Investigating the role of the respiratory microbiome and metabolome in idiopathic pulmonary fibrosis

Rachele Invernizzi

Faculty of Medicine

National Heart and Lung Institute

Imperial College London

A thesis submitted for the degree of

Doctor of Philosophy

November 2020

Per aspera ad astra

Abstract

Idiopathic pulmonary fibrosis (IPF) is a progressive and fatal fibrotic lung disease of unknown aetiology. There is growing evidence that the lung microbiota may play a role in IPF. However, no study has investigated the functional impact of the short-chain fatty acids (SCFAs) on disease pathogenesis. The aim of this thesis was to characterise the respiratory microbiome and metabolome in healthy controls and subjects with IPF and to investigate the influence of the SCFAs on human primary bronchial epithelial cells (PBECs) and human lung fibroblasts (HLFs) from these individuals.

Subjects diagnosed with IPF (*n*=201) and healthy controls (*n*=40) were prospectively recruited and underwent bronchoscopy. Bacterial DNA was isolated and 16S rRNA gene sequencing undertaken to characterise bacterial communities. Untargeted ^{1}H nuclear magnetic resonance spectroscopy-based metabolomics and targeted gas chromatography-mass spectrometry captured the metabolic profile of these samples. PBECs, differentiated at air-liquid interface, and HLF from healthy controls and subjects with IPF were either left untreated or exposed to the SCFAs.

Compared to healthy controls, the IPF microbiota was less diverse (P<0.01), and had increased proportions of *Firmicutes* (P<0.01), *Streptococcus* and *Staphylococcus* (P<0.05). Relative concentrations of the SCFAs were increased in IPF compared to healthy controls, and in IPF, propionate positively correlated with the bacterial burden (*rho*=0.47, P=8 x 10^{-5}). Treatment of PBECs from IPF subjects but not healthy controls with the SCFAs led to morphological changes, a dose-dependent release of pro-inflammatory mediators in the cell supernatant, and a decrease in transepithelial electrical resistance (TEER) over time. Compared to baseline, exposure of IPF PBECs to 1 mM of propionate led to a 40% reduction in TEER and a significant increase (P<0.05) in the secretion of IL-6. Treatment of HLFs with 5 mM of propionate or butyrate was found to increase cell proliferation at 72 hours.

Subjects with IPF display an altered microbiome which is associated with a distinct metabolic signature in the lower airways. Differences in bacterial genera and an increased bacterial burden in IPF results in changes in the SCFAs in the airways. *In vitro* work demonstrates the potential of these SCFAs to shape immunological responses in the lung, mediating the pathogenesis of fibrosis.

Statement of Originality

'I certify that this thesis, and the research to which it refers, are the product of my own work, conducted during the three years of my doctoral research degree at Imperial College London. Any ideas or quotations from the work of other people, published or otherwise, or from my own previous work are fully acknowledged in accordance with the standard referencing practices of the discipline. Additionally, I would like to acknowledge the following people for conducting some of this work: Dr Philip Molyneaux and Dr Richard Hewitt (Royal Brompton Hospital, London) for conducting all of the bronchoscopies; Dr Leopoldo Segal, Dr Benjamin Wu, and Dr Yonghua Li (Division of Pulmonary and Critical Care, New York University) for conducting the 16S rRNA gene sequencing of the samples; Professor Anand Devaraj, Dr Bhavin Rawal and Dr Arjun Nair for scoring the computed tomography data in Chapter 3 (Royal Brompton Hospital and University College Hospital, London); Dr Atsushi Suzuki (Tosei General Hospital, Japan) for providing the bronchoalveolar lavage samples for Chapter 6 of this thesis'.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non-Commercial No Derivatives Licence (CC BY-NC-ND). Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgments

There are innumerable people that have, in some way, shape or form, contributed to the scientific progression of my PhD. To each of these individuals, I am grateful.

My gratitude extends, *in primis*, to my supervisors, Dr Philip Molyneaux, Professor Clare Lloyd and Dr Adam Byrne, for this opportunity. Their guidance and support have been instrumental throughout this chapter of my life and in preparation for the upcoming one. In particular, I would like to thank Dr Philip Molyneaux, whose individual style of supervision kept me motivated from start to finish; it was a pleasure working with and alongside him.

Secondly, I am indebted to the colleagues, and collaborators who shared their time and knowledge with me; without them, this work would have not been possible. These include Professor Toby Maher and all of the patients at the Royal Brompton Hospital; Professor Jonathan Swann, and the team at the National Phenome Centre; Dr Leopoldo Segal and his team at New York University; Dr Atsushi Suzuki and the patients at the Tosei General Hospital in Japan. I would also like to thank several other colleagues who were instrumental for my training, including Dr Maria Zarcone, Dr Natasa Giallourou, Dr Benjamin Wu, Arinbjörn Kolbeinsson, and Dr Alexander Cryer. It is through the collective contributions of these people that I have been able to progress to the stage of submitting this thesis.

Thirdly, I would like to thank my friends and family. In particular, I would like to thank my parents, whose sacrifices have allowed me to embark on this path. My gratitude extends especially to my twin brother, Antonio, for his unwavering support and confidence in me.

Finally, I would like to thank the National Heart and Lung Institute for funding my PhD research, allowing me to undertake and complete this endeavour.

Publications

Some of the work presented in this thesis has been published:

Invernizzi R, Molyneaux PL. The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2019; 28 (152).

Invernizzi R, Barnett J, Rawal B, Nair A, Ghai P, Kingston S, Chua F, Wu Z, Wells AU, Renzoni ER, Nicholson AG, Rice A, Lloyd CM, Byrne AJ, Maher TM, Devaraj A, Molyneaux PL. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020; 55 (4).

Invernizzi R, Lloyd CM, Molyneaux PL. Respiratory microbiome and epithelial interactions shape immunity in the lungs. *Immunology* 2020; 160(2): 171-182.

Invernizzi R, Lloyd CM, Molyneaux PL. Interaktionen zwischen respiratorischem Mikrobiom und Epithelzellen formen Immunität in der Lunge. *Kompass Pneumologie* 2020; 8(5): 240-251.

Invernizzi R, Wu BG, Barnett J, Ghai P, Kingston S, Hewitt RJ, Feary J, Li Y, Chua F, Wu Z, Wells AU, Renzoni EA, Nicholson AG, Rice A, Devaraj A, Segal LN, Byrne AJ, Maher TM, Lloyd CM, Molyneaux PL. The Respiratory Microbiome in Chronic Hypersensitivity Pneumonitis is Distinct from that of Idiopathic Pulmonary Fibrosis. *AJRCCM* 2020.

TABLE OF CONTENTS

Abstract 3

Statement of Originality 4

Copyright Declaration 5

Acknowledgments 6

Publications 7

Table of Contents 8

Abbreviations 16

Figure Index 21

Table Index 26

Chapter 1: Introduction 27

1.1 Interstitial Lung Disease 27

1.1.1 Idiopathic Interstitial Pneumonias 28

1.2 Idiopathic Pulmonary Fibrosis 28

1.2.1 Epidemiology 28

1.2.2 Clinical presentation, symptoms and diagnosis 28

1.2.3 Prognosis and outcomes 31

1.2.4 Acute exacerbations 32

1.2.5 Pathogenesis 33

1.2.5.1 Established model to explain IPF pathogenesis 33

1.2.6 Risk factors 35

1.2.6.1 Genetic susceptibility 35

1.2.6.2 Epigenetics 37

1.2.6.3 Environmental risk factors 37

1.2.7 Infection 38

1.2.7.1 The role of viruses in IPF 39

1.2.7.2 The role of viruses in AE-IPF 39

1.2.8 Therapeutic options 41

1.3 The Composition of the Airway Microbiome 41

1.3.1 The concept of lung sterility 41

1.3.2 Development of the lung microbiome at birth 42

1.3.3 The lung microbiome in health and its ecological determinants 43

1.3.4 The lung microbiome in respiratory disease 44
1.3.5 The lung microbiome in IPF and AE-IPF 45
1.3.6 Effects of antibiotic therapy on the airway microbiome 47
1.4 Microbial Metabolites 48
1.4.1 The short-chain fatty acids 48
1.4.2 The gut-lung axis 49
1.4.3 Short-chain fatty acids promote inflammation in the lungs 51
1.5 The Respiratory Epithelium 51
1.5.1 The epithelial barrier 51
1.5.2 The relationship between the lung and its microbiota 54
1.5.3 Recognition of pathogens by airway epithelial cells 55
1.5.4 Infections and the initiation of the epithelial immune response 57
1.6 Summary of findings and unanswered questions 60
1.7 Hypothesis 61
1.8 Aims 61

Chapter 2: Materials and Methods 62

2.1 Study Design 62
2.1.1 Subject recruitment 62
2.1.2 Ethics statement 63
2.1.3 Subject assessment and pulmonary function testing 63
2.1.4 Fibre-optic bronchoscopy with bronchoalveolar lavage 63
2.1.5 BAL differential cell counts 63
2.1.6 *MUC5B* genotyping 63
2.2 Bacterial 16S rRNA Gene Analysis 64
2.2.1 Reagents for bacterial DNA isolation 64
2.2.2 Bacterial DNA isolation and precipitation 64
2.2.3 Bacterial DNA quantification 64
2.2.4 Identification of procedural contaminants 65
2.2.5 16S rRNA sequencing of the V4 region of bacterial 16S rRNA gene 65
2.2.6 Bioinformatics analysis 66
2.2.6.1 Raw sequence reads 66
2.2.6.2 Metadata 67

2.2.6.3 Importing paired-end sequence data 67
2.2.6.4 Demultiplexing and quality control of paired-end sequence data 67
2.2.6.5 Denoising, dereplicating and clustering of reads 67
2.2.6.6 Taxonomic classification and phylogenetic tree construction 68
2.2.6.7 Exporting feature table, phylogenetic tree, taxonomy and metadata 68
2.2.7 Diversity and statistical analyses 68
2.2.8 Differential abundance analysis 69
2.2.9 Summary of microbiome workflow 69
2.2.10 Statistical analysis 70
2.3 16S Gene Quantitative Polymerase Chain Reaction 71
2.3.1 Amplification of *Vibrio natriegens DSM 759* 71
2.3.2 Gel extraction of *Vibrio natriegens DSM 759* amplicons 71
2.3.3 Amplicon cloning 72
2.3.4 Plasmid DNA purification 72
2.3.5 Quantitative 16S rRNA gene PCR 76
2.4 ^{1}H Nuclear Magnetic Resonance Spectroscopy 77
2.4.1 Reagents for ^{1}H NMR spectroscopy 77
2.4.2 D_2O buffer preparation 77
2.4.3 Sample collection and preparation 77
2.4.4 One-dimensional ^{1}H NMR spectroscopy 77
2.4.5 NOESY ^{1}H NMR spectral acquisition of BAL supernatant samples 78
2.4.6 Spectral pre-processing 78
2.4.7 Statistical pre-processing 78
2.4.8 Multivariate data analysis 80
2.4.9 Summary of ^{1}H NMR spectroscopy workflow 82
2.5 Triple Quadrupole Gas Chromatography-Mass Spectrometry 83
2.5.1 Chemicals and reagents 83
2.5.2 Methyl *tert*-butyl ether preparation 83
2.5.3 Calibration curve preparation 83
2.5.4 Triple quadrupole mass spectrometer setup 85
2.5.5 BAL supernatant sample preparation 86
2.5.6 Statistical analysis 86
2.6 Enzyme-linked immunosorbent assay 87

2.6.1 General ELISA protocol 87
2.7 Human Cell Culture 88
2.7.1 Primary human bronchial epithelial cell culture 88
2.7.2 Submerged culture of PBECs 89
2.7.2.1 Crystal violet assay 89
2.7.2.2 Proliferation assay 89
2.7.2.3 Wound healing assay 90
2.7.3 Air-liquid interface culture 90
2.7.3.1 Lactate dehydrogenase assay 91
2.7.3.2 Trans-epithelial electrical resistance 91
2.7.4 Isolation and culture of primary human lung fibroblasts 91
2.7.5 Submerged culture of primary HLs 92
2.7.5.1 Crystal violet assay 92
2.7.5.2 Proliferation assay 92

Chapter 3: The Respiratory Microbiome in Idiopathic Pulmonary Fibrosis 93
3.1 Introduction 93
3.2 Aims 94
3.3 Materials and Methods 95
3.3.1 Study design and sampling 95
3.3.2 Measurement of BAL cells 95
3.3.3 *MUC5B* genotyping 95
3.3.4 Radiology 95
3.3.5 Bacterial DNA isolation, quantification and 16S rRNA gene sequencing 95
3.3.6 Bioinformatics analysis 95
3.3.7 16S rRNA gene qPCR 96
3.3.8 Statistical analysis 96
3.4 Results 97
3.4.1 Study subjects and sampling 97
3.4.2 Survival analysis 98
3.4.2.1 Disease progression 98
3.4.2.2 *MUC5B* genotype 99
3.4.3 Sampling 101

3.4.3.1 DNA yield and sequencing quality 101
3.4.3.2 Mock community 102
3.4.3.3 Contamination 103
3.4.4 Bacterial burden 104
3.4.4.1 Bacterial burden in health and disease 104
3.4.4.2 Bacterial burden and composite physiologic index 105
3.4.4.3 Bacterial burden and disease progression 106
3.4.4.4 Bacterial burden and BAL differential cell counts in IPF 107
3.4.4.5 Bacterial burden and *MUC5B* genotype 108
3.4.4.6 Bacterial burden and radiological disease extent 108
3.4.5 Microbiota composition in health and disease 110
3.4.5.1 Relative abundance 110
3.4.5.2 Differential abundance 112
3.4.5.3 Diversity analyses 113
3.4.6 Microbiota composition and disease progression 114
3.4.7 Microbiota composition and smoking history 117
3.4.8 Microbiota composition and *MUC5B* genotype 118
3.5 Discussion 119
3.6 Conclusions 126

Chapter 4: The Respiratory Metabolome in Idiopathic Pulmonary Fibrosis 127
4.1 Introduction 127
4.2 Aims 128
4.3 Materials and Methods 129
4.3.1 Study design 129
4.3.2 Sample collection and preparation 129
4.3.3 ^{1}H NMR spectroscopy 129
4.3.4 Triple quadrupole GC-MS 129
4.3.5 Statistical analysis 129
4.4 Results 130
4.4.1 Study subjects 130
4.4.2 Metabolome composition in health and disease 130
4.4.2.1 BAL spectral profiles of healthy controls and IPF subjects 130

4.4.2.2 Integration of metabolites detected in BAL 131
4.4.3 Associations between metabolites and clinical variables in IPF 133
4.4.4 Associations between respiratory microbiome and metabolites in IPF 137
4.4.5 Associations between bacterial burden and metabolites in IPF 139
4.4.6 Triple Quadrupole GC-MS targeted analysis of the SCFAs in BAL samples 142
4.5 Discussion 146
4.6 Conclusions 152

Chapter 5: Investigating the Influence of the Short-Chain Fatty Acids on Primary Human Bronchial Epithelial Cells and Primary Human Lung Fibroblasts from Healthy Individuals and Subjects with Idiopathic Pulmonary Fibrosis 153

5.1 Introduction 153
5.2 Aims 154
5.3 Materials and Methods 155
5.3.1 PBECs culture 155
5.3.2 Submerged culture of PBECs 155
5.3.3 ALI culture of PBECs 155
5.3.4 Isolation of primary HLFs 155
5.3.5 Submerged culture of primary HLFs 155
5.3.6 Statistical analysis 155
5.4 Results 156
5.4.1 Patient characteristics 156
5.4.2 Submerged culture of PBECs 156
5.4.2.1 Assessing the cytotoxicity: post-SCFA treatment 156
5.4.2.2 Assessing the pro-inflammatory response: post-SCFA treatment 157
5.4.2.3 Assessing the proliferation: pre-SCFA treatment 158
5.4.2.4 Assessing the proliferation: post-SCFA treatment 159
5.4.2.5 Assessing the wound healing capacity: pre-SCFA treatment 161
5.4.2.6 Assessing the wound healing capacity: post-SCFA treatment 161
5.4.3 ALI culture of PBECs 162
5.4.3.1 Assessing the cytotoxicity: post-SCFA treatment 162
5.4.3.2 Assessing the trans-epithelial electrical resistance: post-SCFA treatment 163
5.4.3.3 Assessing the changes in morphology: post-SCFA treatment 163

5.4.3.4 Assessing the pro-inflammatory response: post-SCFA treatment 166
5.4.4 Submerged culture of primary HLFs 167
5.4.4.1 Assessing the cytotoxicity: post-SCFA treatment 167
5.4.4.2 Assessing the production of collagen, fibronectin and TGF-β: post-SCFA treatment 167
5.4.4.3 Assessing the proliferation: post-SCFA treatment 168
5.5 Discussion 171
5.6 Conclusions 177

Chapter 6: Changes in the Respiratory Microbiome and Metabolome during Acute Exacerbations of Idiopathic Pulmonary Fibrosis 179

6.1 Introduction 179
6.2 Aims 180
6.3 Materials and Methods 181
6.3.1 Study design and sampling 181
6.3.2 Measurement of clinical parameters 181
6.3.3 ELISA 181
6.3.4 Bacterial DNA isolation, quantification and 16S rRNA gene sequencing 181
6.3.5 Bioinformatics analysis 182
6.3.6 16S rRNA gene qPCR 182
6.3.7 Triple quadrupole GC-MS 182
6.3.8 Statistical analysis 182
6.4 Results 183
6.4.1 Study subjects and sampling 183
6.4.2 Sampling 184
6.4.2.1 DNA yield and sequencing quality 184
6.4.2.2 Contamination 185
6.4.2.3 Mock Community 186
6.4.3 Bacterial burden 186
6.4.3.1 Bacterial burden at baseline and during an AE-IPF 186
6.4.3.2 Bacterial burden and time to exacerbation 187
6.4.3.3 Bacterial burden and BAL differential cell counts 188
6.4.3.4 Bacterial burden and antibiotic treatment 188
6.4.4 Microbiota composition at baseline and during an AE-IPF 189

6.4.4.1 Relative abundance 189
6.4.4.2 Differential abundance 190
6.4.4.3 Diversity analyses 192
6.4.4.4 Paired samples 193
6.4.5 Cytokine profiles at baseline and during an AE-IPF 193
6.4.6 Metabolic changes at baseline and during an AE-IPF 195
6.5 Discussion 199
6.6 Conclusions 203

Chapter 7: Discussion 204
7.1 General discussion 204
7.2 Future work 213

8 Bibliography 215
9 Appendix 254

Abbreviations

α-diversity – alpha diversity

β-diversity – beta diversity

16S rRNA – 16S ribosomal RNA

1D NMR – one-dimensional nuclear magnetic resonance

AE-IPF – acute exacerbation of IPF

AEC – alveolar epithelial cell

AEC1 – alveolar epithelial type I cells

AEC2 – alveolar epithelial type 2 cells

AIP – acute interstitial pneumonia

ALI – air-liquid interface

ASM – airway smooth muscle

ASVs – amplicon sequence variants

ATP – adenosine triphosphate

BAL – bronchoalveolar lavage

BECs – bronchial epithelial cells

CF – cystic fibrosis

CLRs – C-type lectin receptors

CMV – cytomegalovirus

COMET – Correlating Outcomes with Biochemical Markers to Estimate Time-Progression

COP – cryptogenic organising pneumonia

COPD – chronic obstructive pulmonary disease

CPI – composite physiologic index

CRP – C-reactive protein

CT – computed tomography

CTAB – hexadecyltrimethylammonium bromide

CTDs – connective tissue diseases

D_2O – deuterium oxide

DADA2 – divisive amplicon denoising algorithm 2

DAMPs – damage-associated molecular patterns

DIABLO – data integration analysis for biomarker discovery

DIO2 – iodothyronine deiodinase 2

DIP – desquamative interstitial pneumonia

DKC – dyskeratosis congenita

DKK – dikkopf

DL_{CO} – diffusion capacity for carbon monoxide

DMEM – Dulbecco's Modified Eagle Medium

DMSO – dimethyl sulfoxide

EBV – Epstein-Barr virus

ECM – extracellular matrix

EGFR – epidermal growth factor receptor

ELISA – enzyme-linked immunosorbent assay

FEV_1 – forced expiratory volume in 1 second

FGF – fibroblast growth factor

FIP – familial interstitial pneumonia

FVC – forced vital capacity

GC-MS – gas chromatography-mass spectrometry

GF – germ-free

GORD – gastro-oesophageal reflux disease

GPCR – G-protein coupled receptors

HCl – hydrochloric acid

HHV – human herpes virus

HLFs – human lung fibroblasts

HRCT – high-resolution computed tomography

IIPs – idiopathic interstitial pneumonias

IL-6 – interleukin 6

IL-8 – interleukin 8

ILDs – interstitial lung disease

IPF – idiopathic pulmonary fibrosis

KH_2PO_4 – potassium dihydrogen phosphate

KL-6 – Krebs von den Lungen-6

KOH – potassium hydroxide

KRM – kreme

LB – luria-bertani

LDA – linear discriminant analysis

LDH – lactate dehydrogenase

LEfSE – linear discriminant analysis coupled with effect size

LIP – lymphocytic interstitial pneumonia

LME – lysing matrix E

LPA – linear polyacrylamide

LRR – leucine-rich repeat

MCT – monocarboxylate transporter

MDD – multidisciplinary discussion

MHV-68 – murine gamma-herpes virus 68

miRNAs – microRNAs

mRNA – messenger RNA

MTBE – methyl *tert*-butyl ether

MTBSTFA – *N*-*tert*-butyldimethylsilyl-*N*-methylfluoroacetamide with 1% *tert*-butyldimethylchlorosilane

