** However, there was a significant decrease in second derivative peak height at 1637 cm^{-1} (p-value <0.05) in the OligoG CF-5/20 treated samples, suggesting potential interaction between the β-sheet structures of the mucin protein backbone (**Figure 3.5c ii, Table 3.2**).

Figure 3.5a. Mean second derivative spectra for OligoG CF-5/20 treated and untreated sputum within the Amide 1 region between 1628 cm^{-1} and 1664 cm^{-1}.

Baseline-corrected, vector-normalised absorbance spectra. The peaks are negative and point down in second derivative IR spectra. Treated sputum (n=13, 3 replicates), solid grey line; untreated sputum (n=13, 3 replicates), black dashed line; OligoG CF-5/20 (n=3), solid black line. A wavenumber shift at 1652 cm-1 related to random coil secondary structure and absorbance change at 1637 cm-1 related to β-sheet structure were observed.
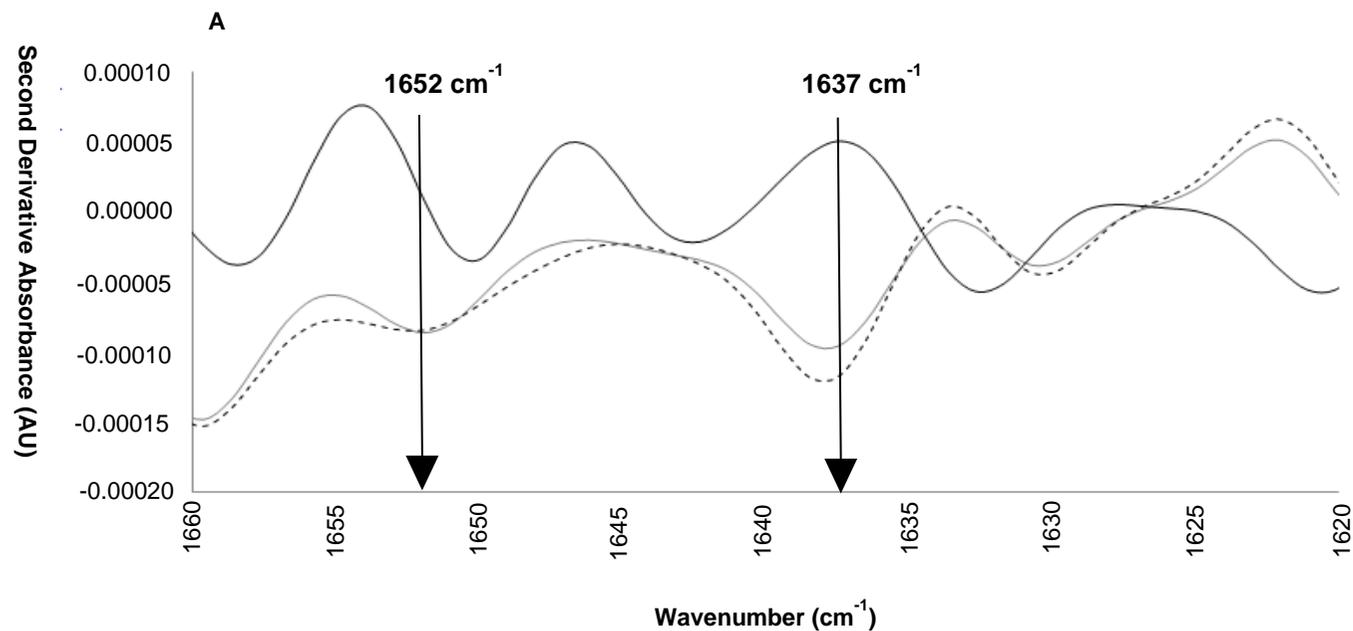
<u>Figure 3.5b. Boxplots showing the distribution of peak positions in untreated and treated samples centred around i)1652 cm^{-1} and ii) 1637 cm^{-1}.</u>

Baseline-corrected, vector-normalised absorbance spectra. Boxplots show the distributions of peak positions in untreated (light grey box) and treated (dark grey box) sputum samples respectively. Mean of n=13 sputum samples for treated and untreated groups; 3 replicates. Statistically significant differences between peak positions were determined using the Mann-Whitney *U* Test. No statistically significant (p-value >0.05) shifts in peak position around 1652cm^{-1} were observed. See overleaf.

<u>Figure 3.5c. Boxplots showing the distribution of second derivative absorbance in untreated and treated samples centred around i)1652 cm^{-1} and ii) 1637 cm^{-1}.</u>

Baseline-corrected, vector-normalised second derivative absorbance spectra. Boxplots show the distribution of second derivative absorbance in untreated (light grey box) and treated (dark grey box) sputum samples respectively. Mean of n=13 sputum samples for treated and untreated groups; 3 replicates. Statistically significant differences between absorbance were determined using the Mann-Whitney *U* Test. There was a significant decrease in second derivative peak height at 1637 cm^{-1} (p-value <0.05) in the OligoG CF-5/20 treated samples, suggesting potential interaction between the β-sheet structures of the mucin protein backbone. See overleaf.

<u>Figure 3.5b.</u> Boxplots showing the distribution of peak positions in untreated and treated samples centred around i)1652 cm[-1] and ii) 1637 cm[-1].

<u>Figure 3.5c.</u> Boxplots showing the distribution of second derivative absorbance in untreated and treated samples centred around i)1652 cm[-1] and ii) 1637 cm[-1].
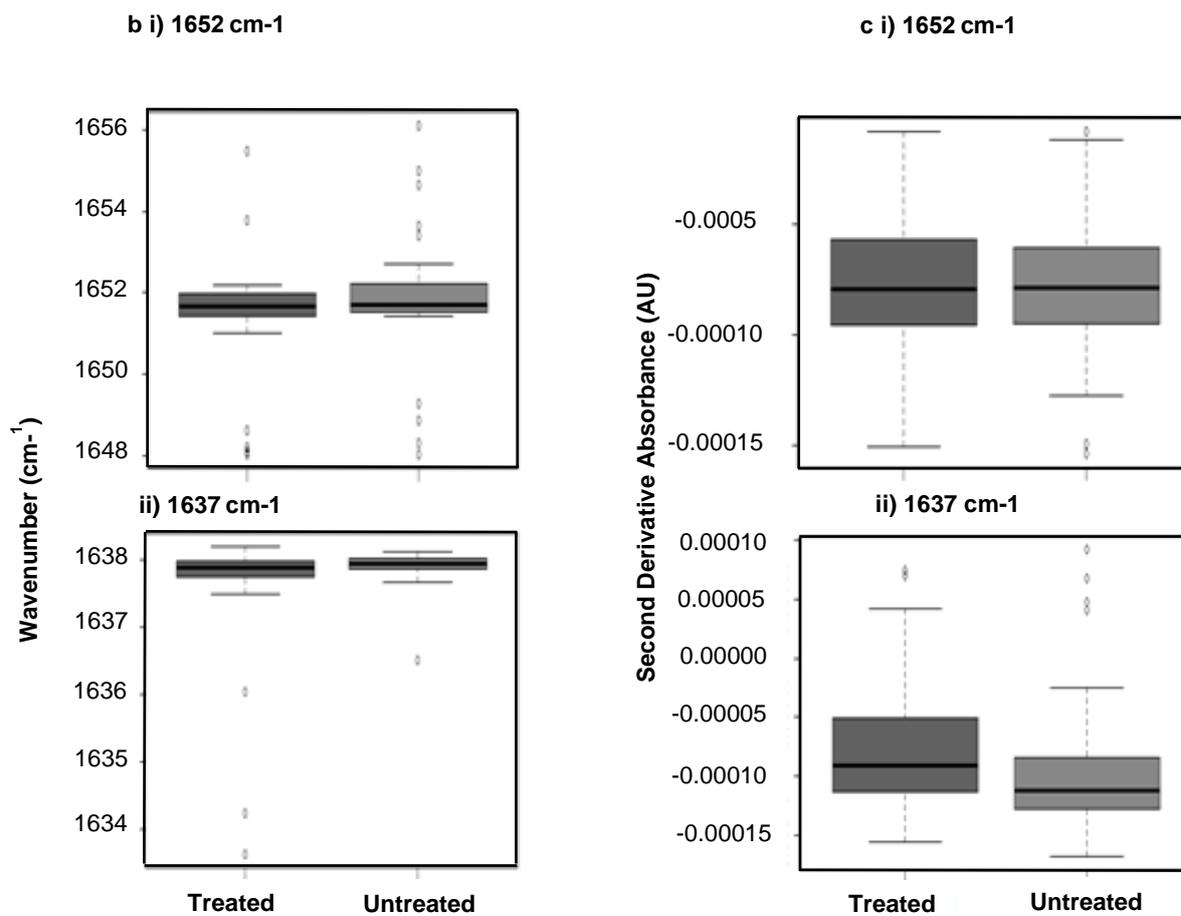
3.5 Discussion

Mucus provides a vital barrier for the human airways, preventing water loss and removing unwanted substances, such as microbes, foreign particles and inflammatory cells (Ma et al. 2018). Mucus stasis, as seen in CF, results in plugging of the smaller airways, chronic bacterial infection and airway inflammation leading to bronchiectasis (Ehre et al. 2014; Taherali et al. 2018). The combination of inadequate airway hydration and defective HCO_3^- -mediated post-secretory mucin expansion is likely to be the cause of airway obstruction seen in CF (Ehre et al. 2014).

A wide range of 'mucoactive' therapies are being developed and approved for use in CF. These agents are compounds which alter the biophysical properties of mucus, but may not act specifically on mucin (Morrison et al. 2019). Mucus represents a key pharmacological target, regardless of the patient's genotype or inflammatory status, therefore making these pharmacotherapies potentially universally useful within the CF population (Morrison et al. 2019). By focusing on altering, and potentially normalising, the properties of CF mucus, it may be possible to reduce the chronic bacterial infections typically seen in this disease, and reduce the need for antimicrobial therapies.

Interaction between alginate and gel-forming mucins has been shown to be electrostatic in nature and can result in weak viscoelastic gels which are thermostable and rheologically reversible (Taylor et al. 2005). Such gels likely contain mucin-mucin and mucin-alginate interactions, with the latter being electrostatic (Taylor et al. 2005). Two potential mechanisms for these interactions and the resultant reduction in network cross-linking are proposed: either direct competitive inhibition of cross links between polymers; or through altering the intramolecular interactions within polymer chains so that interpolymer cross-links are less likely, potentially through conformational changes (Nordgård and Draget 2011). The authors proposed that the introduction of low molecular weight alginate oligomers, such as OligoG CF-5/20, would disrupt these gels by effectively competing for the binding sites on mucins previously occupied by high molecular weight alginates, including those produced by *P. aeruginosa* (Nordgård and Draget 2011). This would modify sputum rheology and potentially improve mucociliary clearance in CF patients (Nordgård and Draget 2011; Sletmoen et al. 2012). It also raises the potential for OligoG CF-5/20 to aid drug

delivery via the respiratory mucosa, using the described mechanisms (Nordgård et al. 2014).

In vitro studies by Pritchard et al. (2016) used a range of techniques to interrogate the ability of OligoG CF-5/20 to disrupt the mucosal barrier. The authors demonstrated modification in mucin surface charge using electrophoretic light scattering (zeta potential), and increased pore size within the mucin network using atomic force microscopy and scanning electron microscopy imaging (Pritchard et al. 2016b). *In silico* Molecular Dynamic (MD) computer simulations were performed to try and predict the interactions between the structure of MUC5AC and a 12 DPn length G-rich alginate oligosaccharide. Results hypothesised that the alginate would bind to both glycan structures and the MUC5AC peptide backbone (Pritchard et al. 2016b). The study demonstrated bonding between the G-alginate oligosaccharide and the mucin's peptide backbone (N- or O-atoms) and between the N-atoms from the amide group in the GalNAc structure. There was also evidence of bonding with hydroxyl groups in the sugar rings of the glycan chains (Pritchard et al. 2016b). This study sought to provide a greater understanding of the exact structural sites at which OligoG CF5/20 may be acting within the sputum samples in order to effect such conformational changes to the mucin structure.

Altered mucin glycosylation patterns in CF patients enables formation of epitopes, which are recognised by bacteria as potential high-affinity binding sites, and theoretically may contribute to pathogenicity and chronic infection (Shori et al. 2001). There is an increased capacity for adhesion, cell-cell/cell-protein interactions, and potentially colonisation by bacteria, such as *P. aeruginosa* (Venkatakrishnan et al. 2015). In addition to bacteria, the host's inflammatory leucocytes may also recognise these sulphated ligands, leading to an inflammatory response and increased pathogenesis of *P. aeruginosa* respectively (Ramphal and Arora 2001; Xia et al. 2005).

The most prevalent epitope in CF sputum is sialyl Lewis x. There is evidence of interaction between this epitope and bacteria and viruses, as well as a correlation with disease severity (Davril et al. 1999; Degroote et al. 1999). The addition of fucose, a key structural component of the Lewis x epitope, alters mucin solubility and mucociliary clearance (Lewis et al. 2013a). This study demonstrated a statistically significant change in the sulphation peak in OligoG CF-5/20 treated sputum samples. Results suggest that OligoG CF-5/20 was able to bind to the sulphate moieties on the

glycans, including the sulphated Lewis x antigen. Clinically, this supports the use of OligoG CF-5/20 to counteract the effects of pathogen recognition and binding, and therefore potentially reducing both acute and chronic infections (Robinson et al. 2012).

Pritchard et al. (2016) also piloted the use of FTIR to further characterise mucin-OligoG interactions. Results suggested there was interaction with the mucin peptide backbone at the site of the carbonyl group within the peptide link. This was demonstrated in the alteration of peak position in the Amide I region of the second derivative spectra (Pritchard et al. 2016b). This provided evidence of changes in the protein's secondary structures, with absorbance shifts at wavenumbers associated with random coil (1652 cm^{-1} to 1650 cm^{-1}) and β-sheets (1637 cm^{-1} to 1634 cm^{-1}) (Pritchard et al. 2016b). FTIR suggested widespread H+ -bonding between OligoG CF-5/20 and the peptide backbone leading to reduced flexibility of the mucin molecule (Pritchard et al. 2016b).

This study did not show a statistically significant shift in peak position at wavenumbers representative of random coil secondary structures (1652cm^{-1}). This may be due to steric hindrance between OligoG CF-5/20 treated sputum and control sputum spectra at this location. However, a significant decrease in second derivative peak height at 1637 cm^{-1} in the OligoG CF-5/20 treated sputum samples was demonstrated. This is hypothesized to be as a result of hydrogen bonding between OligoG CF-5/20 and the peptide backbone, as seen in the previous study, causing a conformational change in the β-sheet secondary structure. The aforementioned study by Pritchard et al. (2016) prepared the sputum specifically to enrich the MUC5AC and MUC5B component, in contrast to this study which used sputum in its entirety (Pritchard et al. 2016b). This may explain the differences in interactions seen between OligoG CF-5/20 and the mucin structures.

Using a previously piloted method (Pritchard et al. 2016b), this study provides further evidence of structural interaction between OligoG CF-5/20 and mucin. An FTIR reference library for monosaccharides and relevant oligosaccharides has been developed to help identify the underlying structures and their interactions (Lewis et al. 2013a). The monosaccharide library spectra indicate potential peaks in glycan structures which may be shifted along or totally absent. The 3D structure of a glycan dictates strongly absorbance and position of some of these wavenumbers due to electrostatic interactions between neighbouring atoms within the structure.

It is important to recognise the limitations of FTIR analysis, even with support from spectral libraries. Structures such as OligoG CF-5/20 can interfere with analysis by absorbing strongly at the same regions as other key structures. For example, glycan-associated structures, such as the C-O bonds in the pentose and hexose rings, absorb strongly between 1280 and 900 cm^{-1}, with numerous second derivative peaks being evident. OligoG CF-5/20 also absorbs strongly within the same region as these mucin glycans, with a number of key peaks identified at the same positions. This study was designed to determine sputum mucin-OligoG CF-5/20 interaction, and not just OligoG CF-5/20 presence in CF sputum. It was therefore necessary to focus the analysis on wavenumbers that showed a difference in absorbance, or the presence of peaks which were not associated with OligoG CF-5/20 peaks, as being indicators of OligoG CF-5/20 interaction.

FTIR cannot determine the exact molecular structure of compounds, and some vibrations detected by FTIR are attributable to multiple chemical groups. There is also a risk with FTIR, as with other high throughput technologies, of false positive findings due to the large amount of data generated (Baker et al. 2016). This risk can be reduced by ensuring careful standardisation of the experimental conditions for specimen preparation, spectral acquisition, data processing and analysis (Baker et al. 2016). This is particularly important when results are being used for clinical decision-making. A recognised limitation of infrared spectroscopy of liquid solutions is the interference of water on the spectrum seen. Strong absorbance of water is seen in the mid-infrared region (near 1645 cm^{-1}), which is also the region of the Amide I band and some side chain bands (Barth 2007). This study included careful visual assessment of all spectra prior to formal analysis to ensure there was no sign of water interference.

It is well-recognised that mucus composition is highly variable, both in terms of mucin concentration, mucin structure and also the other non-mucin components present, such as extracellular DNA (eDNA), which is seen in higher volumes in CF sputum samples (Nordgård and Draget 2011). As a key player in sputum viscoelasticity, it would have been helpful to study interactions with eDNA. However, a previous study using both FTIR spectral analysis and isothermal titration calorimetry showed no interactions between OligoG CF-5/20 and DNA, in terms of changes in peak positions and molar heat effects respectively (Powell et al. 2018). This was in part contributed to by the overlap of OligoG CF-5/20 peaks with the phosphate region. Therefore, this study focused specifically on interactions with mucin. The heterogeneity of CF sputum

means it is particularly important to ensure experiments are undertaken with sufficient replication to identify potential outliers (Bhat et al. 1996). This study showed remarkable intra-sample reproducibility, despite anticipated heterogeneity. Given the relatively small patient sample size, it was not possible to identify any patterns from our clinical data. This included any potential relationship between previously identified bacterial infections or mucolytic therapies and OligoG CF-5/20 interaction with mucin. Results would be enhanced by a larger sample size and a control group, but this was outside the limits of this study.

Many studies have demonstrated the potential use of FTIR as a sensitive diagnostic tool in a range of human cancers (Lewis et al. 2010). Other authors have proposed the use of FTIR as a rapid tool for bacteria identification, including in patients with CF. They have used the infrared spectra provided by intact bacteria and compared these patterns to existing spectral reference libraries (Bosch et al. 2008). This work contributes to the limited number of existing studies on the use of FTIR to analyse sputum. Previously, this technology has been employed to accurately characterise alterations in mucin expression, secretion and glycosylation in respiratory disease (Lewis et al. 2010; Lewis et al. 2013a; Lewis et al. 2013b). This study added additional complexity by using FTIR to analyse interactions of a pharmacotherapy with human mucin. There is limited data regarding the use of FTIR as a tool for identifying drug-biological specimen interactions. However, this simple FTIR-based protocol confirms that the technique can be utilised to detect interaction between OligoG CF-5/20 and sputum mucins.

It is feasible that FTIR analysis could be employed to monitor patient adherence to therapy through simple detection of OligoG CF-5/20 in expectorated sputum. A study assessing such use is underway (patent application number: PCT/EP2018/053347). Using FTIR for such monitoring could be useful in clinical trials for studying washout time. Patients with CF often have sub-optimal adherence to long-term inhaled therapies, but this can be improved through the use of electronic monitoring of compliance (Narayanan et al. 2017). If available as a bedside tool in hospital clinics, it could also enable healthcare providers to give additional support to patients who show poor compliance with essential treatments.

Of relevance to the study of CF sputum, FTIR is particularly useful as it does not place limits on the molecular weight of components found within the sample, such as DNA, which can be present in much higher concentrations in CF sputum compared to

healthy control (Balan et al. 2019). Alterations in glycosylation, including sialylation and sulphation of the Lewis x antigen, are seen in respiratory disease, particularly during infective exacerbations (Lewis et al. 2013a). Spectral patterns representing these structures were easily identifiable within our patient samples, though with the small sample size it was not possible to compare groups. With more samples, the FTIR protocol could be further developed to identify changes in CF sputum by studying 'stable' and 'exacerbating' samples. This could aid CF patient management, including early recognition of pulmonary exacerbations and therefore prompt treatment with antimicrobials including OligoG CF-5/20.

3.6 Conclusion

This work demonstrates the use of FTIR analysis with human CF sputum to study the interaction of OligoG CF-5/20 with respiratory mucin. There is clear evidence of significant structural interactions between the glycan moieties and peptide backbone of mucin molecules. These results support existing literature and provide further evidence for the therapeutic effects of OligoG CF-5/20 on modifying the viscoelastic properties of CF sputum through alginate-mucin interactions.

Chapter Four

Phenotypic and genotypic adaptations in *Pseudomonas aeruginosa* biofilms following long-term exposure to an alginate oligomer inhalation therapy

4.1 Introduction

Cystic fibrosis lung disease is characterised by airway inflammation, viscid mucus with impaired muco-ciliary clearance and abnormal mucosal defence. These features provide an ideal host environment for repeated bacterial colonisation and infection, cycles of inflammation and resultant tissue damage, with eventual respiratory failure and death (Ratjen et al. 2015).

4.1.1 Bacterial evolution in the cystic fibrosis lung environment

Alterations in the CF lung milieu are exploited by CF pathogens, including *Pseudomonas aeruginosa*, which colonise and chronically infect the CF lung (Cullen and McClean 2015). *P. aeruginosa* typically colonises the lungs of CF patients at 8.8 years old (Marvig et al. 2015), where it then faces considerable selective pressures from a range of environmental factors. These include the host immune system, oxidative stress within the CF lung biofilm and the use of chronic antibiotics (Goss and Burns 2007).

The effects of these environmental stressors within the CF lung environment are reflected in the extensive phenotypic and genotypic adaptations observed in CF *P. aeruginosa* (Folkesson et al. 2012). With chronic colonisation, later-stage bacterial isolates are distinctly different from pioneer colonisers, displaying well-defined characteristics such as loss of motility, mucoidy, reduced growth rates, increased antibiotic resistance, and defective quorum sensing (QS) signalling (Goss and Burns 2007). Furthermore, mutations in *P. aeruginosa* to the alginate-producing mucoid phenotype or highly adherent small colony variants (SCVs) occur with increased resistance to host defences (Malone 2015). Eventually, *de novo* mutations lead to evolution of antibiotic resistance in individual patients (Smith et al. 2006a; Yang et al. 2011b), enabling the CF lung to become colonised by multidrug-resistant (MDR) pathogens, particularly *P. aeruginosa* (Folkesson et al. 2012).

4.1.2 Using experimental evolution to evaluate antimicrobial therapies

Antimicrobial discovery has previously been based on laboratory cultures of planktonically-growing bacteria, reflecting effectiveness in acute infections. However, as *P. aeruginosa* is predominantly in sessile form within the CF lung, such models

are unrealistic and likely contribute to the difficulties in finding effective therapies. In contrast, evolutionary biofilm models can recreate, and control for, many of the features seen in the CF lung, such as slow bacterial growth and repeated antimicrobial dosing (Martin et al. 2016). Far fewer evolutionary studies have been completed using populations in biofilms, though a variety of models have been described in the literature (Steenackers et al. 2016).

Chronic infections by *P. aeruginosa* represent a useful opportunity to study persistent bacterial infection and evolution. Authors have explored the genotypic and phenotypic diversity that occurs during long-term CF infections through use of whole genome sequencing and a variety of phenotype profiling on patient samples taken over a period of time, often representing chronic colonisation (Smith et al. 2006a; Yang et al. 2008; Folkesson et al. 2012; Markussen et al. 2014). Such methods have aided the understanding of bacterial evolution *in vivo*, but are less helpful when testing therapeutics to treat such infections.

Poltak and Cooper (2011) described their pioneering study using a novel bead-biofilm model, with cycles of surface colonisation, biofilm formation and dispersal. Their model studied experimental evolution of an initial single clone of *Burkholderia cenocepacia* through changes in colony morphology and growth across six replicate populations. This showed parallelism with three variants developing in the same order across these populations (Poltak and Cooper 2011a). They demonstrated greater diversification in the biofilm environment compared to planktonic culture.

4.1.3 Evolution in the presence of a novel antimicrobial

In previous studies exploring its effect on *P. aeruginosa*, OligoG CF-5/20 has demonstrated anti-biofilm properties (Pritchard et al. 2017a; Powell et al. 2018), and antibiotic potentiation (Khan et al. 2012a; Pritchard et al. 2017a). This novel therapy appears to exert its properties via alteration of bacterial surface charge and motility (Khan et al. 2012a; Powell et al. 2014a), disruption of extracellular polymeric substance (Powell et al. 2018) and inhibition of QS signalling (Jack et al. 2018). Previous work has shown that planktonic subculture of *P. aeruginosa* (PAO1) for 21 days in escalating concentrations of OligoG CF-5/20 did not result in loss of antimicrobial activity (Khan et al. 2012c).

OligoG CF-5/20 has been designed for inhalation as part of regular daily cystic fibrosis therapy. As an anti-biofilm therapy, it is difficult to fully understand its long-term effects *in vivo* without the use of *in vitro* biofilm evolutionary models, such as the aforementioned bead biofilm model (Poltak and Cooper 2011a). It is essential to understand whether treatment with OligoG CF-5/20, either as short- or long-term therapy, will alter bacterial morphotype and genotype. Of particular interest is its impact on the evolution of bacterial phenotypes which are associated with antimicrobial resistance, such as SCVs and mucoid phenotypes, as this would be considered a negative outcome of therapy.

This study will help define the potential implications of long-term treatment with OligoG CF-5/20 for patients with CF. Clinically-relevant benefits of treatment would include a reduction in development of SCVs and mucoid phenotypes. Clearly, a positive impact of selective pressure on biofilm growth, resulting in a negative impact for the host, would be increased numbers of difficult to eradicate pseudomonal phenotypes. These effects would inevitably worsen disease outcomes if translated *in vivo*. As CF patients will typically continue prophylactic antibiotics, as well as intermittent antibiotic treatment courses for acute pulmonary exacerbations, it was deemed useful to also study the effects of OligoG CF-5/20 in the presence of a key CF antibiotic, azithromycin, on *P. aeruginosa* adaptive behaviours.

4.2 Aims

This study aims to experimentally model the effect of long-term exposure of *P. aeruginosa* to OligoG CF-5/20. A bead biofilm model will be used to incorporate regular cycles of surface colonisation, biofilm assembly and dispersal.

The aims of the study will be to determine:
- the phenotypic and genotypic characteristics of the morphotypes isolated in the presence of OligoG CF-5/20 using a range of experimental assays to determine the impact of bacterial evolution;

- the changes in acquisition of resistance to azithromycin (an antibiotic commonly used in the treatment of CF) in the presence and absence of sublethal concentrations of OligoG CF-5/20.

4.3 Materials and methods

4.3.1 Bacterial strains and media

Pseudomonas aeruginosa (PAO1, ATCC 15692) was employed throughout the study. Media included tryptone soy broth (TSB; LabM, Heywood, Lancashire) for overnight cultures and cation-adjusted Mueller-Hinton (MH) broth (LabM) for biofilm growth. Luria-Bertani (LB) agar plates (1% NaCl, 1% tryptone, 0.5% yeast extract, 1.5% agar in high salt agar) were used for the loss of resistance assay. PAO1 was grown overnight (O/N) on blood agar plates (BA; Blood agar base no. 2; LabM; supplemented with 5% horse blood). Fresh O/N cultures (n=4) were prepared in TSB (37°C) and placed on a 'roller mixer' for 16-20 hours.

4.3.2 Alginate oligosaccharide (OligoG CF-5/20)

The process of OligoG CF-5/20 production was described in Chapter Three. For this chapter, 2% OligoG CF-5/20 was used. This concentration value was based on a model previously described, which showed effective release of tethered mucin by OligoG CF-5/20 at this concentration (Ermund et al. 2017).