MUC5B – mucin 5B

Na_2HPO_4 – disodium phosphate

NaCl – sodium chloride

NaH_2PO_4 – monosodium phosphate

NaN_3 – sodium azide

NF – nuclear factor

NIP – nonspecific interstitial pneumonia

NK – natural killer

NLRs – NOD-like receptors

NMR – nuclear magnetic resonance

NOD – nucleotide-binding and oligomerisation domain

NOESY – nuclear overhauser effect spectroscopy

OTUs – operational taxonomic units

PAMPs – pathogen-associated molecular patterns

PBECs – primary bronchial epithelial cells

PC1 – principal component 1

PC2 – principal component 2

PCA – principal component analysis

PCs – principal components

PDGF – platelet derived growth factor

PEG – polyethylene glycol

PERMANOVA – permutational multivariate analysis of variance

PHGDH – phosphoglycerate dehydrogenase

PPFE – pleuroparenchymal fibroelastosis

PRRs – pattern recognition receptors

QCs – quality controls

QIIME – quantitative insights into microbial ecology

qPCR – quantitative polymerase chain reaction

RB-ILD – respiratory bronchiolitis-associated interstitial lung disease

RLRs – retinoic acid-inducible gene-I-like receptors

RSV – respiratory syncytial virus

SCFAs – short-chain fatty acids

SHMT2 – serine hydroxymethyltransferase 2

SNP – single nucleotide polymorphism

SP – surfactant protein

sPLS-DA – sparse partial least-squares-discriminant analysis

TCA – tricarboxylic acid

TEER – transepithelial electrical resistance

TGF-β – transforming growth factor-β

Th2 – T helper 2

TLRs – toll-like receptors

TNF-α: tumour necrosis factor-alpha

Tris-EDTA – tris-ethylenediaminetetracetic acid

TSP – 3-(trimethyl -silyl)-[2,2,3,3,-2H4]-propionic acid

TTV – torque teno virus

UIP – usual interstitial pneumonia

VEGF – vascular endothelial growth factor

ZO – zonular occluden

Figure Index

Figure 1.1 Classification of interstitial lung disease 27
Figure 1.2 Diagnostic approach for idiopathic pulmonary fibrosis 29
Figure 1.3 Typical pattern of usual interstitial pneumonia on high-resolution computed tomography and on histopathology slides 30
Figure 1.4 Potential clinical courses of idiopathic pulmonary fibrosis 31
Figure 1.5 High-resolution computed tomography imaging of idiopathic pulmonary fibrosis 32
Figure 1.6 Proposed mechanism contributing to the pathogenesis of idiopathic pulmonary fibrosis ... 33
Figure 1.7 Ecological determinants of the respiratory microbiome 43
Figure 1.8 Model of intestinal-pulmonary cross-talk during respiratory health and disease 49
Figure 1.9 Composition of the respiratory epithelium in health 52
Figure 1.10 Response of the airway epithelium to infection 58
Figure 2.1 Schematic representation of amplification strategy of 16S V4 region 65
Figure 2.2 Data output in the FASTQ file format 66
Figure 2.3 Community diversity is measured according to species richness and evenness 68
Figure 2.4 Microbiome workflow analysis 70
Figure 2.5 Map of the linearized vector, pCR™2.1-TOPO 74
Figure 2.6 Illustrating the effect of mean centring on the dataset 79
Figure 2.7 Schematic breakdown of principal component analysis 80
Figure 2.8 Detection of outliers using Hotelling's T2 test 81
Figure 2.9 Overall strategy for metabonomic analysis 82
Figure 2.10 Air-liquid interface cell culture schematic 91
Figure 3.1 Kaplan-Meier curve demonstrating survival probability in idiopathic pulmonary fibrosis subjects with stable or progressive disease 99
Figure 3.2 Survival probability of subjects with idiopathic pulmonary fibrosis stratified according to *MUC5B* genotype 100
Figure 3.3 Clinical characteristics stratified according to *MUC5B* genotype 100
Figure 3.4 Distribution of samples sequence depth 101
Figure 3.5 Mock community composition at the genus level from the three sequencing runs 102
Figure 3.6 Comparison between all bronchoalveolar lavage samples and negative controls 104
Figure 3.7 Bronchoalveolar lavage bacterial burden in healthy controls and idiopathic pulmonary fibrosis samples 105
Figure 3.8 Bronchoalveolar lavage bacterial burden in idiopathic pulmonary fibrosis subjects with either stable or progressive disease 105
Figure 3.9 Survival probability of subjects with idiopathic pulmonary fibrosis stratified into tertiles by bacterial burden 106

Figure 3.10 Correlations between bronchoalveolar lavage differential cell counts and bacterial burden of subjects with idiopathic pulmonary fibrosis 107
Figure 3.11 Bronchoalveolar lavage bacterial burden in idiopathic pulmonary fibrosis subjects stratified according to *MUC5B* genotype..... 108
Figure 3.12 Relationship between radiological features and bronchoalveolar lavage bacterial burden 109
Figure 3.13 The respiratory microbiome in idiopathic pulmonary fibrosis is altered compared to healthy controls 110
Figure 3.14 Differentially abundant bacteria within the healthy control and idiopathic pulmonary fibrosis groups 112
Figure 3.15 Alpha-diversity of microbial communities by diagnosis..... 113
Figure 3.16 Beta-diversity of microbial communities by diagnosis 113
Figure 3.17 Increased relative abundance of *Staphylococcus* in the bronchoalveolar lavage of idiopathic pulmonary fibrosis subjects is associated with increased mortality 114
Figure 3.18 The respiratory microbiome in idiopathic pulmonary fibrosis subjects with stable or progressive disease 115
Figure 3.19 Alpha-diversity of idiopathic pulmonary fibrosis subjects with stable or progressive disease 115
Figure 3.20 The respiratory microbiome in idiopathic pulmonary fibrosis subjects stratified according to bronchoalveolar lavage bacterial burden 116
Figure 3.21 Alpha-diversity of idiopathic pulmonary fibrosis subjects stratified according to the bacterial burden..... 116
Figure 3.22 The respiratory microbiome in idiopathic pulmonary fibrosis subjects stratified according to smoking history 117
Figure 3.23 Alpha-diversity of idiopathic pulmonary fibrosis individuals stratified according to smoking history..... 117
Figure 3.24 The respiratory microbiome in idiopathic pulmonary fibrosis subjects based on *MUC5B* genotype 118
Figure 3.25 Alpha-diversity of idiopathic pulmonary fibrosis individuals stratified according to *MUC5B* genotype 118
Figure 4.1 Integration of the metabolites detected by ^{1}H nuclear magnetic resonance spectroscopy in the bronchoalveolar lavage supernatant of healthy controls and subjects diagnosed with idiopathic pulmonary fibrosis 131
Figure 4.2 Multivariate statistical analysis of the bronchoalveolar lavage ^{1}H nuclear magnetic resonance spectroscopy spectra 132

Figure 4.3 Correlation heatmap between metabolites identified in the bronchoalveolar lavage of idiopathic pulmonary fibrosis subjects and bronchoalveolar lavage differential cell counts 133
Figure 4.4 The respiratory metabolome in idiopathic pulmonary fibrosis subjects with stable or progressive disease 134
Figure 4.5 The respiratory microbiome in idiopathic pulmonary fibrosis subjects stratified according to smoking history 135
Figure 4.6 Principal component analysis of the respiratory metabolome in idiopathic pulmonary fibrosis subjects stratified according to forced vital capacity and smoking history 136
Figure 4.7 Associations between microbial genera and metabolites 137
Figure 4.8 Circos plot of horizontal sparse partial least-squares-discriminant analysis displaying correlations between the respiratory metabolome and microbiome datasets. 138
Figure 4.9 Sparse partial least-squares-discriminant analysis depicting the contribution of the selected features in driving the separation between healthy controls and individuals with idiopathic pulmonary fibrosis 139
Figure 4.10 The respiratory microbiome in idiopathic pulmonary fibrosis subjects stratified according to bronchoalveolar lavage bacterial burden tertiles 140
Figure 4.11 Gas chromatography-mass spectrometry of the short-chain fatty acids in the bronchoalveolar lavage supernatant of healthy controls and subjects with idiopathic pulmonary fibrosis 143
Figure 4.12 Inter-individual variability in the short-chain fatty acid landscape of the bronchoalveolar lavage from healthy controls and subjects with idiopathic pulmonary fibrosis 144
Figure 4.13 Hierarchical clustering of the bronchoalveolar lavage microbiome in healthy controls and subjects with idiopathic pulmonary fibrosis 145
Figure 5.1 Cytotoxic effects of the short-chain fatty acids on submerged cultures of human primary bronchial epithelial cells from healthy controls and subjects with idiopathic pulmonary fibrosis 157
Figure 5.2 Short-chain fatty acids induce a dose-dependent IL-6 and TNF-α secretion in submerged cultures of human primary bronchial epithelial cells from idiopathic pulmonary fibrosis subjects 157
Figure 5.3 Proliferation assay of untreated primary bronchial epithelial cells 158
Figure 5.4 Proliferation of human primary bronchial epithelial cells from idiopathic pulmonary fibrosis subjects following exposure to the short-chain fatty acids 159
Figure 5.5 Ki67 staining of human primary bronchial epithelial cells from idiopathic pulmonary fibrosis subjects following exposure to the short-chain fatty acids 160
Figure 5.6 Wound healing assay of untreated primary bronchial epithelial cells 161
Figure 5.7 Wound healing assay of human primary bronchial epithelial cells from idiopathic pulmonary fibrosis subjects following exposure to the short-chain fatty acids 161

Figure 5.8 Cytotoxic effects of the short-chain fatty acids on air-liquid interface cultures of human primary bronchial epithelial cells from healthy controls and subjects with idiopathic pulmonary fibrosis 162
Figure 5.9 Effects of the short-chain fatty acids on the trans-epithelial electrical resistance of human primary bronchial epithelial cells from healthy controls and subjects with idiopathic pulmonary fibrosis differentiated at air-liquid interface 163
Figure 5.10 Morphology of primary bronchial epithelial cells from a healthy control differentiated at air-liquid interface in different culture conditions 164
Figure 5.11 Morphology of primary bronchial epithelial cells from a subject with idiopathic pulmonary fibrosis differentiated at air-liquid interface in different culture conditions 165
Figure 5.12 Secretion of IL-6 in the apical supernatant of human primary bronchial epithelial cells from healthy controls and idiopathic pulmonary fibrosis subjects differentiated at air-liquid interface 166
Figure 5.13 Cytotoxic effects of the short-chain fatty acids on submerged cultures of human lung fibroblasts from subjects with idiopathic pulmonary fibrosis 167
Figure 5.14 Collagen I α1, fibronectin and TGF-β production post-short-chain fatty acid treatment in submerged cultures of human lung fibroblasts from idiopathic pulmonary fibrosis subjects 168
Figure 5.15 Proliferation of human lung fibroblasts from idiopathic pulmonary fibrosis subjects following exposure to the short-chain fatty acids 169
Figure 5.16 Ki67 staining of human lung fibroblasts from idiopathic pulmonary fibrosis subjects following exposure to the short-chain fatty acids 170
Figure 6.1 Distribution of samples sequence depth. Illustrating the total number of reads present in each sample..... 185
Figure 6.2 Comparison between all bronchoalveolar lavage samples and negative controls 186
Figure 6.3 Mock community composition at the genus level from the two sequencing runs 186
Figure 6.4 Bacterial burden in bronchoalveolar lavage of stable and acute exacerbation subjects 187
Figure 6.5 Kaplan-Meier curve comparing the time from diagnosis to an acute exacerbation of IPF .. 187
Figure 6.6 No correlation between bronchoalveolar lavage differential cell counts and bacterial burden of subjects with idiopathic pulmonary fibrosis 188
Figure 6.7 Bacterial burden in bronchoalveolar lavage of subjects with idiopathic pulmonary fibrosis stratified according to exposure to antibiotics 189
Figure 6.8 Mean taxonomic distribution at the phylum level 190
Figure 6.9 Heatmap of bacterial genera in the bronchoalveolar lavage of subjects with idiopathic pulmonary fibrosis 190
Figure 6.10 Differentially abundant bacteria within the baseline-idiopathic pulmonary fibrosis and acute exacerbation of idiopathic pulmonary fibrosis groups 191
Figure 6.11 Alpha-diversity of microbial communities at baseline and during an acute exacerbation 192

Figure 6.12 Beta-diversity of microbial communities at baseline and during an acute exacerbation .. 192
Figure 6.13 Taxonomic distribution at the phylum level in paired samples from two subjects diagnosed with idiopathic pulmonary fibrosis 193
Figure 6.14 Bronchoalveolar lavage supernatant levels of inflammatory cytokines at baseline and during an acute exacerbation of idiopathic pulmonary fibrosis 194
Figure 6.15 Correlation between cytokines detected in bronchoalveolar lavage and bronchoalveolar lavage bacterial burden of subjects with idiopathic pulmonary fibrosis at baseline and during an acute exacerbation 194
Figure 6.16 Correlation between IL-8 detected in bronchoalveolar lavage and percentage of neutrophils in BAL of subjects with an acute exacerbation of idiopathic pulmonary fibrosis 195
Figure 6.17 Gas chromatography-mass spectrometry of the short-chain fatty acids in the bronchoalveolar lavage supernatant of subjects with idiopathic pulmonary fibrosis at baseline and during an acute exacerbation 196
Figure 6.18 Correlation between propionate detected in bronchoalveolar lavage and broncholaveolar lavage bacterial burden of subjects with idiopathic pulmonary fibrosis at baseline and during an acute exacerbation. 197
Figure 6.19 Correlations between cytokines detected in bronchoalveolar lavage and propionate detected in BAL of subjects with idiopathic pulmonary fibrosis at baseline and during an acute exacerbation 197
Figure 6.20 Kaplan-Meier curve comparing the time from diagnosis to an acute exacerbation of idiopathic pulmonary fibrosis according to the concentration of propionate in the bronchoalveolar lavage supernatant at baseline 198

Table Index

Table 1.1 Multidisciplinary guidelines for the diagnosis of idiopathic pulmonary fibrosis are based upon high resolution computed tomography and biopsy patterns 30
Table 1.2 Summary of studies linking the respiratory microbiome with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis 40
Table 1.3 Summary of studies illustrating changes in the respiratory microbiome which occur during chronic lung disease 44
Table 1.4 Summary of studies linking the respiratory microbiome with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis 45
Table 2.1 Inclusion and exclusion criteria for study participants 62
Table 2.2 Nanodrop results of the purified plasmids 73
Table 2.3 Dilutions of working stock used to construct the qPCR standard curve 76
Table 2.4 Illustrating the quantity added for each compound to set up the calibration curve 84
Table 2.5 Illustrating the mother stock containing specific volumes of the respective compounds 84
Table 2.6 Illustrating the dilution series performed to obtain the calibration curve 85
Table 2.7 Retention time for the respective short-chain fatty acids 85
Table 2.8 Illustrating the quantifiers and qualifier ions for each compound 86
Table 2.9 Illustrating commercially available enzyme-linked immunosorbent assay kits utilised 87
Table 3.1 Clinical characteristics and demographics of the study subjects 97
T**able 3.2** Pre-processing of 16S rRNA sequence reads from bronchoalveolar lavage samples 102
Table 3.3 Contaminant genera detected in the sequenced negative controls 103
Table 3.4 Cox proportional hazards model of disease progression in idiopathic pulmonary fibrosis 107
Table 3.5 Radiological features and their prediction of bacterial burden in bronchoalveolar lavage 109
Table 3.6 Spearman's rank correlations (*rho*) between respiratory microbiota (genus level) of idiopathic pulmonary fibrosis subjects and bronchoalveolar lavage cell differentials 111
Table 4.1 ^{1}H NMR chemical shift data for metabolites identified in bronchoalveolar lavage supernatant of controls and idiopathic pulmonary fibrosis subjects 130
Table 4.2 Spearman's rank correlations (*rho*) between the metabolites detected in the bronchoalveolar lavage of idiopathic pulmonary fibrosis subjects and BAL bacterial burden 141
Table 4.3 Clinical characteristics and demographics of the study subjects 142
Table 5.1 Donor characterisation 156
Table 6.1 Baseline clinical characteristics of study subjects 183
Table 6.2 Changes in bronchoalveolar lavage cell differentials and serum marker concentrations 184
Table 6.3 Pre-processing of 16S rRNA sequence reads from bronchoalveolar lavage samples 185

1 **Chapter 1**: Introduction

1.1 Interstitial Lung Disease

Interstitial lung diseases (ILDs) are a diverse group of disorders characterised by inflammation and a dysregulated fibrotic response within the lung parenchyma [1]. Although ILDs commonly display similar clinical, radiological and histological features, each form has a distinct aetiology [2], triggered in response to various stimuli that ultimately elicit a prolonged and exaggerated fibrotic response within the lung interstitium [1]. The excessive deposition of extracellular matrix (ECM) proteins leads to loss of alveolar structure, impairment of gas exchange and ultimately, respiratory failure [3]. Environmental and occupational exposures (e.g. asbestosis, silicosis), certain drugs (e.g. amiodarone) or underlying systemic illnesses such as connective tissue diseases (CTDs), rheumatoid arthritis and scleroderma, may act as triggers of ILD [3]. Nonetheless, some forms of ILDs have no identifiable cause and are classified as idiopathic interstitial pneumonias (IIPs) [4] **(Figure 1.1)**.

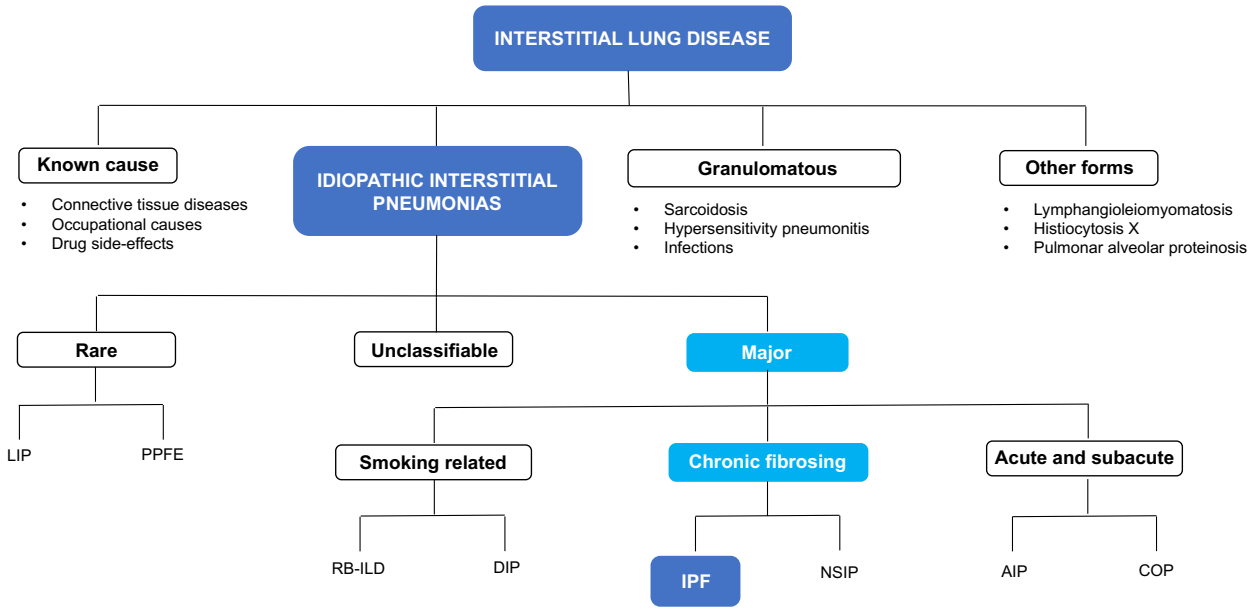


Figure 1.1 Classification of interstitial lung disease.

ILDs can be divided into broad subcategories: ILDs of known cause, the idiopathic interstitial pneumonias (IIPs), the granulomatous diseases including sarcoidosis and hypersensitivity pneumonitis, and a miscellaneous group including other forms such as lymphangiolelomyomatosis and histocytosis X. IIPs are grouped into three groups: major, rare and unclassifiable. Rare IIPs: idiopathic lymphocytic interstitial pneumonia (LIP), idiopathic pleuroparenchymal fibroelastosis (PPFE). Major IIPs are distinguished from rare and unclassifiable cases and are grouped into chronic fibrosing, smoking related and acute/sub-acute IIPs. Chronic fibrosing: idiopathic pulmonary fibrosis (IPF), non-specific interstitial pneumonia (NSIP). Smoking-related: respiratory bronchiolitis-interstitial lung disease (RB-ILD), desquamative interstitial pneumonia (DIP). Acute and subacute: acute interstitial pneumonia (AIP), cryptogenic organising pneumonia (COP). Adapted from [4].

1.1.1 Idiopathic Interstitial Pneumonias

IIPs are ILDs of unknown aetiology that share similar clinical and radiological features [5]. Diagnosis of IIPs is based on history, physical examination, high-resolution computed tomography (HRCT) imaging, pulmonary function tests and lung biopsy. Currently, IIPs are classified into the following 8 histologic subtypes: idiopathic pulmonary fibrosis (IPF), desquamative interstitial pneumonia (DIP), nonspecific interstitial pneumonia (NIP), cryptogenic organising pneumonia (COP), respiratory bronchiolitis-associated interstitial lung disease (RB-ILD), acute interstitial pneumonia (AIP), lymphocytic interstitial pneumonia (LIP) and idiopathic pleuroparenchymal fibroelastosis (PPFE) [2]. These are all characterised by varying degrees of inflammation and fibrosis and all-cause dyspnoea. However, IPF is the most common and severe form of all the idiopathic interstitial pneumonias.

1.2 Idiopathic Pulmonary Fibrosis

1.2.1 Epidemiology

The annual incidence of IPF is rising; worldwide, it has been estimated at 20 per 100,000 people and at 7.94 per 100,000 people in the UK [6, 7]. The lack of a unified definition of the disease until early 2000 has meant that, historically, it has been challenging to accurately assess the epidemiology of the disease [8]. In addition, the overlap of IPF with other chronic respiratory diseases has hindered the ability to determine and quantify the distribution of IPF in time and space. IPF predominantly affects men and cigarette smokers [9]. Advancing age has been shown to be an independent risk factor of IPF, with the onset of the disease typically occurring in patients over 60 years old [9].

1.2.2 Clinical presentation, symptoms and diagnosis

IPF is a chronic, progressive fibrotic lung disease of unknown aetiology for which there is no cure. It is characterised by the progressive deposition of fibrotic tissue and irreversible remodelling of the lung parenchyma [10]. Symptoms are usually insidious and include dry cough, dyspnoea on exertion, fatigue, and gradual decline in the ability to execute daily activities. In advanced disease, clubbing of the fingernails can also be observed [11]. However, these symptoms are not specific to IPF and can occur in other ILDs, making the diagnosis of the disease challenging [9, 12]. IPF is currently defined as a specific form of chronic, progressive, fibrosing interstitial pneumonia of unknown cause, occurring primarily in older adults and limited to the lungs, and defined by the histopathologic and/or radiologic pattern of usual interstitial pneumonia (UIP) [8]. The diagnostic approach to IPF entails the exclusion of known causes of other ILDs along with specific radiographic and/or histopathologic findings [13] **(Figure 1.2)**.