4.3.3 Study design

Figure 4.1 shows a schematic for the experimental design. There were two principal arms to the study. The first arm involved morphotype characterisation using morphotypes from the bead biofilm model grown in MH broth ± 2% OligoG CF-5/20 A range of morphotypic and genotypic assays were performed. The second arm used morphotypes grown in azithromycin ± 2% OligoG CF-5/20 for a cross-resistance study involving antimicrobial susceptibility testing against ten different antibiotics.
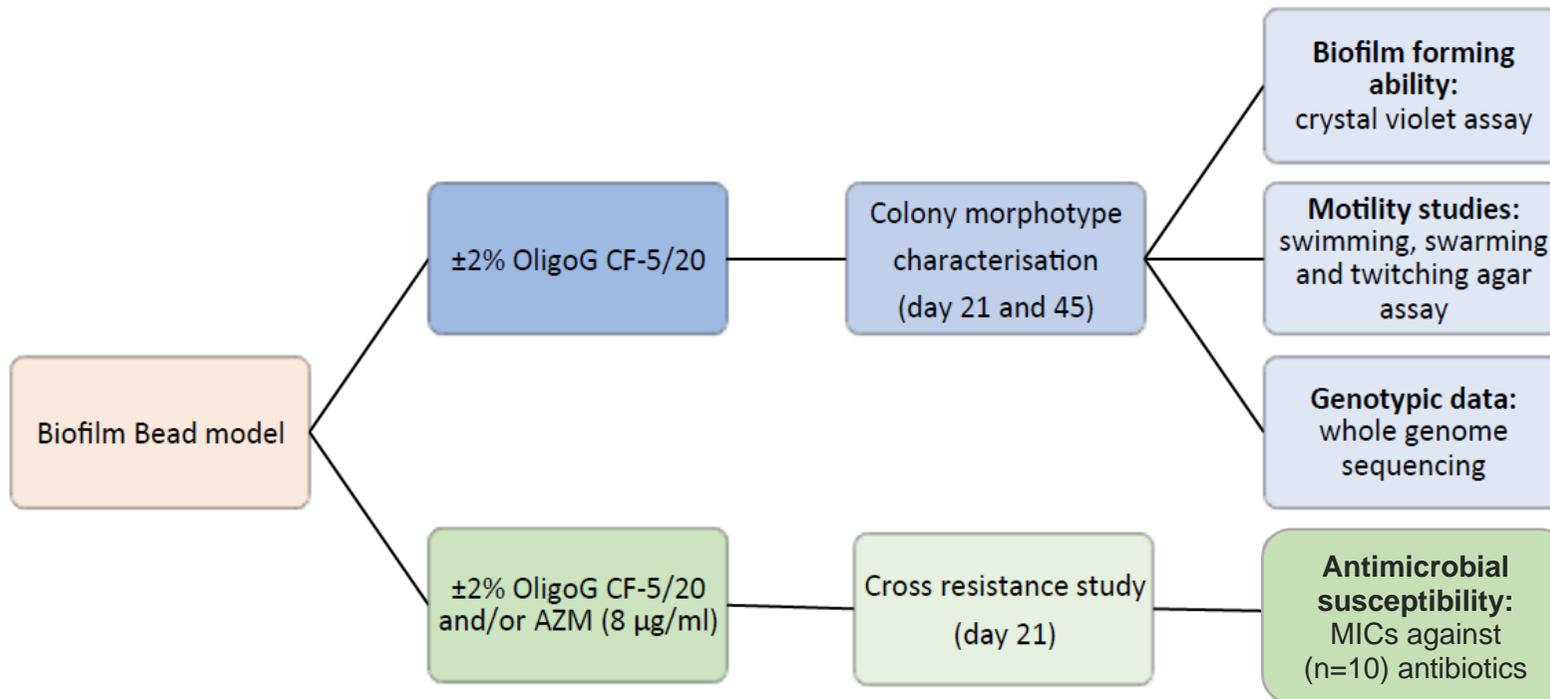
Figure 4.1. Experimental design flowchart.

Bead biofilm model flowchart describing assays according to laboratory conditions. Blue boxes show assays completed using morphotypes grown in control (MH broth) ± 2% OligoG CF-5/20. Green boxes show assays completed using bacteria grown in control (MHB) and 2% OligoG and control (MHB) ± azithromycin (AZM) (8ug/ml).

4.3.4 Bead biofilm evolution model in the presence of OligoG CF-5/20

The bead model used here was adapted from a previous study (Poltak and Cooper 2011a). Biofilms were grown on sterile 7 mm borosilicate glass beads without holes (John F. Allen & Son, Warwick, Rhode Island). Yellow and blue beads were used for alternate transfer days (**Figure 4.2**). Beads were placed in each well of a 24-well microtiter plate containing 1 ml MH broth ± 2% OligoG CF-5/20 (Sigma® cell culture plate, Sigma-Aldrich, Gillingham, United Kingdom). Four biological repeats were conducted in parallel; each repeat also had a sterile control well. Plates were incubated continuously at 37°C (20 rpm) for the duration of the experiment.

On passage days, the biofilm-covered beads were transferred into the corresponding wells of a fresh 24-well plate 3 times/week, which contained fresh medium (± 2% OligoG CF-5/20) alongside a single new sterile bead (in each well). The bacteria from the mature biofilm community would then colonise the new sterile bead, to repeat the transfer sequence again (Oakley et al. 2021). The old beads were then discarded or used for sampling on transfer days 21 and 45, where they were used to inoculate 10ml sterile TSB for overnight growth at 37°C.

At time zero, all wells (except sterile MH broth control wells) were inoculated with 5 μl of overnight growth of Wild type (WT) PAO1 from four independent cultures, standardised to an OD_{600} of 0.05 (10^7 CFU/ml); final cell concentration 5 x 10^4 CFU/ml. It was assumed that planktonic cells within the biofilm wells would decline over time by up to 33% (Poltak and Cooper 2011b) due to the increasing numbers of bacterial cells attaching to the bead. Therefore, the number of generations within the biofilm wells was estimated at 16 x $\log_2$(dilution), giving an estimate for generations over the course of the experiment between 240 to 250. Purity plates were performed weekly on each test well on BA plates. On days 21 and 45, purified morphotypes were frozen at -80°C on microbank beads. These 'monoculture' samples were then used in subsequent morpho- and genotyping experiments.

Fresh overnight cultures from stock PAO1 were used as the control sample, for all experiments, hereafter referred to as Wild Type (WT) PAO1. WT PAO1 represents day 0 in the evolutionary model.

With thanks to Dr Manon Pritchard who helped to maintain bead transfers.
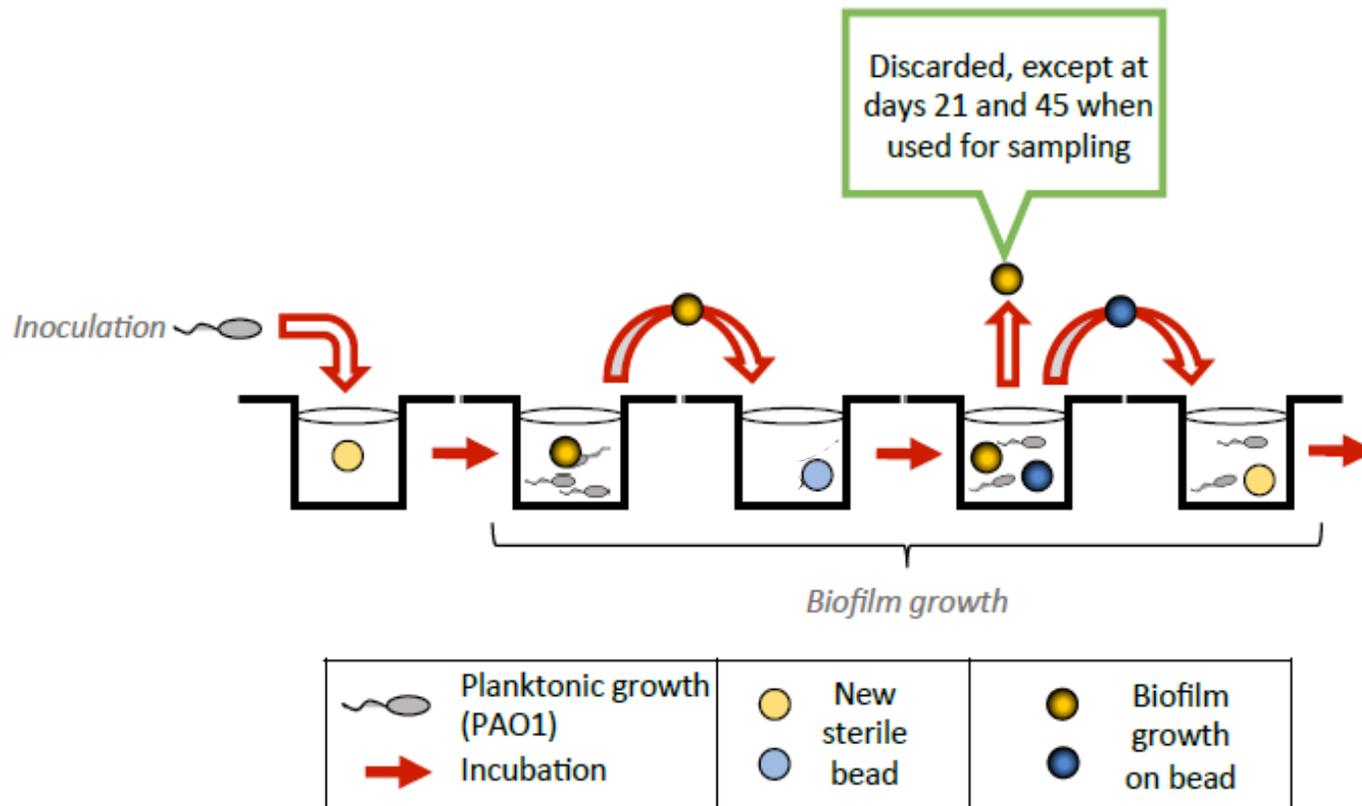
Figure 4.2. Schematic showing the bead biofilm model experimental strategy modelling of prolonged exposure to OligoG CF-5/20 ± azithromycin exposure.

4.3.5 Morphotype characterisation

The old bead was removed, vortexed in fresh TSB and grown overnight. Samples were then plated out onto blood agar plates and incubated for 24-72 hours before purifying individual colonies on fresh BA plates and freezing them for later use. The colony morphotypes isolated (MH ± 2% OligoG CF-5/20) were characterised independently by three researchers (Dr Juliette Oakley/Dr Manon Pritchard/Dr Katja Hill) at both day 21 and day 45 to ensure consistency of morphological identifications. Maintenance of morphotype after freezing was demonstrated by performing further overnight cultures from the freezer stock in TSB (37°C; 120 rpm) and plating cultures onto a fresh BA plate. Any samples demonstrating mixed colonies or contamination after freezing were removed from further analysis.

4.3.6 Phenotypic characterisation

4.3.6.1 Scanning electron microscopy (SEM)

A selection of different morphotypes from the day 21 control biofilm well samples were chosen based on size and surface texture (small ruffled, small studded medium smooth [mucoid], medium studded, and large smooth) plus a WT PAO1 control (n=3) (**Table 4.2**). When multiple samples were available, samples with the best biofilm-producing ability (described below) were selected. Overnight cultures of individual morphotypes were adjusted to 10^6 CFU/ml in MH broth and grown for 24 hours in a 12-well plate (Greiner Bio-One) containing Thermanox glass slides (Agar scientific; 37°C; 20 rpm). Following incubation, the supernatant was removed and each well was immersed in 2.5% (v/v) glutaraldehyde for two hours prior to being washed four times with dH$_2$O and freeze dried. Samples were gold-coated and imaged using a Tescan Vega conventional scanning electron microscope (SEM; 6 kV). With thanks to Dr Manon Pritchard and Dr Lydia Powell for their analysis of the imaging output.

4.3.6.2 Biofilm formation assay

All morphotype samples isolated from day 21 and day 45 were tested for biofilm-forming ability in 96-well polystyrene plates and compared to the WT PAO1 control. Biofilms were grown in a flat-bottomed 96-well plate (5 x 10^5 CFU/ml in MH broth) for 24 h statically (n=3 biological, n=5 technical repeats). An adapted crystal violet (CV) methodology was used to quantify biofilm-formation (O'Toole 2011; Santiago et al. 2016). The supernatant was removed from each well and the plates gently washed in dH$_2$O (x2 for 10 seconds) to remove planktonically-growing cells. 0.1% CV (125 µl)

was then added into each well to stain the attached cells. Following incubation for 15 mins at room temperature, plates were rinsed four times in dH$_2$O and allowed to dry for ≥1 h prior to solubilising the dye in 95% ethanol (200 µl for 30 mins). A 125 µl sample from each well was transferred to a fresh 96-well plate and the absorbance (optical density at 550 nm [OD$_{550}$]) was measured with a spectrophotometer.

4.3.6.3 Confocal scanning laser microscopy (CLSM) imaging of SCVs

Individual isolates representing smooth and ruffled SVC were chosen from the biofilm monocultures obtained on day 21 and day 45 (control and 2% OligoG CF-5/20 plates), including WT PAO1 control (n=3) (**Table 4.2**). As with the SEM protocol, when multiple samples of the same appearance were available, isolates with the best biofilm production were selected. Overnight cultures from freezer stock microbank beads were prepared in TSB (37°C; 120 rpm). CLSM was performed on biofilms grown in Greiner glass-bottomed optical 96-well plates in MH broth using a starting inoculum of 10^6 CFU/ml (37°C; 24 hours; rocking). The supernatant was removed before staining the cells with 6% (v/v) LIVE/DEAD® stain (BacLight^{TM} Bacterial viability kit, Invitrogen) in phosphate-buffered saline (PBS). The plate was wrapped in foil and incubated at room temperature for 10 mins, prior to imaging with a Leica SP5 confocal microscope with x 63 magnification under oil. With thanks to Dr Manon Pritchard and Dr Lydia Powell for their analysis of the imaging output.

Table 4.1. Samples used for SEM

All morphotypes were taken from day 21, using control plate samples only. When multiple samples were available for the morphotype category, the colony with the best biofilm production compared to WT PAO1 was selected (n=3), using data from the biofilm formation crystal violet assay described above.

Key: S; Small, M; Medium, L; large, Ru; Ruffled, St; Studded, Sm; Smooth; Mucoid.

Transfer	Sample well	Morphotype	Biofilm production (% difference from WT PAO1)
Day 21	CB18	S, Ru	63.16
	CC18	S, St	282.37
	CD18	M, Sm, Muc	-5.18
	CD18	M, St	221.56
	CB18	L, Sm	14.58
	WT PAO1		

Table 4.2. Samples used for CLSM

Morphotypes were taken from day 21 and day 45, control and 2% OligoG CF-5/20 plates. When multiple samples were available for the morphotype category, the colony with the best biofilm production compared to WT PAO1 was selected (n=3) using data from the biofilm formation crystal violet assay described above.

Key: S; Small, M; Medium, L; large, Ru; Ruffled, St; Studded, Sm; Smooth.

Transfer	Sample well	Morphotype	Biofilm production (% difference from WT PAO1)
Day 21	CB18	S, Ru	63.16
	CC18	S, St	282.37
	OB18	S, Sm	-47.40
	OB18	S, Ru	220.45
Day 45	CB116	S, Sm	135.84
	OC116	S, Sm	-75.17
	WT PAO1		

<u>4.3.6.4 Motility assays</u>

All morphotypes from days 21 and 45 were tested for swimming, swarming, and twitching ability compared to the WT PAO1 control. Agar plate-based assays were prepared, poured and used immediately once set, as previously described (Deziel et al. 2001). O/N cultures were prepared from freezer stocks in TSB and cells were point inoculated with a sterile toothpick onto the surface of the agar prior to being incubated for 48 hours at 25°C (swimming) or 30°C (swarming) and 72 hours at 30°C (twitching). The widest diameter of bacterial migration (millimetres) was used to measure motility.

4.3.7 Genotypic characterisation of colony morphotypes

Whole-genome sequencing and the subsequent bioinformatic analysis were undertaken entirely by Dr Rebecca Weiser. The methodology and subsequent results have been included in this thesis to detail the complete morphotype characterisation undertaken by this research group. The methods used for whole genome sequencing and bioinformatic analysis are listed in the Appendix.

4.3.8 Cross resistance

<u>4.3.8.1 Acquisition of resistance of *Pseudomonas aeruginosa* in the presence of azithromycin</u>

The rate of acquisition of resistance to azithromycin (AZM) at a sublethal concentration 2-fold lower than the MIC level (AZM; 8 µg/ml) was conducted in the presence and absence of 2% OligoG CF-5/20 using the biofilm bead model (**Figure 4.1**). The AZM concentration used was determined based on pilot studies (data not shown) and established literature MICs (Testing 2020), ensuring bacterial growth and potential development of resistance. Borosilicate beads were placed in 1 ml AZM (8 µg/ml) with or without 2% OligoG CF-5/20 in 24-well plates prior to inoculation with PAO1 cultures (5×10^4 CFU/ml; n = 4) and incubated (80 rpm; 37°C). Beads were transferred into the corresponding wells in a fresh plate 3 times/week (**Figure 4.1**) for 45 days. Purity plates were performed weekly on each test well on BA plates. At days 21 and 45, samples were taken from all wells and frozen at -80°C on microbank beads to be used for cross-resistance studies.

4.3.8.2 Cross resistance to a range of antibiotics

The enriched mixed populations grown in the presence and absence of 2% OligoG CF-5/20 (± AZM) were tested for cross-resistance against a range of antibiotics commonly used (ceftazidime (CEF), ciprofloxacin (CIP), azithromycin (AZM), oxy-tetracycline (Oxy-TET), levofloxacin (LEV), colistin (COL), aztreonam (ATM), meropenem (MER), rifampicin (RIF) and tobramycin (TOB)). These antibiotics were selected to represent a range of mechanisms of action. The day 21 biofilm bead cultures were vortexed in fresh medium and grown overnight (37°C; shaken). Cross-resistance was tested using an MIC assay. MIC assays were performed using a broth microdilution method in MH broth as previously described (Khan et al. 2012c) in accordance with standard guidelines (Jorgensen and Turnidge 2015). In brief, overnight cultures (n=3) were grown in TSB (LabM) and diluted in MHB to McFarland standard 0.5 (approximately 10^8 CFU/ml). Antibiotics included in each assay were added to MHB and serially diluted across the 96-well plate. Bacterial cultures were further diluted 10-fold in MHB and inoculated into the antibiotic serial dilutions (final concentration 5 x 10^5 CFU/ml). Plates were incubated statically (37°C; 16-20 h) and MICs read as the lowest concentration with no visible growth. These were conducted for the 4 biological repeats with WT PAO1 as control.

4.3.8.3 Loss of resistance

Following day 45 of transfer for the AZM study, the stability of AZM resistance was analysed by transferring beads into MH broth or MH broth with 2% OligoG CF-5/20 for a further six transfers. Biofilm growths from each transfer were sub-cultured onto LB agar plates with or without AZM (32 µg/ml), and resistance to AZM was recorded if there was growth on the plates after 24 h.

4.3.9 Statistical analysis

The significance of the crystal violet biofilm assay data was assessed using one-way analysis of variance (ANOVA) followed by Dunnett's multiple-comparison *post hoc* test using Graph Pad Prism 8. The motility assay data were analysed using a nonparametric Kruskal-Wallis test performed on IBM SPSS, followed by a *post hoc* adjustment by the Bonferroni correction. Statistical analyses for the genotypic characterisation were performed using R statistical software (Team 2013) to determine differences between the numbers of mutations in coding regions observed in each population (Control day 21, Control day 45, 2% OligoG CF-5/20-exposed day 21 and 2% OligoG CF-5/20-exposed day 45). As the numbers of mutations in all populations were found to be nonnormally distributed (Shapiro-Wilk test) with equal variances (Levene's test), a nonparametric Kruskal-Wallis test with *post hoc* pairwise comparisons using the Wilcoxon rank sum test and Benjamini-Hochberg adjustment was used to determine differences between the population medians. Genotypic characterisation statistics were analysed by Dr Rebecca Weiser. Differences were considered significant for p-value ≤ 0.05.

4.4 Results

4.4.1 Morphotype characterisation

4.4.1.1 Morphotype characterisation of biofilm-evolved isolates

Individual bacterial colony morphotypes were characterised according to size ([a] small [pin point] <1 mm; [b] medium 1 to 3 mm; and [c] large >3 mm) and surface texture (**Figure 4.3**). Further characterisation of the morphotypes included loss of pigmentation, mucoidy, opacity, halo and margin (**Table 4.3**). As a defined core set of SCV genes has yet to be documented (and with commonality between phenotypic and genotypic changes within different SCV populations not always apparent), for the purposes of this study SCVs were defined as pinpoint colonies formed within 72 h (<1mm in diameter) (Johns et al. 2015; Oakley et al. 2021). The number of colonies with different morphotypes within each well was also determined.

Morphotypic characterisation of the biofilm-evolved monoculture isolates (identified within the agar plates) revealed that 40 different morphotypes were evident across the study. From 4 biological repeats, the morphotypes included: Control day 21: 13 isolates; 2% OligoG CF-5/20 day 21: 12 isolates; Control day 45: 10 isolates; 2% OligoG CF-5/20 day 45: 5 isolates (**Table 4.3**).
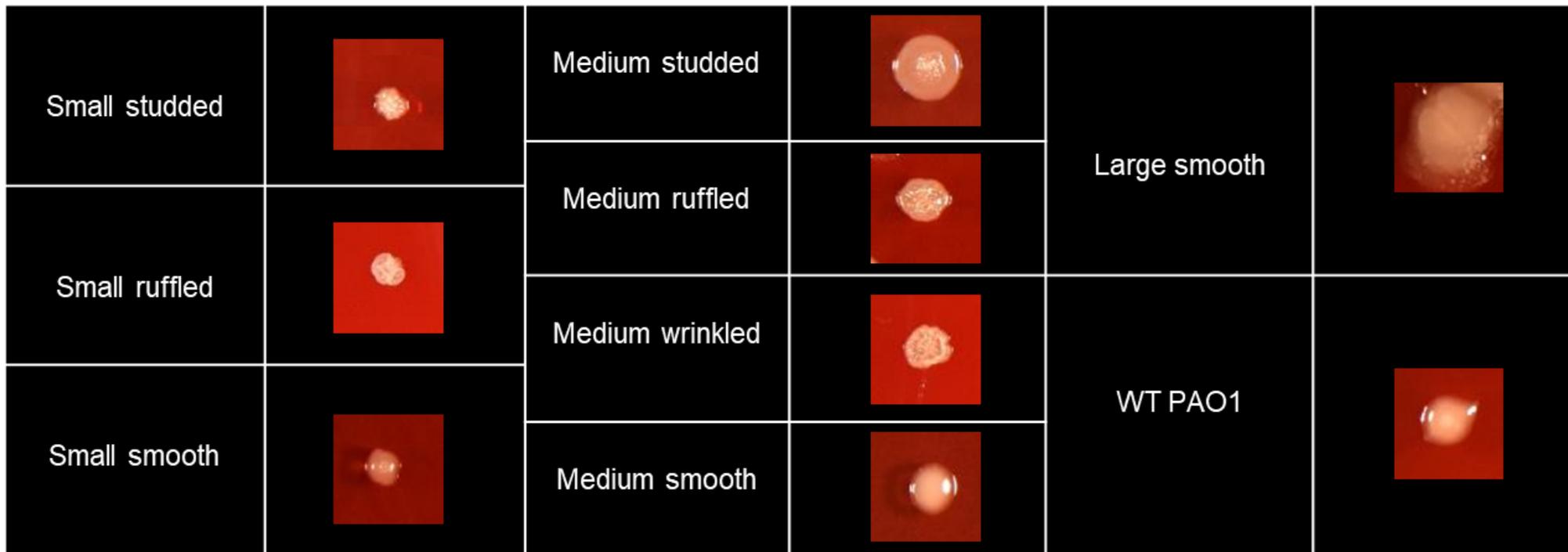
Figure 4.3. Appearance of representative morphologies on blood agar plates from control and 2% OligoG plates, shown to scale.

Morphotypes are described according to size (Small <1 mm, medium 1 to 3 mm, large >3 mm) and texture (Studded/ruffled/wrinkled/smooth). WT PAO1 is wild type PAO1, used as control.

Table 4.3. Morphotypes isolated from the bead biofilm model.

Colony morphology for biofilm colony morphotypes isolated from the wells in the presence and absence of 2% OligoG CF-5/20 at days 21 and 45. Genome sequencing code listed; full sequencing information detailed in Table I. i. in the Appendix.

Key: S; Small, M; Medium, L; large, Y; Yes, N; No, Ru; Ruffled, St; Studded, Sm; Smooth, W; Wrinkled; R, Round margin, NM; No margin.

Transfer	Sample	Size	Loss of pigmentation	Mucoid	Opacity	Surface texture	Halo	Margin	Genome seq. code
Day 21	Control	S	N	N	Y	Ru	N	R	C1a
		S	N	N	Y	Ru	Y	R	C2a
		S	N	N	Y	St	Y	R	C3a
		S	N	N	Y	St	Y	R	C4a
		S	N	N	N	St	Y	R	C5a
		M	N	Y	N	Sm	Y	R	C6a
		M	N	Y	N	Sm	N	R	C7a
		M	N	Y	N	Sm	N	R	C8a
		M	N	N	N	Sm	N	R	C9a
		M	N	N	Y	Sm	N	R	C10a
		M	N	N	Y	St	Y	R	C11a
		M	N	N	N	St	N	R	C12a
		L	N	N	N	Sm	Y	R	C13a
	2% OligoG CF-5/20	S	N	N	Y	Sm	Y	R	O1a
		S	N	N	Y	Ru	Y	R	O2a
		S	N	N	Y	Ru	Y	R	O3a
		S	N	N	Y	Ru	N	R	O4a
		M	N	N	Y	St	Y	R	O5a

		M	N	Y	N	Sm	Y	R	O6a
		M	N	N	N	St	Y	R	O7a
		M	N	N	N	St	Y	R	O8a
		L	N	N	N	Sm	Y	R	O9a
		L	N	N	Y	Sm	Y	R	O10a
		L	N	N	Y	Sm	Y	R	O11a
		L	N	N	N	St	N	R	O12a
Day 45	**Control**	S	N	N	Y	Sm	Y	R	C1b
		S	N	N	Y	Sm	Y	R	C2b
		M	N	N	Y	Sm	Y	R	C3b
		M	N	N	Y	Sm	Y	R	C4b
		M	N	N	Y	St	Y	R	C5b
		M	N	N	Y	Ru	N	R	C6b
		M	N	N	Y	Ru	N	R	C7b
		L	N	Y	N	Sm	Y	NM	C8b
		L	N	Y	N	Sm	Y	NM	C9b
		L	N	Y	N	Sm	Y	NM	C10b
	2% OligoG CF-5/20	S	N	N	Y	Sm	Y	R	O1b
		M	N	N	Y	St	Y	R	O2b
		M	N	N	Y	St	Y	R	O3b
		M	N	N	Y	W	Y	R	O4b
		M	N	N	N	Ru	N	R	O5b

4.4.1.2 Morphotype diversity following prolonged exposure to 2% OligoG CF-5/20

Morphotype diversity decreased between day 21 and day 45. At day 21, the 2% OligoG CF-5/20 samples were divided into three equal groups of small, medium and large colonies. The control group demonstrated 38% SCV ('pin-prick sized' colonies (Johns et al. 2015) and 54% medium colonies. By day 45, SCV represented 20% of both control and 2% OligoG CF-5/20 colonies; medium colonies made up 50% of control samples and there were no longer any large colonies in the 2% OligoG CF-5/20 samples (**Figure 4.4a**). Overall, a decrease in morphotype diversity was evident in both the control and 2% OligoG CF-5/20-treated group by day 45.

4.4.1.3 Characterisation of SCVs

A total of 12 different small colony variant (SCV) morphotypes were isolated from the samples. SCVs were evident across the study, with fewer morphotypes seen from days 21 to 45 in both control and treated groups. One of the original biological repeats did not develop any SCV in either the control or 2% OligoG CF-5/20 wells. The remaining three biological repeats demonstrated a reduction in SCV within the 2% OligoG CF-5/20 treated wells, with only one small smooth colony identified at day 45. The colony morphotype for small colonies was more variable at day 21 compared to day 45 (**Figure 4.4a.**). Small colonies demonstrated ruffled and studded morphologies in the control samples and smooth and ruffled in the 2% OligoG CF-5/20 treated samples at day 21. However, this was lost by day 45 with all clones showing smooth colony morphologies. In contrast, the medium sized colonies demonstrated both smooth and studded textures at day 21 and became more diverse by day 45 for control samples (smooth, ruffled and studded) and 2% OligoG CF-5/20 samples (ruffled, studded and wrinkled). The vast majority of the large colony morphotypes (87%) were smooth at both timepoints (**Figure 4.4b.**).

4.4.1.4 Characterisation of mucoidy morphotypes

There was only one mucoid colony within the 2% OligoG CF-5/20 samples (medium-sized, smooth, day 21), whereas the control samples had six mucoid colonies over all, affecting equal numbers of medium smooth colonies at day 21 and large smooth colonies at day 45 (**Table 4.3**).

Figure 4.4. Comparison of biofilm beads exposed to 0% and 2% OligoG CF-5/20 over 21 and 45 days. A) Categorising numbers of morphotypes (small, medium and large colonies) isolated from each growth condition. B) Subcategorising surface textures of small, medium and large colonies (n =4).

A) Size of morphotypes: Untreated (Control) samples, blue; Treated (2% OligoG CF-5/20), red. B) Texture of morphotypes: smooth, blue; ruffled, red; studded, yellow; wrinkled, green.

A.
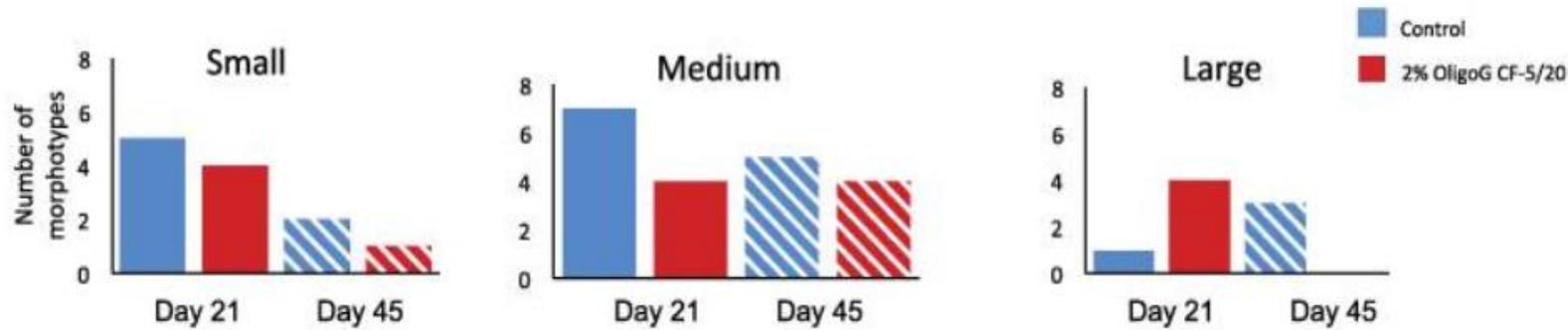
B.
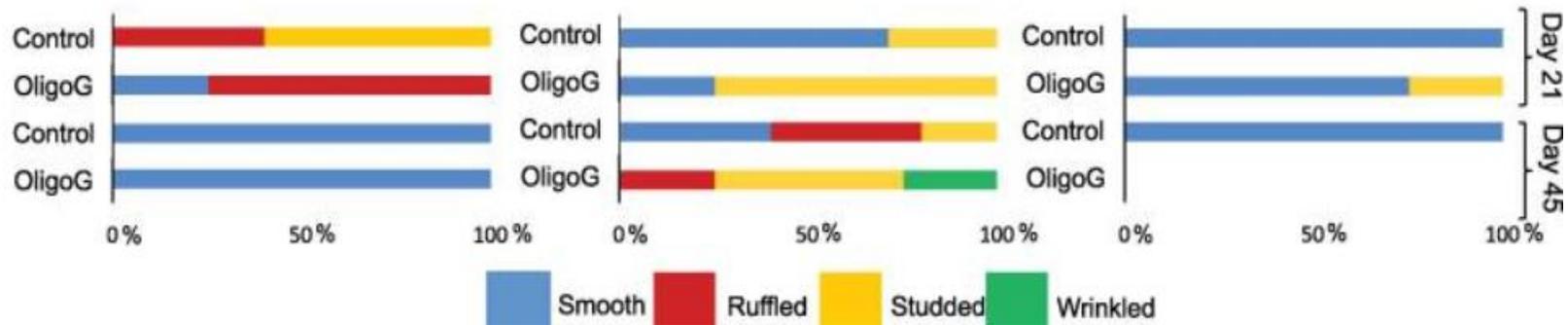
4.4.2 Phenotypic characterisation of the biofilm well colonies

Phenotypic characterisation of different morphotypes evolving during the bead biofilm model was investigated using scanning electron microscopy (SEM), crystal violet (CV) biofilm assays, confocal laser scanning microscopy (CLSM), and motility studies.

4.4.2.1 Scanning electron microscopy

Untreated control samples were visualised under SEM according to their subclassification (at 21 days) (**Table 4.1**). Small, ruffled and studded clones demonstrated formation of medium to large biofilm microcolonies, with the former subtype also forming EPS-encased biofilms bound to the glass surface (**Figure 4.5**). Medium-sized clones formed small clusters of microcolonies, inter-linked with branching cells. The smooth mucoidal subtype attached less firmly to the glass slide, whilst the studded subtype produced a combination of spherical microcolonies and more flat homogenous biofilms attached to the glass surface. Large-sized clones had a similar appearance to the monolayer biofilm produced by the wild-type WT PAO1, although the cells appeared more rounded in shape (**Figure 4.5**).

4.4.2.2 Biofilm formation assay

Results from the crystal violet assay (O'Toole 2011; Santiago et al. 2016) highlighted the inherent variability between the biofilm-forming abilities of the different morphotypes when compared to WT PAO1 (**Figure 4.6**). The majority of SCV (Control and 2% OligoG CF-5/20) had an increased biofilm-forming ability compared to WT PAO1. However, the only SCV found in the day 45 2% OligoG CF-5/20 treated samples actually demonstrated a statistically significant reduction in its biofilm forming ability (p <0.05). The biofilm-forming abilities of medium-sized colonies in the control samples was greater than those from 2% OligoG CF-5/20 treated samples, both at day 21 and day 45. There was no significant different in the biofilms formed by large colonies when compared to WT PAO1 (p>0.05) (**Figure 4.6**).

4.4.2.3 Confocal Laser scanning microscopy

Confocal laser scanning microscopy (CLSM) was performed on 24-hour biofilms grown using a selection of SCVs from the biofilm wells (MHB ± 2% OligoG CF-5/20) (**Table 4.2**). Using LIVE/ DEAD© stain, CLSM demonstrated a variability between the biofilm-forming abilities of different SCVs (**Figure 4.6**). Imaging supported the findings

of the CV assay, with small smooth morphotypes at day 45 showing less dense and more sparse biofilms in 2% OligoG CF-5/20-treated samples compared to the untreated control.
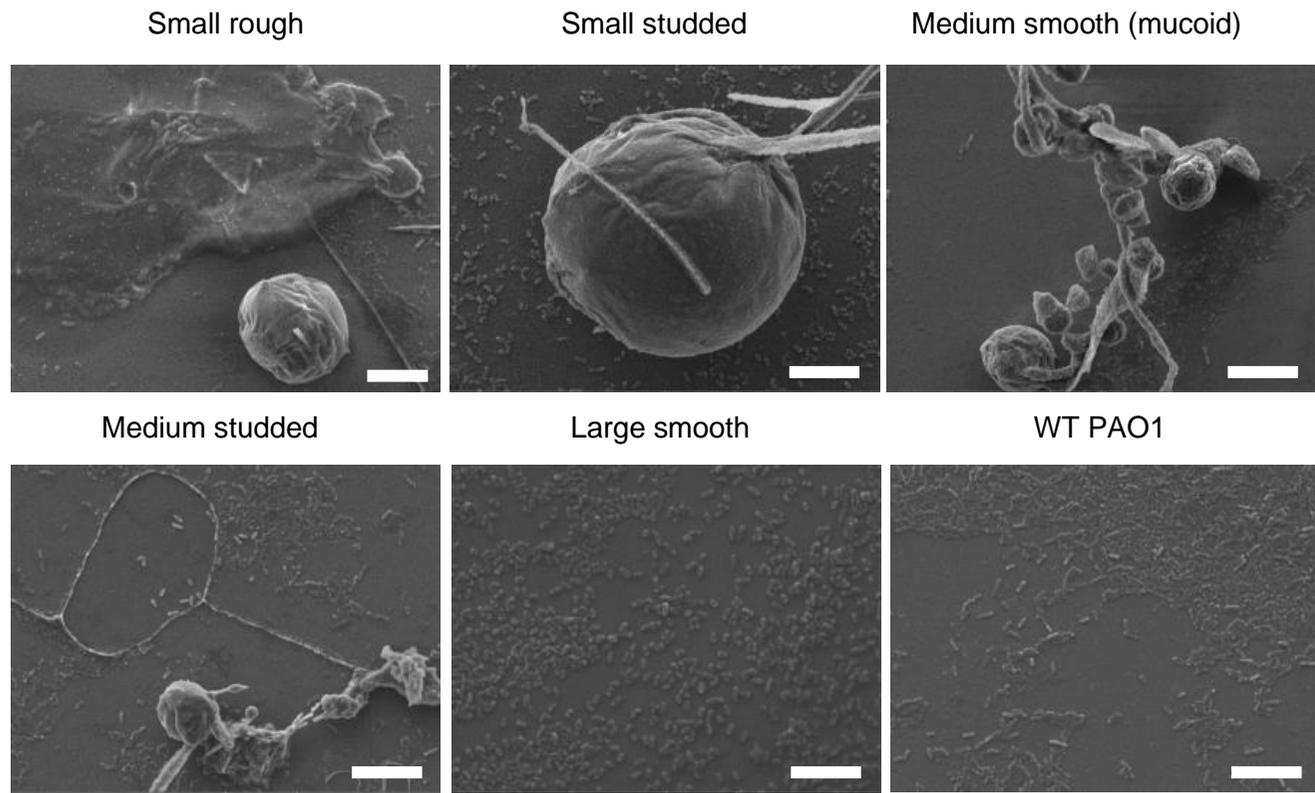
Figure 4.5. Biofilm formation ability of selected sample of morphotypes from control samples using scanning electron microscopy

Day 21; scale bar: 10 µm. Small ruffled and studded clones demonstrated formation of medium to large biofilm microcolonies. Medium-sized clones formed small clusters of microcolonies. Large-sized clones had a similar appearance to the monolayer biofilm produced by the wild-type WT PAO1.

Figure 4.6. Crystal violet quantification of biofilm-forming ability of all morphotypes from control and 2% OligoG CF-5/20 biofilm wells (24 h) with example confocal laser microscopy of small colony variants biofilms.

* Significant differences compared to WT PAO1; p-value <0.05; n=3. CLSM: LIVE/DEAD® stained; scale bar: 20 µm. Live cells (with an intact cell membrane) are green/yellow in colour due to staining with the intensely fluorescent calcein provided by the LIVE/DEAD staining. Dying and dead cells within the biofilm are red due to taking up the bright red fluorescent component of the LIVE/DEAD staining kit.
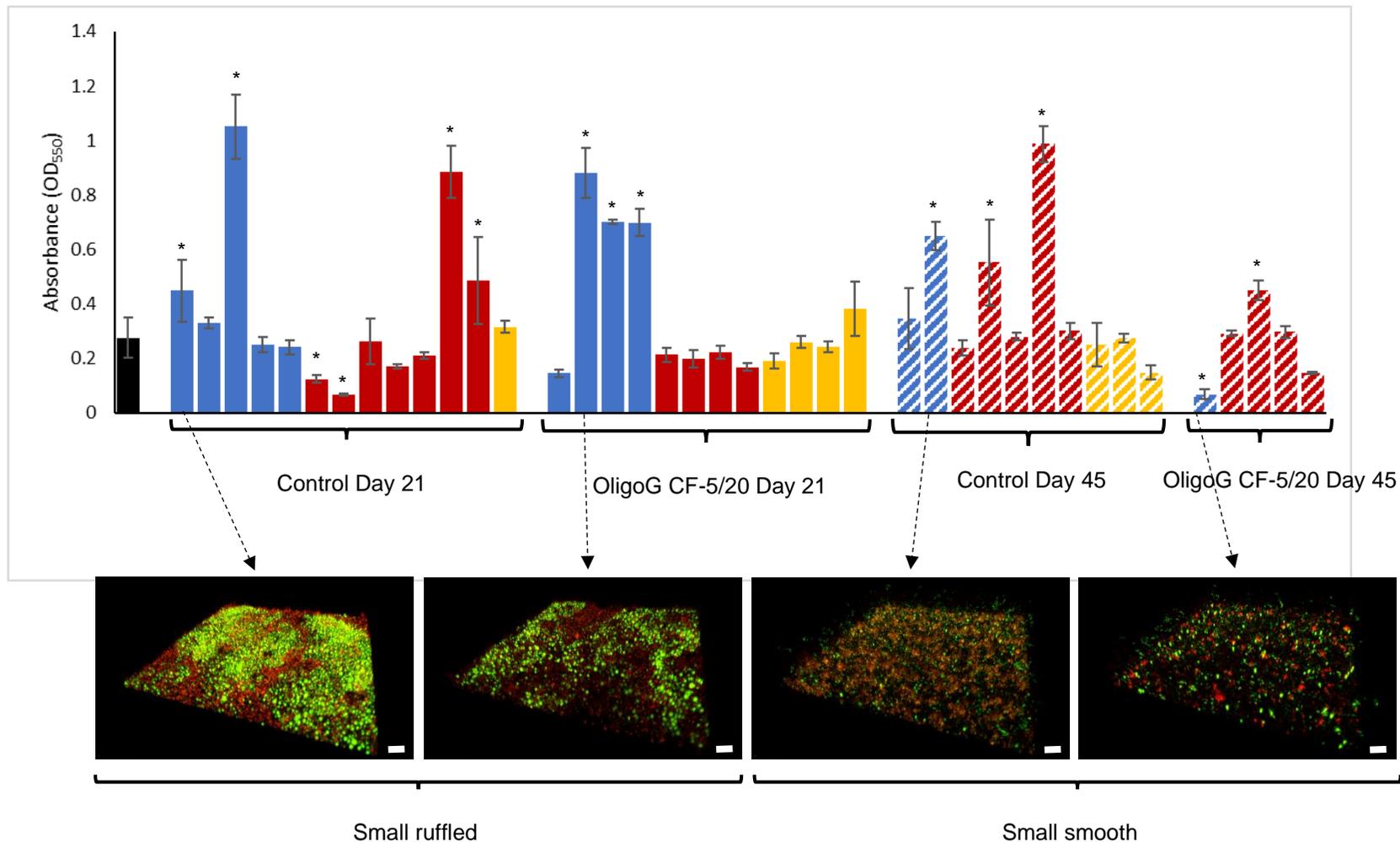
4.4.2.3 Altered motility profiles following exposure to 2% OligoG CF-5/20

Motility of all morphotypes isolated was measured using twitching, swarming and swimming assays, and compared to the WT PAO1. Radar charts provide an overview of the median twitching, swarming and swimming motility (millimetres) compared to WT control for day 21 (**Figure 4.7a**) and day 45 (**Figure 4.7b**).

A.
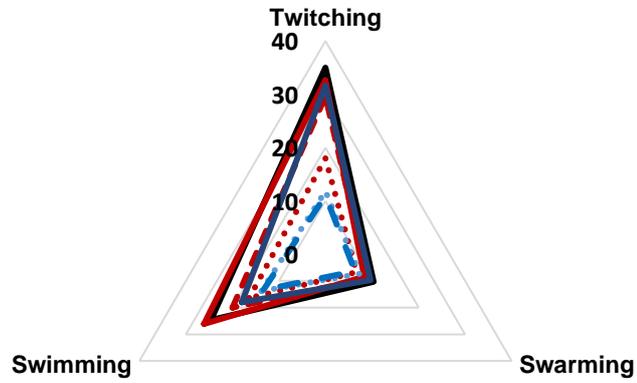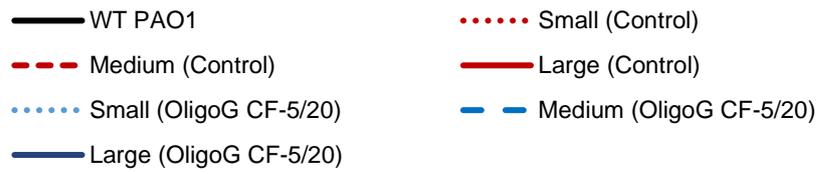
B.
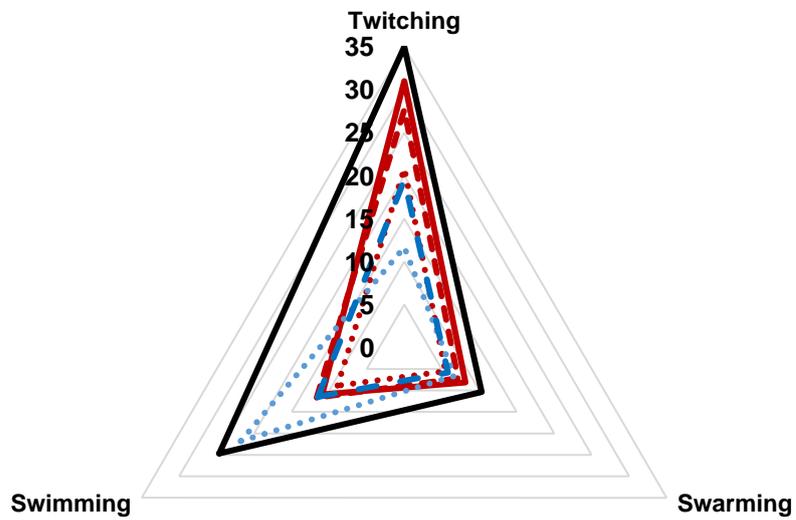
Figure 4.7. Radar charts demonstrating median twitching, swimming, and swarming motility (millimetres) compared to WT control. A. At day 21; B. at day 45. Charts demonstrate maximal motility in WT control at both time points.

4.4.2.3.1 Twitching

Figure 4.8 shows the mean diameter for all morphotypes (millimetres) as measured on agar plates. Comparisons are made to WT PAO1. All SCV morphotypes had less twitching motility than WT PAO1 in both control samples (range: -10% to -80% of WT PAO1) and 2% OligoG CF-5/20-treated samples (range: -27% to -66% of WT PAO1) at day 21 and day 45. Medium and large colonies also demonstrated a decreased twitching ability compared to WT PAO1, with the large colonies demonstrating the greatest twitching ability (range +3% to -20%) when compared to WT PAO1. Examples of PAO1 (wild type and mutants) motility on twitching assay agar plates are shown in **Figure 4.9**.
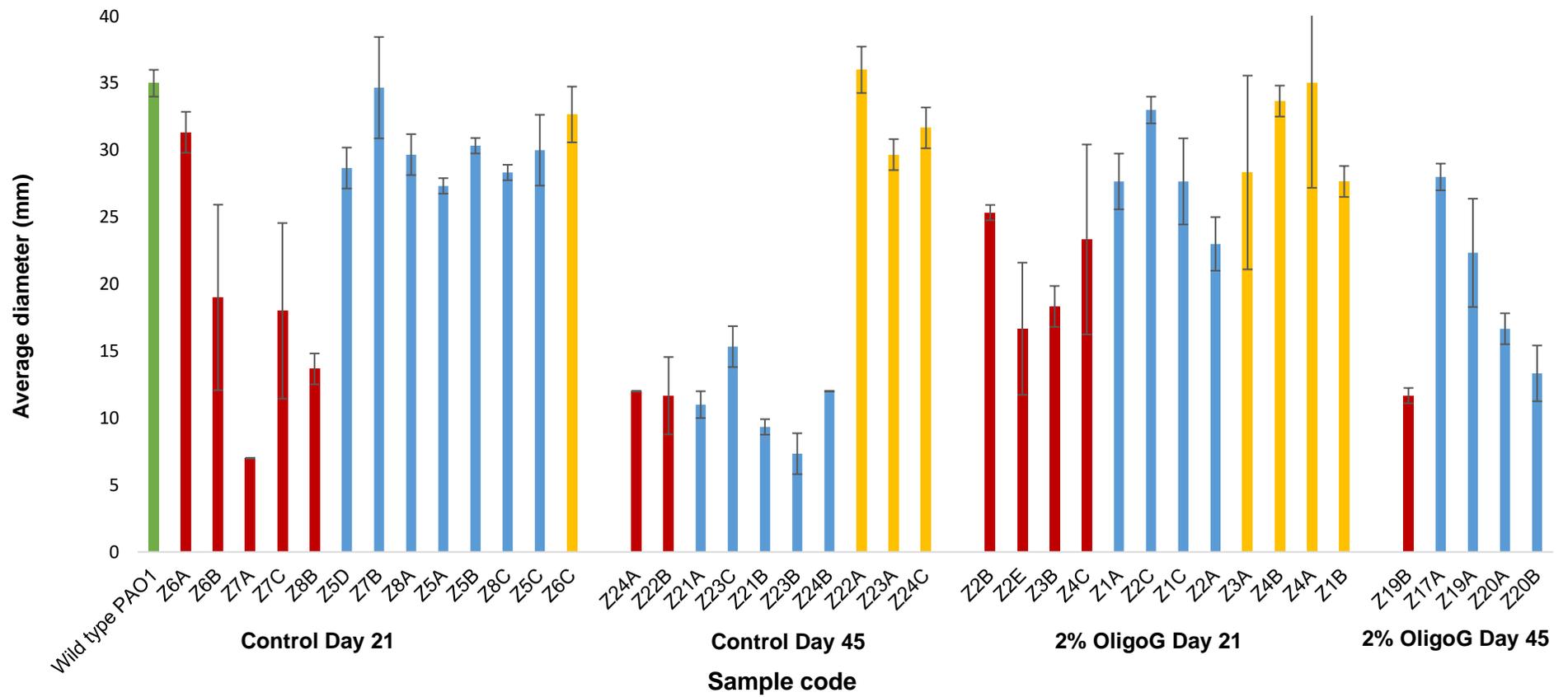
Figure 4.8. Twitching assay in all morphotypes from biofilm wells, plus WT PAO1 (Control and 2% OligoG; days 21 and 45).

Mean diameter (mm) (n=3) with standard deviation.
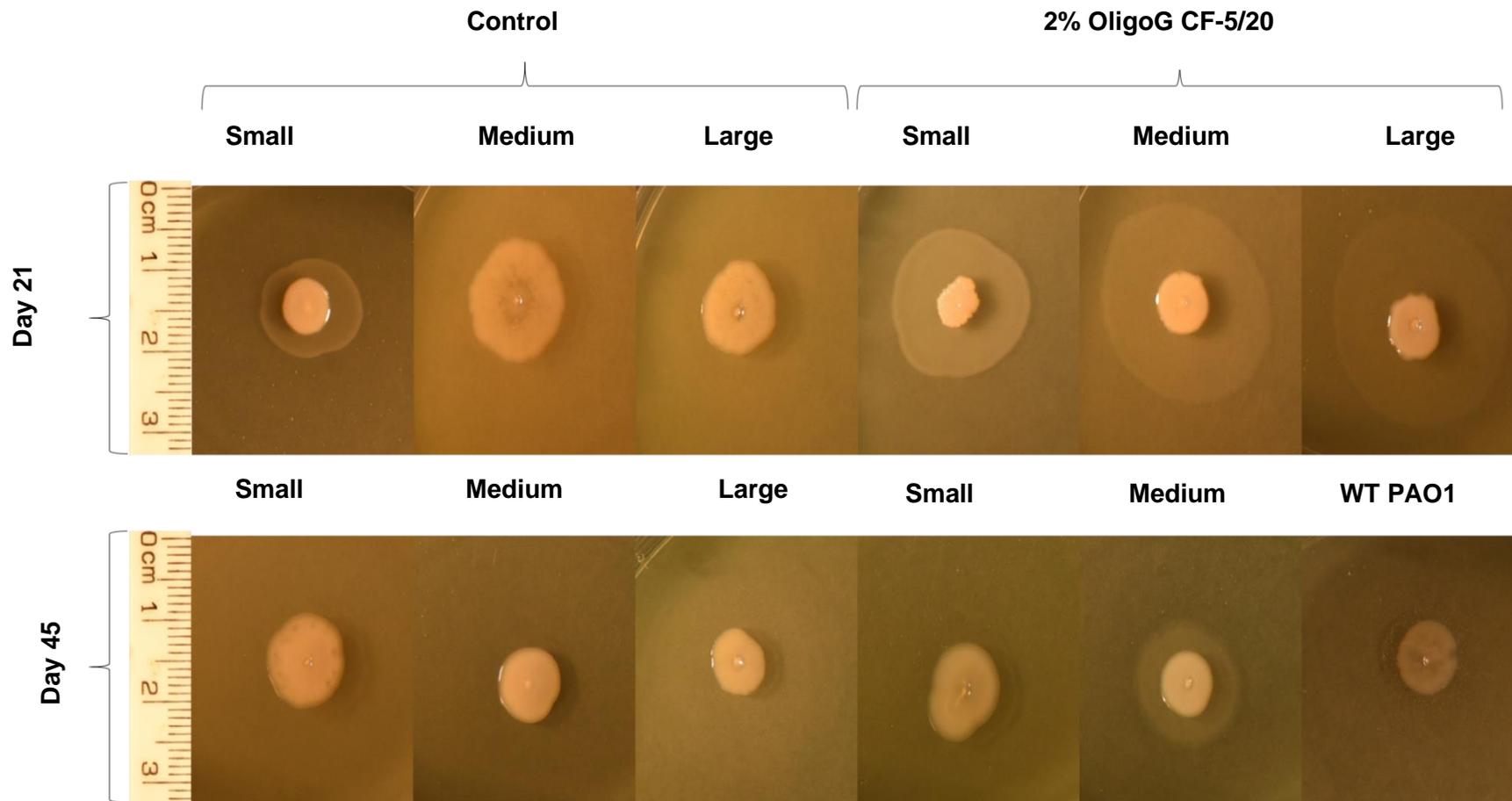
Figure 4.9. Examples representing average pattern of twitching according to size of colony for Control and 2% OligoG biofilm wells, plus WT PAO1 (wild type PAO1 representing day 0).

Images represent the average diameter (mm) of each morphotype according to size (small, medium and large); Scale bar as shown.

<u>4.4.2.3.2 Swarming</u>

Figure 4.10 shows the mean diameter for all morphotypes (millimetres) as measured on agar plates. Comparisons are made to WT PAO1. Similar trends were demonstrated for swarming as seen in the twitching assay, with the small colonies having the least ability to swarm, and the larger colonies swarming the most (**Figures 4.10** and **Figure 4.11**). Notably, of the 5 samples with a greater swarming ability than WT PAO1, 80% were mucoid and only 20% were from 2% OligoG CF-5/20 treated wells (day 21). Examples of WT PAO1 and biofilm well morphotypes' motility on swarming assay agar plates are shown in **Figure 4.11**.
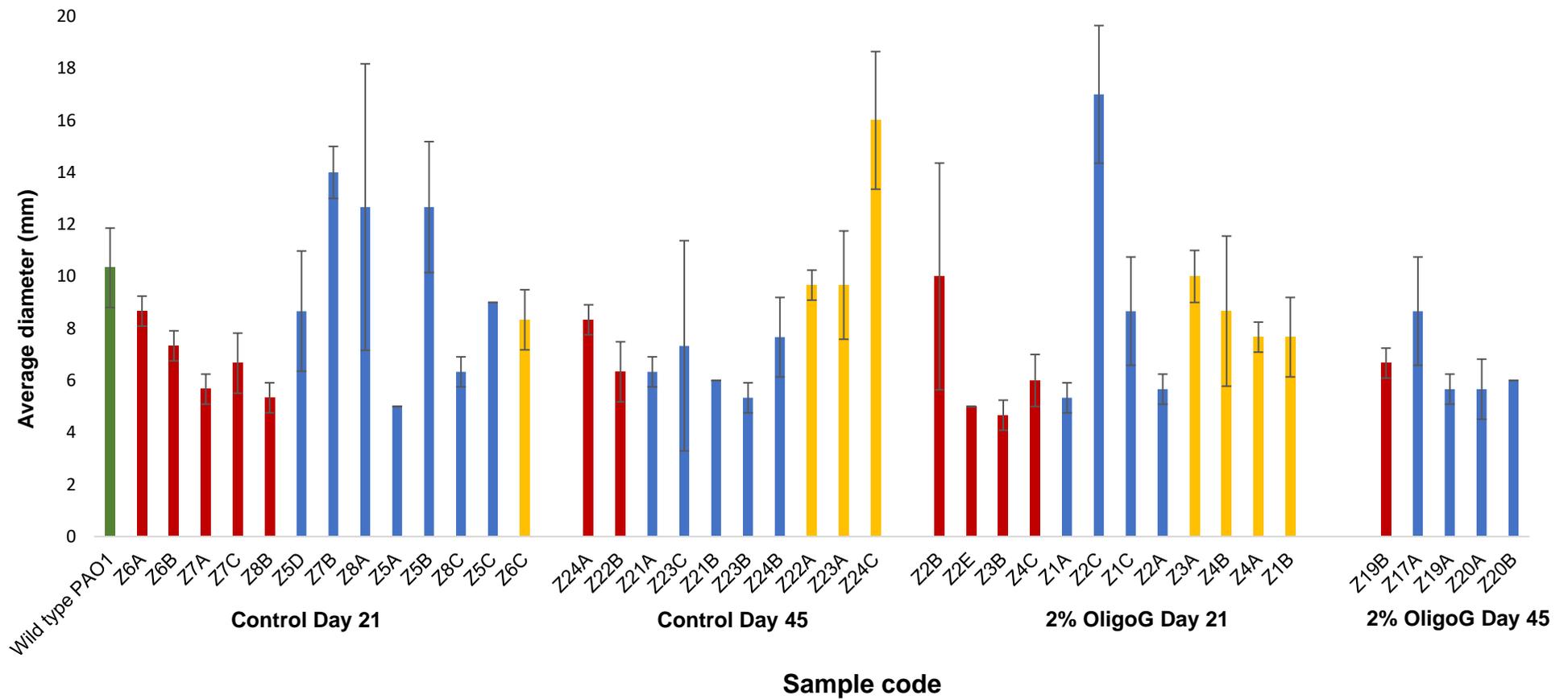
Figure 4.10. Swarming assay in all morphotypes from biofilm wells (Control and 2% OligoG CF-5/20; days 21 and 45), plus WT PAO1

Mean diameter (mm) (n=3) with standard deviation.

Figure 4.11. Examples of swarming assay plates for Control and 2% OligoG CF-5/20 biofilm wells shown according to size, plus WT PAO1 (wild type PAO1 representing day 0).