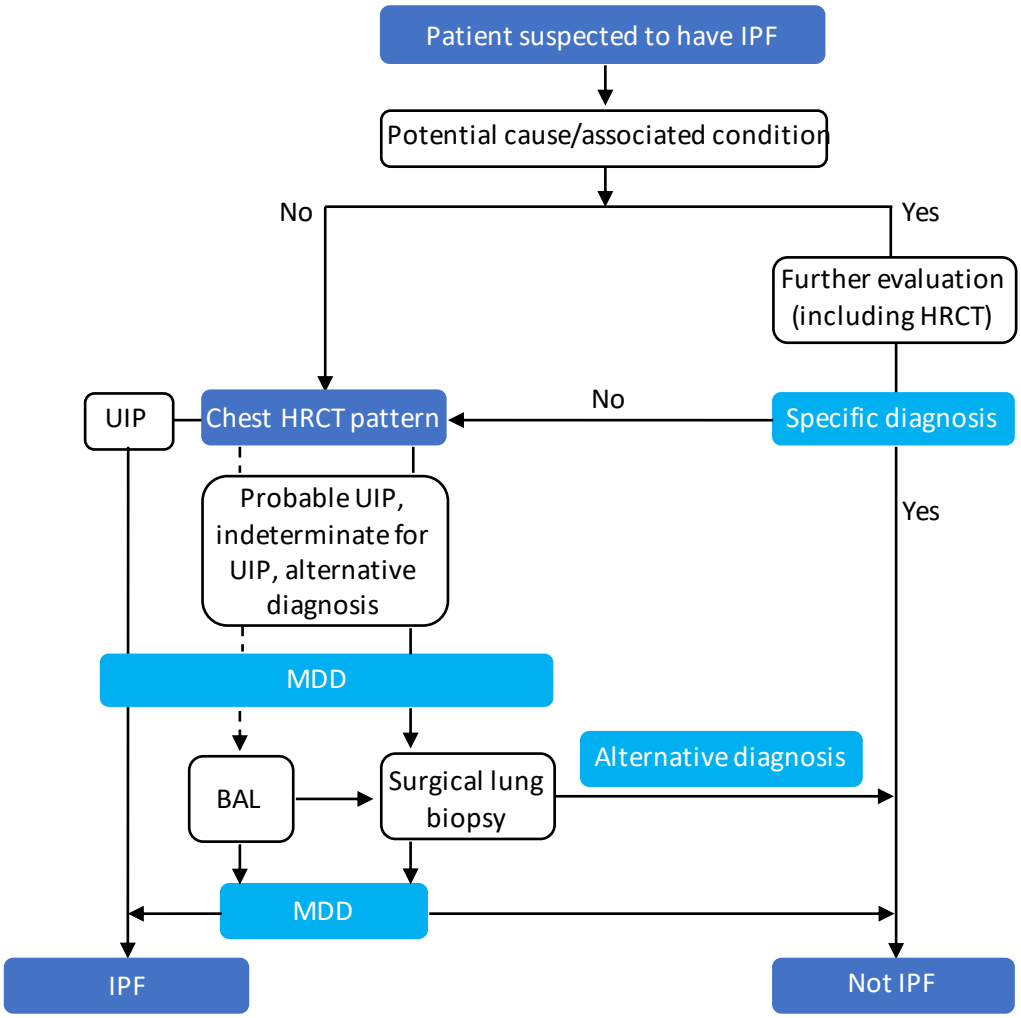


Figure 1.2 Diagnostic approach for idiopathic pulmonary fibrosis.

Patients with suspected IPF are evaluated for potential and/or identifiable causes of ILD, including domestic and environmental exposures, connective tissue disease (CTD), or drug toxicity. In this absence of this, IPF can be diagnosed through the evaluation of HRCT images of the chest and histopathological findings. Surgical lung biopsy is not recommended in patients at high risk for intra-, peri-, or postoperative complications. A MDD is recommended for a more accurate diagnosis. IPF: idiopathic pulmonary fibrosis; HRCT: high-resolution computed tomography; UIP: usual interstitial pneumonia; MDD: multidisciplinary discussion among ILD experts. Adapted from [13].

Currently, the diagnosis of IPF is based on the constellation of clinical, radiologic and histopathologic features [14]. Given the complexity of the disease, a multidisciplinary discussion (MDD) among ILD experts is advised to improve the accuracy of the diagnosis of IPF. Part of the diagnostic approach includes obtaining a thorough history, information about comorbidities, medication use along with serologic studies to identify predisposing domestic and occupational environmental exposures. Furthermore, information about family history can be informative as up to 10% of cases of IIP have been estimated to be familial [15]. In the absence of an identifiable cause for ILD, the radiologic finding of UIP on HRCT combined with a lack of a clinically identifiable trigger in a patients' history is often enough to secure a diagnosis of IPF **(Table 1.1)**.

Table 1.1 Multidisciplinary guidelines for the diagnosis of idiopathic pulmonary fibrosis (IPF) are based upon high-resolution computed tomography (HRCT) and biopsy patterns.

IPF suspected		Histopathology pattern			
		UIP	***Probable UIP***	***Indeterminate UIP***	***Alternative Diagnosis***
HRCT pattern	***UIP***	IPF	IPF	IPF	Non-IPF dx
	Probable UIP	IPF	IPF	IPF (likely)	Non-IPF dx
	Indeterminate for UIP	IPF	IPF (likely)	Indeterminate for IPF	Non-IPF dx
	Alternative diagnosis	IPF (likely)	Non-IPF dx	Non-IPF dx	Non-IPF dx

IPF: idiopathic pulmonary fibrosis; HRCT: high-resolution computed tomography; UIP: usual interstitial pneumonia; dx: diagnosis. Adapted from [13].

However, UIP is simply an endpoint of a fibrotic process characterised by the presence of basal and peripheral reticular opacities and honeycombing, with or without traction bronchiectasis [14] **(Figure 1.3)**. Indeed, several diseases with known aetiologies including asbestosis, CTDs, and sarcoidosis can develop this pattern of UIP, complicating the diagnostic process [9, 12].

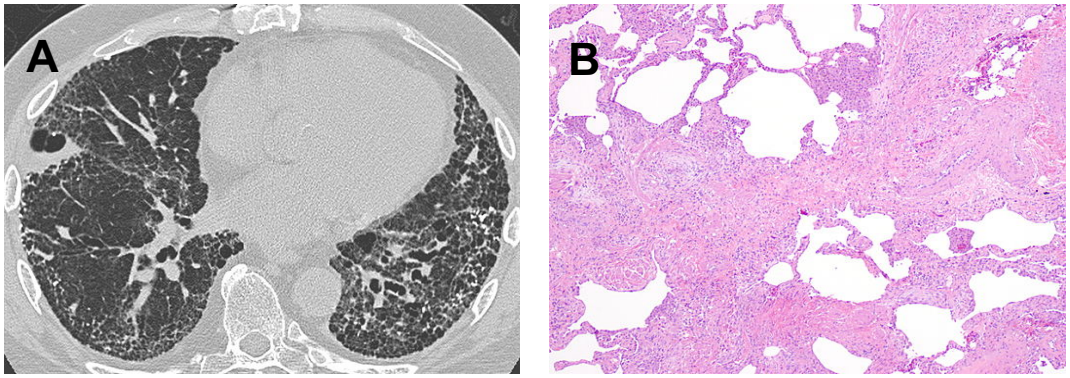


Figure 1.3 Typical pattern of usual interstitial pneumonia (UIP) on high-resolution computed tomography (HRCT) and on histopathology slides.

(A) HRCT shows changes consistent with the UIP pattern (i.e. spatial heterogeneity, honeycombing changes with basal predominance). **(B)** The histological pattern of UIP shows areas of active fibrosis with fibroblastic foci and extracellular matrix deposition, alternating with areas of relatively normal lung tissue. Copyright © *American Journal of Respiratory and Critical Care Medicine* [12].

1.2.3 Prognosis and outcomes

The prognosis of IPF is poor, with a median survival of 3 to 5 years post-diagnosis [9]. The course of IPF can be both variable and unpredictable, with some patients progressing rapidly, while others follow a more indolent pathway [16] **(Figure 1.4)**.

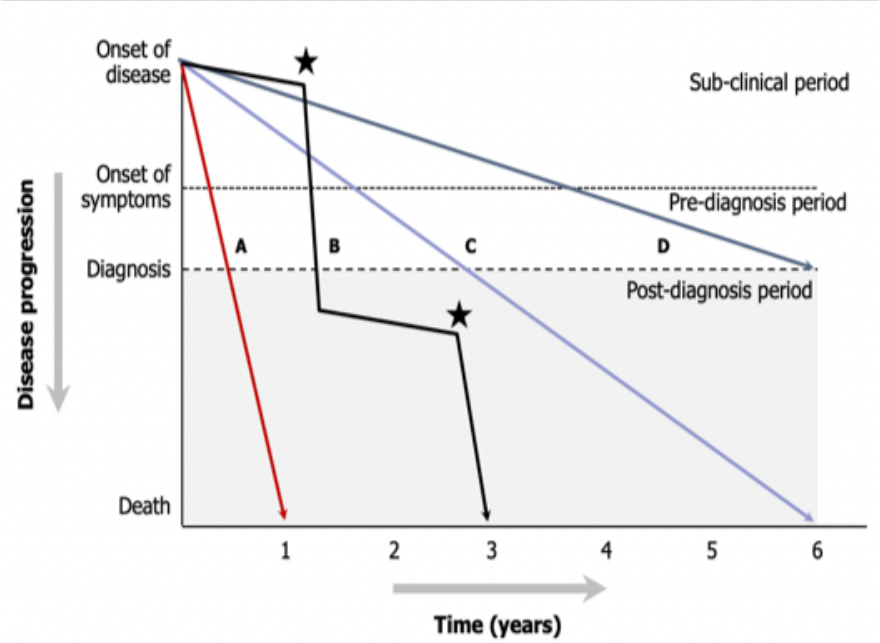


Figure 1.4 Potential clinical courses of idiopathic pulmonary fibrosis (IPF).

As the disease progresses, subjects with IPF may experience **(A)** rapid progression, **(B)** mixed rates of progression with periods of stability interspersed with **(★)** periods of acute deteriorations, or **(C)** slow, stable decline. Copyright © *American Thoracic Society* [16].

Despite current advances, the clinical course of the disease remains challenging to predict. Nonetheless, at baseline, several factors have been identified to predict the risk of mortality in patients with IPF, including a low diffusion capacity for carbon monoxide (DL_{CO}) [17], desaturation during the 6-minute walk test [18], the extent of honeycombing on HRCT [19], increasing age [20], and the presence of pulmonary hypertension at diagnosis [21]. Longitudinal factors which have been associated with an increased risk of mortality in IPF include a decrease in forced vital capacity (FVC) by ≥ 10% [17], a decline in DL_{CO} by ≥ 15% [17], worsening fibrosis on HRCT [22], and increase in the level of dyspnoea [17]. The majority of subjects with IPF have multiple comorbidities which contribute to an increased risk of morbidity and mortality [23]. Common comorbidities among patients with IPF include pulmonary hypertension, gastro-oesophageal reflux disease (GORD), chronic obstructive pulmonary disease (COPD), lung cancer, pulmonary embolism, obstructive sleep apnoea, ischemic heart disease, and diabetes [24].

1.2.4 Acute exacerbations

The clinical course of IPF is both variable and unpredictable, with significant heterogeneity among individual patients [16]. Whatever trajectory the disease adopts it can be punctuated by unpredictable periods of acute deterioration; episodes termed acute exacerbations of IPF (AE-IPF). These are defined as acute, clinically significant respiratory deterioration with no identifiable cause, characterised by new, widespread alveolar abnormalities [25] **(Figure 1.5)**.

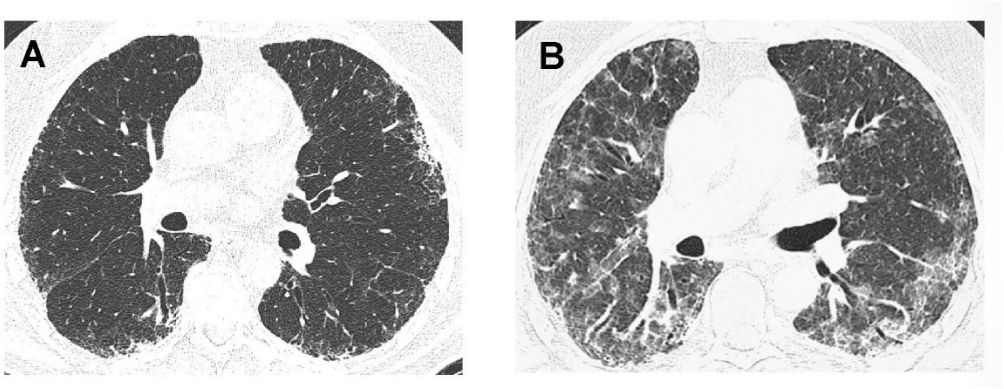


Figure 1.5 High-resolution computed tomography (HRCT) imaging of idiopathic pulmonary fibrosis (IPF).

(A) HRCT image of a patient at the time of IPF. **(B)** HRCT image of the same patient in acute exacerbation (AE) showing extensive ground-glass abnormalities against a background of basilar honeycombing consistent with AE-IPF. Copyright © *American Journal of Respiratory and Critical Care Medicine* [25].

Sometimes, these episodes may be attributed to identifiable causes such as heart failure, pulmonary embolism, pneumonia or mechanical procedures, including lung surgery [26]. However, in most instances, they are unpredictable, and the causes leading to their initiation are unknown. The exact incidence of AEs is unclear but seems to range between 4-20% per year with these episodes conferring substantial morbidity and mortality (often > 50%) [27, 28]. Furthermore, AE-IPF carry a significantly higher mortality rate compared to acute exacerbations occurring in other respiratory diseases, including asthma or COPD [29]. The symptoms with which patients usually present are insidious in onset and include worsening dyspnoea, cough, and fever [29]. Invasive sampling is recommended as a diagnostic approach to exclude infection; however, bronchoscopy is often not feasible as it carries a significant risk in this acute setting. As a result, serological testing and non-invasive sampling are often employed. Nonetheless, our current understanding of the pathophysiology of AE-IPF is limited, and it remains unclear whether AE-IPF represent an accelerated phase of the underlying fibrotic condition or a response to an occult trigger. Therefore, there is a pressing need to elucidate the mechanisms involved in these events.

1.2.5 Pathogenesis

1.2.5.1 Established model to explain IPF pathogenesis

To date, the aetiology of IPF remains elusive. Nonetheless, significant strides have been made in elucidating the pathomechanisms underlying the disease. Recently, the theory that fibrosis in IPF results from an ongoing inflammation in response to an exogenous stimulus has been replaced by the notion of IPF as an epithelial-driven disease. The current paradigm suggests that IPF arises in genetically susceptible ageing individuals as a consequence of repeated or prolonged exposure to environmental triggers which cause recurrent alveolar epithelial cell (AEC) injury and/or AEC death [30]. In the long term, this process results in an aberrant wound-healing response, which leads to fibrosis and causes the progressive structural destruction of the lung [31] **(Figure 1.6)**.

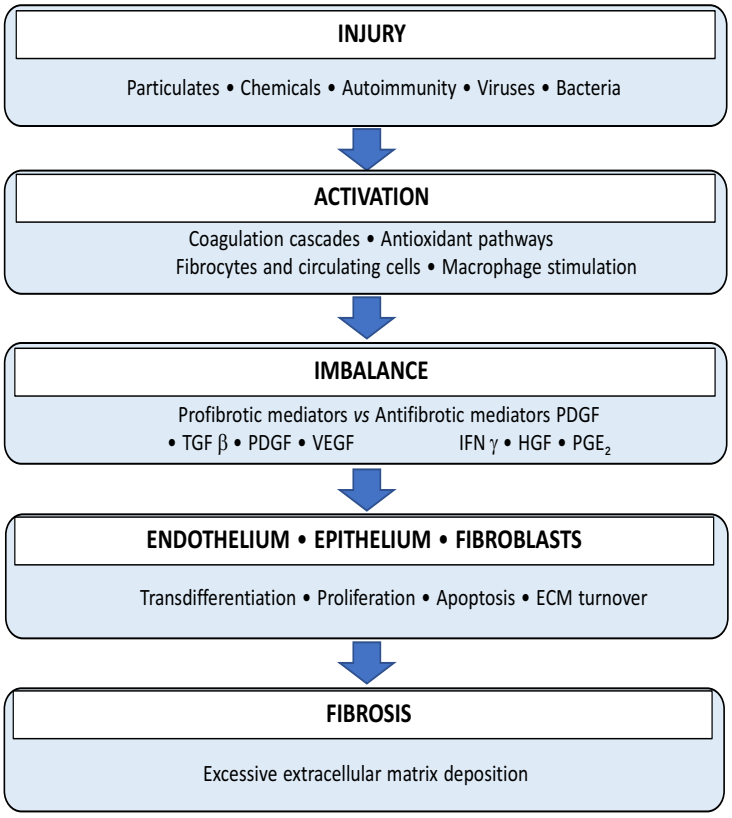


Figure 1.6 Proposed mechanism contributing to the pathogenesis of idiopathic pulmonary fibrosis (IPF).

In genetically susceptible individuals, repetitive alveolar injury activates several inflammatory, cell signalling and repair pathways. Activation of these pathways results in an imbalance of profibrotic and antifibrotic mediators. Consequently, multiple cell types are activated by these mediators, resulting in changes in cellular functioning and cell-cell interactions, which eventually culminate in progressive fibrosis. TGF-β: transforming growth factor beta; PDGF: platelet derived growth factor; VEGF: vascular endothelial growth factor; IFNγ: interferon gamma: PGE_2: prostaglandin E2. Adapted from [31].

Numerous genetic and environmental triggers, as well as infective agents, have been proposed as potential causative factors; however, the initial stimulus remains unidentified [32]. Following an insult, AECs proliferate and migrate to the site of injury to repair the damaged epithelium. Dysfunctional AECs produce several mediators (e.g. cytokines, chemokines, matrix metalloproteinases and growth factors) which promote the migration of cells, including monocytes and fibroblasts as well as induce their differentiation into fibrogenic cell types such as myofibroblasts, which can form fibroblastic foci [32, 33]. It is also thought that AECs may act as a direct source of fibroblasts and myofibroblasts by undergoing epithelial to mesenchymal transition [34]. These cells secrete several inflammatory mediators and enzymes, resulting in further alveolar epithelial barrier and basement membrane damage, as well as an increased deposition of ECM components. Consequently, accumulation of fibrous connective tissue leads to the progressive and irreversible destruction of the lung parenchyma which results in the classical radiologic and histologic patterns of UIP [35]. This irreversible destruction of the lung parenchyma causes loss of alveolar structure, an impairment in gas exchange and ultimately culminates in respiratory failure. Collectively, this hypothetical model suggests that the host and the environment may be acting synergistically in modulating the pathogenesis of IPF. As well as alveolar epithelial cells in the alveolar compartment, bronchial epithelial cells which populate the conducting airways are involved in IPF pathogenesis. Although the causative event that initiates the fibrotic cascade is unknown, apoptosis of epithelial cells has been proposed as a potential trigger [36]. Indeed, a study has shown that IPF patients carry an increased number of apoptotic cells in the bronchial epithelium [37]. The bleomycin mouse model supports this concept by demonstrating that inhibition of epithelial cell apoptosis prevents the development of the disease [38]. Although this model is widely used in IPF research and is representative of the histological features of the fibrotic lung, it does not fully replicate the extent of human disease [39]. Unlike in subjects diagnosed with IPF where the scarring of the lungs is progressive and irreversible, the bleomycin mouse model fails to recapitulate the pathology observed in IPF since this model induces acute lung injury but the mice will recover from the challenge and return to normal.

1.2.6 Risk factors

1.2.6.1 Genetic susceptibility

There is clear evidence that an individual's genetic background plays a role in IPF pathogenesis, with common genetic variants responsible for approximately one-third of the risk of developing IPF [40]. Genome-wide association studies have highlighted the importance of genes involved in telomere maintenance (*TERT*, *TERC*, *OBFC1*), host defence (*MUC5B*, *TOLLIP*) and epithelial cell integrity (*DSP*, *9-DDP*) in IPF susceptibility [41, 42]. These findings support the notion that loss of normal pulmonary physiology and aberrant immunity due to genetic alterations, may play an aetiological role in IPF, and support the concept that infectious agents may act as a source of AEC injury, driving disease initiation and progression.

For several years, case series of pulmonary fibrosis occurring in families suggested a genetic predisposition to the disease. These heritable pulmonary fibrosis forms are called familial interstitial pneumonia (FIP) and only account for 3.7% of all IPF cases [43-45]. These familial forms have been extensively studied to provide a greater understanding of the genetic factors underlying the pathogenesis of IPF. FIP cases have been linked to mutations in the *SFTPC* gene, encoding surfactant protein-C (SP-C) as well as mutations in *SFTPA1*, encoding SP-A1, and *SFTPA2*, encoding SP-A2 [46-48]. Surfactant proteins are collagen-containing lectins essential for pulmonary alveolar immunity and homeostasis [49]. Specifically, SP-A and SP-D are secreted into the alveolar space by type II AECs and have carbohydrate recognition domains that allow them to bind to bacteria and viruses, promoting phagocytosis by macrophages [49, 50]. Mutations in the *SFTPA2* and *SFTPC* genes have been shown to lead to the aggregation of the proteins in the endoplasmic reticulum, leading to increased endoplasmic reticulum stress and preventing their secretion into the alveolar space [47, 51, 52]. SP mutations account for only 5% of sporadic IPF [53]. Nonetheless, given that both SP-A and SP-C are produced exclusively by type 2 AECs, these studies still support a role for endoplasmic reticulum stress in IPF pathogenesis.

Polymorphisms in the promoter region of the *MUC5B* gene have also been linked to IPF in bronchial epithelial cells [41] suggesting that epithelial defects, not only confined to alveolar epithelial cells, may be affecting the onset of disease. Furthermore, a common single nucleotide polymorphism (SNP) in the promoter region of *MUC5B* has been shown to be associated with the development of both familial and sporadic IPF [54, 55]. This gain-of-function variant (*rs35705950*), increases the expression of the mucin 5B (MUC5B) in the distal airways of the IPF lung [55]. Paradoxically, subjects carrying this variant show improved survival compared to individuals without it. However, it remains

unclear how increased MUC5B expression promotes pulmonary fibrosis. In mice, *Muc5b* is essential for normal macrophage function and effective mucociliary clearance of bacteria [56]. Specifically, it has been shown that *Muc5b* overproduction in the distal lungs of mice is associated with mucociliary clearance dysfunction and the induction of an exaggerated fibrotic response which can be prevented by treatment with a mucolytic agent [57]. These findings highlight a causative role for *Muc5b* in murine lung fibrosis and therefore support the hypothesis that, in humans, excess production of MUC5B reduces mucociliary clearance of inhaled particles resulting in prolonged and recurrent exposure, which triggers injury to the alveolar epithelium and eventually leads to the development of fibrosis.

Pulmonary fibrosis has been diagnosed in some patients with dyskeratosis congenita (DKC), an autosomal dominant disorder caused by mutations in *DKC1*, a component of the telomerase complex, although mutations in the *TERT* and *TERC* genes have also been reported [58]. Heterozygous mutations in *TERT* and *TERC* have been identified in 8% of patients in a cohort with FIP, with mutation carriers exhibiting shorter telomeres, indicating that mutations in these two genes can cause pulmonary fibrosis [59]. Furthermore, studies have shown that, in IPF, abnormal telomere shortening is not specific to patients with telomerase mutation-associated pulmonary fibrosis [59-61]. Recently, an early-phase clinical trial conducted in subjects diagnosed with IPF who had short telomeres, revealed that telomeres can be lengthened via the oral administration of the synthetic androgen danazol [62].