Images represent the average diameter (mm) of each morphotype according to size (small, medium, large).
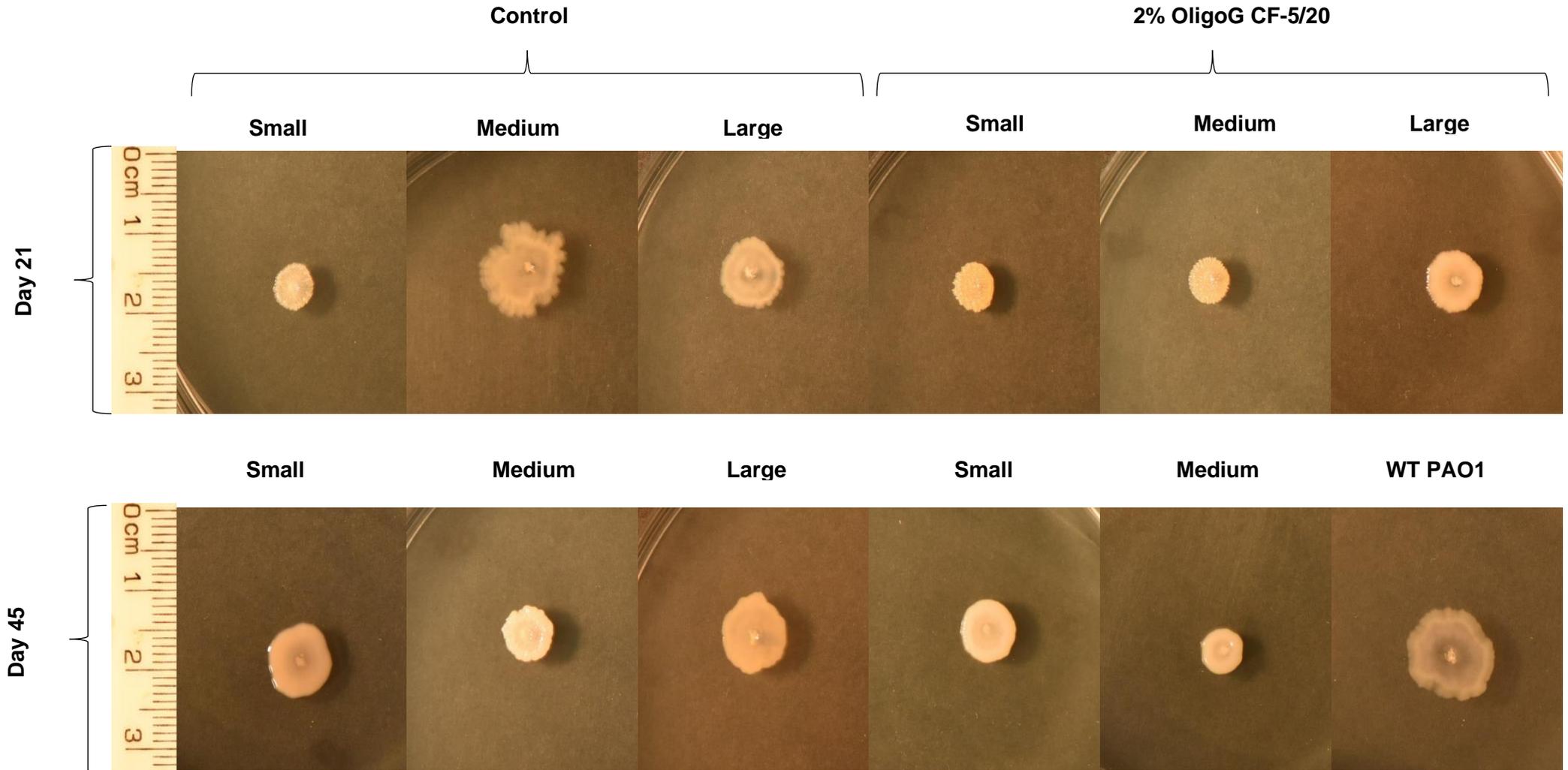
<u>4.4.2.3.2 Swimming</u>

Figure 4.12 shows the mean diameter for all morphotypes (millimetres) as measured on agar plates. Comparisons are made with WT PAO1. For all colony sizes, 2% OligoG CF-5/20-treated mutants were less able to swim at day 21 compared to the control (**Figure 4.12**). Notably, the large colony control sample demonstrated a similar swimming ability to WT PAO1. This trend was lost by day 45, with control samples of all sizes becoming less motile. By day 45, the 2% OligoG CF-5/20-treated small colony morphotype showed a swimming ability similar to that of WT PAO1, whilst the medium colonies' ability remained similar between days 21 and 45. Examples of PAO1 (WT PAO1 and biofilm well morphotypes) motility on swimming assay agar plates are shown in **Figure 4.13**.
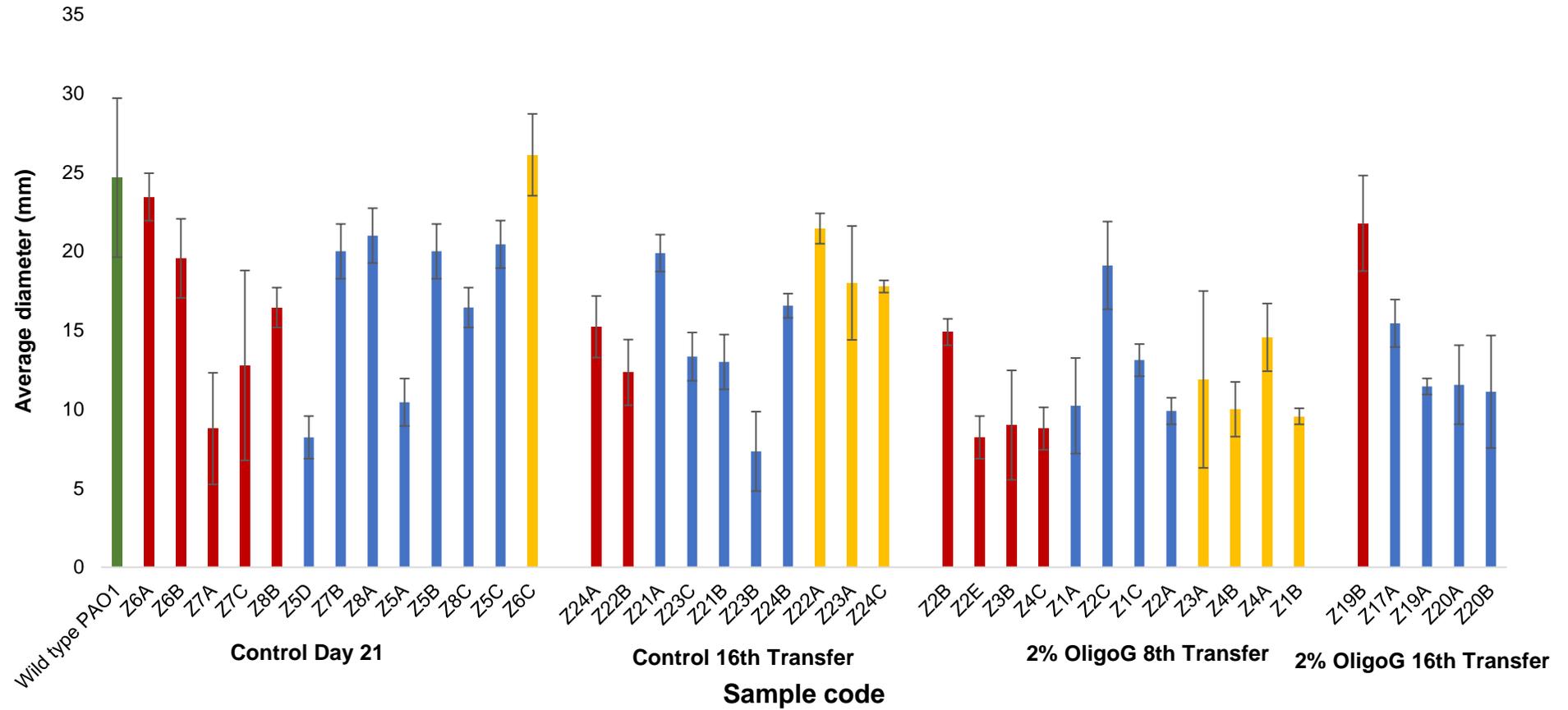
Figure 4.12. Swimming assay in all morphotypes from biofilm wells (Control and 2% OligoG CF-5/20; days 21 and 45), plus WT PAO1

Mean diameter (mm) (n=3) with standard deviation.

Figure 4.13. Examples representing average pattern of swimming according to size of colony for Control and 2% OligoG CF-5/20 biofilm wells, plus WT PAO1 (wild type PAO1 representing day 0).

Images represent the average diameter (mm) of each morphotype according to size (small, medium, large).
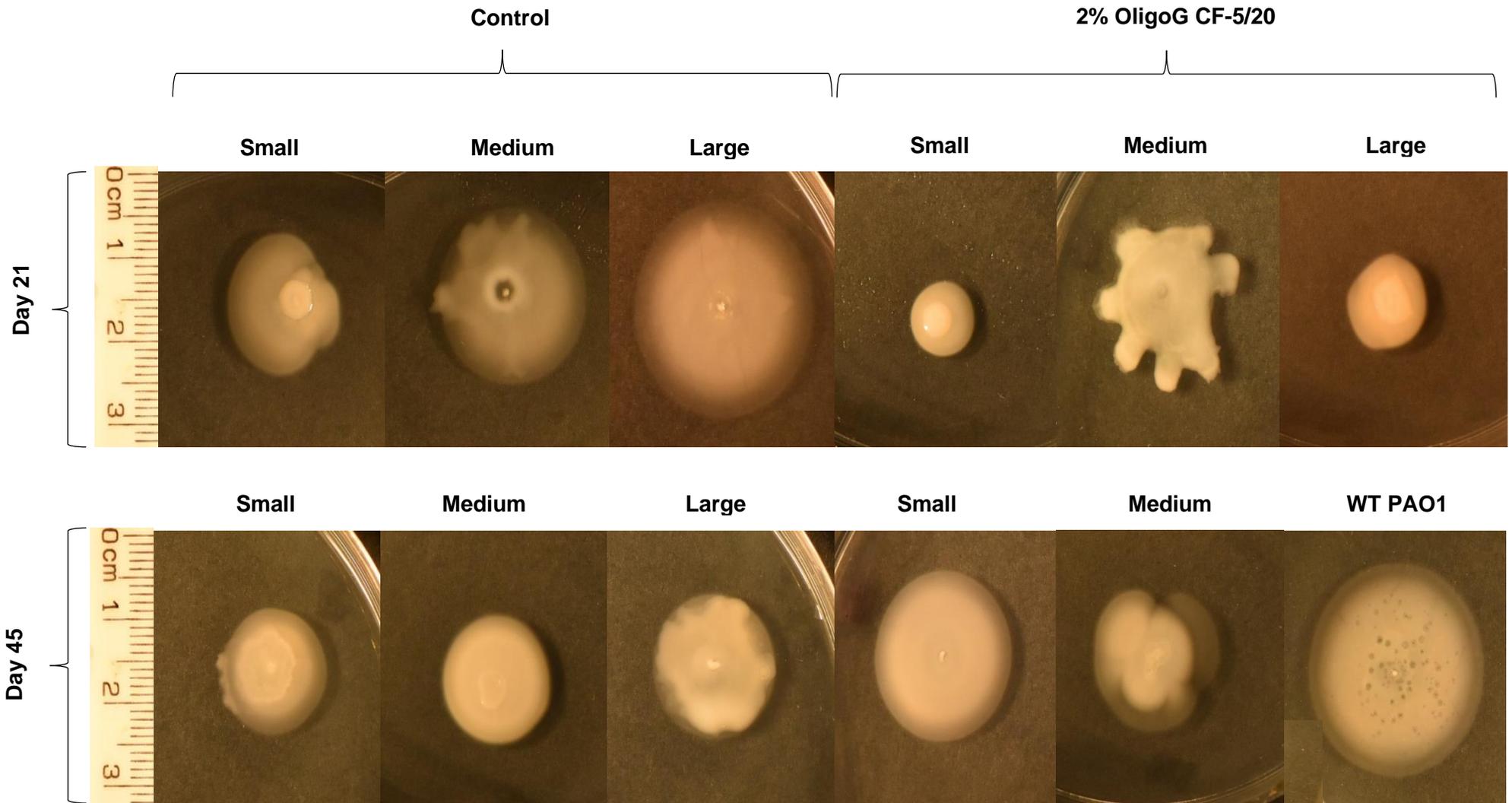
4.4.3 Genetic diversity of PAO1 isolates evolved in 2% OligoG CF-5/20

The full details of whole genome sequencing and genetic diversity resulting from PAO1 bacterial evolution in the presence or absence of 2% OligoG CF-5/20 are described in Appendix I. This data was analysed by Dr Rebecca Weiser.

Genome resequencing was performed to determine genetic changes associated with the evolution of PAO1 biofilm populations, with or without exposure to 2% OligoG CF-5/20. To summarise, there was a significantly higher number of mutations in the day 45 transfer isolates than the day 21 transfer isolates. However, there was no difference seen in the number of mutations occurring between control and 2% OligoG CF-5/20-exposed isolates. There was no evidence of adverse selective pressure to PAO1 isolates evolved in 2% OligoG CF-5/20.

4.4.4 Chronic exposure to 2% OligoG CF-5/20 and azithromycin

In the second arm of the study, the effect of the antibiotic azithromycin (AZM) (in the presence and absence of 2% OligoG CF-5/20) was examined in the bead biofilm model over 45 days (**Figure 4.1** and **Figure 4.2**). Azithromycin is a macrolide antibiotic commonly prescribed for prolonged use in patients colonised with chronic *P. aeruginosa* due to its anti-inflammatory and antibacterial effects (Nichols et al. 2020).

The enriched mixed populations (using the day 21 biofilm beads cultured for 24 hours in fresh medium) were tested for cross-resistance against a range of antibiotics commonly used in CF or representatives of different classes of antibiotics with different mechanisms of action (n=10). Antibiotics targeting key CF pathogens (including *P. aeruginosa)* used within cross-resistance study were comprised of azithromycin (macrolide), ciprofloxacin and levofloxacin (quinolones), colistin (polymyxin E), and aztreonam (monobactam) and tobramycin (aminoglycoside). As cephalosporins, carbapenem, rifamycin, and tetracyclines are also routinely used for CF treatment, ceftazidime, meropenem, rifampicin, and oxytetracycline, respectively, were also employed in the cross-resistance studies (Döring et al. 2012).

Prolonged exposure of PAO1 (grown in the presence and absence of 2% OligoG CF-5/20) demonstrated no change in resistance to azithromycin, both values being 32 µg/ml. However, subculturing at sub-MIC levels of AZM (8 µg/ml) in both the presence and absence of 2% OligoG CF-5/20 resulted in an increase in MIC values from 32 µg/ml to 256 µg/ml at day 21 (**Table 4.4**). This resistance to AZM (± 2% OligoG CF-5/20) was retained for up to 6 subsequent transfers (with no antibiotics) following completion of the 45-day experiment. Samples grown in MH broth only or 2% OligoG CF-5/20 demonstrated no change in resistance to other classes of antibiotics (**Table 4.5**). Biofilms grown in AZM with 2% OligoG CF-5/20 demonstrated a decrease (up to 3-fold) in resistance (MIC) to antibiotics such as aztreonam and oxytetracycline compared to biofilms growth in AZM alone (**Table 4.4**).

Table 4.4. The effect of OligoG CF-5/20 on the acquisition of resistance to azithromycin (AZM) on the whole bacterial population at day 21

MIC values (µg/ml) of cross resistance to other classes of antibiotics. Day 0 and day 21 MIC values listed for control and 2% OligoG for comparison. Antibiotics used: CAZ, ceftazidime; CIP, ciprofloxacin; AZM, Azithromycin; Oxy-TET, Oxy-tetracycline; LEV, levofloxacin; COL, colistin; ATM, aztreonam; MER, meropenem; RIF, rifampicin; TOB, tobramycin. Decrease in antimicrobial activity (MIC) is indicated by blue shaded areas, with a reduction of 2-fold considered clinically relevant.

	AZM	CAZ	CIP	Oxy-TET	LEV	COL	ATM	MER	RIF	TOB
Control (Day 0)	32	16	0.125	4	0.25	1	4	0.5	32	1
2%OligoG (Day 0)	32	16	0.0625	4	0.25	1	4	0.5	32	1
Control (Day 21)	32	16	0.0625	8	0.5	0.5	4	1	64	1
2%OligoG (Day 21)	64	32	0.125	8	0.5	1	4	1	64	1
AZM (Day 21)	256	512	8	128	8	1	8	2	64	0.5
AZM and 2% OligoG (Day 21)	256	128	4	16	4	0.25	1	0.5	64	1

Table 4.5. Cross-resistance of *P. aeruginosa* in the presence and absence of OligoG CF-5/20 in the bead biofilm model.

Whole-population samples at days 21 and 45 tested against a range of antibiotics (CAZ, ceftazidime; CIP, ciprofloxacin; AZM, azithromycin; oxy-TET, oxytetracycline; LEV, levofloxacin; COL, colistin; ATM, aztreonam; MER, meropenem; RIF, rifampicin; TOB, tobramycin). Samples grown in MH broth only or 2% OligoG CF-5/20 demonstrated no change in resistance to other classes of antibiotics.

MIC µg/ml	Transfer (day)	CAZ	CIP	AZM	Oxy-TET	LEV	COL	ATM	MER	RIF	TOB
Control	0	16	0.125	32	4	0.25	1	4	0.5	32	1
	21	16	0.0625	32	8	0.5	0.5	4	1	64	1
	45	32	0.5	64	8	0.5	0.5	4	2	32	0.5
2% OligoG CF-5/20	0	16	0.0625	32	4	0.25	1	4	0.5	32	1
	21	32	0.125	64	8	0.5	1	4	1	64	1
	45	16	0.125	32	8	0.5	0.25	4	2	32	0.25

4.5 Discussion

Experimental evolution enables the study of a select population and the changes that occur as a result of specific laboratory conditions (Steenackers et al. 2016). Such an approach is well suited to bacteria as they have short generation times and resultant isolates can be frozen and stored for later experiments, as used in this study (Steenackers et al. 2016). Previous large-scale evolutionary studies have clearly demonstrated that *P. aeruginosa* undergoes a transition from opportunistic pathogen to primary pathogen within the CF lung (Yang et al. 2011b). Colonising clonal populations show phenotypic and genetic adaptation over time enabling persistence, predominantly through biofilm formation (Boles et al. 2004; Goss and Burns 2007; Bjarnsholt et al. 2013).

Studies suggest that the development of diversity within bacterial communities is of paramount importance, as it enables survival despite environmental stressors (Boles et al. 2004). Incorporating specialised subpopulations with features such as differing motility, increased biofilm formation, being able to detach from biofilms more easily (hyper-detachment) and auxotrophy may be useful to communities where the environmental conditions fluctuate or are subject to change (Boles et al. 2004).

This study aimed to determine the effect of novel and existing antimicrobial therapies, including 2% OligoG CF-5/20 and the commonly-used antibiotic azithromycin, on experimentally-evolved biofilms, to analyse ecological succession. A study design adapted from Poltak and Cooper (2011) was successfully employed, using the well-characterised reference strain *P. aeruginosa* PAO1, which was originally isolated from a wound and is the most frequently used strain in laboratory studies of *Pseudomonas* genetics and phenotyping analyses (Klockgether et al. 2010). This model incorporated features such as regular cycles of surface colonisation, biofilm assembly and dispersal in the presence of the different treatment strategies, to mimic the long-term adaptive process *in vitro* (Poltak and Cooper 2010).

Bacterial adaptation as a result of evolution in the bead biofilm model was expected, but the exact effect of differing conditions, such as the presence of 2% OligoG CF-5/20, on the evolutionary patterns of this strain were to be determined. Analysis considered whether 2% OligoG CF-5/20 placed any selective pressure on biofilm growth. The bead biofilm model effectively demonstrated evolution of phenotypic and

genotypic characteristics of the monocultures and the constituent morphotypes (Steenackers et al. 2016).

Previous *in vitro* biofilm models have shown extensive diversification from a single progenitor within a week (Boles et al. 2004). Our study demonstrated morphotype diversification within the bead biofilm model at day 21, followed by decreased morphotype diversity by day 45, suggesting diversification was not evident over the longer term. Lack of sustained diversity more closely reflects patterns seen in *in vivo* studies of the evolving *P. aeruginosa* lineage (Yang et al. 2011b). It has been hypothesised that bacteria within biofilm communities have less exposure to environmental stressors, such as antimicrobials, than bacteria growing planktonically, which may reduce mutation rates (Ahmed et al. 2018). It is possible that, by enabling formation of an effective surface biofilm on the bead, our model reduced the need for ongoing diversity and therefore more closely reflected the *in vivo* environment seen in the CF lung.

Small colony variants (SCVs), sometimes called 'pin-prick-sized colonies' or 'minis', have been shown to emerge rapidly from *P. aeruginosa* biofilms (Boles et al. 2004; Johns et al. 2015). SCVs have been associated with poor clinical outcome in CF patients (Häußler et al. 1999) and the persistence of infection in animal models (Malone et al. 2010). SCVs are characterised by slow growth, often requiring longer incubation periods, antibiotic resistance and formation of visible aggregates in liquid culture (Drenkard and Ausubel 2002), as well as the aforementioned hyper-detachment phenotype (Boles et al. 2004). They can also have a lack of pigmentation, low virulence potential, reduced haemolytic and coagulase activity and reduced utilisation of carbohydrate (Johns et al. 2015). Boles et al. (2004) showed emergence of SCVs within a 5-day biofilm grown in a drip flow reactor, and described a subpopulation with hyper-detachment. They showed SCV's biofilms detached at a four-fold higher rate than that of wild-type PAO1, and suggested this may allow such colonies to relocate complex biofilm communities to newly colonised sites at times of stress (Boles et al. 2004). We demonstrated SCVs within 21 days, though monocultures containing SCVs had reduced by day 45. Within our population, most SCVs from control wells (days 21 and 45) demonstrated greater biofilm forming abilities than the wild-type progenitor, but 2/3 of those from 2% OligoG CF-5/20 wells grew less dense biofilms.

Authors previously suggested that the phenotypic changes seen in SCVs may be transient and bacteria may revert back to wild type when the environmental stressor, such as an antibiotic, has been removed (Drenkard and Ausubel 2002). However, it appears the process is more complicated, with some SCVs undergoing permanent genetic changes, some reverting to the WT phenotype and a third group, known as the revertant phenotype, developing a phenotype distinct from the progenitor and SCV on sub-culturing (Johns et al. 2015).

SCVs are often present in CF sputum samples. They are considered significant contributors to the overall population's antibiotic resistance, and therefore associated with difficult-to-treat *P. aeruginosa* infections (Drenkard and Ausubel 2002). Due to the slow growth rate, SCVs are often missed in standard culturing techniques, resulting in cessation of antibiotic treatment before the infection has completely cleared and persistence of chronic infection (Johns et al. 2015). Given the increased severity of infections associated with SCVs, it was reassuring to see that colonies grown in the presence of 2% OligoG CF-5/20 had fewer SCVs compared to the control at both time points. Though sample numbers are small, this reduction may have significant clinical advantages for the CF population if results are translated *in vivo*.

Within this study, SCVs demonstrated a similar trend in surface texture between day 21 and day 45 timepoints, both in control and 2% OligoG CF-5/20 treated biofilm samples. Interestingly, many medium and large colonies developed altered textures by day 45, which may be reflective of the emergence of the ruffled and wrinkled variants which also appeared later in the Poltak and Cooper (2011) experiment.

Biofilm drip-flow reactor model studies, using different wild-type variants such as PA14 and clinical CF isolates, demonstrated similar colony variant emergence. Authors hypothesised that this was related to cell-cell signalling within the biofilm system (Boles et al. 2004). Previous studies have demonstrated that OligoG CF-5/20 can modify the *lasI-lasR* and *rhlI-rhlR* QS system, perhaps providing an explanation for the different morphotypes formed in treatment wells compared to MH broth alone (Jack et al. 2018).

A number of studies have demonstrated changes in colony motility within biofilm models (Deziel et al. 2001; Boles et al. 2004). In particular, Boles et al. (2004) showed variability in swimming motility between different morphotypes compared to the wild

type. The authors hypothesised that such variation must be caused by multiple genetic changes induced as early as 5 days into biofilm growth (Boles et al. 2004). Similarly, Workentine et al. (2013) also found similar numbers of motile and nonmotile isolates with a large collection of patient samples in a chronically infected individual with CF (Workentine et al. 2013). Bragonzi et al. (2009) demonstrated loss of twitching and swimming motility in isolates from chronic infection. Using a strain reference panel, Cullen et al. (2015) showed a higher proportion of CF strains were non-motile when compared to non-CF strains, using twitching, swarming and swimming motility assays. They also demonstrated loss of motility over time of colonisation in CF strains and only two of the 42 strains demonstrated 'true' swarming abilities at all (Cullen et al. 2015).

A study using 1,030 *P. aeruginosa* isolates taken from 20 patients with CF over 10 years found that isolates from early colonisation were highly motile, and expressed flagellin and pilin. However, those from chronically infected patients were non-motile and did not express flagellin (Mahenthiralingam et al. 1994). When compared to environmental isolates, CF isolates were much more likely to be non-motile (1.4% vs. 39%) (Mahenthiralingam et al. 1994). Authors have hypothesised that the non-motile phenotype may persist due to its ability to resist phagocytosis, reduce energy expenditure and evade the host immune system (Mahenthiralingam et al. 1994; Amiel et al. 2010).

This study showed reduced motility in all pseudomonal SCVs at both day 21 and day 45, when compared to WT PAO1. This is reflective of results seen by Wang et al. (2015), which demonstrated that *Pseudomonas chlororaphis* SCVs had almost completely defective motility when compared to the WT (Wang et al. 2015) However, interestingly, the 2% OligoG CF-5/20 treated SCVs in this study maintained a swimming ability close to that of WT PAO1 at day 45, suggesting this colony was not behaving like a 'typical' SCV. It appears that prolonged growth in this biofilm model led to reduced motility in most colonies, regardless of growth in 2% OligoG CF-5/20 or MH broth alone.

Switching to the mucoidy phenotype is considered one of the hallmarks of *P. aeruginosa* adaptation to the lower airway environment (Workentine et al. 2013). The production of exopolysaccharide/alginate causes mucoidy and is associated with antibiotic resistance and resistance to host defences by phagocytosis (Li et al. 2005). Evidence supports initial colonisation with the non-mucoid phenotype, followed often

years later by demonstrable conversion to mucoid *P. aeruginosa* (Li et al. 2005). Li et al. (2005) showed 92% of their paediatric cohort grew the mucoidy phenotype by the age of 16 years old, with the median age of development being 13 years old, compared to the median age of 1 year for acquisition of non-mucoid *P. aeruginosa* (Li et al. 2005).

This study showed only 1 mucoidy isolate of 17 isolates arising from the 2% OligoG CF-5/20 biofilm wells over whole study (day 21 and day 45), compared to 26% (6 of 23) of the control biofilm well isolates (day 21 and day 45). Control well mucoidy samples had variable motility, but generally results were more similar to the motility seen in WT PAO1 than for non-mucoid samples. The 2% OligoG CF-5/20 mucoid morphotype demonstrated better swarming than WT PAO1 (64% greater movement), but similar twitching and swimming abilities. Interestingly, previous studies have shown that mucoidy does not necessarily correlate with loss of motility (Workentine et al. 2013). Biofilm forming ability was worse in all mucoid phenotypes, from both control and 2% OligoG CF-5/20 wells. This study clearly demonstrated a lower incidence of mucoid development in PAO1 when exposed to 2% OligoG CF-5/20. Unfortunately, it is not possible to draw any firm conclusions with such small numbers. However, the clinical implications for patients with CF are significant, with potential for lower rates of mucoid *P. aeruginosa* infection and resulting improved lung function and overall life expectancy if the results of this *in vitro* experiment were to be demonstrated *in vivo* (Li et al. 2005; Workentine et al. 2013).

Workentine et al. (2013) identified 169 clonal isolates from 34 sputum samples taken from a single patient over one year, at timepoints including clinical stability and exacerbations. Even when considering isolates with the same colony morphotype from one clinical sample, they showed a large degree of phenotypic variation (Workentine et al. 2013). This included variable antibiotic sensitivity, motility and protease activity, and similar results were seen across a further three patients whose sputum samples were subsequently analysed (Workentine et al. 2013). Similarly, Clark et al. (2015) showed that *in vivo* CF-evolved morphotypes were unreliable predictors of phenotype, with significant variability seen in features such as antimicrobial susceptibility. Therefore, phenotypic and genotypic experiments are essential to fully grasp the composition of the evolved population, as simply observing morphotype is apparently insufficient.

Using biofilm well-evolved isolates for whole genome sequencing demonstrated a spectrum of mutations consistent with those previously described in *P. aeruginosa*. Mutations involved biofilm formation, motility, chemotaxis and quorum sensing pathways. As a result of adaptation to the biofilm lifestyle, all but one of the observed mutations were nonsynonymous (Traverse et al. 2013; McElroy et al. 2014). Similar nonsynonymous mutations have been noted *in vivo* in *Burkholderia dolosa* where adaptive evolution in parallel was also demonstrated in multiple individuals (Lieberman et al. 2011).

The majority of acquired signal transduction gene mutations (*bifA*, *yfiR*, *wspA*, *wspF* and *morA*) are linked to intracellular cyclic-di-GMP levels. These are associated with the transition from a motile (planktonic) to sessile (biofilm) lifestyle (Kim and Harshey 2016). The genes *wspF* and *morA* have been previously implicated in *P. aeruginosa* biofilm adaptation and colony morphology changes, including development of wrinkled colonies and SCVs (Wong et al. 2012). Gene mutations directly related to motility were only seen in day 45 isolates. These included mutations in type IV pili (*pilY1*, *pilM* and *pilT*), which have been described previously in *P. aeruginosa* PAO1 biofilm models (McElroy et al. 2014; Ahmed et al. 2018). Overall, the mutations identified as a result of this bead-biofilm model reassuringly reflect those of previous biofilm model and CF clinical sample studies (Smith et al. 2006b; Winstanley et al. 2016b).

Ahmed et al. (2018) used an experimental evolutionary model including biofilm and planktonic *P. aeruginosa* cultures to observe development of antibiotic resistance to sub-inhibitory levels of ciprofloxacin. They reported a number of genomic mutations across their population, with ciprofloxacin-evolved biofilm cultures showing greater reduction in type IV-pilus-dependent motility (twitching), as seen in our study (Ahmed et al. 2018). Wong et al. (2012) investigated populations modelled in synthetic CF sputum medium and described parallel evolution of antibiotic resistance genes. Though there were a number of common genes affected, additional mutations were identified which were specific to individual experimental isolates (Wong et al. 2012). These studies found mutations in genes specifically linked to ciprofloxacin exposure (Mex efflux systems, DNA gyrases), and genes associated with adaptation to a CF-like environment (QS, motility, cyclic-di-GMP signalling) (Wong et al. 2012). This study demonstrated mutations in the *mexT* gene, a multidrug efflux transcriptional regulator (Tian et al. 2009) across all populations regardless of 2% OligoG CF-5/20 exposure.

Perhaps unsurprisingly, there was a significant increase in the number of mutations observed between day 21 and day 45 transfer populations. However, numbers were similar in control and 2% OligoG CF-5/20-treated isolates. Furthermore, there were no mutated genes in functional gene categories exclusive to 2% OligoG CF-5/20-exposed isolates alone, with the same categories being affected in both control and 2% OligoG CF-5/20-exposed populations. Therefore, this evolutionary model has demonstrated that 2% OligoG CF-5/20 did not drive mutations in specific genes, during the adaption of *P. aeruginosa* PAO1 to biofilm growth.

Antibiotics are grouped in classes according to the basis of their chemical structure. Members of the same class typically share the same target in a cell, such as protein synthesis for aminoglycosides, and also the same mechanisms of resistance (Périchon et al. 2019). It would be expected that resistance may be seen across multiple antibiotics within the same class. As such, antibiotics selected for this experimental model represent all major antibiotic classes. Antibiotic resistance can be intrinsic or acquired. *P. aeruginosa* demonstrates high innate resistance to antibiotics due to its low permeability of its outer membrane due to multidrug efflux systems and chromosomally-encoded β-lactamase. Acquired resistance can occur through various mechanisms, such as alteration of antimicrobial targets and upregulation of efflux pumps (Périchon et al. 2019). Genetic alterations can be seen following a mutation or horizontal transfer of genetic information.

Antibiotic resistance is common when bacteria are chronically exposed to antibiotics. The effect of prolonged exposure to commonly utilized antibiotics at sub-MIC levels positively selects for resistant bacteria (Gullberg et al. 2011) and is a common adaptive mechanism seen in *Pseudomonas aeruginosa* populations within the CF lung. Survival in an environment with prolonged, and often intensive, exposure to antimicrobials is achieved by a variety of phenotypic changes and genetic mutations. This experiment confirmed biofilm growth in 2% OligoG CF-5/20 alone did not alter the acquisition of resistance to a range of ten different antibiotics. Though there was evidence of genetic mutations associated with antimicrobial resistance in both control and 2% OligoG CF-5/20 mixed colonies, it is important to recognise that cross resistance may not be entirely predictable, with variables such as the antimicrobial generation/level of activity and susceptibility of individual host bacteria playing an important role in resistance (Périchon et al. 2019). In-depth cross-resistance studies using all individual colony samples with cross reference to genetic mutations would be an interesting area to explore for future work.

Azithromycin is a macrolide antibiotic with the ability to inhibit bacterial protein synthesis, quorum-sensing and reduce biofilm formation (Parnham et al. 2014). It has been shown to have such efficacy against *P. aeruginosa* virulence factor production and biofilm formation (Parnham et al. 2014). Interestingly, macrolides have also been shown to augment the *in vitro* activity of other anti-pseudomonal microbials, supporting their use as part of combined therapies in cystic fibrosis (Lutz et al. 2012). Mutants formed in the presence of azithromycin and 2% OligoG CF-5/20 actually improved *P. aeruginosa* susceptibility to antibiotics from other classes, including aztreonam and oxytetracycline. This reflects results from planktonic models which have demonstrated antibiotic potentiation in the presence of 2% OligoG CF-5/20 (Khan et al. 2012a; Pritchard et al. 2017a). The mutations incurred can also compensate for fitness cost associated with acquisition of resistance. The altered susceptibility to other classes of antibiotics for mutants formed in the presence of azithromycin and 2% OligoG CF-5/20 highlights a possible alternative mode of action for mutations arising in the presence of 2% OligoG CF-5/20.

Azithromycin resistance can occur when bacteria change the target/binding site, either by mutation of some ribosomal components or methylation of key rRNA nucleotides; or alternatively by efflux pump activity which reduces intra-bacterial accumulation (Parnham et al. 2014). Mutations in the multi-drug resistance regulatory gene, *mexT*, have been shown to enable antimicrobial resistance in clinical isolates of *P. aeruginosa* (Horna et al. 2018). Such mutations affect the MexEF-OprN system as well as the regulation of QS-associated factors (Liu et al. 2022). This study demonstrated genetic mutations affecting *mexT* occurring in bacteria from control and 2% OligoG CF-5/20 conditions. However, there was no difference seen in resistance patterns to azithromycin on day 21 or day 45. This may be a reflection of the experimental design, which did not test antimicrobial resistance patterns for individual bacterial colonies and rather reflects the overall response of colonies with differing genotypes. This would be useful progression of the data for future studies.

Previous studies have shown variation in antibiotic susceptibility profiles and virulence factor production following antibiotic exposure (Fothergill et al. 2010; Cullen et al. 2015). Clark et al. (2015) reported similar cross-resistance between antibiotic classes. Exposure to aztreonam therapy led to increased resistance to aztreonam and ceftazidime; whilst discontinuation of treatment increased susceptibility to both antibiotics plus ciprofloxacin in a selection of isolates (Clark et al. 2015). A parallel study of laboratory adaptation of *Escherichia coli* to various antibiotics over 90 days

demonstrated that acquisition of resistance to a single antibiotic can alter susceptibility to a range of other antibiotics (Suzuki et al. 2014). Whilst cross-resistance Given the likelihood of 2% OligoG CF-5/20 use as part of a patient's polypharmacy, it is critical to minimise negative drug interactions. In particular, demonstrating no antibiotic potentiation, and increased antibiotic susceptibility with a number of antibiotic classes, would be extremely beneficial clinically.

It is important to recognise that this *in vitro* experiment lacked the highly selective and heterogenous local environmental conditions seen in the CF lung, such as oxidative and osmotic-stress (Pestrak et al. 2018). Such stressors are key determinants of the bacterial adaptive behaviour following *P. aeruginosa* colonisation of the lower respiratory tract (Folkesson et al. 2012) and therefore results from this study may not fully reflect how bacteria would truly behave within the human host. This study used a laboratory isolate, PAO1, rather than a clinical isolate as the ancestral clone as it is a well-characterised bacterial strain. Interestingly, other authors have suggested that diversity may be limited when laboratory isolates are used, due to a higher level of preadaptation to the laboratory environment (McElroy et al. 2014; Steenackers et al. 2016). However, Cullen et al. (2015) demonstrated a similar degree of phenotypic diversity across their *P. aeruginosa* strain panel, which included clinical and environmental strains. There were some differences seen, with transmissible CF strains showing less virulence and pyocyanin production, and lack of O antigen, particularly in isolates obtained later in infection (Cullen et al. 2015). The advantage of laboratory models such as the bead biofilm model used here is therefore apparent, being able to produce reproducible biofilms under controlled conditions.

Future studies could consider using clinical specimens for a more representative sample. However, analysis of clinical samples is notoriously challenging due to sampling variability and microbial diversity within the CF lung compartmentalisation (Ronchetti et al. 2018b). Though *in vivo* experiments clearly provide the most 'realistic' results, Steenackers et al. eloquently describe the numerous limitations associated with using clinical samples. These include difficulties in reproducing data, lack of control on heterogenous conditions and being unable to impose and follow evolutionary dynamics (Steenackers et al. 2016).

In future experiments it would be interesting to investigate the overall population structure and the potential interaction between different morphotypes (Workentine et al. 2013). Also, due to practical constraints, it was not possible to undertake all of the

phenotypic tests described in the literature, including tests for auxotrophy, quorum sensing, virulence factor production and antibiotic susceptibility for each morphotype. However, to extend the study further, it would be useful perform these assays on both the individual morphotypes and the biofilm wells' monoculture samples.

This controlled *in vitro* study confirmed no negative selective pressure was placed on the phenotypic and genotypic diversification during prolonged exposure to 2% OligoG CF-5/20. As 2% OligoG CF-5/20 is currently undergoing phase IIb clinical trials, it would be useful to study its phenotypic and genotypic effects on patient samples in the future. This would hopefully improve understanding of this novel therapy's impact on *in vivo* bacterial adaptation, and the potential interplay between resultant isolates and the host and other bacteria within the CF lung environment.

4.6 Conclusion

This study effectively utilised an *in vitro* bead biofilm model to explore *P. aeruginosa* evolution in the presence of 2% OligoG CF-5/20. It demonstrated that growth in the presence of 2% OligoG CF-5/20 did not alter the acquisition of resistance to azithromycin within these biofilms when assessed at day 21. However, mutants formed in the presence of azithromycin and 2% OligoG CF-5/20 demonstrated increased susceptibility to other classes of antibiotics. Findings also demonstrated that bacteria growing in vitro in the presence of 2% OligoG CF-5/20 have fewer colonies with MDR-associated phenotypes and improved antibiotic susceptibilities, which may afford significant clinical benefits in patients treated with this novel anti-biofilm therapy.

Chapter Five

Discussion

5.1 General Discussion

5.1 The evolving climate in cystic fibrosis

People with CF can now expect to live longer lives with better quality of health than ever before (Registry 2018). Burgel et al. (2015) used European registry data to predict that the number of adults with CF would increase by up to 78% by 2021, while the number of children with CF would only increase by 20%, demonstrating increasing survival into adulthood (Burgel et al. 2015).

Since beginning this research in 2016, there have been significant changes in the treatment options available to many people with CF. Prior to the development of CFTR modulator therapies, CF management focused on targeted symptom management, with mucolytics, antibiotics, physiotherapy and nutritional support being the principal treatments available. There are now four single or combination CFTR modulators available on the market, with clear evidence to support their use in patients as highly effective treatments for people with the most common genetic mutation, *F508 del*, as well as numerous other mutations (Paterson et al. 2020; Dagenais et al. 2021; Middleton and Taylor-Cousar 2021). These treatments have shown improvements in lung function (FEV1), reduction in pulmonary exacerbations and improved nutritional status, with the realistic potential to significantly change the landscape of this chronic disease for the better (Ramos et al. 2021). Even in patients with advanced lung disease, there is clear evidence of benefit from CFTR modulator therapy. Burgel et al. (2021) showed rapid clinical improvement in patients aged >12 years with advanced lung disease, with a number of patients no longer requiring lung transplantation after starting the modulator combination of elexacaftor-tezacaftor-ivacaftor (Burgel et al. 2021).

Monitoring of CF pulmonary disease is now more advanced, with complex lung function testing available to children of all ages, in addition to improved imaging techniques (Dournes et al. 2021). Monitoring of lung clearance index (LCI) is a sensitive marker for small airway disease in CF and is often incorporated into clinical trials as a standard surveillance technique (Frauchiger et al. 2021; Stanojevic et al. 2021).

Despite new therapies and improved disease monitoring, there will be many patients with existing infection and severe lung disease in whom antimicrobials will continue to be a life-saving aspect of their treatment. There is also a sub-group of people with CF whose genotypes are not compatible with existing CFTR modulator therapy. These patients will also continue to require proactive treatment with both prophylactic and treatment antibiotics, as well as mucolytic therapies. For most patients with CF, the treatment burden is considerable (Herbert et al. 2021). Many patients will spend up to 2 hours a day on treatment, therefore it is important to ensure that prescribed medications are optimised and rationalised as much as possible and new therapies are only introduced if they offer clear benefit to a patient's health and well-being (Davies et al. 2020). Reducing the treatment burden in the new era of CFTR modulators is the main research question in the CF STORM trial (Trial reference number: 138613).

5.2 Airway sampling and infection

The management of respiratory health in CF is based on two key interventions: early identification and effective treatment of infection; and management of respiratory secretions. Infections represent a major contributor to CF lung disease and use of targeted antibiotic therapy offers to eradicate bacterial pathogens and prevent long-term lung damage and eventual respiratory failure (Forton 2019). The difficulties of sampling the airway in young children and non-expectorating adults have been discussed at length in this work. The introduction of sputum induction has been supported by numerous studies (Al-Saleh et al. 2010; Blau et al. 2014; Forton 2015; D'Sylva et al. 2017; Ronchetti et al. 2018b). The CF-SpIT study demonstrated the use of sputum induction as an infection-diagnostic in children aged 6 months to 18 years, with greater sensitivity in pathogen yield for IS compared to the gold-standard two-lobe BAL. The greatest yield was seen when six-lobe BAL and IS were combined, suggesting compartmentalisation of pulmonary infections even in younger children with CF (Ronchetti et al. 2018b).

Through the use of culture-independent approaches to microbiology, it is clear that culture-dependent methods only provide limited information regarding the bacterial environment within the CF lung. There is clear evidence of alterations in the respiratory microbiota between states of health and disease. In CF, the evolution from a balanced community to a pathogen-dominated community within the first two

decades of life has been well-described (Zemanick et al. 2017; Zhao et al. 2020). Weiser et al. (2022) explored the use of sputum induction for microbiota analysis, building on the success of IS as a comparable sampling method for culture-dependent pathogen yield demonstrated by the CF-SpIT trial (Ronchetti et al. 2018b). The study demonstrated adequate yield of IS when compared to the gold-standard of bronchoalveolar lavage. The aim of **Chapter 2** was to undertake an in-depth study of the whole CF-SpIT IS dataset, representing the largest IS microbiota analysis for children with CF. The work successfully characterised signatures of disease, including clinical parameters and treatment received and demonstrated clear correlations with microbial diversity. These patterns reflected those previously described using expectorated sputum and BAL samples, therefore supporting the use of IS as a valid tool for sampling of the CF respiratory microbiota.

Further work is required to understand the causative nature of the correlations between clinical features and microbial diversity described in **Chapter 2**. In particular, it remains unclear whether evolution of the microbiota is the principal driver of disease or whether disease progression leads to changes in the respiratory microbiota through alterations in the ecological niche and changes in selective pressure (Forton 2019). Linnane et al. (2021) used paediatric BAL samples from CF and non-CF patients to study diversity over 5 years, supporting reducing diversity in CF patients but ongoing increasing diversity in non-CF children (Linnane et al. 2021). A longitudinal study using IS samples from our patient cohort, ideally from infancy to early adulthood, would provide an opportunity to track microbiota evolution within patients. This would require a large dataset, with multiple sampling points. Sputum induction offers the best sampling method for such a study, as it could be performed at routine outpatient visits regularly throughout the year.

The data in **Chapter 2** showed evidence of pathogen dominance in some children and CF-related pathogens present in almost all samples. The loss of diversity and development of a pathogen-dominated microbiota is associated with worsening clinical parameters, as seen in previous studies (Blainey et al. 2012; Boutin and Dalpke 2017). Gaining a better understanding of the effects of pathogen-dominance versus microbial diversity may also inform clinicians about timings of transition to the 'disease-state microbiota'. There may be an optimal time for the introduction of focused treatment interventions, which will likely vary according to the individual (Forton 2019). Microbiota analysis may represent a potential biomarker for disease,

but further studies are required before this can be incorporated into routine clinical practice.

5.3 The COVID-19 pandemic

This thesis was written during the COVID-19 pandemic. The effects of this global pandemic have been wide-reaching. Children with CF were considered 'clinically extremely vulnerable' in the early months of the pandemic and therefore advised to shield at home. Traditional hospital-based CF care, whereby patients attended outpatient clinics for spirometry, cough swabs or sputum for microbiological culture, growth measurements and clinical reviews by multi-disciplinary team (MDT) members, were no longer possible (Dixon et al. 2021). This prompted rapid introduction of telemedicine into many CF clinics.

Telemedicine and remote care delivery models for CF have become a significant focus of interest over the past 2 years. Reported benefits have included similar access to the full MDT, improved personalised approach, enhanced convenience for the patient and their family, gains for the environment with less travel, and reduced infection risk (Dixon et al. 2021). Potential difficulties related to remote care delivery included lack of ability to perform physical examinations, reliance on home spirometers for lung function measurements, lack of access to more complex investigations and imaging, and barriers to effective communication (Dixon et al. 2021).

Of particular relevance to this work, the ability to ensure adequate airway sampling was a clear challenge for patients and their teams. Prior to the pandemic, integration of IS as a valid and reliable sampling method was increasingly commonplace in clinical practice (Weiser et al. 2022). However, sputum-induction traditionally requires a standard process involving nebuliser therapy and a physiotherapist (Ronchetti et al. 2018b). As a natural development of the CF-SpIT study (Ronchetti et al. 2018b), Dafydd et al. (2020) systematically evaluated home sputum-induction testing (HomeSpIT) within their paediatric CF clinic. The study showed that home sputum-induction was successful in children >5 years, with comparable outcomes to clinic sputum-induction procedures (Dafydd et al. 2020). Whilst particularly useful within the context of fewer face-to-face clinical appointments during the pandemic, this study also supports the transitioning to sputum-induction as the primary routine sampling

method over cough swabs. For many patients, it should be possible to produce their samples prior to clinic visits, thus reducing aerosol-generating procedures in clinic and enabling physiotherapists to focus their expertise on younger children who will require greater support to obtain an adequate IS sample (Dafydd et al. 2020).

The results of **Chapter 2** demonstrated the use of IS for microbiota diversity and the HomeSpIT study (Dafydd et al. 2020) further supports the incorporation of IS for routine culture-dependent and culture-independent airway microbiological surveillance. The ability to produce IS samples at home may enable more frequent sampling whilst minimising the negative impact on the patient. Future studies could use home sputum-induction for regular microbiota sampling, perhaps for understanding microbiome alterations around the time of pulmonary exacerbations.

The technique could also be incorporated into monitoring of response to both standard treatments but also new clinical therapies within clinical research trial settings. The incorporation of microbiota analysis into the OligoG phase 2b clinical trials has been reported by Weiser et al. (2021). The study demonstrated highly concordant microbiota profiles for paired sputum samples taken 2 hours apart, suggesting that a single collection is sufficient to capture microbiota diversity at relevant trial time-points. The samples also showed stable microbiota profiles for individual patients over time, though there was significant inter-patient sample heterogeneity and variations in pathogen dominance (Weiser et al. 2021). The authors suggested that understanding patient's microbiota profiles may be beneficial when designing future trials. Perhaps trials should consider pre-trial microbiota profiling to be of equal, if not greater, importance to understanding patients' culture-dependent microbiology status, and this should be incorporated into the trial protocol. This would be of particular relevance for anti-infective clinical trials. Weiser et al. (2021) used expectorated sputum samples from adult patients for their analysis. However, given the concordance of microbiota analysis between BAL and IS samples in (Weiser et al. 2022), plus the data shown in **Chapter 2**, there is clear evidence that IS could be incorporated into study designs as an alternative method. This may make it easier to include children in future clinical trials, which would clearly be very beneficial for providing data for this patient group.

5.4 Antimicrobial resistance

Antimicrobial resistance (AMR) is a major healthcare challenge across the globe. Infection-control processes must be stringent for patients with CF, with strict rules in place to minimise patient contact and potential cross-infection (Brown et al. 2021). A key risk factor for selecting antimicrobial resistance is repeated bacterial exposure to suboptimal antibiotic concentrations (Castagnola et al. 2021). The efficacy of antibiotics can be significantly affected in CF as a result of patients' pathophysiological conditions, such as changes in lean body mass and glomerular hyperfiltration for example. These conditions can influence drug pharmacokinetic and pharmacodynamic parameters and increase the likelihood of suboptimal antibiotic concentrations being reached at the target organ (Castagnola et al. 2021). Therefore, cystic fibrosis represents a 'perfect storm' of recurrent infections, increased need for regular and long-term antibiotics, and poor achievement of therapeutic drug delivery, leading to increasing rates of AMR. To tackle the problem of evolving AMR, novel antimicrobial therapies are needed.

5.5 Development of novel therapies

Numerous potential therapeutic drugs are in clinical trial stages and thereafter being approved for patients with CF (Kotnala et al. 2021). The Cystic Fibrosis Drug Development Pipeline lists drugs in development according to their role. These categories include restoration or correction of CFTR protein function (modulator therapies), therapies to improve mucociliary clearance, anti-inflammatory, anti-infective and nutritional formulations (Kotnala et al. 2021).

CFTR mutations are associated with depletion of airway surface liquid, retention of thick and viscous mucus, increased infection and inflammation (Lopes-Pacheco 2016). The viscoelastic properties of mucus are predominantly affected by mucins. As a result of CFTR dysfunction, mucin production is abnormal, with clear upregulation of certain respiratory mucins particularly seen during pulmonary exacerbations (Ostedgaard et al. 2017; Kotnala et al. 2021). Interactions between mucins and CF therapies is an obvious area of interest. OligoG CF-5/20 is a novel therapy targeting improved mucociliary clearance. **Chapter 3** aimed to identify possible structural interactions between OligoG CF-5/20 and *ex-vivo* CF sputum using Fourier Transform Infrared Spectroscopy (FTIR). Samples from the CF SpIT

trial were used to analyse the mechanism of action of OligoG CF-5/20 on the sputum of patients with CF. Results showed clear evidence of interaction between OligoG CF-5/20 and the respiratory mucins within CF sputum samples, particularly with glycan moieties and the peptide backbone of mucin. This data has made an important contribution to our understanding of the mechanism of action of OligoG CF-5/20 and provides a potential mechanism whereby this novel therapy can modify the viscoelastic properties of CF sputum.

The ultimate goal for OligoG CF-5/20 is to incorporate this novel therapy into routine patient use. As with all new therapeutics, there are strict processes which must be followed to demonstrate safety and efficacy from laboratory to clinical use. OligoG CF-5/20 is currently at the stage of Phase 2b trials, with a number of trials completed and others due to start over the next year. It now has Orphan Drug designation from both the European Medicines Agency and the United States Food and Drug Administration (Algipharma.com). **Chapter 4** aimed to use a novel *in vitro* bead biofilm model to represent the CF lung environment and study the effects of prolonged exposure to 2% OligoG CF-5/20 on *P. aeruginosa*. A range of assays were utilised to characterise potential phenotypic and genotypic alterations that may arise in *P. aeruginosa* colonies. This was an important study as clinical use of OligoG CF-5/20 will be delivered as either repeated treatment courses or for prolonged patient use over weeks to months and understanding the possible effects on bacterial pathogenicity is essential for clinical use. The data from **Chapter 4** demonstrated a reduction in colonies with multi-drug resistant-associated phenotypes, such as SCVs. Also, 2% OligoG CF-5/20 did not drive mutations in specific genes during the adaption of *P. aeruginosa* to biofilm growth. Interestingly, the study showed that isolates grown in the presence of 2% OligoG CF-5/20 and azithromycin had altered susceptibility to other classes of antibiotics, with greater susceptibility to antibiotics such as aztreonam and oxytetracycline. Overall, extrapolation of these findings to clinical use would suggest significant clinical benefits to patients. However, clearly further clinical trials are required to see whether such changes within an *in vitro* model are experienced in CF patients.

Since commencing this research, the body of evidence supporting the use of OligoG CF-5/20 in patients with CF has expanded significantly. Peer-reviewed publications have demonstrated the ability of OligoG CF-5/20 to alter the visco-elastic properties of mucus through both direct and indirect effects on mucin (Nordgård and Draget 2011; Pritchard et al. 2016a; Vitko et al. 2016; Ermund et al. 2017). There have also

been numerous studies, particularly focusing on *P. aeruginosa*, demonstrating its ability to disrupt bacterial biofilm formation and growth (Powell et al. 2013; Powell et al. 2014b; Pritchard et al. 2017a; Jack et al. 2018; Powell et al. 2018) and potentiate the effect of antibiotics (Khan et al. 2012a; Pritchard et al. 2017a). Data has also shown the potential to combine OligoG CF-5/20 with colistin as a conjugate to improve treatment of MDR gram-negative bacterial infections (Stokniene et al. 2020). This raises the possibility of using this novel therapy in combination with existing antimicrobials to maximise efficacy in difficult-to-treat infections.

Clinical trial data has confirmed that repeated inhalation of OligoG CF-5/20 dry powder (DPI) was safe in adults, but the study was unable to show significant treatment benefit with OligoG CF-5/20 compared to placebo (van Koningsbruggen-Rietschel et al. 2020). Future phase 2B clinical studies, which are currently being performed under the framework of HORIZON2020, are exploring the use of OligoG CF-5/20 DPI at a lower dose. A number of other clinical trials are ongoing, aiming to provide further safety and efficacy data supporting the use of OligoG CF-5/20 in patients with CF.

5.6 Conclusion

Significant developments in CF management over the past decade should be celebrated. However, as a scientific and medical community, we must not become complacent and should continue to seek new therapies to optimise the health of children and adults with CF. The maintenance and progression of the CF antimicrobial development pipeline is essential and novel therapies such as OligoG CF-5/20 require ongoing investment of expertise and finances to ensure that the data required to enable safe and effective clinical use is produced and shared with patients, clinicians and the wider research community. This research has demonstrated the use of induced sputum sampling for microbiota analysis within the paediatric population. Future use for both routine clinical surveillance and integration into clinical trial methodology seems a natural progression, but further studies will be needed to really understand the role of sputum induction in disease monitoring, prediction of outcomes and use in targeting appropriate therapies for patients.

Appendix I:

Chapter Two: Induced sputum samples can be used to investigate microbial diversity in children with cystic fibrosis

I.I Parent information sheet and consent form

Children's Hospital for Wales

Bwrdd Iechyd Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board

NISCHR

AHSC

Ariennir gan **Lywodraeth Cymru** Funded by **Welsh Government**

The use of induced sputum in monitoring infection in children with Cystic Fibrosis

STUDY INFORMATION SHEET FOR PARENTS

Dear Parents

Invitation

- Your child is being invited to take part in a research study. As the child's parent/legal guardian we are asking you to read this leaflet as you will be asked to make the decision on behalf of your child. We also have an information leaflet suitable for children should they wish to have one. Before you make a decision it is important to understand why this research is being carried out and what it will involve. Please take time to read this leaflet carefully and speak to others about it if you wish.

Background
What is the purpose of this study?

- In Cystic Fibrosis (CF), lung infection is common and needs to be treated aggressively with antibiotics even if it isn't causing many symptoms. Doctors need to use different antibiotics for different types of infection. In order to identify which bacteria is causing the infection, your child is often asked to give us a cough swab, so that a sample of your child's airway liquids can be sent to the lab to see if anything grows.
- Cough swabs are relatively easy to obtain but are not as good a test as bronchoscopy, where a fibre-optic camera is put down into the large airway of the lungs so that mucous samples can be taken directly from the lower airway. Obviously having a bronchoscopy is a much larger procedure than having a cough swab, but sometimes it is necessary. Some CF centres feel everyone with CF should have a bronchoscopy every year.
- The purpose of this research study relates to a third way of getting samples from the airway called "induced sputum". This is a little bit more complicated than a cough swab but much less complicated than having a bronchoscopy. It involves your child inhaling a fine mist of salt water and getting some physiotherapy. The salt water inhalation causes the phlegm (or sputum) to loosen up so that it can be more easily coughed up from the lower airway. We plan to compare the induced sputum to a cough swab and to a throat swab and to a nasal swab. If your child is going to have a bronchoscopy because your doctor feels he/she needs one, then we will compare these samples to the results of the bronchoscopy as well.

- In this research project we want to find out just how beneficial induced sputum really is, if we do it just once in the year, as part of the annual review, over and above the many cough swabs that are taken over the year anyway.

- Induced sputum can be done in the outpatient clinic or in the hospital ward and takes about 30 minutes. The technique is safe and used routinely in children with other respiratory illnesses. However, this kind of research needs to be done before we can start using the induced sputum technique routinely in patients with CF, as we need to be very sure that the procedure is well tolerated and also that it makes a worthwhile contribution to improving health care.

Why has my child been chosen?

- We are asking all children with CF over age 6 months who receive full care in Cardiff if they would like to be included in the study

Does my child have to take part?

- No. It is up to you and your child whether you decide to participate or not. If you do, you will be given this information sheet to keep and will be asked to sign a consent form on behalf of your child. You are still free to withdraw at any time and without giving a reason. A decision not to take part, or to withdraw at any time, will not affect the standard of care your child receives. Any samples or data files relating to your involvement in the study will then be destroyed immediately.

What will happen to my child if he/she decides to take part?

- One of the CF physiotherapists will take a cough swab and a throat swab and a nasal swab and then start the procedure for induced sputum.
- Oxygen saturation and heart rate will be monitored for the procedure. If your child is over 7 years then he/she will be asked to do lung function before the procedure starts.
- The procedure will involve a salty nebuliser (hypertonic saline) which will last about 15 minutes. After each 5 minute period, the physiotherapist will make an assessment of the chest and give appropriate physiotherapy or guide your child through breathing exercises to try and mobilise secretions. Any secretions will be collected either into a pot, or by suction from the back of the throat.
- Lung function will be taken again at the end of the procedure if appropriate.
- The final step is to take another cough swab.

What does my child have to do?

- Your child will be guided through the procedure as outlined above. The procedure will take place at about 12 o'clock and it is best if your child does not have anything to eat after 10am that morning

What are the other possible disadvantages and risks of taking part?

- All of the procedures being used in this study are already used by doctors in the treatment of children. Sometimes the salty nebuliser can make you cough and some children can wheeze. Generally it is well tolerated in all age groups.

Induced sputum in children with cystic fibrosis.
Project ID 11/RPM/5216.
Protocol Amendment version 2.1_29/2/2015
Julian Forton

215

What are the possible benefits to taking part in this trial?

- The main benefit of this research is for the CF community as a whole, as we explore whether induced sputum should become part of routine care. There are no immediate benefits to the patient from taking part in this study, but should any organisms be identified from the samples taken, then appropriate treatment will be prescribed.

What happens when the research study stops?

- This research study is planned to run over five years.
- Once we have obtained a sample from your child, you and your child have made your contribution to the study.
- As the study lasts for three years it may be that you are approached again at subsequent annual reviews to go through the procedure again. There is no obligation for you to repeat the procedure just because you enrolled in the past. It would be entirely up to you if you wanted to contribute again.