Mitochondrial dysfunction and accelerated senescence also potentially contribute to the pathogenesis of IPF [63]. Studies have reported increased levels of free mitochondrial DNA in the plasma and bronchoalveolar lavage of subjects with IPF [64] as well as shown that in the IPF lungs, AECs exhibit large numbers of damaged mitochondrial DNA [65]. Specifically, decreased expression of the signalling molecule PINK1 in type 2 AECs has been observed in IPF lungs and has been shown to correlate with the accumulation of dysmorphic mitochondria as well as increased AEC apoptosis [65]. In a separate study, it was shown that the activity and expression of iodothyronine deiodinase 2 (DIO2), an enzyme that activates the thyroid hormone, were higher in the lungs from patients with IPF [66]. *Dio2* knockout mice exhibited enhanced bleomycin-induced pulmonary fibrosis and delivery of aerosolised thyroid hormone resolved fibrosis and improved survival. Furthermore, treatment with thyroid hormone promoted the expression of PINK1, inducing the restoration of normal mitochondrial function and rescue from mitochondria-regulated apoptosis [66]. The authors concluded that the antifibrotic properties of the thyroid hormone are associated with the protection

of alveolar epithelial cells and restoration of mitochondrial function and that this hormone may therefore represent a potential therapy for pulmonary fibrosis [66]. An increased risk of IPF has also been linked to other genetic variants in *DSP*, *AKAP13*, *CTNNA and DPP9*, which are involved in cell adhesion, integrity and mechanotransduction [67]. Collectively, these variants and genetic mutations may predispose individuals to AEC dysfunction following epithelial injury at the initiation of fibrosis [53].

1.2.6.2 Epigenetics

Epigenetics is the study of heritable changes in gene expression that do not involve changes to the underlying DNA sequence [68]. Epigenetic modifications can be influenced by environmental stimuli and may represent an important mechanism in IPF. These changes encompass a wide range of molecular modifications to both DNA and chromatin and are mediated primarily by DNA methylation, microRNAs (miRNAs) and histone modifications [68]. In IPF, a distinct pattern of differential DNA methylation in the lung tissue has been observed compared with healthy controls [69]. Changes in DNA methylation have been shown to correspond to altered messenger RNA (mRNA) expression of a number of genes, suggesting that DNA methylation may play an important role in the pathogenesis of IPF [70]. miRNAs are post-transcriptional regulators that control gene expression by binding to untranslated regions of mRNA, which can either lead to mRNA degradation or inhibition of protein translation [71]. Studies have shown that approximately 10% of the microRNAs are significantly altered in IPF lungs [72]. Furthermore, miRNAs may be involved in the development of IPF as miRNA expression in IPF lung has been correlated with disease severity [73].

1.2.6.3 Environmental risk factors

Although genetic risk factors have been recognised to play a role in IPF, they explain only a minority of the risk for developing the disease. Epidemiological studies have also implicated exposures to inhaled environmental or occupational agents to an increased risk of IPF, including exposure to cigarette smoke, wood dust, silica, farming and livestock [74, 75]. However, their role in disease pathogenesis is unclear. Smoking is the most prevalent risk factor that has been linked to the development of IPF. Smoking history has been associated with decreased survival in subjects diagnosed with IPF and has been implicated as the strongest risk factor for the development of FIP, suggesting that a history of smoking can contribute to the development of the disease [20, 76-78]. However, given that IPF also occurs in non-smokers, smoking cannot be attributed as the only trigger.

GORD is a condition which develops when the reflux of the stomach contents enters the oesophagus and causes unwanted symptoms and complications [79]. GORD is one of the most common comorbid conditions in IPF and given the increased prevalence, it has been hypothesised that microaspiration of gastric material may play a role in IPF pathogenesis. GORD has been previously associated with the development of post-transplant bronchiolitis obliterans [80]. However, research is still ongoing to determine whether the aspiration of acid refluxate causes the parenchymal damage associated with IPF.

1.2.7 Infection

The hypothesis that infection may play a role in the pathogenesis of IPF is supported empirically through observations that active infection and immunosuppression with prednisolone and azathioprine in combination with N-acetylcysteine increase the risk of death and hospitalisation in patients with IPF [81, 82]. In contrast, use of the antibiotic co-trimoxazole has been shown to reduce the frequency of respiratory tract infections and improve the overall quality of life in patients with IPF [83]. Genetic studies have highlighted the importance of disordered host defence and thus susceptibility to infection in contributing to disease progression in IPF [41, 54, 84]. In mice, *Muc5b* appears essential for normal macrophage function and effective mucociliary clearance of bacteria, suggesting that infection may be an environmental trigger in susceptible individuals [56, 85].

1.2.7.1 The role of viruses in IPF

Extensive research investigating the role of infections in the pathogenesis of ILDs and as triggers of AE in IPF have focused primarily on viruses [86-88]. It has been hypothesised that chronic infection with viruses may provide a persistent antigenic stimulus in susceptible hosts, culminating in fibrosis [87]. There is growing evidence, both from animal and human studies, supporting a mechanistic role for these pathogens in the initiation ad progression of IPF. The human herpes virus (HHV) family, which includes cytomegalovirus (CMV), Epstein-Barr virus (EBV), HHV-7 and HHV-8, has received considerable attention as either an aetiologic or exacerbating agent of IPF [88]. The first serological study to suggest an association between HHV and IPF found that 12 out of 13 patients were seropositive for EBV [89]. Consistent with this finding, several studies reported an increased incidence of EBV in bronchoalveolar lavage (BAL) samples and lung biopsies from subjects with IPF compared to controls [90, 91]. Furthermore, in support of these human studies, a mouse model studying murine gamma-herpes virus 68 (MHV-68), a virus closely related to EBV, revealed that mice latently infected with MHV-68 prior to bleomycin exposure developed an exaggerated fibrotic response compared to non-infected animals [92]. According to a study by Lawson and colleagues, HHV infection was also found in a majority of cases with both sporadic and familial forms of IPF but not in the healthy control group [93]. Viral presence in these subjects closely correlated with markers of endoplasmic reticulum stress, suggesting that AEC stress as a result of occult viral infections, may promote aberrant repair processes, leading to architectural remodelling of the lung. Furthermore, at-risk, asymptomatic relatives of patients with familial IPF displayed interstitial changes under HRCT scan, had increased levels of HHV DNA in BAL and showed evidence of HHV antigen expression in AECs, which closely associated with ER stress markers [94]. Collectively, these findings suggest that the presence of subclinical infection and virus-mediated AEC dysfunction prior to the development of any symptoms, strongly suggests that occult infection may be a factor involved in the pathogenesis of the disease.

1.2.7.2 The role of viruses in AE-IPF

Emerging evidence supports the hypothesis that occult viral infections may play a pathogenic role as co-factors for the development of AE-IPF [95] **(Table 1.2)**. Using molecular analysis of viral particles and BAL, Wotton and colleagues found infection in 19 of 43 cases of AE-IPF, including 12 cases of torque teno virus (TTV) in the exacerbation cohort with no viruses detected in the stable-IPF group [87].

Table 1.2 Summary of studies linking the respiratory microbiome with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis.

Virus	Main conclusions
HHV	EBV detected in serum of 12/13 subjects with IPF but not in patients with other forms of ILD
	MHV-68 triggers an exaggerated fibrotic response in mice
	Increased incidence of EBV in BAL and lung biopsies of IPF subjects compared to controls
	EBV detected in BAL of 2/43 AE-IPF subjects
	CMV detected in BAL of 2/43 AE-IPF subjects
	A total of 38% of AE-IPF subjects exhibited evidence of CMV infection
	HHV detected in nasopharyngeal swabs of 15/30 AE-IPF subjects and 4/30 individuals with stable-IPF
TTV	TTV detected in BAL of 12/43 AE-IPF subjects
	No evidence of TTV in BAL of stable-IPF subjects
	Increased mortality in IPF subjects with presence of TTV-DNA in serum compared to IPF subjects with no TTV-DNA
Influenza A	Influenza A detected in nasopharyngeal swabs of 12/30 AE-IPF subjects but not in individuals with stable-IPF
	A case of AE-IPF was reported following pandemic influenza A vaccination

HHV: human herpes virus; TTV: torque teno virus; EBV: Epstein-Barr virus; ILD: interstitial lung disease; MHV-68: murine gamma-herpes virus-68; BAL: bronchoalveolar lavage; AE-IPF: acute exacerbation of IPF; CMV: cytomegalovirus. Copyright © *European Respiratory Review* [96].

A separate study which evaluated the relationship between TTV and IPF found that there was a lower survival rate in IPF subjects who were seropositive for TTV-DNA compared to those who were not [96]. Concordant with Wootton's findings, Ushiki and colleagues examined BAL from subjects with an AE-IPF and detected respiratory syncytial virus (RSV) and CMV by polymerase chain reaction in one and two cases, respectively [97]. In another study which employed immunohistochemical techniques to detect viral antigens, 38% of IPF subjects experiencing an AE exhibited evidence of measles or CMV infection [98]. More recently, Weng and colleagues investigated viral sequences in nasopharyngeal swabs of 30 stable and 30 AE-IPF patients [99]. A total of 57 different viruses were detected, with virus-positive samples highest in the AE-IPF group (60%) compared to the stable-IPF group (43.3%). HHV, which has already been shown to be associated with IPF and to cause exacerbation of pulmonary fibrosis in animal models, was the most prominent virus in the AE-IPF group [100, 101]. Influenza A was also found to dominate the AE-IPF group. Given that AE-IPF have

already been reported following influenza A vaccination, its presence suggests that cold-associated infection may potentially trigger these events [102]. Collectively, these findings highlight that AE may result from multiple causes, including viral infections. However, most cases do not show such infections, and overall, the data supporting a mechanistic role for airborne viruses is insufficient to prove their causal role in AE-IPF. One potential explanation for this is that AE occur as a result of a previous viral infection that is no longer present at the time of symptom onset.

1.2.8 Therapeutic options

The lack of pathogenic understanding combined with the diagnostic difficulties mean there is a paucity of therapies for the treatment of IPF. In 2014, two antifibrotic drugs were approved for the treatment of IPF: pirfenidone, an antifibrotic and anti-inflammatory agent, and nintedanib, a tyrosine-kinase inhibitor which inhibits several key receptors including those for platelet-derived growth factor (PDGF), vascular endothelial growth factor (VEGF), and fibroblast growth factor (FGF) [103, 104]. Results from clinical trials have demonstrated that both drugs are effective in slowing disease progression by reducing the decline in lung function in patients with IPF [105-107]. However, they are unable to improve or even stabilise the disease, or more importantly, the symptoms experienced by the patient [9]. Despite these advances, however, there remains no evidence-based or licenced therapies for the treatment of AE-IPF [108].

1.3 The Composition of the Airway Microbiome

1.3.1 The concept of lung sterility

Until recently, the lungs were considered to be sterile, given the inability to isolate bacteria from the healthy lungs using conventional culture techniques. The emergence of culture-independent techniques has been instrumental in identifying bacterial species and shaping our understanding of the role of bacteria in health and disease. The “lung microbiome” is defined as the commensal, symbiotic and pathogenic community of microorganisms that inhabit the epithelial surfaces of the lower respiratory tract. The most commonly used approach to the study of bacterial communities exploits high-throughput sequencing of amplicons of the 16S ribosomal RNA (16S rRNA) gene, a highly conserved locus in the bacterial genome [109]. While some regions share the same sequences across all bacteria, the 16S rRNA locus also contains hypervariable regions which provide species-specific signatures useful for phylogenetic classification [109]. Groups of bacteria sharing similar gene sequences are clustered together and identified by comparison to 16S rRNA reference databases. Bioinformatics analysis of the data produced on platforms such as Illumina MiSeq allows quantifying the total and relative abundance of these bacterial groups [110, 111]. High-throughput

sequencing techniques have shown that the epithelial surfaces of the human respiratory tract are colonised by a complex and dynamic microbial community, including bacteria, fungi, phages and viruses. Compared to conventional culture techniques which detected bacteria in 39.1% of BAL samples, high-throughput bacterial 16S rRNA sequencing has been shown to identify bacterial DNA in 95.7% of BAL samples [112]. These studies have highlighted differences in bacterial communities in health and disease, which may contribute to explain the pathogenesis of these respiratory diseases [113-117].

1.3.2 Development of the lung microbiome at birth

Until recently, our understanding of how the airway microbiome forms has been hindered by the long-held view that the lungs are sterile. The foetus develops within a sealed amniotic fluid sac, and the current dogma is that the womb is sterile [118]. Although there is no evidence of *in utero* colonisation [119], recent studies have reported the presence of bacterial DNA in the placenta and amniotic fluid [120, 121]. A wide range of microorganisms have been detected in the upper respiratory tract of healthy neonates following birth [122, 123]. Aagard and colleagues [119] demonstrate that the lower respiratory tract microbiota forms within the first two years of birth and that the mode of delivery determines the microbiota constituents in preterm, but not term, births. Despite contrasting findings between studies, several environmental factors, including mode of delivery, antibiotic treatment, and infant feeding, have been reported to affect the microbial composition of the upper airways [124-126]. Only a very limited number of studies have addressed the timing and colonisation pattern of the lower airways shortly after birth given the ethical and technical limitations of obtaining samples. Mouse studies have shown that the lung microbiome develops within the first days of life, following which microbial load and diversity increase. Shortly after birth, the lower airways of mice are dominated by bacteria of the *Firmicutes* and *Proteobacteria* phyla, followed by an increase in the *Bacteroidetes* phylum prior to weaning, up to adulthood [127]. Studies using germ-free (GF) mice have shown that the microbiome is crucial for the development of the host immune responses, particularly in the context of allergy; the absence of a microbiome in GF mice results in their exaggerated responsiveness to allergens [128]. Growing evidence from both animal and human studies supports the concept that the microbiome in very early life has associations with long-term lung health.

1.3.3 The lung microbiome in health and its ecological determinants

The respiratory microbiome is a complex and adaptive ecosystem; in health, the composition of the lung microbiome is determined by three factors: microbial immigration, microbial elimination, and the relative reproduction rates of its members [129] **(Figure 1.7)**.

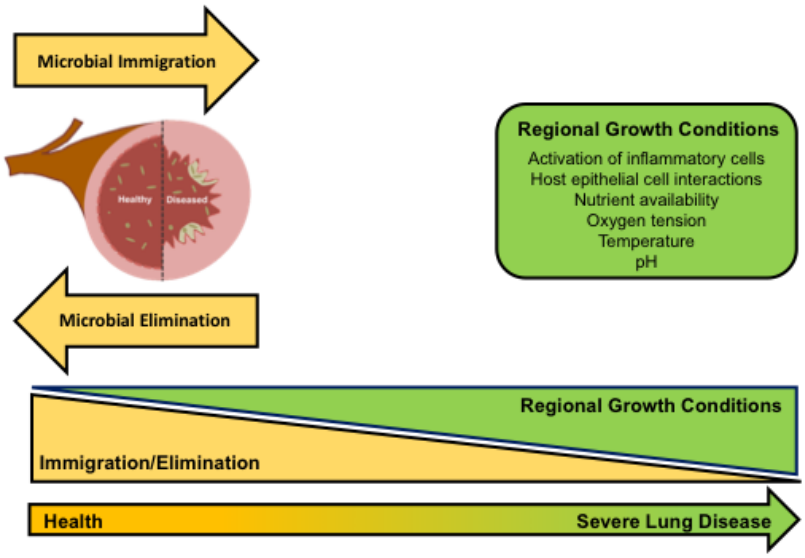


Figure 1.7 Ecological determinants of the respiratory microbiome.

The total bacterial burden and the relative community constitution of the lung microbiome is determined by the balance of three factors: microbial immigration, microbial elimination, and the relative reproduction rates of its members. In health, the composition of the microbiome is mostly determined by the balance of immigration and elimination; in severe lung disease, community membership is mainly determined by the regional growth conditions. Adapted from [126].

The similarities in community composition between the oral and lung microbiota suggest that micro-aspiration from the digestive tract is likely to be the dominant route of immigration [130]. It has also been proposed that direct mucosal dispersion of microbes from the nasal and oral cavities into the lungs, as a result of their anatomical continuity, may represent another route for microbial immigration to the lower airways [129, 131].

The quantity and composition of the microbiome are influenced by the balance between the rate of immigration and clearance. Microbial elimination is an active process which is achieved through mucociliary clearance, coughing and host immune defences which selectively recognise, kill and clear the microbes [132]. The airway microbiota in healthy lungs is dominated by the *Bacteroidetes*, *Firmicutes*, and *Proteobacteria* phyla [133]. The dominant genera found in the lower airways are *Prevotella*, *Veillonella* and *Streptococcus* [134]. Unlike the gastrointestinal tract, the bacterial load in the lungs is very low, and the bacterial species inhabiting the lungs must adapt to survive in highly selective conditions. The environmental conditions necessary for microbial growth within the respiratory tract (e.g. temperature, pH, nutrient availability, oxygen tension and activation of host

inflammatory cells) differ significantly from the upper respiratory tract to the alveoli [129, 131]. This implies than in a single healthy lung, considerable regional variation in bacterial community composition can be observed. Nonetheless, differences in the microbial composition between individuals is far greater than within an individual, thus rendering the identification of an optimal healthy respiratory microbiome rather challenging [135]. Several protective mechanisms act synergistically with the local conditions in the lungs to limit bacterial infection, including; the epithelial barrier, a highly reactive mucosal immune system, antimicrobial peptide production, pulmonary surfactant production, and mucous secretion [129, 133, 136].

1.3.4 The lung microbiome in respiratory disease

The discovery of the lung microbiome prompted the exploration of its role in a variety of respiratory conditions [114, 115, 137-144] **(Table 1.3)**.

Table 1.3 Summary of studies illustrating changes in the respiratory microbiome which occur during chronic lung disease.

Disease	Taxonomic changes at the genus level*	Sample type	Cohort size
Asthma	Increased abundance of *Haemophilus*	Bronchoscopy	11
	Increased abundance of *Haemophilus*, *Neisseria*, *Fusobacterium*, and *Porphyromonas*	Bronchoscopy	42
	Increased abundance of *Streptococcus*	Nasopharyngeal	234
COPD	Increased abundance of *Lactobacillus*.	Lung tissue	24
	Increased abundance of *Fusobacteria*, *Leptotrichia*, and *Fusobacterium*	BAL	32
CF	Increased abundance of *Burkholderia*, *Streptococcus* and *Staphylococcus*	Sputum	23
	Increased abundance of *Pseudomonas*, *Staphylococcus*, *Stenotrophomonas*, and *Achromobacter*	Sputum	17
	Increased abundance of *Staphylococcus*, *Streptococcus*, and *Pseudomonas*	BAL	95
IPF	Increased abundance of *Haemophilus*, *Streptococcus*, *Neisseria*, and *Veillonell*a	BAL	65
	Increased abundance of *Staphylococcus* and *Streptococcus*	BAL	55
	Increased abundance of *Streptococcus*, *Prevotella*, *Veillonella*, *Haemophilus*, and *Pseudomonas*	BAL	35

COPD: chronic obstructive pulmonary disease, CF: cystic fibrosis; IPF: idiopathic pulmonary fibrosis; BAL: bronchoalveolar lavage. *Changes in the microbiota in cases compared to controls. Copyright © *Immunology* [138].

It is now well recognised that host-microbial communities play a fundamental role in shaping healthy lung function, and that disturbances in their composition are commonly observed in respiratory diseases such as asthma, COPD, cystic fibrosis (CF), and IPF [113-115, 117, 141, 144-146]. In these chronic respiratory diseases, the balance between immigration and elimination is lost, leading to alterations in the lung microbiota, including a decrease in overall richness and the overgrowth of certain bacteria which have been associated with disease progression.

1.3.5 The lung microbiome in IPF and AE-IPF

The first study to investigate the presence of bacterial growth in the lower airways of subjects with IPF was conducted in 22 individuals and confirmed the presence of known pathogens including *Haemophilus*, *Pseudomonas* and *Streptococcus* in 36% of the cases [99, 114, 115, 144, 145, 147-149] **(Table 1.4)**.

Table 1.4 Summary of studies linking the respiratory microbiome with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis.

Diagnosis	Main conclusions	[Ref.]
IPF	Positive BAL cultures in 8/22 stable IPF subjects: *Haemophilus influenzae* (2/22); *Haemophilus parainfluenzae* (2/22); *Moraxella catarrhalis (*1/22); *Pseudomonas aeruginosa* (1/22); *Proteus mirabilis* (1/22); *Streptococcus pneumonia* (1/22)	[147]
	Increased abundance of *Streptococcus* OTU1345 and *Staphylococcus* OTU1348 is associated with a significant reduction in progression-free survival in IPF	[114]
	Streptococcus pneumoniae triggers progression of pulmonary fibrosis through pneumolysin in two different mouse models	[148]
	Increased bacterial burden in IPF subjects compared with COPD and healthy controls	[115]
	Higher bacterial burden at the time of diagnosis predicts disease progression in IPF	[149]
	GF mice protected from mortality following bleomycin exposure	
AE-IPF	Four-fold increase in bacterial burden in AE-IPF subjects compared to stable-IPF	
	Increased abundance of *Campylobacter* and *Stenotrophomonas* and decreased abundance of *Veillonella* in AE-IPF compared to stable-IPF	[144]
	Positive sputum cultures in 9/48 AE-IPF subjects: *Klebsiella pneumoniae* (2/48), *Mycobacterium tuberculosis* (4/48), *Pseudomonas aeruginosa* (1/48), *Loffi Acinetobacter* (1/48), Other (1/48)	[99]

IPF: idiopathic pulmonary fibrosis; OTU: operational taxonomic unit; AE-IPF: acute exacerbation of idiopathic pulmonary fibrosis. Copyright © *European Respiratory Review* [96].

Over the past five years, advances in high-throughput molecular sequencing techniques have been used to characterise the microbial composition in the lower airways of IPF patients. These studies have investigated the role of the lung microbiome in the airways in IPF and suggest that a perturbed microbial state may be linked to disease outcome.

It has been hypothesised that these bacterial communities, present outside the context of overt infection, can act as a continued driver of epithelial injury. Specifically, an increased bacterial burden, loss of microbial diversity due to the outgrowth of certain pathogenic bacteria have now been implicated with disease progression in IPF [114, 115]. Furthermore, these studies have also reported that baseline bacterial burden can predict the rate of disease progression and risk of death [115]. The Correlating Outcomes with Biochemical Markers to Estimate Time-Progression (COMET) study was the first study to evaluate the lung microbiome in patients with IPF [114]. Following adjustment for confounding factors, the authors found a significant association between the abundance of specific microbial community members and disease progression. Having an increased abundance of *Streptococcus* or *Staphylococcus* species was associated with a reduced progression-free survival time. However, the increased presence of these two bacteria was observed in less than half of the cohort, making it unlikely that these organisms alone can explain the disease pathogenesis. In a prospective study, Molyneaux and colleagues evaluated the microbial composition in BAL in patients with IPF, COPD and age-matched healthy controls [115]. Subjects with IPF were found to have an increased bacterial burden which was associated with disease progression, as well as an increased expression of genes related to host defence. Furthermore, they displayed significantly less diverse bacterial communities and were found to harbour a greater abundance of potentially pathogenic *Haemophilus*, *Neisseria*, and *Streptococcus* species. In a separate cohort, the same authors found that the bacterial burden was also increased in AE-IPF subjects compared to patients with stable disease [144]. Furthermore, despite showing no evidence of infection in clinical cultures, patients experiencing AE-IPF showed distinct changes in microbial community composition, such as an increased relative abundance of potentially pathogenic *Stenotrophomonas* and *Campylobacter* species, compared to stable disease. These organisms are potential respiratory pathogens and the presence of *Campylobacter*, a bacterium usually confined to the gastrointestinal tract, led the authors to suggest that micro-aspirations and increased bacterial immigration to the lungs, may play a role in AE-IPF. Although this study offers an insight into the potential aetiology of AE-IPF, the patient cohort was relatively small, and there were only two paired samples from the same individuals when stable and experiencing an exacerbation. Therefore, future studies comparing samples from the same individual at baseline and during an

acute exacerbation will allow for a more accurate insight into the biological changes underlying these events.

Collectively, these studies highlight that alterations in the respiratory microbiome exist in IPF, but they do not provide any insight into the functional mechanisms underlying the host-microbial interactions that may drive IPF pathogenesis and progression. O'Dwyer and colleagues have recently tried to overcome this issue in an attempt to move from correlation to causation [149]. The authors provided a link between microbial alterations in the lung, disturbed alveolar inflammation and IPF progression. Specifically, they found that alveolar inflammation was perturbed in patients with IPF and that a reduction in bacterial diversity promoted proinflammatory and profibrotic signalling profiles in the airways of these patients.

Using a GF mouse model of fibrosis, the authors set out to explore the role of bacteria in alveolar inflammation and fibrinogenesis. Interestingly, they found that the absence of microbiota in GF mice was associated with altered humoral and cellular immunity and reduced mortality compared to conventional mice, despite exhibiting similar levels of fibrosis. This led the authors to suggest that IPF subjects not only die from progressive fibrosis but also from distinct inflammatory causes of acute-on-chronic respiratory failure, including respiratory infections and AE-IPF [145]. Overall, this study provides the first causal evidence that the microbiome is involved in the pathogenesis and mortality of fibrotic lung disease and that preclinical GF models are a useful tool to investigate potential mechanisms of host-microbiota interactions. However, the potential mechanism of the respiratory microbiome in the pathogenesis of IPF remains unclear, and there is a pressing need for functional and longitudinal which will allow to determine more targeted treatments for lung diseases including IPF. Additionally, the microbiome has not been fully characterised, and future studies should also characterise organisms other than bacteria and viruses, including fungi.

1.3.6 Effects of antibiotic therapy on the airway microbiome

Increasing evidence has emphasised the use of oral antibiotics and their effect on lung immunity. Antibiotic treatment could affect the lung microbiome either directly or indirectly through its impact on the gut microbiome, which subsequently affects lung immunity. Indeed, murine studies provide evidence that depletion of certain species within the gut microbiota due to antibiotic intake influences lung diseases and allergic inflammation [150]. In Finnish pre-school children, microbial disturbances due to the use of macrolides in early life correlated with an increased risk of asthma [151]. It is evident that to better understand the impact of antibiotics on lung health, we need to study the interactions between the gut and airway microbiota in greater detail. Management of

subjects with IPF is arduous given that treatment options are aimed at reducing disease progression but not mortality. The use of antibiotics in IPF was first reported in 1996 when a patient with advanced fibrotic lung disease received oral co-trimoxazole and showed clinical improvement. In a later double-blind, randomised, placebo-controlled pilot study in 20 patients with progressive fibrotic lung disease, Varney and colleagues validated this finding by showing that treatment with co-trimoxazole improved exercise capacity, breathlessness, and symptom scores [152].

Similarly, in a double-blind follow-up study of 181 patients with fibrotic idiopathic interstitial pneumonia, co-trimoxazole treatment resulted in an overall improvement of health-related quality of life, and a reduction in the percentage of patients requiring an increase in oxygen therapy [83]. However, the treatment was not well tolerated and resulted in high drop-out rates, and there was no detectable difference in the primary outcome of change in forced vital capacity. Nonetheless, a reduction of infection and mortality was noted in patients who tolerated the treatment. However, it is difficult to conclude whether this observation may be a result of co-trimoxazole's antimicrobial effects, given its simultaneous anti-inflammatory actions. Ultimately, a more targeted antibiotic approach may be a promising avenue for the treatment of IPF.

1.4 Microbial Metabolites

1.4.1 The short-chain fatty acids

The airways consist of diverse habitats that shape the composition of the lung microbiota. The microbial communities that reside in the lungs differ in metabolite production. Strikingly, even bacteria of the same species, but belonging to distinct communities, produce different metabolites [153]. The relationship between microbiome and host is regulated by a complex immune-metabolic-neuroendocrine crosstalk, which appears to be mostly mediated by microbial metabolites [154]. The short-chain fatty acids (SCFAs), including acetate (CH_3COO^-), propionate ($C_2H_5COO^-$), and butyrate ($C_3H_7COO^-$) are produced from the bacterial fermentation of non-digestible carbohydrates by intestinal bacteria in the gut lumen [153]. In the large intestine, SCFAs can be found at concentrations ranging from 30 to 150 mM [155]. These metabolites have widespread effects on the gut but have also been shown to exert their effects in peripheral tissues such as lung [156]. SCFAs act as a link between the microbiota and the immune system by interacting with immune cells and modulating their recruitment, differentiation, activation and survival at different tissues [157]. They have been implicated in a plethora of biochemical processes including digestion, nutrient absorption, hormonal secretion, cell proliferation and differentiation, neurological functionality, host metabolism and energy production [158, 159]. Furthermore, there is evidence of

SCFA being implicated in several chronic inflammatory diseases, including type 2 diabetes, gastrointestinal and neurological disorders, obesity, heart disease and respiratory disorders [160]. However, there is contrasting evidence on whether these SCFAs have protective or causative effects in the context of inflammation [161, 162].

1.4.2 The gut-lung axis

Cumulating studies have provided evidence of the crosstalk which occurs between the intestinal microbiota and the lung, termed the 'gut-lung axis'. It is now well recognised that the intestinal microbiota influences the microbial composition within the lung either by direct seeding of bacteria into the airways or by the distribution of SCFAs [163] **(Figure 1.8)**.

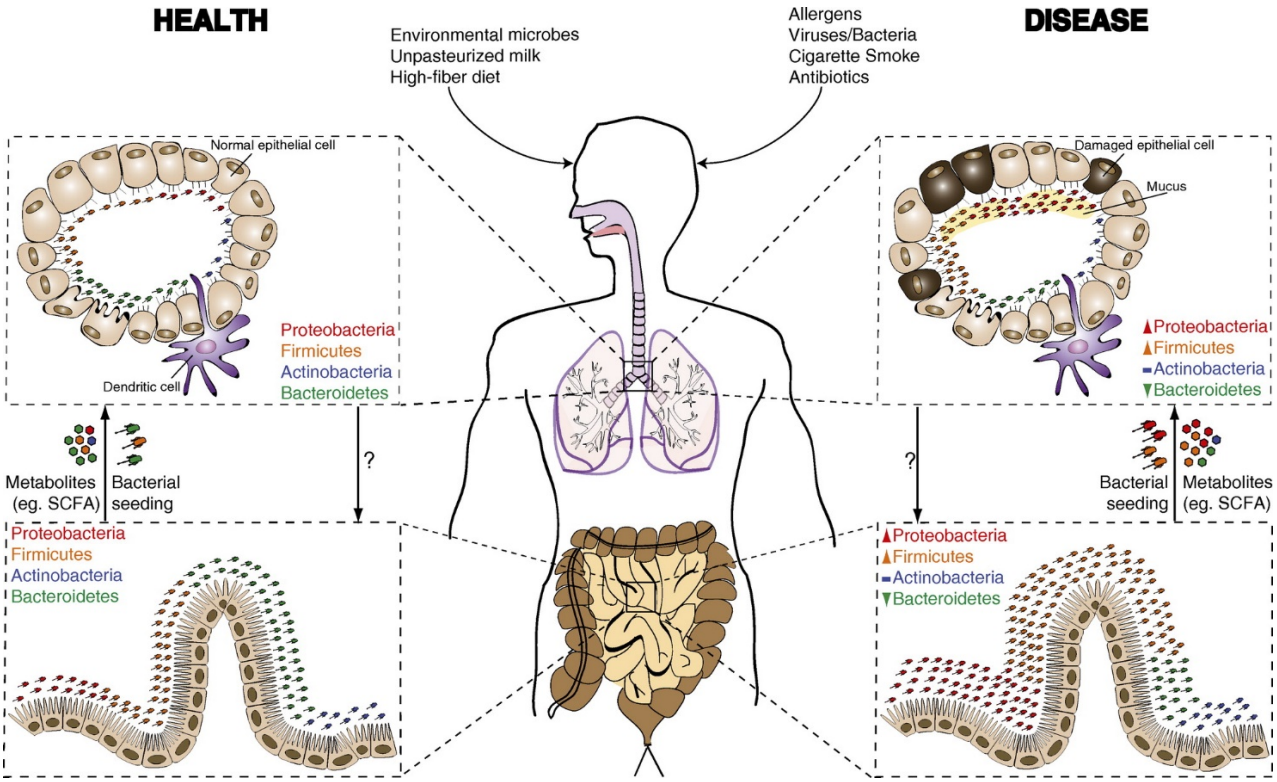


Figure 1.8 Model of intestinal-pulmonary cross-talk during respiratory health and disease.

In health, the intestinal and airway microbiota is dominated by bacteria belonging to the *Bacteroidetes* and *Firmicutes* phyla. During respiratory disease, a shift in community composition of both the intestinal and airway microbiota has been reported, with an outgrowth of *Proteobacteria* and *Firmicutes*. A cross-talk between these two mucosal sites has been proposed. The intestinal microbiota influences pulmonary microbial composition and immune responses in two ways: by the direct seeding of the respiratory tract with bacteria and by the distribution of bacterial metabolites, such as the short-chain fatty acids (SCFAs), which promote the growth of certain SCFA-producing bacteria (e.g. *Bacteroidetes*) and may act directly as immunomodulatory molecules. Copyright © *Annals American Thoracic Society* [163].

These SCFAs have been shown to exert their downstream effects via two main mechanisms: by i) activating G-protein coupled receptors (GPCR41, GPCR43 and GPCR109a), which are expressed on several cell types (e.g. epithelial cells, endothelial cells, adipocytes, neurones, hepatocytes, and

immunocytes) and by ii) inhibiting histone deacetylases [161, 164]. While the majority of SCFAs are produced and utilised within the gut, the widespread expression of SCFA receptors implies that the SCFAs may also influence host physiology by acting at peripheral sites, including the lungs [165, 166].

Disturbances in the intestinal microbiota has been implicated in several lung diseases and respiratory infections [156, 167]. The predominance of *Bacteroidetes fragilis* and total anaerobes in early life, along with decreases in the total abundances of *Escherichia coli*, *Faecalibacterium*, *Lachnospira*, *Rothia*, and *Veillonella* bacterial genera have been implicated in an increased risk of asthma [168, 169]. Furthermore, murine studies suggest that mice devoid of their intestinal microbiota during early life are at an increased risk of developing allergic airway disease [169, 170]. Mice fed a peptide-based enteral diet led to an increased production of SCFAs, which attenuated elastase-induced inflammation and emphysema in the lung [171]. Similarly, the relationship between fibre intake and increased production of SCFAs has also been studied in humans. A study by Thornburn and colleagues [172] found that high dietary fibre intake during late pregnancy was associated with higher acetate levels in humans and protected the infants against the development of airway disease.

Additionally, in a mouse model of allergic airway disease, administration of a high-fibre diet caused alterations in the gut microbiota by increasing proportions of bacteria belonging to the *Bacteroidetes* phylum, which are known as potent fermenters of fibre into SCFAs [173]. These changes were accompanied by increases in circulating SCFAs [167]. Mice fed a fibre-rich diet showed reduced levels of IL-4, IL-5, IL-13 and IL-17A expression in lung tissue, reduced eosinophil infiltration into the lung, and improved airway hyperresponsiveness. These effects on allergic inflammation were found to be dependent on GPCR41 but not GPCR43. Collectively, these results illustrate that dietary fermentable fibre and SCFAs can shape the immunological environment in the lung and influence the severity of allergic inflammation.

This axis is bidirectional and respiratory infection can drive perturbations in the gut microbiome [169, 170]. A study reported that intratracheal administration of lipopolysaccharide disrupted the airway microbiota and led to the translocation of airway bacteria belonging to the *Clostridium* and *Lachnospiraceae* genera into the bloodstream which altered the gut microbiota within 24 hours and led to an increase in total bacterial load [174]. Furthermore, intranasal delivery of a nonabsorbable tracer into the nasal cavity of mice was found in the gastrointestinal tract shortly after [170]. Collectively, these studies illustrate the existence of a bidirectional communication occurring between these two mucosal sites. However, what still remains unclear is whether these changes

observed in the microbiota are the cause or the effect of disease and further studies are urgently needed to validate these findings.

1.4.3 Short-chain fatty acids promote inflammation in the lungs

In the lungs, the pro-inflammatory roles of the SCFAs have also been reported. In 2015, it was discovered that anaerobic bacteria identified in the lungs of CF patients produced millimolar concentrations of SCFAs, which induced a dose-dependent release of IL-8 from bronchial epithelial cells (BECs) of CF patients [175]. Mirković and colleagues observed that this effect was reduced in BECs from control subjects compared to CF and that GPCR41, but not GPCR43, was upregulated on the surface of CF BECs compared to BECs from control subjects. This observation was further validated by showing that suppression of the GPR41 gene, resulted in a significant reduction in SCFA-mediated IL-8 release by CF BECs [175]. Similarly, Ghorbani and colleagues studied the effects of the SCFAs acetate, propionate and butyrate on airway inflammatory responses using epithelial cell lines and primary cultures. In this study, they provide evidence that SCFAs stimulate the release of proinflammatory cytokines, including IL-8, granulocyte colony-stimulating factor and, IL-6. The secretion of SCFAs was associated with increased neutrophil recruitment, suggesting that SCFAs may drive innate immunity in CF airways. Recently, treatment of primary human lung fibroblasts and airway smooth muscle (ASM) cells with supraphysiological concentrations of propionate or butyrate in combination with TNF-α resulted in significantly greater IL-6 and IL-8 release from human lung fibroblasts and ASM cells than challenge with TNF-α alone [176]. In ASM cells, treatment with acetate in combination with TNF-α induced IL-6 release only. In agreement with previous CF studies, the authors found that upregulation of IL-6 and IL-8 was mediated through the activation of GPCR41, but not GPCR43 [176]. These findings suggest that SCFAs may be acting synergistically with other endogenous molecules such as TNF-α, to promote airway inflammation. Therefore, SCFAs appear to play a key role in the regulation of inflammation through the activation of GPCRs. Nonetheless, their effects are context-dependent, and may either result in the stimulation or dampening of inflammatory processes. However, the relevance of SCFAs in chronic lung diseases such as IPF is still unknown.

1.5 The Respiratory Epithelium

1.5.1 The epithelial barrier

The human respiratory tract acts as the first line of defence against the continuous immigration of a vast array of airborne microorganisms, including microbes and allergens. The surface of the lungs is composed of a continuous layer of epithelial cells, which represent a physical and biological

barrier for potentially harmful environmental stimuli and pathogens [137] **(Figure 1.9)**. The respiratory system is divided into the upper respiratory tract including the nasal cavity, pharynx, larynx and the lower respiratory tract which encompasses the conducting airways (trachea and bronchi), the small airways (bronchioles) and the respiratory zone (the alveoli). All of these compartments within the respiratory system contain distinct cell types which are required for maintenance and immune homeostasis. Defects in the respiratory epithelium can lead to the development of several inflammatory disorders of the airways and lungs [177]. The core function of the lungs is gas exchange between the lungs and the blood, and the total surface area of the lungs which is of approximately 70 m^2 allows for this gas exchange to occur [178].

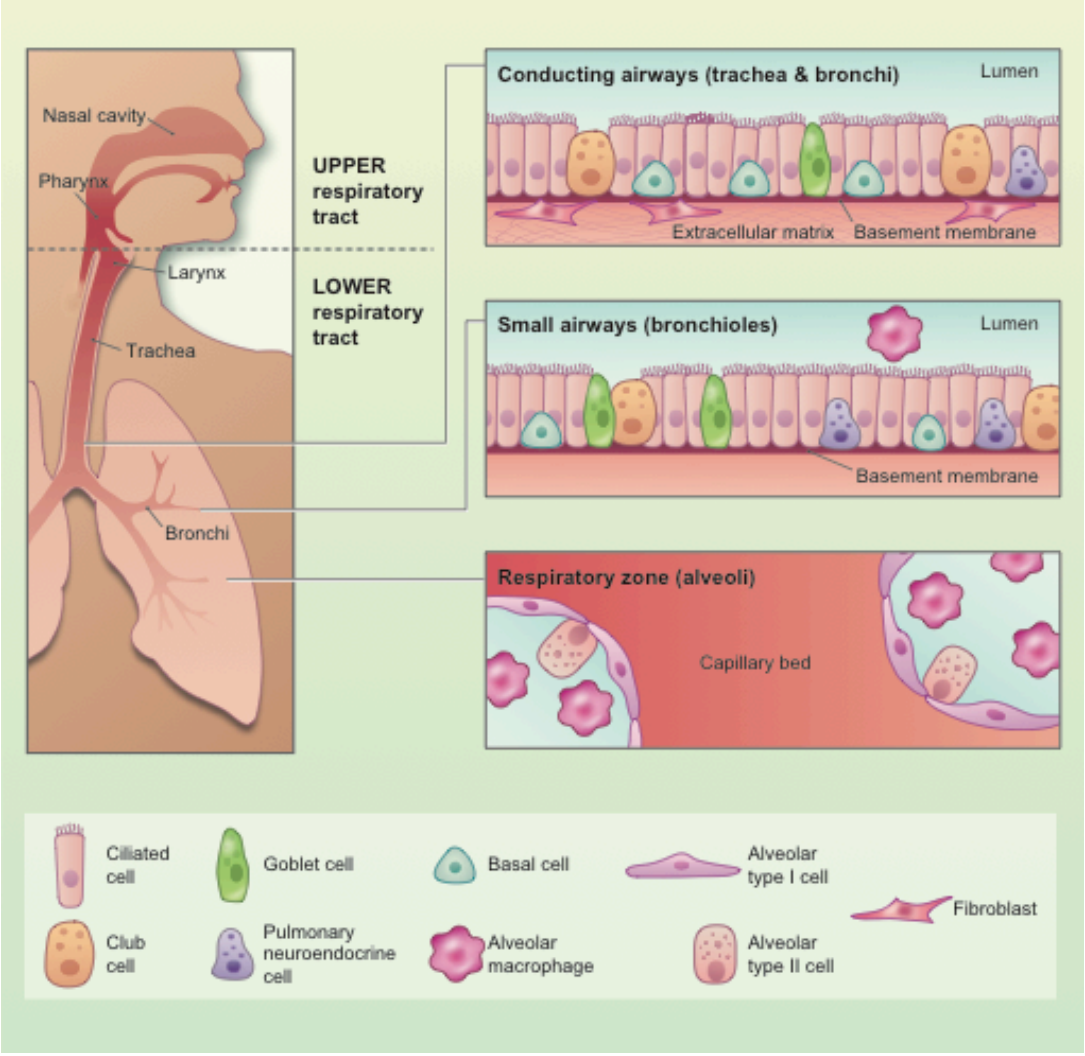


Figure 1.9 Composition of the respiratory epithelium in health.

The human respiratory tract is divided into the upper respiratory tract including the nasal cavity, pharynx, larynx and the lower respiratory tract which comprises the conducting airways (trachea and bronchi), the small airways (bronchioles) and the respiratory zone (the alveoli). The trachea and large airways are composed of ciliated cells, mucus-secreting goblet cells, neuroendocrine cells and undifferentiated basal cells. Stromal cells such as fibroblasts provide ligands and extracellular matrix which modulate airway epithelial cell turnover. The alveoli are comprised of alveolar macrophages and alveolar type I and type II cells. Copyright © *Immunology* [138].

The respiratory tract harbours several bacterial communities which vary in density and diversity; with the highest bacterial load observed in the upper respiratory tract. It comprises different anatomical sites, and the types of epithelial cells which constitute these sites are varied both in composition and in structure, reflecting the distinct functions of the airway epithelial cells lining the conducting airways and the respiratory zone of the lungs. The lungs are constantly exposed to diverse communities of microbes, and the respiratory epithelium in the lower respiratory tract acts as the first line of defence acting as a barrier against potentially harmful environmental stimuli [177]. The pseudostratified epithelium of the conducting airways is designed to facilitate effective mucociliary clearance of particles and microbes and is comprised of the following cell types: ciliated cells, mucus-secreting goblet cells, club cells, neuroendocrine cells and basal cells which secrete surfactant [179]. The cilia on the apical surface of ciliated cells shuttle the inhaled particulates and mucus out of the airways. Scattered throughout the cilia are goblet and secretory cells which secrete mucus and trap inhaled particulates and microorganisms. Basal cells, also known as progenitor cells, are capable of self-renewal and act as resident stem cells for the trachea and proximal airways. These cells are responsible for repopulating the differentiated epithelium during homeostasis and following lung injury [180]. The trachea and bronchi are underlined with smooth muscle and interstitial fibroblasts which act as niche-supporting cells aiding in the regulation of the regenerative response of the airway epithelium after injury [180]. In the alveoli, the epithelial lining consists of flat alveolar epithelial type I cells (AEC1) which form a continuous layer allowing gas exchange, and cuboidal alveolar epithelial type 2 cells (AEC2) which act as progenitor cells and secrete pulmonary surfactant [181, 182]. The presence of intracellular tight junctions and adherens junctions is a common feature shared by airway epithelial cells and alveolar epithelial cells. Tight junctions are specialised sites of adhesion between cells and are localised at the apical surface; they are essential for epithelial adhesion and barrier function, ensuring that the cells adhere together to create semipermeable barriers between tissue compartments [183].

This intracellular adhesion complex is composed of a branching network of sealing strands, each strand acting independently from the others. The three major transmembrane proteins are occludin, claudins and transmembrane junctional adhesion molecules. These associate with different proteins, including zonular occluden (ZO)-1, -2, and -3 located on the intracellular side of the plasma membrane. The adherens junctions are positioned immediately below the tight junctions and are composed of β-catenin and E-cadherin, which mechanically connect the adjacent cells and initiate proliferation and differentiation [183, 184]. The complexes can be influenced by a

variety of factors including cellular differentiation, exposure to a range of stimuli including allergens and pollutants and may vary in disease conditions, contributing to immunopathology [185-188].