If you are interested in taking part please read on for more details

What will happen if I don't want my child to carry on with the study?

- You are free to withdraw your child from the study at any time, including while he/she is performing the induced sputum test, and this will not affect the future care of your child.

Will my child taking part in the study be kept confidential?

- Yes. All information which is collected about your child will be kept strictly confidential. Access to data which may identify you will be limited to the research team, who have a duty of confidentiality. Any information about you which is used outside the hospital/surgery will have your name and address removed and will only be identified by a code. Nothing which could reveal your child's identity will be disclosed outside the research site.

What will happen to the samples that my child provides?

- The samples will be sent to the labs to see what organisms can be identified. This will be done in the routine way, but also using a new state of the art approach using bacterial genetics, so that we can see if the new way works even better. These tests are not used routinely at the moment but may be in the future, and we need to see how well they work.
- These samples are very valuable to scientists as they are difficult to obtain. Here at the CF Unit in Cardiff we work with scientists who are very interested in inflammation in the lung and how the body responds to infection. We will collaborate with them in studies to look at inflammation.
- All samples will be supplied anonymously to researchers; only Dr Forton and members of his research group will be able to identify which samples your child donated. The recipients of the samples will not be supplied with your child's name or any other identifiable information and will not be able to identify your child from the samples.
- Any residual samples at the end of the study will be stored under a license from the Human Tissue Authority. License no: 12422.
- Your child's samples may be retained at the end of this study for use in future research within the UK and abroad. At this stage we do not know what the research will involve but some of it could include more bacterial genetic research and further research on lung inflammation. On the consent form you will be given the option to exclude your child's

samples from these areas of research. Your child's samples will not be sold and will not be used in human genetic research, animal research or the commercial sector.

- **Current Use of samples in this study**
  Participation in this study is voluntary and you are free to withdraw at any time without giving a reason and without your medical care or legal rights being affected.
  If you do withdraw your consent your child's samples will not be used further in this study and will be destroyed according to locally approved practices. Any samples, or results derived from the samples, that have already been used prior to the withdrawal of consent will continue to be used in this study.
  Future Use of samples in other related studies
  You may withdraw your consent for the storage and future use of your child's samples at any point. If you do withdraw your consent, your child's samples will not be used in any subsequent studies and will be destroyed according to locally approved practices. Any samples already distributed for use in research prior to the withdrawal of consent will continue to be used in that study and any samples remaining at the end of the study will be destroyed.

What will happen to the results of this study?

- We intend to publish the study in a peer reviewed medical journal so that it can be seen by the rest of the scientific and medical community. The findings may also be presented at conferences. Such reporting is normal practice among researchers. Your child will not be personally identified in any report or publication.

Who is organising and funding the research?

- This study has been funded by a grant to Dr Julian Forton, consultant in Paediatric Respiratory Medicine, from the National Institute for Social Care and Health Research, Wales (NISCHR). It is being coordinated by Dr. Julian Forton here in Cardiff.

Who has reviewed the study?

- All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee (REC), to protect your interests. This study has been reviewed and given a favourable opinion by the Wales REC2

What if there is a problem?

- If you have a concern about any aspect of this study, you should ask to speak to the researchers who can be contacted on the numbers listed below. They will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure]. Details can be obtained from the University Hospital for Wales
- In the event that something does go wrong and you are harmed during the research and this is due to someone's negligence then you may have grounds for a legal action for compensation against Cardiff and Vale UHB but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.
- CONTACT DETAILS for researchers: 02920743530 or 02920744891

Induced sputum in children with cystic fibrosis.
Project ID 11/RPM/5216.
Protocol Amendment version 2.1_29/2/2015
Julian Forton

217

Thank you for taking time to read this leaflet. Please do not hesitate to ask a member of the research team if you would like to discuss anything further.

Dr Julian Forton
MA(Hons) MB BChir (Cantab) MRCPCH Ph.D

12th December 2014

If you have any concerns please do not hesitate to contact
Dr Julian Forton
Cystic Fibrosis/Respiratory Unit,
University Hospital of Wales,
Cardiff.
Tel 029 20743530 or 02920744891

Induced sputum in children with cystic fibrosis.
Project ID 11/RPM/5216.
Protocol Amendment version 2.1_29/2/2015
Julian Forton

218

CONSENT FORM

The use of induced sputum in monitoring infection in children with Cystic Fibrosis.

Name of Patient:

Please initial box

1. I confirm that I have read and understand the information sheet dated 11/12/2014 (version 2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my child's participation is voluntary and that I am free to withdraw my child at any time, without giving any reason, without my child's medical care or legal rights being affected.

3. I understand that relevant sections of any of my child's medical notes and data collected during the study, may be looked at by responsible individuals from the research team, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.

4. I understand that samples being collected in this research may be stored and used for future research. Any residual material will be stored under a license from the Human Tissue Authority. License no: 12422.

5. I understand that I am able to withdraw my consent for this and /or other future studies study at any time. Any samples my child has donated which are still being stored will be destroyed at this stage

6. I agree for my child to take part in the above study.

Name of Parent   **Date**   **Signature**

Researcher   **Date**   **Signature**

Appendix II:

Chapter Four

Phenotypic and genotypic adaptations in

Pseudomonas aeruginosa biofilms following long-

term exposure to an alginate oligomer inhalation

therapy

II.I Materials and methods for genotypic characterisation

II.I.I Genotypic characterisation of colony morphotypes

Whole-genome sequencing was performed following genomic DNA extraction from wild-type P. aeruginosa PAO1 and evolved PAO1 isolates using the Maxwell instrument and Maxwell 16 tissue DNA purification kits (Promega) according to the manufacturer's instructions. Briefly, 3ml of fresh overnight culture was pelleted by centrifugation, and the pellet was resuspended in 300ml 4 M UltraPure guanidine isothiocyanate (ThermoFisher Scientific) and added directly into the DNA purification kits. Eluted DNA was stored at 220°C. Whole-genome sequencing was performed at the Cardiff School of Biosciences Genomics Research Hub. DNA was prepared for sequencing using the NEBNext Ultra II DNA library prep kit for Illumina and NEBNext multiplex oligonucleotides for Illumina (New England BioLabs Inc.). Sequencing was carried out on an Illumina NextSeq500 using a NextSeq 500/550 Mid Output v2 kit (300 cycles), giving, on average, 135-bp paired-end reads. Approximately 2.9 million reads (range: 2.5 to 3.4 million) were yielded per sample, corresponding to an average coverage depth of approximately 125x (range: 106 to 142x) (Oakley et al. 2021).

II.I.II Bioinformatic analysis of whole-genome sequencing data

Bioinformatic analysis was carried out on a virtual machine, hosted by the Cloud Infrastructure for Microbial Bioinformatics (CLIMB) consortium (Connor et al. 2016). Quality control and Illumina adapter trimming of the raw sequencing reads were performed using FastQC v0.11.5 (Andrews 2010) and Trim Galore! v0.4.3 (Krueger) for paired-end reads. Genome assembly for the PAO1 WT was achieved using Unicycler v0.4.7 (Wick et al. 2017) with SPAdes v3.11.0 (Bankevich et al. 2012) and the option for short-read assembly. Assembly quality was visualised with Bandage assembly graphs (Wick et al. 2015) and with QUAST v4.6.3 (Gurevich et al. 2013) to determine that the PAO1 wild type shared >98.8% genomic DNA with the *P. aeruginosa* PAO1 ATCC 15692 sequence (GenBank accession number GCA_001729505.1). Contig ordering to the *P. aeruginosa* PAO1 ATCC 15692 sequence was performed using ABACAS v1.3.1 (Assefa et al. 2009). The resulting draft genome sequence was annotated with Prokka v1.12 (Seemann 2014). Polymorphic sites in the evolved PAO1 isolates were identified using Snippy v3.2 (Seemann 2015) with the draft genome sequence of the wild-type PAO1 as the reference. For variant calling, the default parameters of minimum base quality of 20, minimum read coverage of 10x, and 90% read concordance at each locus were used.

Only variants in the annotated coding regions were included in the analysis. Variants identified in the wild-type PAO1 sequence reads were subtracted from all other evolved PAO1 isolates. Correct annotation of coding sequences containing variants was confirmed using the BLASTN search tool of the Pseudomonas Genome Database (Winsor et al. 2016) against the *P. aeruginosa* PAO1 reference sequence. Functional information for coding sequences was derived from the *Pseudomonas* Genome Database (Oakley et al. 2021).

Sequence data supporting the genomic analysis have been deposited in the European Nucleotide Archive with the accession code PRJEB36146 (ERP119298). With thanks to Dr Rebecca Weiser who completed the whole-genome sequencing and bioinformatic analysis.

II.II.I Results for genotypic characterisation

This data was analysed and reported by Dr Rebecca Weiser. The results are provided to report the complete dataset produced from the bead biofilm model isolates (Oakley et al. 2021).

Genome resequencing was performed to determine genetic changes associated with the evolution of PAO1 biofilm populations, with or without exposure to 2% OligoG CF-5/20. Overall, 96 mutations (single nucleotide polymorphisms [SNPs], insertions, deletions, duplications) were identified across 38 bead biofilm-evolved isolates, with two day 21 transfer isolates (C10a and O6a) having no evidence of genomic mutation. Eight mutations were in noncoding regions, while 88 were in coding regions. The 8 mutations in noncoding regions were all identified in control day 45 transfer isolates (C1b, C2b, and C3b) and excluded from the overall "functional" analysis. The 88 mutations in coding regions represented 39 unique changes, affecting 21 coding regions (**Table II.i**). Only five mutations were synonymous, 1 was found in the tssL1 gene in the isolate C3b, and two mutations were found in a gene encoding a hypothetical bacteriophage-associated protein in two isolates, C1b and C2b. Analysis of the distribution of the 88 mutations revealed that there was a significantly higher number of mutations in the day 45 transfer isolates than the day 21 transfer isolates (control, day 21 versus day 45, $P = 0.02$; OligoG CF-5/20, day 21 versus day 45, $P = 0.02$). There was no difference, however, in the numbers of mutations observed between control and OligoG CF-5/20- exposed isolates (control versus OligoG CF-5/20, day 21, $P = 0.28$; day 45, $P = 0.90$) (Oakley et al. 2021).

Table II.i Mutation detection by whole-genome resequencing of evolved biofilm isolates

Gene with mutation	Contig in reference	Locus in reference	Mutational effect	Mutation class	Evolved isolate (condition)	PGD functional classification*	Pathways*
Signal transduction							
bifA cyclic-Di-GMP phosphodiesterase	PAO1_108	PAO1_04439	missense_variant c.1796A>T p.Lys599Met	Non-synonymous; missense	C6a, C9a, C3b, C5b	Cell wall/LPS/capsule; Motility and attachment	Biofilm formation
yfiR protein	PAO1_93	PAO1_03876	missense_variant c.404T>A p.Val135Glu	Non-synonymous; missense	C12a	Cell wall/LPS/capsule	
			missense_variant c.281T>A p.Val94Glu	Non-synonymous; missense	C11a		
wspA probable chemotaxis inducer	PAO1_30	PAO1_01217	missense_variant c.992C>T p.Ser331Leu	Non-synonymous; missense	C1a, C2a, C2b	Motility and attachment; Adaptation, protection; chemotaxis	Two-component system; Chemotactic transducer (MCP); Chemosensory; Biofilm formation
wspF probable methylesterase	PAO1_30	PAO1_01222	frameshift_variant c.688delA p.Ile230fs	Deletion	C3a, C4a, C6b, C4b, O3a	Chemotaxis; transcriptional regulators; motility and attachment	Two-component system; Chemotaxis; Chemosensory; Biofilm formation
			frameshift_variant c.880_911delACCATCGCCCAGGACCAGGCCAGTTGCGCAGT p.Thr294fs	Deletion	C5a		
			frameshift_variant c.445delG p.Ala149fs	Deletion	C1b, C7b		
			missense_variant c.61G>C p.Ala21Pro	Non-synonymous; missense	O5a, O12a, O7a, O2b		
			missense_variant c.842C>A p.Ala281Asp	Non-synonymous; missense	O8a		
			stop_gained c.658C>T p.Gln220*	Non-synonymous; stop	O1a		
			conservative_inframe_deletion c.37_63delGTCGAGGCGCTGCGCCGCGCGCTGGCC p.Val13_Ala21del	Deletion	O2a		
			missense_variant c.821C>T p.Thr274Ile	Non-synonymous; missense	O4a		
			stop_gained c.418C>T p.Gln140*	Non-synonymous; stop	O3b, O1b		
Motility regulator *morA*	PAO1_108	PAO1_04681	missense_variant c.3464T>A p.Leu1155Gln	Non-synonymous; missense	O11a, O10a, O4b, O5b	Membrane proteins	

Gene	PAO1_	Locus	Variant	Type	Serotype(s)	Function category	Pathway
Glucose transport sensor *gtrS*	PAO1_36	PAO1_01743	missense_variant c.1014C>G p.His338Gln	Non-synonymous; missense	O5b	Two-component regulatory systems	Two-component system
Secretion							
tssL1 membrane protein	PAO1_11	PAO1_00042	synonymous_variant c.1137G>A p.Pro379Pro	Synonymous	C3b	Hypothetical, unclassified, unknown; Protein secretion/export apparatus	HCP secretion island (HIS-I) type VI secretion system; Bacterial secretion system
Translation							
Elongation factor G *fusA1*	PAO1_21	PAO1_00659	missense_variant c.953C>T p.Ser318Leu	Non-synonymous; missense	C3a	Translation, post-translational modification, degradation	
	PAO1_21	PAO1_00659	missense_variant c.1546G>A p.Gly516Ser	Non-synonymous; missense	C8a, C10b		
Transcription							
Transcriptional regulator *mvfR*	PAO1_93	PAO1_04000	conservative_inframe_deletion c.109_120delTCGGCGGTCAGC p.Ser37_Ser40del	Deletion	O11a, O10a, O4b, O5b	Transcriptional regulators; Biosynthesis of cofactors, prosthetic groups and carriers	Quorum sensing; biofilm formation
			frameshift_variant c.782_785dupGCGG p.Ile263fs	Duplication	C3b, C5b		
			missense_variant c.101C>T p.Ala34Val	Non-synonymous; missense	O9a		
			missense_variant c.112G>A p.Ala38Thr	Non-synonymous; missense	O1b		
			missense_variant c.440T>C p.Ile147Thr	Non-synonymous; missense	C4a, C4b		
			missense_variant c.527A>C p.His176Pro	Non-synonymous; missense	C6b		
Transcriptional regulator *mexT*	PAO1_41	PAO1_02477	conservative_inframe_insertion c.389_390insCCT p.Val130_Leu131insLeu	Insertion	C12a, C6a, C2a, C13a, C7a, C5a, C5b, C2b, C9b, C6b, C4b, C1b, C7b, O5a, O7a, O8a, O1a, O2a, O11a, O10a, O4a, O2b, O3b, O1b, O4b	Transcriptional regulators	
Glycerol-3-phosphate regulon repressor *glpR*	PAO1_30	PAO1_01342	missense_variant c.169G>A p.Ala57Thr	Non-synonymous; missense	C8b	Transcriptional regulators	
Transcriptional regulator *lasR*	PAO1_93	PAO1_03564	missense_variant c.628T>C p.Phe210Leu	Non-synonymous; missense	C6b	Transcriptional regulators; Adaptation, protection	Quorum sensing; biofilm formation

	PAO1_20	PAO1_00624	frameshift_variant c.594dupG p.Leu200fs	Duplication	O5b	Transcriptional regulators	Two-component system; Quorum sensing; Biofilm formation
Transcriptional regulator *vfr*							

Motility

			stop_gained c.1528C>T p.Gln510*	Non-synonymous; stop	O4b		
fimL protein	PAO1_88	PAO1_03161	conservative_inframe_insertion c.1340_1341insCCTGGC p.Gly447_Leu448insLeuAla	Insertion	C5b	Motility and attachment	
Type 4 fimbrial biogenesis protein *piLY1*	PAO1_108	PAO1_04632	stop_gained c.2993C>A p.Ser998*	Non-synonymous; stop	C6b	Motility and attachment	Pilin biosynthesis
Type 4 fimbrial biogenesis protein *pilM*	PAO1_122	PAO1_05151	frameshift_variant c.670delG p.Gly224fs	Deletion	C4b	Motility and attachment	Pilin biosynthesis
Twitching motility protein *pilT*	PAO1_19	PAO1_00365	conservative_inframe_deletion c.970_984delGTCGCCAAGGGCCTG p.Val324_Leu328del	Deletion	O1b	Cell wall/LPS/capsule; Motility and attachment	Pilin biosynthesis

Unknown

	PAO1_79	PAO1_02747	conservative_inframe_deletion c.868_888delTTTGAGACTGCTATTTCCCAG p.Phe290_Gln296del	Deletion	O3b		
Hypothetical protein	PAO1_88	PAO1_03081	missense_variant c.523T>C p.Phe175Leu	Non-synonymous; missense	C9b		
	PAO1_120	PAO1_04769	frameshift_variant c.281dupC p.Leu95fs	Duplication	C2b		
			frameshift_variant c.716delC p.Pro239fs	Deletion	C1b		

Unknown (phage related)

	PAO1_100	PAO1_04289	synonymous_variant c.246G>T p.Gly82Gly	Synonymous	C2b, C1b		
Hypothetical protein from bacteriophage Pf1			synonymous_variant c.198T>C p.Ser66Ser	Synonymous	C2b, C1b		

Footnotes; PDG, Pseudomonas Genome database; * Functional classifications and pathways according to the Pseudomonas Genome Database

226

Genes that had acquired mutations in the evolved isolates were clustered according to function (**Table II.ii**). The majority of mutations occurred in genes associated with signal transduction (n = 32) or transcription (n = 39) and genes encoding hypothetical or bacteriophage-associated hypothetical proteins (n = 8), with smaller numbers of mutations linked to motility (n = 5), secretion (n = 1), and translation (n = 3). Several genes in the signal transduction and transcription functional categories were found in pathways involved in biofilm formation, chemotaxis, motility, and QS. Apart from the mvfR and mexT transcriptional regulators and the wspF methyltransferase involved in signal transduction, no other genes were found to have mutations in isolates from all four populations. Notably, acquired mutations in motility genes were present only, in both control and 2% OligoG CF-5/20-treated isolates, at day 45 (Oakley et al. 2021).

Table II. ii. Distribution of mutations in evolved genotypes

The heatmap indicates the numbers of mutations in genes belonging to different functional groups, as per the *Pseudomonas* Genome Database. The colour intensity reflects the frequency of mutations in each population (actual values also given inside the boxes). Populations are indicated at the top of the figure: control and OligoG CF-5/20 at transfer days 21 and 45. Links to functional pathways are given in parentheses next to gene identifications: BF, biofilm; CH, chemotaxis; M, motility; QS, quorum sensing; and MDR, multidrug resistance.

Table II. ii. (Description overleaf).

Population		Control		OligoG CF-5/20	
Transfer day		**21**	**45**	**21**	**45**
Number of isolates		13	10	12	5
Signal transduction	*bifA* (BF)	2	2		
	yfiR	2			
	wspA (CH;BF)	2	1		
	wspF (CH;BF)	3	4	8	3
	morA (M)			2	2
	gtrS				1
Secretion	*tssL1*		1		
Translation	*fusA1*	2	1		
Transcription	*mvfR* (QS;BF)	1	4	3	3
	mexT (MDR)	6	7	8	4
	glpR		1		
	lasR (QS:BF)		1		
	vfr (QS:BF)				1
Motility	*fimL*		1		1
	pilY1		1		
	pilM		1		
	pilT				1
Unknown			3		1
Unknown (phage related)			4		
Total mutations in coding regions		18	32	21	17
Mean (mutations/isolate)		1.38	3.2	1.75	3.4
Median (mutations/isolate)*		1	4	2	4

References

Cystic fibrosis: Genetics and pathogenesis. Available at: [Accessed.

<DeVries mucoid to non mucoid conversion.pdf>.

The clinical and function translation of CFTR. Available at: https://www.cftr2.org [Accessed: 20th February].

Acosta, N. et al. 2018. Sputum microbiota is predictive of long-term clinical outcomes in young adults with cystic fibrosis. *Thorax* 73(11), pp. 1016-1025.

Acosta, N., Whelan, F. J., Somayaji, R., Poonja, A., Surette, M. G., Rabin, H. R. and Parkins, M. D. 2017. The evolving cystic fibrosis microbiome: a comparative cohort study spanning 16 years. *Annals of the American Thoracic Society* 14(8), pp. 1288-1297.

Ahmed, M. N., Porse, A., Sommer, M. O. A., Hoiby, N. and Ciofu, O. 2018. Evolution of Antibiotic Resistance in Biofilm and Planktonic Pseudomonas aeruginosa Populations Exposed to Subinhibitory Levels of Ciprofloxacin. *Antimicrob Agents Chemother* 62(8), doi: 10.1128/aac.00320-18

Al-Saleh, S., Dell, S. D., Grasemann, H., Yau, Y. C. W., Waters, V., Martin, S. and Ratjen, F. 2010. Sputum induction in routine clinical care of children with cystic fibrosis. *The Journal of pediatrics* 157(6), pp. 1006-1011.

Amiel, E., Lovewell, R. R., O'Toole, G. A., Hogan, D. A. and Berwin, B. 2010. Pseudomonas aeruginosa evasion of phagocytosis is mediated by loss of swimming motility and is independent of flagellum expression. *Infection and immunity* 78(7), pp. 2937-2945.

An, S. q., Warris, A. and Turner, S. 2018. Microbiome characteristics of induced sputum compared to bronchial fluid and upper airway samples. *Pediatric pulmonology* 53(7), pp. 921-928.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Armstrong, D. S. et al. 2005. Lower airway inflammation in infants with cystic fibrosis detected by newborn screening. *Pediatric pulmonology* 40(6), pp. 500-510.

Assefa, S., Keane, T. M., Otto, T. D., Newbold, C. and Berriman, M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* 25(15), pp. 1968-1969.

Baker, M. J. et al. 2016. Developing and understanding biofluid vibrational spectroscopy: a critical review. *Chemical Society Reviews* 45(7), pp. 1803-1818.

Balan, V., Mihai, C.-T., Cojocaru, F.-D., Uritu, C.-M., Dodi, G., Botezat, D. and Gardikiotis, I. 2019. Vibrational spectroscopy fingerprinting in medicine: From molecular to clinical practice. *Materials* 12(18), p. 2884.

Ballok, A. E. and O'Toole, G. A. 2013. Pouring salt on a wound: Pseudomonas aeruginosa virulence factors alter Na+ and Cl− flux in the lung. *Journal of bacteriology* 195(18), pp. 4013-4019.

Bankevich, A. et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology* 19(5), pp. 455-477.

Bansil, R. and Turner, B. S. 2006. Mucin structure, aggregation, physiological functions and biomedical applications. *Current opinion in colloid & interface science* 11(2-3), pp. 164-170.

Barth, A. 2007. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1767(9), pp. 1073-1101.

Bassis, C. M. et al. 2015. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *MBio* 6(2), pp. e00037-00015.

Bell, S. C. et al. 2020. The future of cystic fibrosis care: a global perspective. *The Lancet Respiratory Medicine* 8(1), pp. 65-124.

Bhagirath, A. Y., Li, Y., Somayajula, D., Dadashi, M., Badr, S. and Duan, K. 2016. Cystic fibrosis lung environment and Pseudomonas aeruginosa infection. *BMC pulmonary medicine* 16(1), pp. 1-22.

Bhat, P. G., Flanagan, D. R. and Donovan, M. D. 1996. Drug diffusion through cystic fibrotic mucus: Steady-state permeation, rheologic properties, and glycoprotein morphology. *Journal of pharmaceutical sciences* 85(6), pp. 624-630.

Biesbroek, G., Bosch, A. A. T. M., Wang, X., Keijser, B. J. F., Veenhoven, R. H., Sanders, E. A. M. and Bogaert, D. 2014. The impact of breastfeeding on nasopharyngeal microbial communities in infants. *American journal of respiratory and critical care medicine* 190(3), pp. 298-308.

Bjarnsholt, T., Ciofu, O., Molin, S., Givskov, M. and Hoiby, N. 2013. Applying insights from biofilm biology to drug development - can a new approach be developed? *Nature Reviews Drug Discovery* 12(10), pp. 791-808. doi: 10.1038/nrd4000

Blainey, P. C., Milla, C. E., Cornfield, D. N. and Quake, S. R. 2012. Quantitative analysis of the human airway microbial ecology reveals a pervasive signature for cystic fibrosis. *Science translational medicine* 4(153), pp. 153ra130-153ra130.

Blaser, M. J. and Falkow, S. 2009. What are the consequences of the disappearing human microbiota? *Nature Reviews Microbiology* 7(12), p. 887.

Blau, H. et al. 2014. Induced sputum compared to bronchoalveolar lavage in young, non-expectorating cystic fibrosis children. *Journal of Cystic Fibrosis* 13(1), pp. 106-110.

Blössner, M., Siyam, A., Borghi, E., Onyango, A. and De Onis, M. 2009. WHO AnthroPlus for personal computers manual: software for assessing growth of the world's children and adolescents. *World Health Organization: Geneva, Switzerland*,

Bogaert, D., de Groot, R. and Hermans, P. W. M. 2004. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *The Lancet infectious diseases* 4(3), pp. 144-154.

Boles, B. R. and Singh, P. K. 2008. Endogenous oxidative stress produces diversity and adaptability in biofilm communities. *Proceedings of the National Academy of Sciences* 105(34), pp. 12503-12508.

Boles, B. R., Thoendel, M. and Singh, P. K. 2004. Self-generated diversity produces "insurance effects" in biofilm communities. *Proceedings of the National Academy of Sciences of the United States of America* 101(47), pp. 16630-16635. doi: 10.1073/pnas.0407460101

Bordenstein, S. R. and Theis, K. R. 2015. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS biology* 13(8), p. e1002226.

Borowitz, D. 2015. CFTR, bicarbonate, and the pathophysiology of cystic fibrosis. *Pediatric pulmonology* 50(S40), pp. 2S4-S30.

Bosch, A. et al. 2008. Fourier transform infrared spectroscopy for rapid identification of nonfermenting gram-negative bacteria isolated from sputum samples from cystic fibrosis patients. *Journal of clinical microbiology* 46(8), pp. 2535-2546.

Bosch, A. A. T. M. et al. 2016. Development of upper respiratory tract microbiota in infancy is affected by mode of delivery. *EBioMedicine* 9, pp. 336-345.

Boutin, S. and Dalpke, A. H. 2017. Acquisition and adaptation of the airway microbiota in the early life of cystic fibrosis patients. *Molecular and cellular pediatrics* 4(1), pp. 1-9.

Bragonzi, A. et al. 2009. Pseudomonas aeruginosa microevolution during cystic fibrosis lung infection establishes clones with adapted virulence. *American journal of respiratory and critical care medicine* 180(2), pp. 138-145.

Bray, J. R. and Curtis, J. T. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs* 27(4), pp. 325-349.

Brennan, S. 2020. A patient's experience of cystic fibrosis care. *The Lancet Respiratory Medicine* 8(1), pp. 14-16.

Brennan, S., Gangell, C., Wainwright, C. and Sly, P. D. 2008. Disease surveillance using bronchoalveolar lavage. *Paediatric respiratory reviews* 9(3), pp. 151-159.

Brockhausen, I., Dowler, T. and Paulsen, H. 2009. Site directed processing: role of amino acid sequences and glycosylation of acceptor glycopeptides in the assembly of extended mucin type O-glycan core 2. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1790(10), pp. 1244-1257.

Brown, N., Buse, C., Lewis, A., Martin, D. and Nettleton, S. 2021. Pathways, practices and architectures: Containing antimicrobial resistance in the cystic fibrosis clinic. *Health* 25(2), pp. 196-213.

Brown, P. S. et al. 2014. Directly sampling the lung of a young child with cystic fibrosis reveals diverse microbiota. *Annals of the American Thoracic Society* 11(7), pp. 1049-1055.

Burgel, P.-R., Bellis, G., Olesen, H. V., Viviani, L., Zolin, A., Blasi, F. and Elborn, J. S. 2015. Future trends in cystic fibrosis demography in 34 European countries. *European Respiratory Journal* 46(1), pp. 133-141.

Burgel, P.-R. et al. 2021. Rapid improvement after starting elexacaftor-tezacaftor-ivacaftor in patients with cystic fibrosis and advanced pulmonary disease. *American Journal of Respiratory and Critical Care Medicine* (ja),

Burgel, P.-R., Montani, D., Danel, C., Dusser, D. J. and Nadel, J. A. 2007. A morphometric study of mucins and small airway plugging in cystic fibrosis. *Thorax* 62(2), pp. 153-161.

Burns, J. L. et al. 2001. Longitudinal assessment of Pseudomonas aeruginosa in young children with cystic fibrosis. *The Journal of infectious diseases* 183(3), pp. 444-452.

Bäckhed, F. et al. 2012. Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell host & microbe* 12(5), pp. 611-622.

Carmody, L. A. et al. 2013. Changes in cystic fibrosis airway microbiota at pulmonary exacerbation. *Annals of the American Thoracic Society* 10(3), pp. 179-187.

Castagnola, E., Cangemi, G., Mesini, A., Castellani, C., Martelli, A., Cattaneo, D. and Mattioli, F. 2021. Pharmacokinetics and pharmacodynamics of antibiotics in cystic fibrosis: a narrative review. *International Journal of Antimicrobial Agents*, p. 106381.