1.5.2 The relationship between the lung and its microbiota

Airway epithelial cells contribute significantly to the barrier function of the airway epithelium by acting as a physical barrier between the external environment and internal milieu. The three main contributors of this barrier function of the airway epithelium are: mucociliary escalators which trap and remove inhaled microbes and foreign particles from the airways, intracellular tight and adherens junctions which regulate epithelial paracellular permeability, and secreted antimicrobial products which kill inhaled pathogens [177]. The key players that contribute to the mucociliary function of the airway epithelium are mucus and cilia. Specifically, the submucosal gland in the large airways harbours the mucus and serous cells which contribute to the production of most airway mucus [189]. Towards the small airways (bronchioles), goblet and club cells produce mucus whereas in the respiratory zone (alveoli) AEC1 and AEC2 secrete pulmonary surfactant [190-192]. Mucus is a gel composed of approximately 95% water and the remaining 5% composed of mucins, salts, lipids, proteins and immunological factors [193]. The major gel-forming mucins secreted in the airways are MUC5AC from goblet cells and MUC5B from submucosal glands [194, 195]. Mucus plays an important role in lubricating, hydrating and protecting epithelial surfaces. It also houses the majority of the human microbiome and facilitates interactions between human cells and commensal microorganisms through the formation of a thin mobile layer which is supported by the periciliary layer covering the cilia [196, 197]. A functioning mucous barrier is essential for providing an effective defence against epithelial injury. In contrast, mucus dysregulation is associated with increased susceptibility to viral infections and respiratory allergies as well as contributing to obstruction in several respiratory diseases such as asthma, COPD, CF and pneumonia [198, 199]. For example, mucus overproduction can hamper the ability of the cilia to transport mucus out of the lungs and may result in the establishment of certain pathogenic bacteria in this obstructed niche [200].

The intracellular tight and adherens junctions are located at the apicolateral border of airway epithelial cells and provide adhesive contacts between neighbouring epithelial cells [201]. Tight junctions regulate the transport of solutes and ions across epithelia, whilst the adherens junctions mediate cell to cell adhesion and initiate the formation and promotion of tight junctions [183, 184]. During disease, disruptions in these junctional complexes can lead to increased epithelial permeability and pathogen entry [185-188].

In addition to acting as a physical barrier, the airway epithelium also acts as a biochemical barrier against invading pathogens. Airway epithelial cells release a wide variety of humoral factors such as antimicrobial peptides, which exert broad-spectrum antimicrobial activity and kill inhaled pathogens [202]. Amongst these, human β-defensins are the most abundant antimicrobial peptides expressed on the surface of the airway epithelium and target a wide range of respiratory pathogens. It has been shown that their expression in airway epithelia may be either constitutive or inducible by bacterial products or pro-inflammatory cytokines [203].

Another factor which shapes the maintenance and selection of the lung microbiota is nutrient availability in a specific niche. Mucins, cytokines, defensins and lactoferrins are all host compounds which act as a nutrient source for bacteria in the lungs [177, 204]. Along with temperature and nutrient availability, changes in oxygen tension can also influence microbial growth rates. Significant regional variation in oxygen tension can be observed within the lungs and can alter population dynamics by affecting bacterial proliferation [111]. For example, chronic pulmonary diseases are characterised by an increased mucus production which promotes bacterial growth and leads to the formation of zones with low oxygen concentration and high temperature favouring the selection and maintenance of particular bacteria [129]. Taken together, these differences in the biotic (e.g. cell layers) and abiotic (e.g. temperature, mucus, surfactant, and oxygen tension) characteristics of the lungs have a major impact on the installation and persistence of the bacterial communities.

1.5.3 Recognition of pathogens by airway epithelial cells

The airway epithelium plays a key role in the host defence system not only simply providing a structural barrier. Airway epithelial cells also express several pattern recognition receptors (PRRs) and secrete antimicrobial molecules and mucins, which aid in the defence against invading pathogens [179]. PRRs are evolutionarily conserved receptors which sense pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) [205-207]. These transmembrane receptors are classified into four families: the toll-like receptors (TLRs), the cytoplasmic proteins retinoic acid-inducible gene-I-like receptors (RLRs), the NOD-like receptors (NLRs) and the C-type lectin receptors (CLRs) [208]. Activation of these PRRs results in the initiation of numerous extracellular activating cascades, as well as various intracellular signalling pathways that culminate in inflammatory responses [209].

To date, ten (TLR1-TLR10) have been identified in humans and are characterised by an NH2-terminal extracellular leucine-rich repeat (LRR) domain, which mediates the recognition of their respective PAMPs, and a COOH-terminal intracellular trail containing a conserved region called the Toll/IL-1

receptor homology domain [210]. Two major pathways are initiated following activation of TLRs: the MyD88-dependent pathway, which is used by all TLRs except TLR3 and results in the activation of nuclear factor (NF)-κB and activator protein-1; and the TRIF-dependent pathway, which is initiated by TLRs and TLR4 and results in the activation of type I interferons [211, 212].

TLRs are critical for airway epithelial cell recognition of inhaled pathogens and for innate immunity. In mammals, the main function of TLRs is to recognise distinct PAMPs including components of microbial cell walls (TLR1, TLR2, TLR4, and TLR6), flagellin (TLR5), single-stranded or double-stranded viral RNA (TLR3, TLR7, and TLR8) or CpG-containing DNA (TLR9) [213-216]. TLR3, TLR7, TLR8, and TLR9 are localised in the endosome for ligand recognition in the lumen of the endosome [216-220]. In contrast, TLR10 is the only PRR in humans without known ligand specificity and biological function [221]. TLRs are expressed on immune cells including macrophages, dendritic cells and B cells as well as nonimmune cells such as endothelial cells and epithelial cells. Signalling through these TLRs is pivotal in early inflammatory responses to bacteria through the activation and expression of host cytokines, chemokines and antimicrobial peptides. In humans, bronchial epithelial cells express most TLRs, including TLR2, TLR3, TLR5, and TLR6 [222]. Allergens and viruses can also trigger the activation of the PRR system [223-225]. The PRRs embedded in the epithelium can recognise allergen-associated danger signals via two pathways: protease dependent pathways which either require direct cleavage of cell-surface molecules or the activation of G-protein coupled protease-activated receptors and protease independent pathways [226-229].

RLRs are expressed in the cytoplasm in a variety of cells and recognise the presence of genomic RNA or double-stranded RNA intermediates from single-stranded RNA viruses. For example, melanoma differentiation-associated gene-5 is expressed in human bronchial epithelial cells and is involved in the detection of rhinovirus and other viruses [230].

NRLs are a group of cytoplasmic PRRs that are characterised by a central nucleotide-binding and oligomerisation domain (NOD) and C-terminal leucine-rich repeats [231]. The major role of NRLs is to recognise cytoplasmic microbial PAMPs and endogenous danger signals, however, the role of NOD function in airway epithelial cells still remains to be fully elucidated. The role of airway epithelial cells in immunity to fungi is less well studied. CLRs are PRRs which belong to the dectin family and recognise fungal cell wall components. For example, dectin-1 is a type II transmembrane natural killer (NK)-cell-receptor-like C-type lectin with a single extracellular C-type lectin-like domain which is expressed in airway epithelial cells and can mediate various cellular responses, including cytokine and chemokine production [232-234]. Dectin-1 recognises β-glucans; these are

carbohydrate polymers that are found in the cell walls of fungi, but also in plants and some bacteria [167, 235-237]. Fungal spores have been shown to be potent allergens which are associated with allergic and chronic diseases of the respiratory tract such as asthma; therefore, their involvement should not be overlooked [238].

1.5.4 Infections and the initiation of the epithelial immune response

Infections play an important role in the induction and exacerbation of a multitude of chronic respiratory diseases, including asthma and COPD [239]. These infections may involve bacteria (e.g. *Streptococcus pneumoniae*, *Haemophilus influenzae, Chlamydia pneumoniae*) or viruses (e.g. RSV, influenza) and represent a major cause of morbidity, mortality and exacerbations in these diseases [240, 241]. The initiation of infection usually occurs through the exposure of the respiratory tract to bacterial or viral particles **(Figure 1.10)**. Bacteria and viruses bind to specific cellular receptors, including intracellular adhesion molecule 1, sialic acid residues, platelet-activating factor and mannose-6 phosphate receptors [242]. The internalisation of these infectious agents by airway epithelial cells induces the stimulation of the PRRs and the secretion of antimicrobial peptides, reactive oxygen species and reactive nitrogen species. In addition, production of cytokines, chemokines and other mediators facilitate the recruitment of cells of the adaptive and innate immune system that are crucial for host defence [177].

In healthy conditions, this combined response results in pathogen clearance. In contrast, in patients with chronic respiratory diseases, these defences can be impaired, leading to increased susceptibility to infection that drives inflammatory responses that in turn exacerbate the underlying disease [243]. During chronic lung disease, barrier function has been shown to be impaired, leading to increased epithelial permeability. This loss of barrier function results in the influx of pathogens and other particles and has been linked to triggering of abnormal repair and remodelling [244]. Air pollutants, allergens and respiratory viruses can all inflict damage to the airway epithelium, either directly via the proteolytic degradation of airway epithelial intracellular adhesions or indirectly by binding to PAR-2 expressed on the apical surface of bronchial epithelial cells.

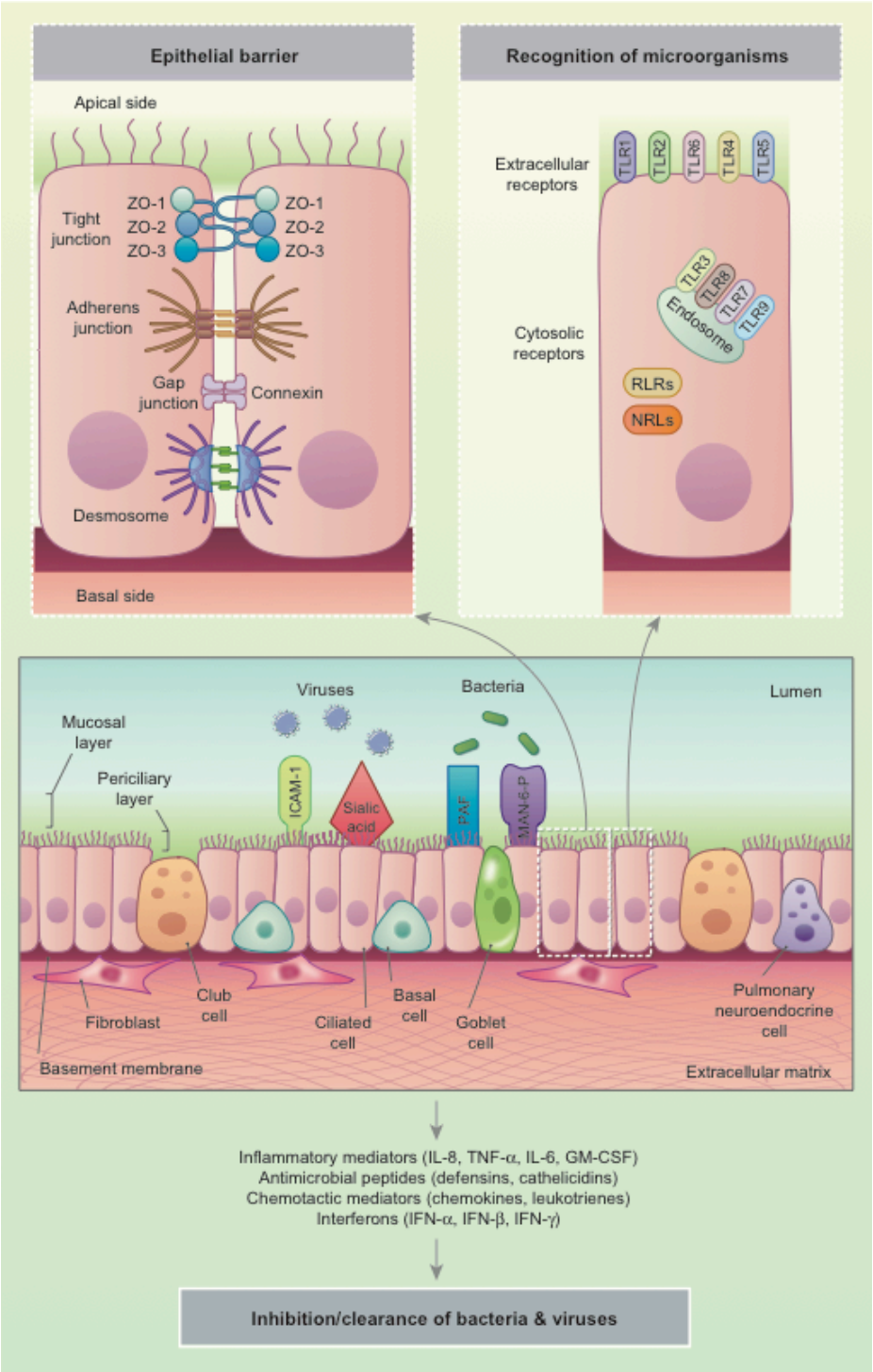


Figure 1.10 Response of the airway epithelium to infection.

Bacteria and viruses bind to cellular receptors, including intercellular adhesion molecule 1 (ICAM-1), sialic acid residues, platelet-activating factor (PAF) and mannose-6-phosphate (Man-6-P). The pathogens are internalised in airway epithelial cells which induce the production of innate immune defensins and the stimulation of intracellular and extracellular immune receptors such as the toll-like receptors (TLR), retinoic acid-inducible gene-I-like receptor (RLRs), and NOD-like receptors (NLRs). In healthy individuals, pathogen clearance is achieved following the generation of pro-inflammatory innate and adaptive immune responses. Copyright © *Immunology* [138].

In vitro studies have shown that infection of cells with rhinovirus results in the loss of ZO-1 from the junctional complex leading to a decrease in transepithelial electrical resistance (TEER), which is a measure of barrier leakiness. Rhinovirus, along with coxsackie virus, have been shown to affect occludins, whilst RSV can cause barrier dysfunction by inducing the disassembly of epithelial junctional complexes [245-247]. Bronchial epithelial barrier function has been shown to be compromised in asthmatic subjects, with decreases in ZO-1 and TEER correlating with asthma severity [248]. Furthermore, decreases in E-cadherin and β-catenin have been observed in bronchial epithelial cells from patients with asthma, and reduced expression of occludin and claudins has been reported in COPD [249, 250]. *In vivo*, leakage of proteins, including albumin into the alveolar space has been reported following influenza infection [251]. Viruses, pollutants and allergens have all been reported to associate with junction dysfunction and exacerbated clinical symptoms across the spectrum of these diseases [245-247, 252, 253].

Similarly, defects in barrier structure and function have also been reported in primary epithelial cell cultures from asthmatics differentiated at air-liquid interface for several weeks, indicating that reduced barrier function is a stable property of these cells [254]. Barrier dysfunction is also induced by the type 2 cytokines IL-4 and IL-13 which inhibit the expression of ZO-1, occludin, E-cadherin and β-catenin [255]. Furthermore, it has been hypothesised that junctional disruption could be a T helper 2 (Th2) promoting signal in the airway to restore mucosal integrity [256]. This concept is corroborated by the fact that siRNA knockdown of E-cadherin results in epidermal growth factor receptor (EGFR) dependent production of Th2 promoting cytokines thymus- and activation-regulated chemokine and thymic stromal lymphopoietin by airway epithelial cells which justifies the association between barrier dysfunction and Th2 driven pulmonary disease [257]. Furthermore, E-cadherin is also a ligand for CD103, which is expressed on innate and adaptive immune cells and a subset of dendritic cells. It binds to KLRG1, which is expressed by NK cells, regulatory T cells and group 2 innate lymphoid cells.

In asthmatics, E-cadherin and occludin are downregulated [248, 258, 259], and the genes for E-cadherin and protocadherin-1 are known to be associated with airway hyperresponsiveness [260, 261]. E-cadherin prevents the activation of the EGFR receptor by anchoring it within the adherens junctions [262-264]. As a result of reduced expression of E-cadherin, EGFR can redistribute to the luminal surface of the cell leading to its activation [265]. However, excessive EGFR activation can lead to cell proliferation and goblet cell hyperplasia/metaplasia [266]. Indeed, increased expression of EGFR has been observed in asthma and correlated with impaired mucosal barrier function [248].

Following dissociation from E-cadherin, β-catenin can translocate to the nucleus, activating Wnt signalling which results in cell proliferation [267]. It has been shown that dikkopf (DKK) proteins regulate Wnt signalling via interaction with Kreme (KRM) receptors [268] and the Wnt/β-catenin pathway has been implicated in IPF [36, 269]. For example, DKK1 expression is increased in lung and bronchoalveolar lavage fluid of IPF patients. Specifically, the highest level of expression has been observed in basal bronchial epithelial cells. Moreover, KRM1 expression is also elevated in the lung tissue of IPF patients, particularly in areas of bronchiolisation [269]. Interestingly, the phosphorylation signatures of β-catenin in IPF are the same as in normal lung development, again indicating that repair of lung injury involves reactivation of developmental programmes [270]. Wnt/β-catenin signalling is also implicated in the pathogenesis of COPD. In contrast to what has been observed in IPF, signalling in the mesenchymal compartment is inactivated rather than activated [271]. It has been hypothesised that a shift from canonical Wnt signalling (β-catenin dependent) to non-canonical pathways may drive IPF pathology [272].

Basal cells are an important population of self-renewing, progenitor cells which differentiate into ciliated and secretory cells to repair the injured airway epithelium and are found in highest concentration in the trachea, with numbers diminishing further down the bronchial tree to the bronchioles. When damage occurs, these cells upregulated p63; a basal cell restricted transcription factors, which has been implicated in epithelial-mesenchymal transition [273]. To further support the involvement of bronchial derived cells in the pathogenesis of IPF, a study by Jonsdottir and colleagues has shown that basal epithelial cells overlying fibroblastic foci in IPF acquire increased reactivity, determined by p63 expression, and gain a partial mesenchymal phenotype [274].

1.6 Summary of findings and unanswered questions

Systematic changes in bacterial community composition, such as increased bacterial burden and abundance of potentially pathogenic bacteria may drive disease progression, acute exacerbations, and mortality in IPF. It is likely that the dialogue between specific bacterial communities and the host is a dynamic phenomenon. The cellular components within the lungs, as well as the abiotic environment (e.g. temperature, pH, mucus, pulmonary surfactant, nutrient availability, and oxygen tension), may affect the installation and location of the bacterial communities within the lungs. Indeed, in several chronic lung diseases, loss of barrier has been shown to influence the composition of the respiratory microbiota by allowing the entry of pathogens and other particles. Yet, in IPF, the functional role of these bacterial communities in the airway compartment remains unclear with the functional interactions between the host and its respiratory microbiome unstudied. In particular,

the precise mechanisms underlying how the lung microbiome interacts with lung resident inflammatory cells and the airway epithelium, are yet to be elucidated. The focus of research in IPF has been on type 2 cells and alveolar epithelial cells, however, the microbiome and the anti-pathogen defences described in **sections 1.5.2-1.5.4** are generally active higher up in the proximal and distal airways where bronchial epithelial cells reside. Currently, the bleomycin-induced lung injury model remains the most widely used model to study IPF. However, this model fails to fully recapitulate the human IPF disease state, and no drugs have been successfully developed using this model [275]. Therefore, for the purpose of this thesis, the human *in vitro* model was used to study the impact of the respiratory microbiome and metabolome on primary bronchial epithelial cells (PBECs) and human lung fibroblasts (HLFs) from healthy controls and subjects with IPF.

It is clear that more information is needed to better understand the nature of the dialogues occurring between the respiratory microbiota and epithelial cells in the lungs to further elucidate the function and mechanisms through which the microbiota shapes the development of fibrosis in IPF. Further studies are urgently needed to validate these findings, address the crucial question of causality and potentially shape future interventional trials.

1.7 Hypothesis

The overarching hypothesis is that increased bacterial burden and shifts in bacterial communities within the lungs of IPF subjects results in aberrant production of metabolites which drive inflammation, lung structural remodelling and ultimately result in fibrosis.

1.8 Aims

1. To phenotype the airway microbiome in patients with IPF and healthy controls.
2. To phenotype the airway metabolic profile in patients with IPF and healthy controls by untargeted ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy and by targeted gas-chromatography-mass spectrometry (GC-MS).
3. To investigate the effect of microbial metabolites, specifically the SCFAs acetate, propionate and butyrate on airway epithelial cells and fibroblasts.

2 **Chapter 2**: Materials and Methods

2.1 Study Design

2.1.1 Subject recruitment

Subjects with a clinical diagnosis of IPF according to the current ATS/ERS/JRS/ALAT guideline definitions were recruited from the Interstitial Lung Disease Unit at the Royal Brompton Hospital (London, United Kingdom), between May 2014 and May 2018 [276]. Age-matched healthy control subjects were also recruited at this time and included smokers and non-smokers. No *a priori* power calculation was performed given the exploratory nature of the study and based on feasibility of sampling. Inclusion and exclusion criteria for these subjects are detailed in **Table 2.1**.

Table 2.1 Inclusion and exclusion criteria for study participants.

Diagnosis	Inclusion Criteria	Exclusion criteria
IPF	Aged > 18 years Clinical diagnosis of IPF according to the latest ATS/ERS/JRS/ALAT guidelines [276].	Co-existent conditions known to be associated with development of fibrotic lung disease (i.e. connective tissue disease, asbestosis, sarcoidosis, drug induced lung disease). Subjects with auto-immune profile considered diagnostic for a specific connective tissue disease even in the absence of systemic symptoms. Subjects with a history of self-reported upper or lower respiratory tract infection, an AE-IPF or antibiotic use in the three months prior to bronchoscopy. Subjects involved in clinical trials for therapeutic agents for IPF.
Controls	Age > 18 years No history or clinical diagnosis of COPD, asthma or allergic rhinitis. Post-bronchodilator FEV_1 > 80% predicted for age and height. Post-bronchodilator FEV_1/FVC ratio > 70%.	Subjects with a history of self-reported upper or lower respiratory tract infection or antibiotic use in the three months prior to bronchoscopy. Subjects with any clinically relevant or significant systemic disease.

IPF: idiopathic pulmonary fibrosis; AE-IPF: acute exacerbation of idiopathic pulmonary fibrosis; HRCT: high-resolution computed tomography; BAL: bronchoalveolar lavage; COPD: chronic obstructive pulmonary disease; FEV_1: forced expiratory volume in 1 second; FVC: forced vital capacity.

2.1.2 Ethics statement

Ethical approval for the study was given by the Research Ethics Committee (10/HO720/12 and 15/SC/0101), and all patients provided informed written consent.

2.1.3 Subject assessment and pulmonary function testing

As part of the routine clinical workup, a full history and clinical examination was performed on all subjects. Lung function testing was performed at baseline and every six months. Testing was conducted according to established protocols by NHS staff at the Royal Brompton Hospital (London, United Kingdom).

2.1.4 Fibre-optic bronchoscopy with bronchoalveolar lavage

Fibreoptic bronchoscopy with bronchoalveolar lavage (BAL) was performed via the oropharyngeal route as per international guidelines [277]. Sixty ml aliquots of buffered warm saline were separately instilled into a segment of the right middle lobe to a total volume of 240 ml and then aspirated by manual suction. Negative bronchoscopy controls were also collected at the time of sampling to control for any potential contamination at the time of bronchoscopy. These samples were collected by aspirating the same buffered saline through the bronchoscope suction channel prior to bronchoscopy and processed in parallel with the samples. Two bronchial brushes (Olympus; Cat #: BC-202D-3010) for cell culture were obtained for each patient in the same location as the bronchoscopy and placed directly in Dulbecco's Modified Eagle Medium (DMEM, Gibco, Cat #: 41966029) transport media. The BAL samples and bronchial brushes were immediately placed on ice in a sealed container and transported to Imperial College London for further processing.