Castellani, C. et al. 2018. ECFS best practice guidelines: the 2018 revision. *Journal of Cystic Fibrosis*,

Caverly, L. J., Zhao, J. and LiPuma, J. J. 2015. Cystic fibrosis lung microbiome: opportunities to reconsider management of airway infection. *Pediatric pulmonology* 50(S40), pp. S31-S38.

Charlson, E. S. et al. 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. *American journal of respiratory and critical care medicine* 184(8), pp. 957-963.

Charman, S., Connon, R., Cosgriff, R., Lee, A. and Carr, S. 2018. *UK Cystic Fibrosis Registry 2017 Annual Data Report*. London: Trust, U.C.F.

Christensen, B. E. 2011. Alginates as biomaterials in tissue engineering. *Carbohydrate chemistry: chemical and biological approaches* 37, pp. 227-258.

Clark, S. T. et al. 2015. Phenotypic diversity within a Pseudomonas aeruginosa population infecting an adult with cystic fibrosis. *Scientific reports* 5, p. 10932.

Coburn, B. et al. 2015. Lung microbiota across age and disease stage in cystic fibrosis. *Scientific reports* 5(1), pp. 1-12.

Connor, T. R. et al. 2016. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. *Microbial genomics* 2(9),

Consortium, T. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486(7402), pp. 207-214.

Cox, M. J. et al. 2010. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. *PloS one* 5(6), p. e11044.

Cullen, L. and McClean, S. 2015. Bacterial adaptation during chronic respiratory infections. *Pathogens* 4(1), pp. 66-89.

Cullen, L. et al. 2015. Phenotypic characterisation of an international Pseudomonas aeruginosa reference panel: Strains of cystic fibrosis origin show less in vivo virulence than non-CF strains. *Microbiology-SGM* 161(10), pp. 1961-1977.

Cuthbertson, L. et al. 2020. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* 8, pp. 1-13.

Cystic Fibrosis Trust. 2019, 2nd February-a. Available at: https://www.cysticfibrosis.org.uk/what-is-cystic-fibrosis [Accessed.

Cystic Fibrosis Trust. 2019, 2nd February-b. Available at: https://www.cysticfibrosis.org.uk/what-is-cystic-fibrosis [Accessed: 20th February].

D'Argenio, D. A. et al. 2007. Growth phenotypes of Pseudomonas aeruginosa lasR mutants adapted to the airways of cystic fibrosis patients. *Molecular microbiology* 64(2), pp. 512-533.

D'Sylva, P. et al. 2017. Induced sputum to detect lung pathogens in young children with cystic fibrosis. *Pediatric pulmonology* 52(2), pp. 182-189.

Dafydd, C., Saunders, B., Edmondson, C., Doull, I., Thia, L., Dhillon, R. and Forton, J. T. 2020. Home sputum-induction testing in children with cystic fibrosis. *Pediatric Pulmonology*, pp. 164-164.

Dagenais, R. V. E., Su, V. C. and Quon, B. S. 2021. Real-world safety of CFTR modulators in the treatment of cystic fibrosis: A systematic review. *Journal of Clinical Medicine* 10(1), p. 23.

Daniels, T. W. V. et al. 2013. Impact of antibiotic treatment for pulmonary exacerbations on bacterial diversity in cystic fibrosis. *Journal of Cystic Fibrosis* 12(1), pp. 22-28.

Davies, G. et al. 2020. Characterising burden of treatment in cystic fibrosis to identify priority areas for clinical trials. *Journal of Cystic Fibrosis* 19(3), pp. 499-502.

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. and Callahan, B. J. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6(1), pp. 1-14.

Davril, M. et al. 1999. The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection. *Glycobiology* 9(3), pp. 311-321.

de Agüero, M. G. et al. 2016. The maternal microbiota drives early postnatal innate immune development. *Science* 351(6279), pp. 1296-1302.

De Blic, J. et al. 2000. Bronchoalveolar lavage in children. ERS Task Force on bronchoalveolar lavage in children. European Respiratory Society. *European Respiratory Journal* 15(1), pp. 217-231.

De Boeck, K. and Amaral, M. D. 2016. Progress in therapies for cystic fibrosis. *The Lancet Respiratory Medicine* 4(8), pp. 662-674.

de Koff, E. M., de Winter-de Groot, K. M. and Bogaert, D. 2016. Development of the respiratory tract microbiota in cystic fibrosis. *Current opinion in pulmonary medicine* 22(6), pp. 623-628.

Degroote, S., Ducourouble, M.-P., Roussel, P. and Lamblin, G. 1999. Sequential biosynthesis of sulfated and/or sialylated Lewis x determinants by transferases of the human bronchial mucosa. *Glycobiology* 9(11), pp. 1199-1211.

Delmotte, P., Degroote, S., Lafitte, J.-J., Lamblin, G., Perini, J.-M. and Roussel, P. 2002. Tumor necrosis factor α increases the expression of glycosyltransferases and sulfotransferases responsible for the biosynthesis of sialylated and/or sulfated Lewis x epitopes in the human bronchial mucosa. *Journal of Biological Chemistry* 277(1), pp. 424-431.

DeVries, C. A. a. O., D.E. 1994. Mucoid-to-non-mucoid conversion in alginate-producing *pseudomonas aeruginosa* often results from spontaneous mutations in *algT,* encoding a putative alternate sigma factor, and shows evidence for autoregulation. *Journal of Bacteriology*, pp. 6677-6687.

Deziel, E., Comeau, Y. and Villemur, R. 2001. Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *Journal of Bacteriology* 183(4), pp. 1195-1204. doi: 10.1128/jb.183.4.1195-1204.2001

Dickson, R. P., Martinez, F. J. and Huffnagle, G. B. 2014. The role of the microbiome in exacerbations of chronic lung diseases. *The Lancet* 384(9944), pp. 691-702.

Dixon, E. et al. 2021. Telemedicine and cystic fibrosis: Do we still need face-to-face clinics? *Paediatric respiratory reviews*,

Dournes, G. et al. 2021. The clinical use of lung MRI in cystic fibrosis: what, now, how? *Chest* 159(6), pp. 2205-2217.

Drenkard, E. and Ausubel, F. M. 2002. *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* 416(6882), pp. 740-743. doi: 10.1038/416740a

Döring, G., Flume, P., Heijerman, H., Elborn, J. S. and Consensus Study, G. 2012. Treatment of lung infection in patients with cystic fibrosis: current and future strategies. *Journal of Cystic Fibrosis* 11(6), pp. 461-479.

Edmondson, C. et al. 2018. Cystic fibrosis newborn screening: outcome of infants with normal sweat tests. *Archives of disease in childhood* 103(8), pp. 753-756.

Ehre, C., Ridley, C. and Thornton, D. J. 2014. Cystic fibrosis: an inherited disease affecting mucin-producing organs. *The international journal of biochemistry & cell biology* 52, pp. 136-145.

Ehre, C. et al. 2012. Overexpressing mouse model demonstrates the protective role of Muc5ac in the lungs. *Proceedings of the National Academy of Sciences* 109(41), pp. 16528-16533.

Ermund, A. et al. 2017. OligoG CF-5/20 normalizes cystic fibrosis mucus by chelating calcium. *Clin Exp Pharmacol Physiol* 44(6), pp. 639-647. doi: 10.1111/1440-1681.12744

Fleming-Dutra, K. E. et al. 2016. Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011. *Jama* 315(17), pp. 1864-1873.

Flight, W. G. et al. 2015. Rapid detection of emerging pathogens and loss of microbial diversity associated with severe lung disease in cystic fibrosis. *Journal of clinical microbiology* 53(7), pp. 2022-2029.

Floyd, M. et al. 2016. Swimming Motility Mediates the Formation of Neutrophil Extracellular Traps Induced by Flagellated Pseudomonas aeruginosa. *PLoS Pathog* 12(11), p. e1005987. doi: 10.1371/journal.ppat.1005987

Flynn, J. M., Niccum, D., Dunitz, J. M. and Hunter, R. C. 2016. Evidence and role for bacterial mucin degradation in cystic fibrosis airway disease. *PLoS pathogens* 12(8),

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

Forton, J. 2015. Induced sputum in young healthy children with cystic fibrosis. *Paediatric respiratory reviews* 16, pp. 6-8.

Forton, J. T. 2019. Detecting respiratory infection in children with cystic fibrosis: Cough swab, sputum induction or bronchoalveolar lavage. *Paediatric respiratory reviews* 31, pp. 28-31.

Fothergill, J. L., Mowat, E., Ledson, M. J., Walshaw, M. J. and Winstanley, C. 2010. Fluctuations in phenotypes and genotypes within populations of Pseudomonas aeruginosa in the cystic fibrosis lung during pulmonary exacerbations. *Journal of medical microbiology* 59(4), pp. 472-481.

Frauchiger, B. S., Binggeli, S., Yammine, S., Spycher, B., Krüger, L., Ramsey, K. A. and Latzin, P. 2021. Longitudinal course of clinical lung clearance index in children with cystic fibrosis. *European respiratory journal* 58(1),

Gellatly, S. L. and Hancock, R. E. W. 2013. Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. *Pathogens and disease* 67(3), pp. 159-173.

Gollwitzer, E. S. et al. 2014. Lung microbiota promotes tolerance to allergens in neonates via PD-L1. *Nature medicine* 20(6), p. 642.

Goodrich, J. K. et al. 2014. Conducting a microbiome study. *Cell* 158(2), pp. 250-262.

Goss, C. H. and Burns, J. L. 2007. Exacerbations in cystic fibrosis center dot 1: Epidemiology and pathogenesis. *Thorax* 62(4), pp. 360-367. doi: 10.1136/thx.2006.060889

Goulding, J. et al. 2007. Respiratory infections: do we ever recover? *Proceedings of the American Thoracic Society* 4(8), pp. 618-625.

Griffen, A. L. et al. 2011. CORE: a phylogenetically-curated 16S rDNA database of the core oral microbiome. *PloS one* 6(4), p. e19051.

Gullberg, E., Cao, S., Berg, O. G., Ilbäck, C., Sandegren, L., Hughes, D. and Andersson, D. I. 2011. Selection of resistant bacteria at very low antibiotic concentrations. *PLoS pathogens* 7(7), p. e1002158.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29(8), pp. 1072-1075.

Ha, D.-G., Kuchma, S. L. and O'Toole, G. A. 2014. Plate-based assay for swimming motility in Pseudomonas aeruginosa.*Pseudomonas Methods and Protocols*. Springer, pp. 59-65.

Ha, D. G. and O'Toole, G. A. 2015. c-di-GMP and its Effects on Biofilm Formation and Dispersion: a Pseudomonas Aeruginosa Review. *Microbiol Spectr* 3(2), pp. Mb-0003-2014. doi: 10.1128/microbiolspec.MB-0003-2014

Harris, J. K. et al. 2007. Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic fibrosis. *Proceedings of the National Academy of Sciences* 104(51), pp. 20529-20533.

Hay, I. D., Gatland, K., Campisano, A., Jordens, J. Z. and Rehm, B. H. 2009a. Impact of alginate overproduction on attachment and biofilm architecture of a supermucoid Pseudomonas aeruginosa strain. *Appl Environ Microbiol* 75(18), pp. 6022-6025. doi: 10.1128/AEM.01078-09

Hay, I. D., Gatland, K., Campisano, A., Jordens, J. Z. and Rehm, B. H. A. 2009b. Impact of alginate overproduction on attachment and biofilm architecture of a supermucoid Pseudomonas aeruginosa strain. *Appl. Environ. Microbiol.* 75(18), pp. 6022-6025.

Hay, I. D., Rehman, Z. U., Moradali, M. F., Wang, Y. and Rehm, B. H. A. 2013a. Microbial alginate production, modification and its applications. *Microbial biotechnology* 6(6), pp. 637-650.

Hay, I. D., Ur Rehman, Z., Moradali, M. F., Wang, Y. and Rehm, B. H. 2013b. Microbial alginate production, modification and its applications. *Microb Biotechnol* 6(6), pp. 637-650. doi: 10.1111/1751-7915.12076

Hay, I. D., Wang, Y., Moradali, M. F., Rehman, Z. U. and Rehm, B. H. 2014. Genetics and regulation of bacterial alginate production. *Environ Microbiol* 16(10), pp. 2997-3011. doi: 10.1111/1462-2920.12389

Hengzhuang, W., Song, Z., Ciofu, O., Onsoyen, E., Rye, P. D. and Hoiby, N. 2016. OligoG CF-5/20 Disruption of Mucoid Pseudomonas aeruginosa Biofilm in a Murine Lung Infection Model. *Antimicrob Agents Chemother* 60(5), pp. 2620-2626. doi: 10.1128/AAC.01721-15

Henke, M. O., John, G., Germann, M., Lindemann, H. and Rubin, B. K. 2007. MUC5AC and MUC5B mucins increase in cystic fibrosis airway secretions during

pulmonary exacerbation. *Am J Respir Crit Care Med* 175(8), pp. 816-821. doi: 10.1164/rccm.200607-1011OC

Henke, M. O., Renner, A., Huber, R. M., Seeds, M. C. and Rubin, B. K. 2004. MUC5AC and MUC5B mucins are decreased in cystic fibrosis airway secretions. *American journal of respiratory cell and molecular biology* 31(1), pp. 86-91.

Henwood, C. J., Livermore, D. M., James, D., Warner, M. and Pseudomonas Study Group, t. 2001. Antimicrobial susceptibility of Pseudomonas aeruginosa: results of a UK survey and evaluation of the British Society for Antimicrobial Chemotherapy disc susceptibility test. *Journal of Antimicrobial Chemotherapy* 47(6), pp. 789-799.

Herbert, S. et al. 2021. Exploring the challenges of accessing medication for patients with cystic fibrosis. *Thorax*,

Hewer, S. C. L. and Smyth, A. R. 2017. Antibiotic strategies for eradicating Pseudomonas aeruginosa in people with cystic fibrosis. *Cochrane Database of Systematic Reviews* (4),

Hoboth, C. et al. 2009. Dynamics of adaptive microevolution of hypermutable Pseudomonas aeruginosa during chronic pulmonary infection in patients with cystic fibrosis. *The Journal of infectious diseases* 200(1), pp. 118-130.

Hoffman, L. R., Kulasekara, H. D., Emerson, J., Houston, L. S., Burns, J. L., Ramsey, B. W. and Miller, S. I. 2009. Pseudomonas aeruginosa lasR mutants are associated with cystic fibrosis lung disease progression. *Journal of Cystic Fibrosis* 8(1), pp. 66-70.

Hogardt, M. and Heesemann, J. 2010. Adaptation of Pseudomonas aeruginosa during persistence in the cystic fibrosis lung. *International Journal of Medical Microbiology* 300(8), pp. 557-562.

Horna, G., López, M., Guerra, H., Saénz, Y. and Ruiz, J. 2018. Interplay between MexAB-OprM and MexEF-OprN in clinical isolates of Pseudomonas aeruginosa. *Scientific reports* 8(1), pp. 1-11.

Horsley, A., Rousseau, K., Ridley, C., Flight, W., Jones, A., Waigh, T. A. and Thornton, D. J. 2014. Reassessment of the importance of mucins in determining sputum properties in cystic fibrosis. *Journal of Cystic Fibrosis* 13(3), pp. 260-266.

Huse, H. K., Kwon, T., Zlosnik, J. E. A., Speert, D. P., Marcotte, E. M. and Whiteley, M. 2010. Parallel evolution in Pseudomonas aeruginosa over 39,000 generations in vivo. *MBio* 1(4), pp. e00199-00110.

Häussler, S. 2010. Multicellular signalling and growth of Pseudomonas aeruginosa. *International Journal of Medical Microbiology* 300(8), pp. 544-548.

Häußler, S., Tümmler, B., Weißbrodt, H., Rohde, M. and Steinmetz, I. 1999. Small-colony variants of Pseudomonas aeruginosa in cystic fibrosis. *Clinical infectious diseases* 29(3), pp. 621-625.

Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S. and Ciofu, O. 2010. Antibiotic resistance of bacterial biofilms. *International Journal of Antimicrobial Agents* 35(4), pp. 322-332. doi: 10.1016/j.ijantimicag.2009.12.011

Jack, A. A. et al. 2018. Alginate oligosaccharide-induced modification of the *lasI-lasR* and *rhlI-rhlR* quorum-sensing systems in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 62, pp. e02318-02317. doi: doi.org/10.1128/AAC.02318-17

Jain, M. et al. 2004. Type III secretion phenotypes of Pseudomonas aeruginosa strains change during infection of individuals with cystic fibrosis. *Journal of Clinical Microbiology* 42(11), pp. 5229-5237.

Jakobsson, H. E., Jernberg, C., Andersson, A. F., Sjölund-Karlsson, M., Jansson, J. K. and Engstrand, L. 2010. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PloS one* 5(3), p. e9836.

Jayaraman, S., Joo, N. S., Reitz, B., Wine, J. J. and Verkman, A. S. 2001. Submucosal gland secretions in airways from cystic fibrosis patients have normal [Na+] and pH but elevated viscosity. *Proceedings of the National Academy of Sciences* 98(14), pp. 8119-8123.

Jesaitis, A. J. et al. 2003. Compromised host defense on Pseudomonas aeruginosa biofilms: characterization of neutrophil and biofilm interactions. *The Journal of Immunology* 171(8), pp. 4329-4339.

Johns, B. E., Purdy, K. J., Tucker, N. P. and Maddocks, S. E. 2015. Phenotypic and genotypic characteristics of small colony variants and their role in chronic infection. *Microbiology insights* 8, pp. MBI-S25800.

Johnson, J. S. et al. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature communications* 10(1), pp. 1-11.

Jorgensen, J. H. and Turnidge, J. D. 2015. Susceptibility test methods: dilution and disk diffusion methods. *Manual of clinical microbiology*, pp. 1253-1273.

Jorth, P. et al. 2015. Regional isolation drives bacterial diversification within cystic fibrosis lungs. *Cell host & microbe* 18(3), pp. 307-319.

Jovel, J. et al. 2016. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Frontiers in microbiology* 7, p. 459.

Katkin, J. 2019a. *Cystic fibrosis: Clinical manifestations and diagnosis*. Available at: https://www-uptodate-com.abc.cardiff.ac.uk/contents/cystic-fibrosis-clinical-manifestations-and-diagnosis [Accessed: February].

Katkin, J. 2019b. *Cystic fibrosis: genetics and pathogenesis*. Available at: https://www-uptodate-com.abc.cardiff.ac.uk/contents/cystic-fibrosis-genetics-and-pathogenesis?search=cystic%20fibrosis%20microbiome&topicRef=6371&source=see_link [Accessed: 20th February].

Khan, S. et al. 2012a. Overcoming drug resistance with alginate oligosaccharides able to potentiate the action of selected antibiotics. *Antimicrobial Agents and Chemotherapy* 56(10), pp. 5134-5141.

Khan, S. et al. 2012b. Overcoming drug resistance with alginate oligosaccharides able to potentiate the action of selected antibiotics. *Antimicrob Agents Chemother* 56(10), pp. 5134-5141. doi: 10.1128/AAC.00525-12

Khan, S. et al. 2012c. Overcoming drug resistance with alginate oligosaccharides able to potentiate the action of selected antibiotics. *Antimicrobial agents and chemotherapy* 56(10), pp. 5134-5141.

Kidd, T. J. et al. 2018. Defining antimicrobial resistance in cystic fibrosis. *Journal of Cystic Fibrosis*,

Kidd, T. J. et al. 2015. Pseudomonas aeruginosa genotypes acquired by children with cystic fibrosis by age 5-years. *Journal of Cystic Fibrosis* 14(3), pp. 361-369.

Kim, B.-R. et al. 2017. Deciphering diversity indices for a better understanding of microbial communities. *Journal of Microbiology and Biotechnology* 27(12), pp. 2089-2093.

Kim, H. K. and Harshey, R. M. 2016. A Diguanylate Cyclase Acts as a Cell Division Inhibitor in a Two-Step Response to Reductive and Envelope Stresses. *mBio* 7(4), pp. e00822-00816. doi: 10.1128/mBio.00822-16

King, M. and Rubin, B. K. 2002. Pharmacological approaches to discovery and development of new mucolytic agents. *Adv Drug Deliv Rev* 54(11), pp. 1475-1490. doi: 10.1016/s0169-409x(02)00156-4

Kipnis, E., Sawa, T. and Wiener-Kronish, J. 2006. Targeting mechanisms of Pseudomonas aeruginosa pathogenesis. *Med Mal Infect* 36(2), pp. 78-91. doi: 10.1016/j.medmal.2005.10.007

Kirisits, M. J., Prost, L., Starkey, M. and Parsek, M. R. 2005. Characterization of colony morphology variants isolated from Pseudomonas aeruginosa biofilms. *Appl. Environ. Microbiol.* 71(8), pp. 4809-4821.

Klockgether, J. et al. 2010. Genome diversity of Pseudomonas aeruginosa PAO1 laboratory strains. *Journal of bacteriology* 192(4), pp. 1113-1121.

Klockgether, J. and Tümmler, B. 2017. Recent advances in understanding Pseudomonas aeruginosa as a pathogen. *F1000Research* 6,

Kocincova, D. and Lam, J. S. 2011. Structural diversity of the core oligosaccharide domain of Pseudomonas aeruginosa lipopolysaccharide. *Biochemistry (Moscow)* 76(7), pp. 755-760.

Konstan, M. W. and Flume, P. A. 2020. Clinical care for cystic fibrosis: preparing for the future now. *The Lancet Respiratory Medicine* 8(1), pp. 10-12.

Kotnala, S., Dhasmana, A., Kashyap, V. K., Chauhan, S. C., Yallapu, M. M. and Jaggi, M. 2021. A bird eye view on cystic fibrosis: An underestimated multifaceted chronic disorder. *Life Sciences* 268, p. 118959.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. and Schloss, P. D. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology* 79(17), pp. 5112-5120.

Krueger, F. Trim Galore! A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D. and Knight, R. 2012. Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics* 13(1), pp. 47-58.

Kunin, V., Engelbrektson, A., Ochman, H. and Hugenholtz, P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental microbiology* 12(1), pp. 118-123.

Kunzelmann, K., Schreiber, R. and Hadorn, H. B. 2017. Bicarbonate in cystic fibrosis. *Journal of Cystic Fibrosis* 16(6), pp. 653-662.

Lai, S. K., Wang, Y. Y., Wirtz, D. and Hanes, J. 2009. Micro- and macrorheology of mucus. *Adv Drug Deliv Rev* 61(2), pp. 86-100. doi: 10.1016/j.addr.2008.09.012

Lamblin, G. et al. 2001. Human airway mucin glycosylation: A combinatory of carbohydrate determinants which vary in cystic fibrosis. *Glycoconjugate Journal* 18(9), pp. 661-684. doi: 10.1023/A:1020867221861

Lang, G. I. and Desai, M. M. 2014. The spectrum of adaptive mutations in experimental evolution. *Genomics* 104(6), pp. 412-416.

Lemos, L. N., Fulthorpe, R. R., Triplett, E. W. and Roesch, L. F. W. 2011. Rethinking microbial diversity analysis in the high throughput sequencing era. *Journal of microbiological methods* 86(1), pp. 42-51.

Lenski, R. E. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME journal* 11(10), p. 2181.

Lethem, M. I., James, S. L. and Marriott, C. 1990. The role of mucous glycoproteins in the rheologic properties of cystic fibrosis sputum1-3. *Am Rev Respir Dis* 142, pp. 1053-1058.

Lewis, A. T., Jones, K., Lewis, K. E., Jones, S. and Lewis, P. D. 2013a. Detection of Lewis antigen structural change by FTIR spectroscopy. *Carbohydr Polym* 92(2), pp. 1294-1301. doi: 10.1016/j.carbpol.2012.09.078

Lewis, P. D. et al. 2010. Evaluation of FTIR spectroscopy as a diagnostic tool for lung cancer using sputum. *BMC cancer* 10(1), p. 640.

Lewis, S. P., Lewis, A. T. and Lewis, P. D. 2013b. Prediction of glycoprotein secondary structure using ATR-FTIR. *Vibrational Spectroscopy* 69, pp. 21-29.

Li, Z. et al. 2005. Longitudinal development of mucoid Pseudomonas aeruginosa infection and lung disease progression in children with cystic fibrosis. *Jama* 293(5), pp. 581-588.

Lieberman, T. D. et al. 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature genetics* 43(12), p. 1275.

Lillehoj, E. P., Kato, K., Lu, W. and Kim, K. C. 2013. Cellular and molecular biology of airway mucins. *International review of cell and molecular biology*. Vol. 303. Elsevier, pp. 139-202.

Lim, M. Y., Yoon, H. S., Rho, M., Sung, J., Song, Y.-M., Lee, K. and Ko, G. 2016. Analysis of the association between host genetics, smoking, and sputum microbiota in healthy humans. *Scientific reports* 6, p. 23745.

Linden, S. K., Sutton, P., Karlsson, N. G., Korolik, V. and McGuckin, M. A. 2008. Mucins in the mucosal barrier to infection. *Mucosal Immunol* 1(3), pp. 183-197. doi: 10.1038/mi.2008.5

Linnane, B. et al. 2021. The lung microbiome in young children with cystic fibrosis: a prospective cohort study. *Microorganisms* 9(3), p. 492.

Liu, Y. et al. 2022. Microevolution of the mexT and lasR Reinforces the Bias of Quorum Sensing System in Laboratory Strains of Pseudomonas aeruginosa PAO1. *Frontiers in microbiology*,

Lloyd-Price, J., Abu-Ali, G. and Huttenhower, C. 2016. The healthy human microbiome. *Genome medicine* 8(1), p. 51.

Lopes-Pacheco, M. 2016. CFTR modulators: shedding light on precision medicine for cystic fibrosis. *Frontiers in pharmacology* 7, p. 275.

Lopez-Causape, C., Rojo-Molinero, E., Macia, M. D. and Oliver, A. 2015. The problems of antibiotic resistance in cystic fibrosis and solutions. *Expert Rev Respir Med* 9(1), pp. 73-88. doi: 10.1586/17476348.2015.995640

Ludwig, J. A., Quartet, L., Reynolds, J. F. and Reynolds, J. F. 1988. *Statistical ecology: a primer in methods and computing*. John Wiley & Sons.

Lund-Palau, H. et al. 2016. Pseudomonas aeruginosa infection in cystic fibrosis: pathophysiological mechanisms and therapeutic approaches. *Expert review of respiratory medicine* 10(6), pp. 685-697.

Lutz, L., Pereira, D. C., Paiva, R. M., Zavascki, A. P. and Barth, A. L. 2012. Macrolides decrease the minimal inhibitory concentration of anti-pseudomonal agents against Pseudomonas aeruginosa from cystic fibrosis patients in biofilm. *BMC microbiology* 12(1), pp. 1-7.

Ma, J., Rubin, B. K. and Voynow, J. A. 2018. Mucins, mucus, and goblet cells. *Chest* 154(1), pp. 169-176.

Mahenthiralingam, E., Campbell, M. E. and Speert, D. P. 1994. Nonmotility and phagocytic resistance of Pseudomonas aeruginosa isolates from chronically colonized patients with cystic fibrosis. *Infection and immunity* 62(2), pp. 596-605.

Malone, J. G. 2015. Role of small colony variants in persistence of Pseudomonas aeruginosa infections in cystic fibrosis lungs. *Infection and drug resistance* 8, p. 237.

Malone, J. G. et al. 2010. YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in Pseudomonas aeruginosa. *PLoS pathogens* 6(3), p. e1000804.

Man, W. H., de Steenhuijsen Piters, W. A. A. and Bogaert, D. 2017. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nature Reviews Microbiology* 15(5), p. 259.

Marchesi, J. R. and Ravel, J. 2015. The vocabulary of microbiome research: a proposal. Springer.

Markussen, T. et al. 2014. Environmental heterogeneity drives within-host diversification and evolution of Pseudomonas aeruginosa. *MBio* 5(5), pp. e01592-01514. doi: 10.1128/mBio.01592-14

Marsh, R. L., Kaestli, M., Chang, A. B., Binks, M. J., Pope, C. E., Hoffman, L. R. and Smith-Vaughan, H. C. 2016. The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. *Microbiome* 4(1), p. 37.

Martin, M., Hölscher, T., Dragoš, A., Cooper, V. S. and Kovács, Á. T. 2016. Laboratory evolution of microbial interactions in bacterial biofilms. *Journal of bacteriology* 198(19), pp. 2564-2571.

Marvig, R. L., Sommer, L. M., Molin, S. and Johansen, H. K. 2015. Convergent evolution and adaptation of Pseudomonas aeruginosa within patients with cystic fibrosis. *Nature genetics* 47(1), p. 57.

Matsui, H., Grubb, B. R., Tarran, R., Randell, S. H., Gatzy, J. T., Davis, C. W. and Boucher, R. C. 1998. Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. *Cell* 95(7), pp. 1005-1015.

Mayer-Hamblett, N. et al. 2014. Pseudomonas aeruginosa in vitro phenotypes distinguish cystic fibrosis infection stages and outcomes. *American journal of respiratory and critical care medicine* 190(3), pp. 289-297.

Maziarz, M., Pfeiffer, R. M., Wan, Y. and Gail, M. H. 2018. Using standard microbiome reference groups to simplify beta-diversity analyses and facilitate independent validation. *Bioinformatics* 34(19), pp. 3249-3257.

McElroy, K. E. et al. 2014. Strain-specific parallel evolution drives short-term diversification during Pseudomonas aeruginosa biofilm formation. *Proceedings of the National Academy of Sciences* 111(14), p. E1419.

McGinn, S. and Gut, I. G. 2013. DNA sequencing–spanning the generations. *New biotechnology* 30(4), pp. 366-372.