2.1.5 BAL differential cell counts

Post-collection, unfiltered, and unprocessed BAL samples were immediately placed on ice and cells were counted (300 cells counted). BAL cell differentials were performed to assess counts of macrophages, lymphocytes, and neutrophils. The remaining BAL fluid aliquots were frozen at -80°C.

2.1.6 *MUC5B* genotyping

Genotypes of the *MUC5B* SNP *rs35705950* were determined using the Taqman SNP genotyping assay (ThermoFisher, UK). Reactions were carried out in 96-well plates and acquired using the Viia 7 Real-Time PCR System (ThermoFisher, UK).

2.2 Bacterial 16S rRNA Gene Analysis

2.2.1 Reagents for bacterial DNA isolation

Sodium chloride (NaCl), monosodium phosphate (NaH_2PO_4), disodium phosphate (Na_2HPO_4), hexadecyltrimethylammonium bromide (CTAB), polyethylene glycol 6000/NaCl (PEG6000/NaCl), Tris-ethylenediaminetetracetic acid (Tris-EDTA), linear polyacrylamide (LPA, GenElute™), chloroform:isoamyl alcohol (24:1, v/v), phenol:chloroform:isoamyl alcohol (25:24:1, v/v), and 70% ethanol were purchased from Sigma-Aldrich (UK). All solutions were autoclaved prior to use.

2.2.2 Bacterial DNA isolation and precipitation

Two aliquots of BAL (2 ml) were centrifuged at 21,000 *x g* for 30 minutes to pellet cell debris and bacteria. Pellets were resuspended in 100 µl of supernatant and transferred to a lysing matrix E (LME) tube (MP Biomedicals, USA). The remaining supernatants were stored in CryoPure tubes (Sarstedt, UK) at -80 °C for further analysis. An aliquot of 500 µl of CTAB extraction buffer (10% w/v CTAB in 0.5 M phosphate-buffered NaCl) was added to each LME tube and incubated at room temperature for 15 minutes. Then, 500 µl of phenol:chloroform:isoamyl alcohol was added, after which samples were homogenised using a FastPrep®-24 homogeniser (CY: 24 x 2, 5.5 m s^{-1}, 60 seconds, MP Biomedicals, USA) and centrifuged at 16,000 *x g* for 5 minutes. Subsequently, the supernatant was transferred into a 1.5 ml microcentrifuge tube and samples were centrifuged at 2,000 *x g* for 2 minutes. The aqueous phase was recovered, to which an equal volume of chloroform:isoamyl alcohol was added. Samples were mixed by inversion and centrifuged again at 2,000 *x g* for 2 minutes. The aqueous phase was transferred into a fresh microcentrifuge tube containing 1 µl of LPA to which 2 volumes of precipitation solution (30% w/v PEG6000 in NaCl) were added. Samples were left overnight at 4°C to precipitate. The following day, DNA was pelleted (16,000 *x g*, 4 °C, 20 minutes), and the precipitation solution was aspirated. To remove any salts pellets were washed three times with ice-cold 70% ethanol and resuspended in 50 µl of Tris-EDTA.

2.2.3 Bacterial DNA quantification

The quality and quantity of the isolated DNA were measured using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, UK). The ratio of absorbance at 260 nm and 280 nm was used to assess the purity of DNA; a ratio of approximately 1.8 is considered pure for DNA [278]. If this reading is lower it may indicate the presence of contaminants which absorb at that specific wavelength. It may also indicate the presence of a protein or chemicals such as the phenol used in the DNA isolation step. The 260/230 nm ratio is used as an additional measure of nucleic acid purity. A value in the range of 2.0-2.2 is considered pure for DNA. Lower values may indicate the presence

of contaminants such as the elution buffer Tris-EDTA, which absorbs at 230 nm. Following quantification, the DNA was stored at -80°C until further use.

2.2.4 Identification of procedural contaminants

Contamination presents a critical problem when evaluating samples with low microbial biomass, such as the lungs. The low amount of starting material may be swamped by the amplification of reagent-derived contaminants and result in erroneous representations of the microbial community. To determine the level of background microbial contamination negative control samples were collected by aspirating buffered saline through the bronchoscope suction channel before bronchoscopy. Reagent controls including the buffers used in the DNA extraction protocol were also included in the study. Mock community standards were analysed as quality controls in each sequencing run. Samples were randomised prior to processing to avoid batch effect.

2.2.5 16S rRNA sequencing of the V4 region of bacterial 16S rRNA gene

Illumina MiSeq paired-end sequencing of the hypervariable V4 region of the 16S rRNA gene with 515 Forward (5'-GTGCCAGCMGCCGCGGTAA-3') and 806 Reverse (5'-GGACTACHVGGGTWTCTAAT-3') primers was performed at the New York University School of Medicine (USA) [279] **(Figure 2.1)**. The choice of target marker gene region was based on previous studies which have shown that the V4 region generates the most accurate descriptions in low biomass samples.

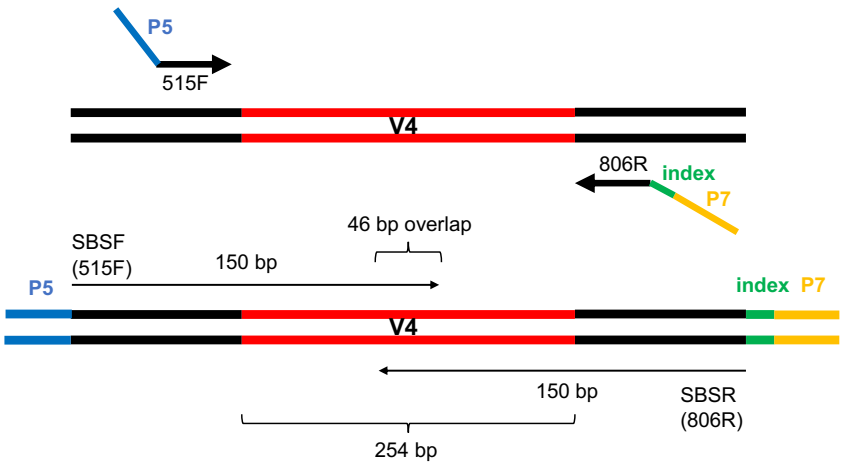


Figure 2.1 Schematic representation of amplification strategy of 16S V4 region.

The V4 region was amplified (150bp read length, paired-end protocol) from each sample using primer set 515F/806R tailed with sequences (P5 and P7) to incorporate Illumina adapters with indexing barcodes [279].

Each unique barcoded amplicon was generated in pairs of 25 ml reactions using the following reaction conditions: 11 ml PCR-grade H_2O, 10 ml Hot MasterMix (Beckman Coulter, Brea, CA), 2 ml of forward and reversed barcoded primer (5 mM) and 2 ml template DNA. Reactions were run on a C1000 Touch Thermal Cycles (Bio-Rad, Hercules, CA) with the following reaction conditions: initial denaturing at 94°C for 3 minutes; 35 cycles of denaturation at 94°C for 45 seconds; annealing at 58°C for 1 minute; extension at 72°C for 90 seconds; and final extension of 10 minutes at 72°C. Amplicons were quantified using Agilent 2200 TapeStation system and pooled. Purification was performed using Ampure XT (Beckman Coulter, Brea, CA) as specified in the manufacturer's instructions. Sequencing was performed using an Illumina MiSeq instrument (Illumina, San Diego, CA) to produce 150 base-paired reads [279].

2.2.6 Bioinformatics analysis

Post-sequencing, bioinformatic analysis of bacterial 16S rRNA amplicon data was conducted using the Quantitative Insights Into Microbial Ecology (QIIME, http://qiime.org; version 2) pipeline [280] running on the Imperial College Codon cluster computer. The code used for each step detailed below is included in **Appendix 1**. The following steps (2.3.1.1-2.3.1.7) were performed:

2.2.6.1 Raw sequence reads

The barcodes, forward and reverse raw sequence reads were stored in the format of paired .fastq files. The standard storage format for raw nucleotide sequences is the .fastq text-based file format which also holds identifying information along with the corresponding Phred quality score [281] for each base. A FASTQ file is divided into 4 lines per sequence; Line 1: begins with '@' followed by the sequence identifier and associated description. Line 2: contains the raw sequence letters. Line 3: starts with '+' and may optionally contain the sequence identifier. Line 4: contains the quality values for the sequence listed in Line 2 and must be of the same length as the sequence. The quality score reflects the probability of sequencing error at that nucleotide position **(Figure 2.2)**. Sequences are available via the National Centre for Biotechnology Information Sequence Read Archive (accession number PRJNA609242).