Micah, H., Claire, F.-L. and Rob, K. 2007. The Human Microbiome Project: Exploring the Microbial Part of Ourselves in a Changing World. *Nature* 449(7164), pp. 804-810.

Middleton, P. G. and Taylor-Cousar, J. L. 2021. Development of elexacaftor–tezacaftor–ivacaftor: Highly effective CFTR modulation for the majority of people with Cystic Fibrosis. *Expert Review of Respiratory Medicine*, pp. 1-13.

Mogayzel Jr, P. J. et al. 2013. Cystic fibrosis pulmonary guidelines: chronic medications for maintenance of lung health. *American journal of respiratory and critical care medicine* 187(7), pp. 680-689.

Moreau-Marquis, S., Stanton, B. A. and O'Toole, G. A. 2008. Pseudomonas aeruginosa biofilm formation in the cystic fibrosis airway. *Pulmonary pharmacology & therapeutics* 21(4), pp. 595-599.

Morrison, C. B., Markovetz, M. R. and Ehre, C. 2019. Mucus, mucins, and cystic fibrosis. *Pediatr Pulmonol* 54 Suppl 3, pp. S84-s96. doi: 10.1002/ppul.24530

Mott, L. S. et al. 2012. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 67(6), pp. 509-516.

Mowat, E. et al. 2011. Pseudomonas aeruginosa population diversity and turnover in cystic fibrosis chronic infections. *American journal of respiratory and critical care medicine* 183(12), pp. 1674-1679.

Mrsny, R., Daugherty, A., Short, S., Widmer, R., Siegel, M. and Keller, G.-A. 1996. Distribution of DNA and alginate in purulent cystic fibrosis sputum: implications to pulmonary targeting strategies. *Journal of drug targeting* 4(4), pp. 233-243.

Mulcahy, L. R., Burns, J. L., Lory, S. and Lewis, K. 2010. Emergence of Pseudomonas aeruginosa Strains Producing High Levels of Persister Cells in Patients with Cystic Fibrosis. *Journal of Bacteriology* 192(23), p. 6191. doi: 10.1128/JB.01651-09

Muzzey, D., Evans, E. A. and Lieber, C. 2015. Understanding the basics of NGS: from mechanism to variant calling. *Current genetic medicine reports* 3(4), pp. 158-165.

Narayanan, S., Mainz, J. G., Gala, S., Tabori, H. and Grossoehme, D. 2017. Adherence to therapies in cystic fibrosis: a targeted literature review. *Expert review of respiratory medicine* 11(2), pp. 129-145.

Nichols, D. P. et al. 2020. Pulmonary outcomes associated with long-term azithromycin therapy in cystic fibrosis. *American journal of respiratory and critical care medicine* 201(4), pp. 430-437.

Nivens, D. E., Ohman, D. E., Williams, J. and Franklin, M. J. 2001. Role of alginate and its O acetylation in formation of Pseudomonas aeruginosa microcolonies and biofilms. *J Bacteriol* 183(3), pp. 1047-1057. doi: 10.1128/JB.183.3.1047-1057.2001

Noller, H. F. 1984. Structure of ribosomal RNA. *Annual review of biochemistry* 53(1), pp. 119-162.

Nordgård, C. T. and Draget, K. I. 2011. Oligosaccharides as modulators of rheology in complex mucous systems. *Biomacromolecules* 12(8), pp. 3084-3090.

Nordgård, C. T., Nonstad, U., Olderøy, M. Ø., Espevik, T. and Draget, K. I. 2014. Alterations in mucus barrier function and matrix structure induced by guluronate oligomers. *Biomacromolecules* 15(6), pp. 2294-2300.

O'Toole, G. A. 2011. Microtiter dish biofilm formation assay. *JoVE (Journal of Visualized Experiments)* (47), p. e2437.

Oakley, J. L. et al. 2021. Phenotypic and Genotypic Adaptations in Pseudomonas aeruginosa Biofilms following Long-Term Exposure to an Alginate Oligomer Therapy. *Msphere* 6(1),

Oliver, A. 2004. Carbapenem resistance and Acinetobacter baumannii. *Enfermedades infecciosas y microbiologia clinica* 22(5), p. 259.

Oliver, A., Cantón, R., Campo, P., Baquero, F. and Blázquez, J. 2000. High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. *Science* 288(5469), pp. 1251-1253.

Oliver, A., Mulet, X., Lopez-Causape, C. and Juan, C. 2015. The increasing threat of Pseudomonas aeruginosa high-risk clones. *Drug Resistance Updates* 21, pp. 41-59.

Organisation, W. H. 2014. *Antimicrobial resistance: global report on surveillance.* World Health Organisation: [Accessed: 16t[h] February 2017].

Ostedgaard, L. S. et al. 2017. Gel-forming mucins form distinct morphologic structures in airways. *Proceedings of the National Academy of Sciences* 114(26), pp. 6842-6847.

Overhage, J., Bains, M., Brazas, M. D. and Hancock, R. E. W. 2008. Swarming of Pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. *Journal of bacteriology* 190(8), pp. 2671-2679.

O'Neill, J. 2016. Tackling Drug-resistant Infections Globally: Final Report and Recommendations–The Review on Antimicrobial Resistance Chaired by Jim O'Neill. *Wellcome Trust and HM Government, London*,

Parnham, M. J., Haber, V. E., Giamarellos-Bourboulis, E. J., Perletti, G., Verleden, G. M. and Vos, R. 2014. Azithromycin: mechanisms of action and their relevance for clinical applications. *Pharmacology & therapeutics* 143(2), pp. 225-245.

Paterson, S. L., Barry, P. J. and Horsley, A. R. 2020. Tezacaftor and ivacaftor for the treatment of cystic fibrosis. *Expert review of respiratory medicine* 14(1), pp. 15-30.

Peix, A., Ramírez-Bahena, M.-H. and Velázquez, E. 2009. Historical evolution and current status of the taxonomy of genus Pseudomonas. *Infection, Genetics and Evolution* 9(6), pp. 1132-1147.

Pestrak, M. J. et al. 2018. *Pseudomonas aeruginosa* rugose small-colony variants evade host clearance, are hyper-inflammatory, and persist in multiple host environments. *Plos Pathogens* 14(2), doi: 10.1371/journal.ppat.1006842

Pettigrew, M. M., Laufer, A. S., Gent, J. F., Kong, Y., Fennie, K. P. and Metlay, J. P. 2012. Upper respiratory tract microbial communities, acute otitis media pathogens, and antibiotic use in healthy and sick children. *Appl. Environ. Microbiol.* 78(17), pp. 6262-6270.

Pitt, T. L., Sparrow, M., Warner, M. and Stefanidou, M. 2003. Survey of resistance of Pseudomonas aeruginosa from UK patients with cystic fibrosis to six commonly prescribed antimicrobial agents. *Thorax* 58(9), pp. 794-796.

Poltak, S. R. and Cooper, V. S. 2010. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *The Isme Journal* 5, p. 369. doi: 10.1038/ismej.2010.136
https://www.nature.com/articles/ismej2010136#supplementary-information

Poltak, S. R. and Cooper, V. S. 2011a. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *Isme Journal* 5(3), pp. 369-378. doi: 10.1038/ismej.2010.136

Poltak, S. R. and Cooper, V. S. 2011b. Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *Isme j* 5(3), pp. 369-378. doi: 10.1038/ismej.2010.136

Powell, L. C. et al. 2014a. A nanoscale characterization of the interaction of a novel alginate oligomer with the cell surface and motility of *Pseudomonas aeruginosa*. *American Journal of Respiratory Cell and Molecular Biology* 50(3), pp. 483-492.

Powell, L. C. et al. 2014b. A nanoscale characterization of the interaction of a novel alginate oligomer with the cell surface and motility of Pseudomonas aeruginosa. *American journal of respiratory cell and molecular biology* 50(3), pp. 483-492.

Powell, L. C. et al. 2018. Targeted disruption of the extracellular polymeric network of Pseudomonas aeruginosa biofilms by alginate oligosaccharides. *npj Biofilms and Microbiomes* 4, doi: 10.1038/s41522-018-0056-3

Powell, L. C. et al. 2013. The effect of alginate oligosaccharides on the mechanical properties of Gram-negative biofilms. *Biofouling* 29(4), pp. 413-421.

Prevaes, S. M. P. J. et al. 2016. Development of the nasopharyngeal microbiota in infants with cystic fibrosis. *American journal of respiratory and critical care medicine* 193(5), pp. 504-515.

Pritchard, M. F. et al. 2019. Mucin structural interactions with an alginate oligomer mucolytic in cystic fibrosis sputum. *Vibrational Spectroscopy* 103, p. 102932.

Pritchard, M. F. et al. 2017a. A low molecular weight alginate oligosaccharide disrupts pseudomonal microcolony formation and enhances antibiotic effectiveness. *American Agents and Chemotherapy*, doi: doi: 10.1128/AAC.00762-17

Pritchard, M. F. et al. 2017b. The antimicrobial effects of the alginate oligomer OligoG CF-5/20 are independent of direct bacterial cell membrane disruption. *Scientific reports* 7(1), pp. 1-12.

Pritchard, M. F. et al. 2016a. A New Class of Safe Oligosaccharide Polymer Therapy To Modify the Mucus Barrier of Chronic Respiratory Disease. *Mol Pharm* 13(3), pp. 863-872. doi: 10.1021/acs.molpharmaceut.5b00794

Pritchard, M. F. et al. 2016b. A new class of safe oligosaccharide polymer therapy to modify the mucus barrier of chronic respiratory disease. *Molecular pharmaceutics* 13(3), pp. 863-872.

Périchon, B., Courvalin, P. and Stratton, C. W. 2019. Antibiotic Resistance☆. In: Schmidt, T.M. ed. *Encyclopedia of Microbiology (Fourth Edition)*. Oxford: Academic Press, pp. 127-139.

Qin, Y. 2008. Alginate fibres: an overview of the production processes and applications in wound management. *Polymer International* 57(2), pp. 171-180. doi: 10.1002/pi.2296

Quanjer, P. H. et al. 2012. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respiratory Soc.

Quinton, P. M. 2008. Cystic fibrosis: impaired bicarbonate secretion and mucoviscidosis. *The Lancet* 372(9636), pp. 415-417.

R-Core-Team. 2013. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Ramos, K. J., Pilewski, J. M. and Taylor-Cousar, J. L. 2021. Challenges in the use of highly effective modulator treatment for cystic fibrosis. *Journal of Cystic Fibrosis*,

Ramphal, R. and Arora, S. K. 2001. Recognition of mucin components by Pseudomonas aeruginosa. *Glycoconjugate journal* 18(9), pp. 709-713.

Ratjen, F., Bell, S. C., Rowe, S. M., Goss, C. H., Quittner, A. L. and Bush, A. 2015. Cystic fibrosis. *Nat Rev Dis Primers* 1, p. 15010. doi: 10.1038/nrdp.2015.10

Registry, C. F. F. P. 2018. *2017 Annual Data Report*. Bethesda, Maryland: Available at: https://www.cff.org/Research/Researcher-Resources/Patient-Registry/2017-Patient-Registry-Annual-Data-Report.pdf [Accessed: 16th August 2019].

Rehm, B. H. 2010. Bacterial polymers: biosynthesis, modifications and applications. *Nat Rev Microbiol* 8(8), pp. 578-592. doi: 10.1038/nrmicro2354

Relman, D. A. 2012. The human microbiome: ecosystem resilience and health. *Nutrition reviews* 70(suppl_1), pp. S2-S9.

Ricotta, C. and Podani, J. 2017. On some properties of the Bray-Curtis dissimilarity and their ecological meaning. *Ecological Complexity* 31, pp. 201-205.

Riordan, J. R. et al. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 245(4922), pp. 1066-1073.

Robinson, C. V., Elkins, M. R., Bialkowski, K. M., Thornton, D. J. and Kertesz, M. A. 2012. Desulfurization of mucin by Pseudomonas aeruginosa: influence of sulfate in the lungs of cystic fibrosis patients. *Journal of medical microbiology* 61(12), pp. 1644-1653.

Rogers, G. B., Carroll, M. P., Serisier, D. J., Hockey, P. M., Jones, G. and Bruce, K. D. 2004. Characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16S ribosomal DNA terminal restriction fragment length polymorphism profiling. *Journal of clinical microbiology* 42(11), pp. 5176-5183.

Ronchetti, K. et al. 2018a. The CF-Sputum Induction Trial (CF-SpIT) to assess lower airway bacterial sampling in young children with cystic fibrosis: a prospective internally controlled interventional trial. *The Lancet Respiratory Medicine* 6(6), pp. 461-471.

Ronchetti, K. et al. 2018b. The CF-Sputum Induction Trial (CF-SpIT) to assess lower airway bacterial sampling in young children with cystic fibrosis: a prospective internally controlled interventional trial. *Lancet Respiratory Medicine* 6(6), pp. 461-471. doi: 10.1016/s2213-2600(18)30171-1

Rose, M. C. and Voynow, J. A. 2006. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* 86(1), pp. 245-278. doi: 10.1152/physrev.00010.2005

Rosenfeld, M. et al. 1999. Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. *Pediatric pulmonology* 28(5), pp. 321-328.

Rosenfeld, M. et al. 2001. Early pulmonary infection, inflammation, and clinical outcomes in infants with cystic fibrosis. *Pediatric pulmonology* 32(5), pp. 356-366.

Rowland, R. E. and Nickless, E. M. 2000. Confocal microscopy opens the door to 3-dimensional analysis of cells. *Bioscene* 26(3), pp. 3-7.

Roy, M. G. et al. 2014. Muc5b is required for airway defence. *Nature* 505(7483), pp. 412-416.

Saint-Criq, V. and Gray, M. A. 2017. Role of CFTR in epithelial physiology. *Cellular and Molecular Life Sciences* 74(1), pp. 93-115.

Salter, S. J. et al. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology* 12(1), p. 87.

Santiago, A. J. et al. 2016. Inhibition and dispersal of Pseudomonas aeruginosa biofilms by combination treatment with escapin intermediate products and hydrogen peroxide. *Antimicrobial agents and chemotherapy* 60(9), pp. 5554-5562.

Scharfman, A., Delmotte, P., Beau, J., Lamblin, G., Roussel, P. and Mazurier, J. 2000. Sialyl-Lex and sulfo-sialyl-Lex determinants are receptors for P. aeruginosa. *Glycoconjugate journal* 17(10), pp. 735-740.

Schatten, H. 2012. *Scanning electron microscopy for the life sciences*. Cambridge University Press.

Schloss, P. D. 2010. The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS Comput Biol* 6(7), p. e1000844.

Schloss, P. D. et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75(23), pp. 7537-7541.

Schulz, B. L. et al. 2007. Glycosylation of sputum mucins is altered in cystic fibrosis patients. *Glycobiology* 17(7), pp. 698-712.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14), pp. 2068-2069.

Seemann, T. 2015. Snippy: Fast bacterial variant calling from NGS reads [WWW Document]. *URL https://github. com/tseemann/snippy*,

Sender, R., Fuchs, S. and Milo, R. 2016. Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell* 164(3), pp. 337-340.

Shafquat, A., Joice, R., Simmons, S. L. and Huttenhower, C. 2014. Functional and phylogenetic assembly of microbial communities in the human microbiome. *Trends in microbiology* 22(5), pp. 261-266.

Sharp, K. 2014. Section 2 Confocal laser scanning microscopy. *Biofouling Methods*, p. 15.

Shori, D. et al. 2001. Altered sialyl-and fucosyl-linkage on mucins in cystic fibrosis patients promotes formation of the sialyl-Lewis X determinant on salivary MUC-5B and MUC-7. *Pflügers Archiv* 443(1), pp. S55-S61.

Simpson, E. H. 1949. Measurement of diversity. *nature* 163(4148), pp. 688-688.

Sletmoen, M., Maurstad, G., Nordgård, C. T., Draget, K. I. and Stokke, B. T. 2012. Oligoguluronate induced competitive displacement of mucin–alginate interactions: relevance for mucolytic function. *Soft Matter* 8(32), pp. 8413-8421.

Sly, P. D. et al. 2009. Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening. *American journal of respiratory and critical care medicine* 180(2), pp. 146-152.

Sly, P. D. et al. 2013. Risk factors for bronchiectasis in children with cystic fibrosis. *New England Journal of Medicine* 368(21), pp. 1963-1970.

Smith, E. E. et al. 2006a. Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. *Proceedings of the National Academy of Sciences* 103(22), pp. 8487-8492.

Smith, E. E. et al. 2006b. Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. *Proceedings of the National Academy of Sciences* 103(22), p. 8487.

Soberón-Chávez, G., Lépine, F. and Déziel, E. 2005. Production of rhamnolipids by Pseudomonas aeruginosa. *Applied microbiology and biotechnology* 68(6), pp. 718-725.

Sousa, A. and Pereira, M. 2014. Pseudomonas aeruginosa diversification during infection development in cystic fibrosis lungs—a review. *Pathogens* 3(3), pp. 680-703.

Sousa, A. M., Machado, I., Nicolau, A. and Pereira, M. O. 2013. Improvements on colony morphology identification towards bacterial profiling. *Journal of microbiological methods* 95(3), pp. 327-335.

Southern, K. C. i. 2020. *CF START: A RANDOMISED REGISTRY TRIAL.* http://www.cfstart.org.uk/: Available at: http://www.cfstart.org.uk/ [Accessed: 15/10/2020].

Stanojevic, S. et al. 2021. Determinants of lung disease progression measured by lung clearance index in children with cystic fibrosis. *European Respiratory Journal* 58(1),

Stanton, B. A. 2017. Effects of Pseudomonas aeruginosa on CFTR chloride secretion and the host immune response. *American Journal of Physiology-Cell Physiology* 312(4), pp. C357-C366.

Steenackers, H. P., Parijs, I., Foster, K. R. and Vanderleyden, J. 2016. Experimental evolution in biofilm populations. *FEMS microbiology reviews* 40(3), pp. 373-397.

Stokell, J. R., Gharaibeh, R. Z., Hamp, T. J., Zapata, M. J., Fodor, A. A. and Steck, T. R. 2015. Analysis of changes in diversity and abundance of the microbial

community in a cystic fibrosis patient over a multiyear period. *Journal of clinical microbiology* 53(1), pp. 237-247.

Stokniene, J. et al. 2020. Bi-Functional Alginate Oligosaccharide–Polymyxin Conjugates for Improved Treatment of Multidrug-Resistant Gram-Negative Bacterial Infections. *Pharmaceutics* 12(11), p. 1080.

Su, K.-Y. and Lee, W.-L. 2020. Fourier Transform Infrared Spectroscopy as a Cancer Screening and Diagnostic Tool: A Review and Prospects. *Cancers* 12(1), p. 115.

Suzuki, S., Horinouchi, T. and Furusawa, C. 2014. Prediction of antibiotic resistance by gene expression profiles. *Nature communications* 5, p. 5792.

Taherali, F., Varum, F. and Basit, A. W. 2018. A slippery slope: On the origin, role and physiology of mucus. *Advanced drug delivery reviews* 124, pp. 16-33.

Taylor, C., Pearson, J. P., Draget, K. I., Dettmar, P. W. and Smidsrød, O. 2005. Rheological characterisation of mixed gels of mucin and alginate. *Carbohydrate polymers* 59(2), pp. 189-195.

Team, R. C. 2013. R: A language and environment for statistical computing.

Testing, E. C. o. A. S. 2020. Breakpoint tables for interpretation of MICs and zone diameters. *Pseudomonas spp.*

Thomas, S. et al. 2017. The host microbiome regulates and maintains human health: a primer and perspective for non-microbiologists. *Cancer research* 77(8), pp. 1783-1812.

Thornton, D. J., Davies, J. R., Kraayenbrink, M., Richardson, P. S., Sheehan, J. K. and Carlstedt, I. 1990. Mucus glycoproteins from 'normal'human tracheobronchial secretion. *Biochemical journal* 265(1), pp. 179-186.

Thornton, D. J., Rousseau, K. and McGuckin, M. A. 2008. Structure and function of the polymeric mucins in airways mucus. *Annu. Rev. Physiol.* 70, pp. 459-486.

Tian, Z. X. et al. 2009. MexT modulates virulence determinants in Pseudomonas aeruginosa independent of the MexEF-OprN efflux pump. *Microbial Pathogenesis* 47(4), pp. 237-241. doi: 10.1016/j.micpath.2009.08.003

Todar, K. 2006. Todar's online textbook of bacteriology. University of Wisconsin-Madison Department of Bacteriology Madison, Wis, USA.

Toutain, C. M., Zegans, M. E. and O'Toole, G. A. 2005. Evidence for two flagellar stators and their role in the motility of Pseudomonas aeruginosa. *Journal of bacteriology* 187(2), pp. 771-777.

Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R. and Cooper, V. S. 2013. Tangled bank of experimentally evolved Burkholderia biofilms reflects selection during chronic infections. *Proc Natl Acad Sci U S A* 110(3), pp. E250-259. doi: 10.1073/pnas.1207025110

v. Wintzingerode, F., Göbel, U. B. and Stackebrandt, E. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS microbiology reviews* 21(3), pp. 213-229.

Valentini, M. and Filloux, A. 2016. Biofilms and Cyclic di-GMP (c-di-GMP) Signaling: Lessons from Pseudomonas aeruginosa and Other Bacteria. *J Biol Chem* 291(24), pp. 12547-12555. doi: 10.1074/jbc.R115.711507

Van Boeckel, T. P. et al. 2015. Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences* 112(18), pp. 5649-5654.

Van Der Gast, C. J. et al. 2011. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *The ISME journal* 5(5), pp. 780-791.

van Koningsbruggen-Rietschel, S. et al. 2020. Inhaled dry powder alginate oligosaccharide in cystic fibrosis: a randomised, double-blind, placebo-controlled, crossover phase 2b study. *ERJ open research* 6(4),

ven Belkum, A., Renders, N. H. M., Smith, S., Overbeek, S. E. and Verbrugh, H. A. 2000. Comparison of conventional and molecular methods for the detection of bacterial pathogens in sputum samples from cystic fibrosis patients. *FEMS Immunology & Medical Microbiology* 27(1), pp. 51-57.

Venkatakrishnan, V., Thaysen-Andersen, M., Chen, S. C. A., Nevalainen, H. and Packer, N. H. 2015. Cystic fibrosis and bacterial colonization define the sputum N-glycosylation phenotype. *Glycobiology* 25(1), pp. 88-100.

Verdugo, P. 2012. Supramolecular dynamics of mucus. *Cold Spring Harbor perspectives in medicine* 2(11), p. a009597.

Virgin, H. W. and Todd, J. A. 2011. Metagenomics and personalized medicine. *Cell* 147(1), pp. 44-56.

Vitko, M. et al. 2016. A novel guluronate oligomer improves intestinal transit and survival in cystic fibrosis mice. *Journal of Cystic Fibrosis* 15(6), pp. 745-751.

von Götz, F. et al. 2004. Expression analysis of a highly adherent and cytotoxic small colony variant of Pseudomonas aeruginosa isolated from a lung of a patient with cystic fibrosis. *Journal of bacteriology* 186(12), pp. 3837-3847.

Voynow, J. A. and Rubin, B. K. 2009. Mucins, mucus, and sputum. *Chest* 135(2), pp. 505-512.

Wagner, C. E., Wheeler, K. M. and Ribbeck, K. 2018. Mucins and their role in shaping the functions of mucus barriers. *Annual review of cell and developmental biology*,

Wang, D. et al. 2015. Adaptation genomics of a small-colony variant in a Pseudomonas chlororaphis 30-84 biofilm. *Applied and environmental microbiology* 81(3), pp. 890-899.

Webb, C. O., Ackerly, D. D. and Kembel, S. W. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24(18),

Weiser, R. et al. 2022. The lung microbiota in children with cystic fibrosis captured by induced sputum sampling. *Journal of Cystic Fibrosis*,

Weiser, R., Rye, P. D. and Mahenthiralingam, E. 2021. Implementation of microbiota analysis in clinical trials for cystic fibrosis lung infection: Experience from the OligoG phase 2b clinical trials. *Journal of Microbiological Methods*, p. 106133.

West, N. E. and Flume, P. A. 2018. Unmet needs in cystic fibrosis: the next steps in improving outcomes. *Expert review of respiratory medicine* 12(7), pp. 585-593.

Whiteman, S. C., Yang, Y., Jones, J. M. and Spiteri, M. A. 2008. FTIR spectroscopic analysis of sputum: preliminary findings on a potential novel diagnostic marker for COPD. *Therapeutic advances in respiratory disease* 2(1), pp. 23-31.

Whitney, J. C. et al. 2015. Dimeric c-di-GMP is required for post-translational regulation of alginate production in Pseudomonas aeruginosa. *J Biol Chem* 290(20), pp. 12451-12462. doi: 10.1074/jbc.M115.645051

WHO. 2017. Antibiotic Resistance.   [Accessed: 16th February 2017].   doi: http://www.euro.who.int/en/health-topics/disease-prevention/antimicrobial-resistance/antibiotic-resistance

Wick, R. R., Judd, L. M., Gorrie, C. L. and Holt, K. E. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS computational biology* 13(6), p. e1005595.

Wick, R. R., Schultz, M. B., Zobel, J. and Holt, K. E. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31(20), pp. 3350-3352.

Widdicombe, J. G. 1997. Airway liquid: a barrier to drug diffusion? *Eur Respir J* 10(10), pp. 2194-2197. doi: 10.1183/09031936.97.10102194

Widdicombe, J. H. and Wine, J. J. 2015. Airway gland structure and function. *Physiological reviews* 95(4), pp. 1241-1319.

Wimpenny, J., Manz, W. and Szewzyk, U. 2000. Heterogeneity in biofilms. *FEMS microbiology reviews* 24(5), pp. 661-671.

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. and Brinkman, F. S. L. 2016. Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic acids research* 44(D1), pp. D646-D653.

Winstanley, C. and Fothergill, J. L. 2009. The role of quorum sensing in chronic cystic fibrosis Pseudomonas aeruginosa infections. *FEMS microbiology letters* 290(1), pp. 1-9.

Winstanley, C., O'Brien, S. and Brockhurst, M. A. 2016a. Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends Microbiol* 24(5), pp. 327-337. doi: 10.1016/j.tim.2016.01.008

Winstanley, C., O'Brien, S. and Brockhurst, M. A. 2016b. Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends in Microbiology* 24(5), pp. 327-337. doi: 10.1016/j.tim.2016.01.008

Wong, A., Rodrigue, N. and Kassen, R. 2012. Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*. *PLoS Genetics* 8(9), p. e1002928.

Workentine, M. L. et al. 2013. Phenotypic heterogeneity of Pseudomonas aeruginosa populations in a cystic fibrosis patient. *PloS one* 8(4), p. e60225.

Wostmann, B. S. 1981. The germfree animal in nutritional studies. *Annual review of nutrition* 1(1), pp. 257-279.

Wozniak, D. J. and Keyser, R. 2004. Effects of subinhibitory concentrations of macrolide antibiotics on Pseudomonas aeruginosa. *Chest* 125(2), pp. 62S-69S.

Xia, B., Royall, J. A., Damera, G., Sachdev, G. P. and Cummings, R. D. 2005. Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis. *Glycobiology* 15(8), pp. 747-775.

Yang, J.-S., Xie, Y.-J. and He, W. 2011a. Research progress on chemical modification of alginate: A review. *Carbohydrate Polymers* 84(1), pp. 33-39. doi: 10.1016/j.carbpol.2010.11.048

Yang, L., Haagensen, J. A. J., Jelsbak, L., Johansen, H. K., Sternberg, C., Høiby, N. and Molin, S. 2008. In situ growth rates and biofilm development of Pseudomonas aeruginosa populations in chronic lung infections. *Journal of bacteriology* 190(8), pp. 2767-2776.

Yang, L. et al. 2011b. Evolutionary dynamics of bacteria in a human host environment. *Proceedings of the National Academy of Sciences of the United States of America* 108(18), pp. 7481-7486. doi: 10.1073/pnas.1018249108

Yang, N., Garcia, M. A. S. and Quinton, P. M. 2013. Normal mucus formation requires cAMP-dependent HCO_3^- secretion and Ca^{2+}-mediated mucin exocytosis. *The Journal of physiology* 591(18), pp. 4581-4593.

Yun, Y. et al. 2014. Environmentally determined differences in the murine lung microbiota and their relation to alveolar architecture. *PloS one* 9(12), p. e113466.

Zahm, J. M. et al. 1995. Dose-dependent in vitro effect of recombinant human DNase on rheological and transport properties of cystic fibrosis respiratory mucus. *European Respiratory Journal* 8(3), pp. 381-386.

Zampoli, M., Pillay, K., Carrara, H., Zar, H. J. and Morrow, B. 2016. Microbiological yield from induced sputum compared to oropharyngeal swab in young children with cystic fibrosis. *Journal of Cystic Fibrosis* 15(5), pp. 605-610.

Zemanick, E. T. et al. 2013. Inflammation and airway microbiota during cystic fibrosis pulmonary exacerbations. *PloS one* 8(4), p. e62917.

Zemanick, E. T. et al. 2017. Airway microbiota across age and disease spectrum in cystic fibrosis. *European Respiratory Journal* 50(5),

Zhang, Y., Doranz, B., Yankaskas, J. R. and Engelhardt, J. F. 1995. Genotypic analysis of respiratory mucous sulfation defects in cystic fibrosis. *The Journal of clinical investigation* 96(6), pp. 2997-3004.

Zhao, C. Y., Hao, Y., Wang, Y., Varga, J. J., Stecenko, A. A., Goldberg, J. B. and Brown, S. P. 2020. Microbiome data enhances predictive models of lung function in people with cystic fibrosis. *bioRxiv*, p. 656066.