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATATCGATCCAAATAGTAAATCCATTTCCATTTGTTCAA
+
!''*(((**))%%%++) (%%%%).1***-+*''))**55CCF>>>>>>>CCCCCCC65
```

Figure 2.2 Data output in the FASTQ file format.

Illustrating an example of a FASTQ file containing a single sequence.

2.2.6.2 Metadata

Prior to analysis, a mapping file containing the names and barcode sequences used for each sample, the linker and primer sequence used for amplification and the metadata associated with each sample was generated in excel.

2.2.6.3 Importing paired-end sequence data

The paired-end sequences were imported using the *--qiime tools import* command for importing paired-end sequence data.

2.2.6.4 Demultiplexing and quality control of paired-end sequence data

A unique barcode sequence is appended to the end of each sequence; this allows to determine which barcode sequence is associated with each sample and to subsequently demultiplex the sequences. Paired-end reads from the Illumina MiSeq were demultiplexed using the *--qiime demux emp-paired* command. The barcode reads were the reverse complement of those included in the sample metadata file, so the *--p-rev-comp-mapping-barcodes* parameter was included as an additional command to account for this. Following demultiplexing of the reads, the sequence quality based on ten-thousand randomly selected reads were observed and used to determine what trimming parameters were used for denoising the reads (step v).

2.2.6.5 Denoising, dereplicating and clustering of reads

Denoising of reads was conducted using the Divisive Amplicon Denoising Algorithm 2 (DADA2) method. The *--qiime dada2 denoise-paired* method filters out noisy sequences, corrects errors in marginal sequences and filters the sequences by removing chimeras and singletons [282]. Joining of the denoised paired-end reads happens at this stage following which the sequences are dereplicated. Dereplication of the sequences reduces the repetition and file size requirement in downstream steps. The main outputs are the feature table and the corresponding feature sequences. The features produced are termed amplicon sequence variants (ASVs). The set of query sequences were then aligned against the Greengenes database set of reference sequences using the *vsearch* algorithm [283]. This method aligns feature sequences to a set of reference sequences to identify sequences that hit or miss the reference within a specified per cent identity (97% identity). Open-reference clustering of the feature table was performed at 97% identity against the Greengenes reference database [284].

2.2.6.6 Taxonomic classification and phylogenetic tree construction

Taxonomy was assigned using the *--qiime feature-classifier classify-sklearn* command using the Greengenes reference database [285]. A Newick formatted tree was constructed using FastTree [286], and the *--qiime phylogeny midpoint root* command was utilised to construct a rooted tree for downstream diversity analyses.

2.2.6.7 Exporting feature table, phylogenetic tree, taxonomy and metadata

The feature table constructed for downstream analysis along with the representative sequences, rooted phylogenetic tree and mapping file (containing the metadata for each sample) were imported into R (version 3.6.1) and combined using the *phyloseq* package [287]. Details on further analyses carried out are detailed in **section 2.2.7**.

2.2.7 Diversity and statistical analyses

All analyses were performed in the *R* software environment (version 3.6.1) [288]. The number of different taxa present (richness) and the proportions of taxa (evenness) are used to describe a microbial community [289] (**Figure 2.3**). These factors determine the diversity of a specific community. The diversity within a single sample is referred to as alpha diversity (α-diversity).

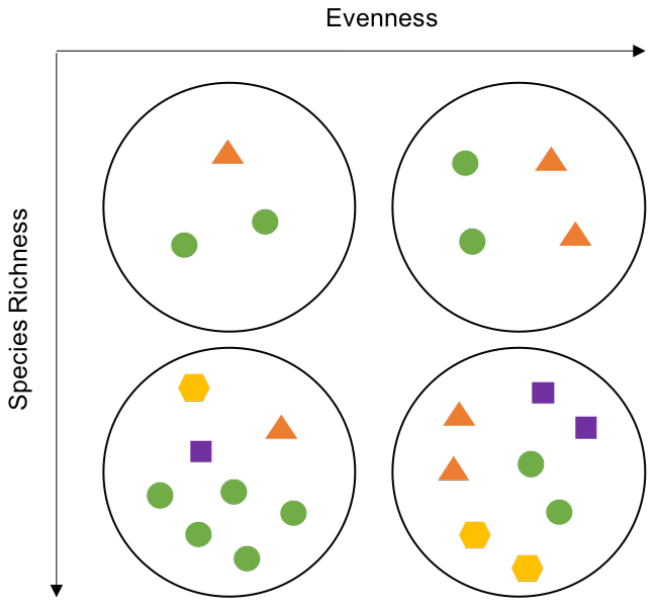


Figure 2.3 Community diversity is measured according to species richness and evenness.

The communities are represented by the four circles and within them are the different taxa illustrated by the shapes and their distinct colours. Adapted from Cox [289]. Copyright © Human Molecular Genetics.

Community structure can be characterised using a variety of metrics. For the purpose of **Chapter 3** and **Chapter 6**, the following α-diversity metrics were calculated using the *phyloseq* package [287]:

- Shannon's index: a measure of entropy which accounts for both the abundance and evenness of the taxa present in a sample [290].
- Chao1 index: estimates diversity from abundant taxa as well as the number of rare taxa missed from undersampling [291].

Beta diversity (β-diversity) refers to the level of dissimilarity between communities within a given group [292]. In order to visualise how different dissimilar communities were to one another, principal component analysis (PCA) was used. PCA was performed using the *phyloseq* package using the Euclidean distance as the measure of the distance between communities. Statistical differences in community between groups were determined by permutational multivariate analysis of variance (PERMANOVA) on a Euclidean distance matrix using the *adonis* function in the vegan package. This test compares the variance within-groups and between-groups and uses permutation to test for significance rather than assuming a normal distribution.

2.2.8 Differential abundance analysis

Differences in taxonomic abundance between groups were tested using the Mann-Whitney U test for comparisons between two groups, the Kruskal-Wallis test for comparisons between more than two groups. Spearman's rank correlation was used for the measurement of correlation with continuous variables. Differences in taxa abundance were restricted to ASV that were present at greater than 1% of the sampled population. A *P*-value of 0.05 was considered statistically significant. Differentially abundant bacteria within each group were identified using linear discriminant analysis coupled with effect size (LEfSE) [293]. Only taxa meeting a linear discriminant analysis threshold > 2.0 were plotted.

2.2.9 Summary of microbiome workflow

An outline of the 16S workflow is illustrated in **Figure 2.4** and can be divided into upstream and downstream stages.

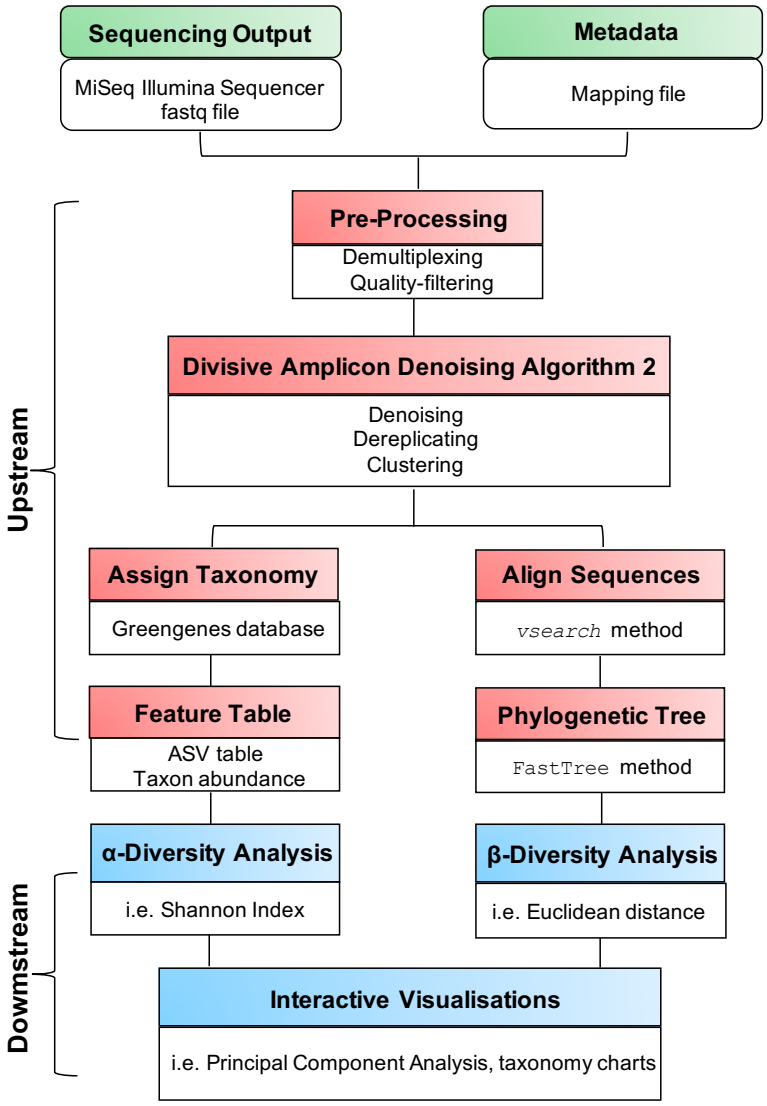


Figure 2.4 Microbiome workflow analysis.

The upstream analysis includes the processing of the sequencing output and generating the feature table and phylogenetic tree for microbial analysis. The downstream analysis uses the feature table and phylogenetic tree to perform diversity analyses, statistics and interactive visualisations of the data.

2.2.10 Statistical analysis

Continuous variables are presented as means (± standard deviation) and categorical variables as proportions. Data analysis was performed using *R* (version 3.4.3) [294]. Differences in BAL differential cell counts, bacterial burden and taxonomic composition between healthy controls and subjects with IPF were statistically tested by Mann-Whitney U test. Differences in bacterial burden according to *MUC5B* genotype were tested using Kruskal-Wallis test with Dunn's multiple comparison test. Spearman's *rho* was used to calculate correlations between continuous variables. Assessment for co-linearity of predictor variables was made using correlation coefficients. Differentially abundant bacteria between the two groups were assessed using LEfSe. Features

significantly discriminating among the groups with a linear discriminant analysis (LDA) score > 2.0 were represented as a cladogram, as produced by LEfSe with default parameters. Changes in microbial biodiversity and richness between healthy controls and IPF subjects were evaluated using the Shannon and Chao1 α-diversity indexes and tested by Mann-Whitney U test. β-diversity of healthy controls and IPF subjects was restricted to ASVs that were present at greater than 1% of the sampled population, calculated using the Euclidean distance, visualised by PCA and significance of differences in community composition determined by PERMANOVA adjusted for age, sex, smoking status, FVC (% predicted) and DL_{CO} (% predicted). To assess survival probability in IPF subjects Kaplan-Meier curves were constructed and assessed using a multivariate Cox proportional hazard model, adjusted for age, sex, smoking status and FVC (% predicted). Data were log-transformed where appropriate. Transformed bacterial burden data was examined using univariate and stepwise multivariate linear regression models against computed tomography (CT) features and cell differentials. In all analyses, a *P* value < 0.05 was considered statistically significant.

2.3 16S Gene Quantitative Polymerase Chain Reaction (qPCR)

2.3.1 Amplification of *Vibrio natriegens DSM 759*

Vibrio natriegens DSM 759 strain DNA was amplified using 16S 27 Forward and 1492 Reverse primers [295]. Triplicate PCR reaction were set up containing 1 µl of Forward primer, 1 µl Reverse primer (both 10 µM stock concentration), 1 µl of template, 0.5 µl dNTP mix (50 mM), 1 µl BSA (Sigma B8667 20 mg/ml), 6.5 µl Glycine betaine (1.3 M), 2.5 µl FastStart 10X Buffer, 0.25 µl FastStart HiFidelity Polymerase (5 U/µl), and 11.25 µl PCR-grade water to a final reaction volume of 25 µl per reaction. A negative control containing 1 µl of PCR-grade water rather than template was also included. Cycling conditions were: one cycle at 95°C for 2 minutes, 30 cycles at 94°C for 20 seconds and 72°C for 5 minutes.

2.3.2 Gel extraction of *Vibrio natriegens DSM 759* amplicons

The products were visualised by electrophoresis on a 1% agarose gel (1g agarose gel per 100 ml 1X Tris/Borate/EDTA stained with 1 µl of GelRed stain to confirm amplification and verify the negative PCR control. Amplicons were excised and purified using the QIAquick® Gel Extraction Kit according to manufacturer's instructions. All centrifugation steps were performed at 17,900 *x g* (Prism™ Microcentrifuge, UK). The column purified PCR product was eluted in 50 µl of final elution buffer.

2.3.3 Amplicon cloning

Amplicons were cloned using the TOPO® TA Cloning® kit (Invitrogen) as follows: a cloning reaction containing 4 μl of PCR product, 1 μl of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), and 1 μl of the TOPO® Vector were combined to a final reaction volume of 6 μl and incubated at room temperature for 5 minutes. One vial of One Shot® *Escherichia coli* (*E.coli*) cells was thawed on ice after which 2 μl of the cloning reaction was added, mixed and incubated on ice for 30 minutes. Subsequently, the cells were heat-shocked for 30 seconds in a water bath at 42°C and immediately transferred on ice. An aliquot of 250 μl of SOC medium was added to the tube and incubated in a shaking incubator (New Brunswick™ I26/126R Incubator, Eppendorf, UK) at 2000 *x g* at 37°C for 1 hour. Volumes of 5 μl, 5 μl and 20 μl were plated onto three pre-warmed Luria-Bertani (LB) agar plates containing X-Gal and 100 μg/ml ampicillin and incubated overnight at 37°C. Following overnight culture, ten white colonies were picked and cultured overnight at 37°C in LB culture medium containing 100 μg/ml of ampicillin.

2.3.4 Plasmid DNA purification

Plasmid DNA was isolated from the transformed and cultured cells using the QIAprep® Spin Miniprep Kit (QIAGEN) as per manufacturer's instructions. An aliquot of 2 ml of the bacterial overnight culture was pelleted by centrifugation at 6800 *x g* for 3 minutes at room temperature. The pelleted bacterial cells were resuspended in 250 μl Buffer P1 and transferred to a microcentrifuge tube. An aliquot of 250 μl Buffer P2 was added, and the contents were mixed thoroughly by inverting the tube 5 times. A volume of 350 μl Buffer N3 was added and contents mixed by inverting the tube a further 5 times. The contents were centrifuged for 10 minutes at 17,900 *x g*. An aliquot of 800 μl supernatant was added to the spin column (QIAprep 2.0) by gently pipetting and centrifuged for 60 seconds, following which the flow-through was discarded. The spin column was washed by adding 0.5 ml Buffer PB and the flow-through discarded after centrifuging for 60 seconds. The spin column was washed an additional time by adding 0.75 ml Buffer PE and the flow-through discarded after centrifuging for 60 seconds. Following this, the spin column was centrifuged for 1 minute to remove residual wash buffer and placed in a clean 1.5 ml microcentrifuge tube. To elute DNA, 50 μl Buffer EB (10 mM Tris Cl, pH 8.5) was pipetted to the centre of the spin column, left to stand for 1 minute and centrifuged for an additional minute. The quantity and quality of plasmid DNA were measured using the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, UK). The ten plasmids yielded high yields of pure DNA **(Table 2.2)**.

Table 2.2 Nanodrop results of the purified plasmids.

Plasmid	ng/ul	A260/280	A260/230
1	342	1.87	2.01
2	311	1.86	1.89
3	302	1.88	2.06
4	315	1.86	1.90
5	215	1.86	1.80
6	326	1.87	2.03
7	342	1.89	2.29
8	280	1.85	1.91
9	273	1.87	2.10
10	293	1.86	1.98

Purified amplicons were sent externally for Sanger sequencing (Eurofins Scientific, UK). Using the recommended external Sp6 Promoter (F) and T7 promoter (R) for sequencing produces a large product of at least 1542 bp (**Figure 2.5**). The aligned sequences had a length of 1492 bp and the consensus sequence aligned exactly to *Vibrio natriegens DSM 759*.

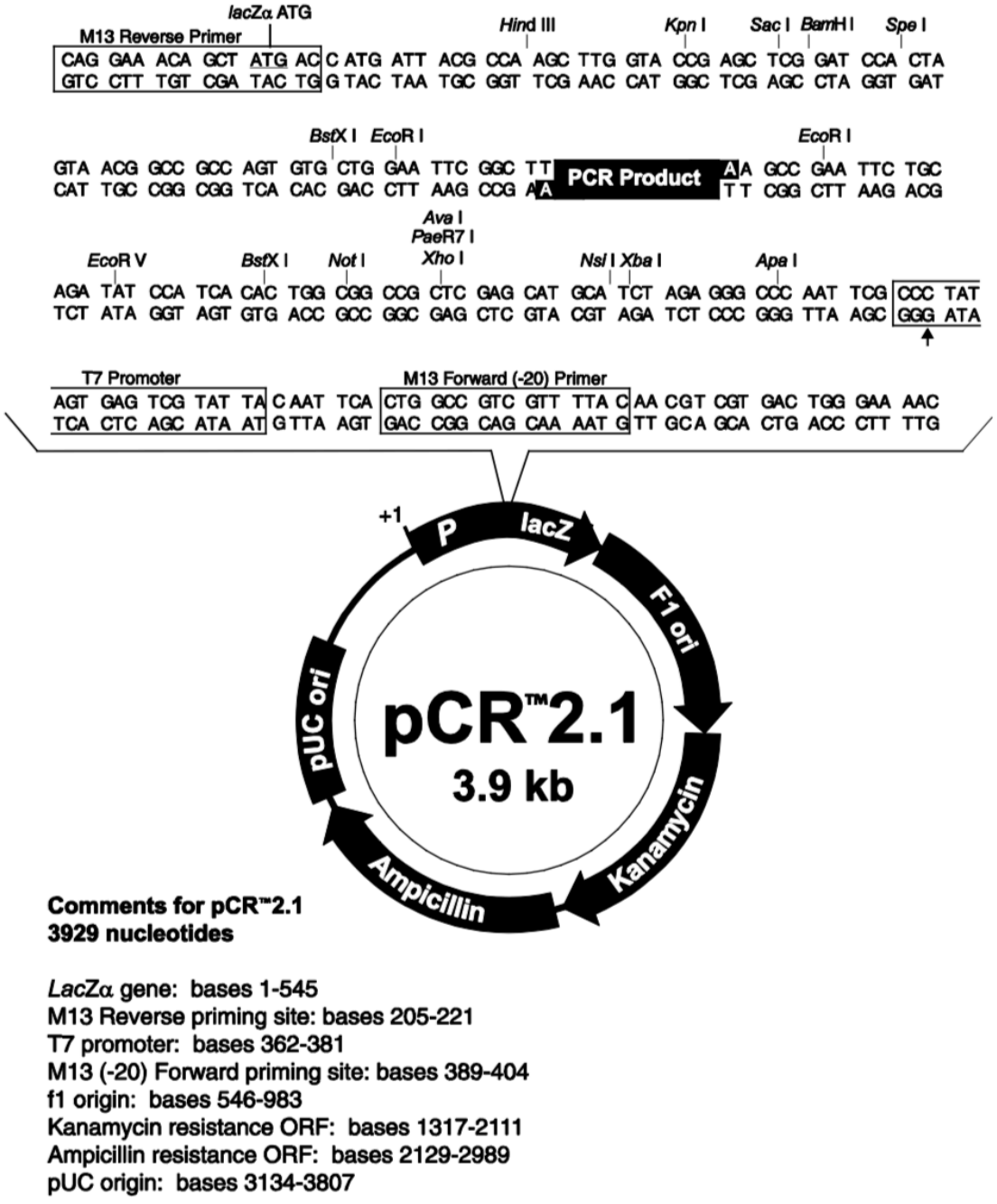


Figure 2.5 Map of the linearized vector, pCR™2.1-TOPO.

Illustrating the start of transcription for the T7 RNA polymerase (↑), restriction- and primer sites and the insertion site of the PCR product. Obtained from the TA Cloning® Kit (Invitrogen) user guide.

The purified PCR product was diluted down to a working concentration based on the NanoDrop readings (**Table 2.3**). To calculate the number of molecules/µl the following calculation was used:

$$\text{Molecules}/\mu\text{l} = \frac{273.1 \text{ x } 6.022 \text{ x } 10^{23}}{656.6 \text{ x } 10^{9} \text{ x } 5396} = 4.64 \text{ x } 10^{10} \text{ mol}/\mu\text{l}$$

- Concentration of plasmid DNA (**Table 2.2**; Plasmid #9): 273.1 ng/µl.
- Avogadro constant: 6.022 x 10^{23}.
- Average weight of a single DNA base pair: 656.6 Da. Multiplied by 10^9 to convert to ng.
- Length of plasmid and insert in base pair: 3931 + (1492-27) = 5396 base pairs.

A 500 µl 10^9 stock of plasmid DNA was calculated as follows:

$$\frac{4.64 \text{ x } 10^{10} \text{ mol}/\mu\text{l}}{1\text{x}10^{9} \text{ mol}/\mu\text{l}} = 46.4 \text{ df}$$

$$\frac{500 \ \mu\text{l}}{46.4 \text{ df}} = 10.77 \ \mu\text{l}$$

$$500 \ \mu\text{l} - 10.77 \ \mu\text{l} = 489.23 \ \mu\text{l}$$

A new 10^9 stock containing 489.23 µl PCR-grade water and 10.77 µl plasmid DNA was prepared and stored at -80°C until further analysis. From here, two serial dilutions (1:10) were performed to produce a stock solution for the standard curve **(Table 2.3)**. Serial dilutions of the standard curve were prepared fresh for each qPCR run.

Table 2.3 Dilutions of working stock used to construct the qPCR standard curve.

Standard	DNA (µl)	Dilution in H_2O* (µl)	Molecules/µl
1	10 (10^9 stock DNA)	90	1 x 10^8
2	10 (dilution)	90	1 x 10^7
3	10 (dilution)	90	1 x 10^6
4	10 (dilution)	90	1 x 10^5
5	10 (dilution)	90	1 x 10^4
6	10 (dilution)	90	1 x 10^3
7	10 (dilution)	90	1 x 10^2

*PCR-grade H_2O (Qiagen, Cat #: 17000).

2.3.5 Quantitative 16S rRNA gene PCR

To assess the quantity of 16S bacteria present in each sample qPCR was performed and acquired using the Viia 7 Real-Time PCR System (ThermoFisher, UK). Reactions were performed in triplicate on a 96-well plate. These were set up to obtain a final reaction volume of 10 µl. Each reaction contained 1 µl of bacterial DNA and 9 µl of Femto™ bacterial qPCR premix (Cambridge bioscience, UK). Each run contained a 10-fold dilution series run in duplicate of the *Vibrio natriegens DSM 759* 16S rRNA gene cloned into a plasmid of known size and a non-template control (PCR-grade water, H_2O). For data acquisition, the following cycling parameters were used: 1 cycle of 95°C for 10 minutes; 40 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute; and 1 cycle of 72°C for 7 minutes. Following the final extension step, a melting curve was acquired at 95°C for 15 seconds, 60°C for 1 minute and 95°C for 15 seconds. The standard curve samples were used to extrapolate the total 16S rRNA copy number from cycle threshold values for the BAL samples.

2.4 ^{1}H Nuclear Magnetic Resonance Spectroscopy

2.4.1 Reagents for ^{1}H NMR

Potassium dihydrogen phosphate (KH_2PO_4, Cat #: P3786), potassium hydroxide (KOH, Cat #: P5958), sodium azide (NaN_3, Cat #: 08591), 3-(trimethyl -silyl)-[2,2,3,3,-2H4]-propionic acid (TSP, Cat #: 269913) and NMR-grade deuterium oxide (D_2O, Cat #: 151882) were purchased from Sigma-Aldrich (UK).

2.4.2 D_2O buffer preparation

A total of 10.2 g KH_2PO_4 were dissolved in 40 ml D_2O. Separately, 50 mg of the internal standard, TSP, and 6.5 mg NaN_3 were dissolved in 3 ml D_2O. The solutions were combined, mixed with a magnetic spinner and topped up with 7 ml D_2O to obtain a final volume of 50 ml. The pH was adjusted to 7.4 with KOH tablets and checked with a pH meter (Mettler Toledo, UK).

2.4.3 Sample collection and preparation

BAL was obtained from patients undergoing bronchoscopy, as previously described in **section 2.1.4**. Samples stored at –80°C were thawed on ice, vortexed and centrifuged at 15,000 x *g* for 30 minutes at 4°C to pellet debris. A total of 540 µl of BAL supernatant was transferred into an Eppendorf to which 60 µl D_2O buffer was added. For each sample, an aliquot of 550 µl of BAL supernatant was transferred to a 5 mm outer diameter NMR tube (Bruker, UK, Cat #: Z112273) for analysis.

2.4.4 One-dimensional ^{1}H NMR spectroscopy

One-dimensional (1D) ^{1}H NMR spectroscopy experiments with water pre-saturation enable the acquisition of the resonance of all ^{1}H-containing molecules in a given sample including signals from small molecules and broader signals from large molecules [296]. 1D NMR spectroscopy has been utilised throughout this work and involves two steps: pulse and acquire. A single pulse sequence employed in 1D NMR experiments generates a free induction decay response which can be acquired. In order to maximise the signal to noise ratio, the pulse sequence can be repeated several times. Subsequently, the background signal to noise ratio is reduced. Lack of sufficient relaxation will prevent the signal intensity to fully recover prior to the subsequent pulse and can result in a less accurate quantitation of the analytes. Therefore, to minimise this, a delay time is applied between consecutive pulses to ensure that complete relaxation occurs.

2.4.5 Nuclear Overhauser Effect Spectroscopy (NOESY) ^{1}H Nuclear Magnetic Resonance (NMR) spectral acquisition of bronchoalveolar lavage (BAL) supernatant samples

^{1}H NMR spectroscopy was performed at 300 K (26.85 °C) using a TXI probe on a Bruker 600 MHz spectrometer (Bruker Biospin, Karlsruhe, Germany). All samples were acquired using the 1D NOESY-pulse sequence, and BAL spectra were referenced to TSP (δ 0.0). The 'NOESY presat' pulse sequence is used for the acquisition of spectra from samples which are characterised by a poor macromolecular content [296]. Samples were measured in 5 mm outer diameter NMR tubes. Typical acquisition parameters were: spectral width: 20 ppm; time-domain points: 65536; acquisition time: 2.7 s; number of scans: 32; number of dummy scans: 4 and line broadening: 0.3 Hz.

2.4.6 Spectral pre-processing

Pre-treatment of raw spectral data is pivotal to generate reliable, interpretable models and to draw accurate comparisons between datasets. Following the spectral acquisition, the frequency has to be adjusted to the internal reference standard, TSP (δ 0.0). This ensures that the chemical shifts of the metabolites are identical across all acquired spectra. TSP absorbs energy at a low frequency and therefore its resonance does not overlap with the resonances of other metabolites in the sample of interest. Molecules are characterised by their chemical shift relative to the reference compound in an experiment. Additional spectral corrections include phasing of the spectra to provide adequate quality peak shape and accurate peak integration as well as baseline corrections which impede the false identification and quantification of NMR resonances [297]. Spectral regions that have no real diagnostic information, such as the 4.7-4.9 ppm region, which includes the water resonance are removed. Therefore, all ^{1}H NMR spectra were manually phased, baseline-corrected and referenced to the TSP singlet (δ 0.0) using Topspin software (version 3.5; Bruker BioSpin GmbH, USA). The spectral resonance of each metabolite was assigned on the basis of the 600-MHz library in Chenomx NMR Suite software (Chenomx Inc., Canada) and using in-house assignment databases. Spectra were imported using an in-house MATLAB (version R2014a, The MathWorks, Inc.; USA) script. Peaks arising from incomplete water suppression were removed (4.7-4.9 ppm) as well as those resonances arising from TSP (δ 0.0).

2.4.7 Statistical pre-processing

Statistical pre-processing is a key step in obtaining a robust and accurate statistical model. Projection-based methods such as PCA are sensitive to the scaling of the variables, implying that variables with the highest value range will dominate and potentially skew the model [298]. In order to adjust for this, prior to conducting the multivariate data analysis, the following methods were

utilised: normalisation, mean centring and scaling to unit variance. These methods accentuate the biological information in the dataset and subsequently improve its interpretability [299]. Normalisation is a mathematical correction which is applied to the dataset in order to minimise the variation induced by biological or technical sources which are unrelated to the effect of interest and may therefore influence the data. This includes factors such as variation in BAL return volume following bronchoscopy. Normalisation scales spectra to the same virtual overall concentration in order to account for the different dilutions of the samples [300]. For the purpose of **Chapter 4**, all BAL spectra were normalised using the probabilistic quotient normalisation method which calculates the most probable dilution by looking at the distribution of the quotients of the amplitudes of a test spectrum with those of a reference spectrum [300]. Following normalisation, mean centring was applied to improve the interpretability of the model. The principle of mean centring is to calculate the mean of each variable and then to subtract this value from the dataset, therefore removing the mean trajectory from the dataset [298]. The resulting data points will have a common reference point based on the spread around the mean [301] **(Figure 2.6)**.

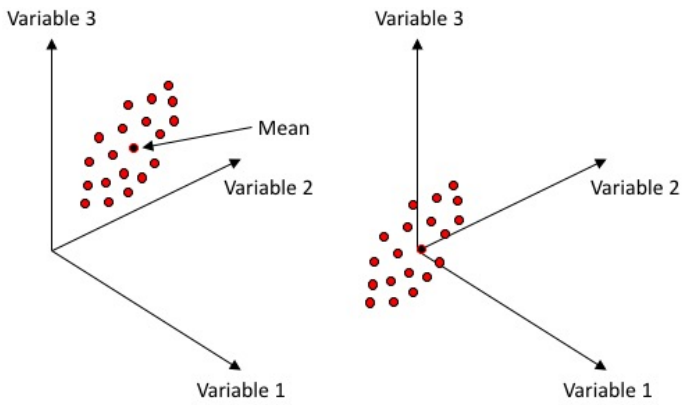


Figure 2.6 Illustrating the effect of mean centring on the dataset.

Adapted from [301].

As such, the first principal component will not be dominated by the mean of each variable but will represent the covariance of all of the data points. In order to develop a robust model, the BAL NMR spectra were scaled to assign equal weight to both low and high concentration metabolites. Unit variance scaling is the most common method used to achieve this, and it consists of dividing the value of each variable by its standard deviation. This ensures that all variables have equal variances and no variable dominates over any other [302].

2.4.8 Multivariate data analysis

Multivariate statistical analysis facilitates data identification and analysis by converting the peaks of the NMR spectra into more easily interpretable coordinates in multidimensional plots of metabolic activity in order to differentiate between groups within the data. For the purpose of **Chapter 4**, data were subjected to multivariate analyses using unsupervised PCA. All multivariate statistical modelling was performed in *R* (version 3.4.3). PCA is a multivariate statistical technique which requires no *a priori* information to construct a model [303]. The work originates from a data matrix which consists of rows and columns that represent the observations (samples) and the variables (metabolites), respectively **(Figure 2.7)**.

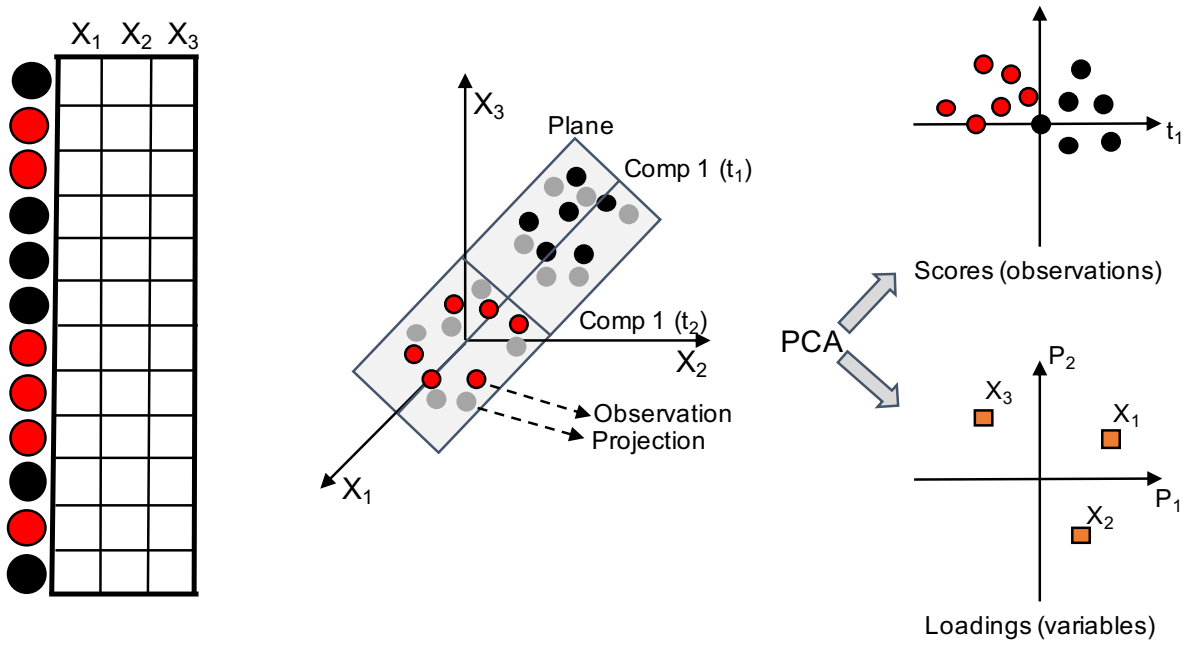


Figure 2.7 Schematic breakdown of principal component analysis (PCA).

The data matrix consists of three variables (X1, X2, X3) and 12 samples which are placed in a lower-dimensional plane. The scores (observations) plot facilitates the visualisation of the relation amongst the different sample groups in the model plane. The loadings (variables) plot illustrates the influence of each of the variables in the model and the relation between them. Adapted from [303].

PCA is an unsupervised pattern recognition method, which reduces the number of observed variables to a smaller number of principal components (PCs), which account for most of the variance of the observed variables [303]. Principal component 1 (PC1), accounts for most of the variance in the data, followed by principal component 2 (PC2), which is uncorrelated to PC1 and accounts for the second-largest amount of variance in the data. The scores plots facilitate the visualisation of the different sample groups whilst the loadings plots explain the contribution of each of the variables in the model. Furthermore, outliers can be detected using the Hotelling's T^2 test, which is a statistical approach to identify outliers outside a 95% confidence interval **(Figure 2.8)** [304].

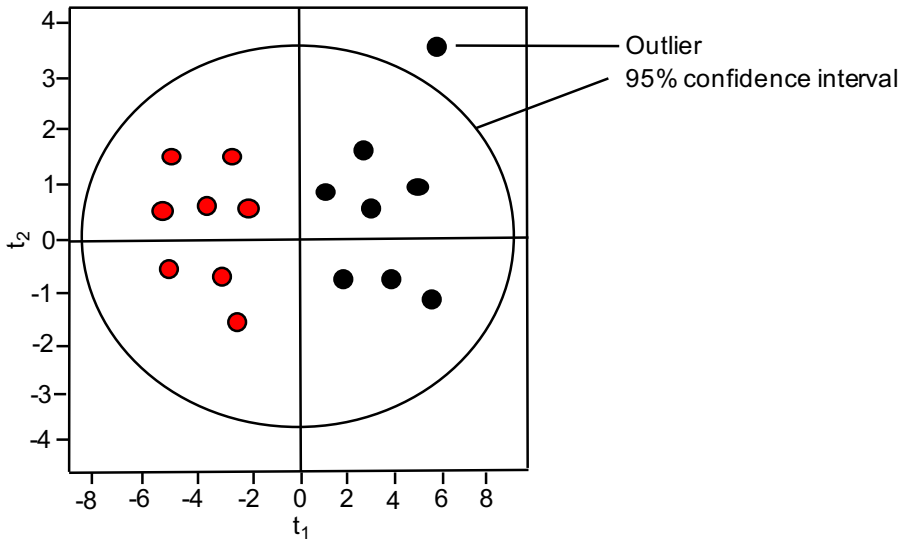


Figure 2.8 Detection of outliers using Hotelling's T2 test.

Illustrating a PCA scores plot with outliers demarcated by the Hotelling's T2 test. The ellipse on the scores plot delineates the 95% confidence interval.

Associations between the top microbial genera in each group, identified in **Chapter 3**, and the metabolites was performed using the Data Integration Analysis for Biomarker discovery using Latent variable approaches for Omics studies (DIABLO) bioinformatics pipeline in the mixOmics R package [305, 306]. The mixOmics block.splsda function, with full weighted design and 10 components, was used to identify the optimal number of components, which was defined in 3 methods using the centroid distance technique. To decide which variables to keep in each component models with 10,5,5 and 5 randomly selected variables were tested for the microbiome and metabolomics datasets, respectively. Different model features were obtained and the results were plotted using mixOmics predefined and ad-hoc functions.

2.4.9 Summary of ^{1}H NMR spectroscopy workflow

The overall strategy employed for metabonomic analysis of the BAL samples is displayed in **Figure 2.9**.

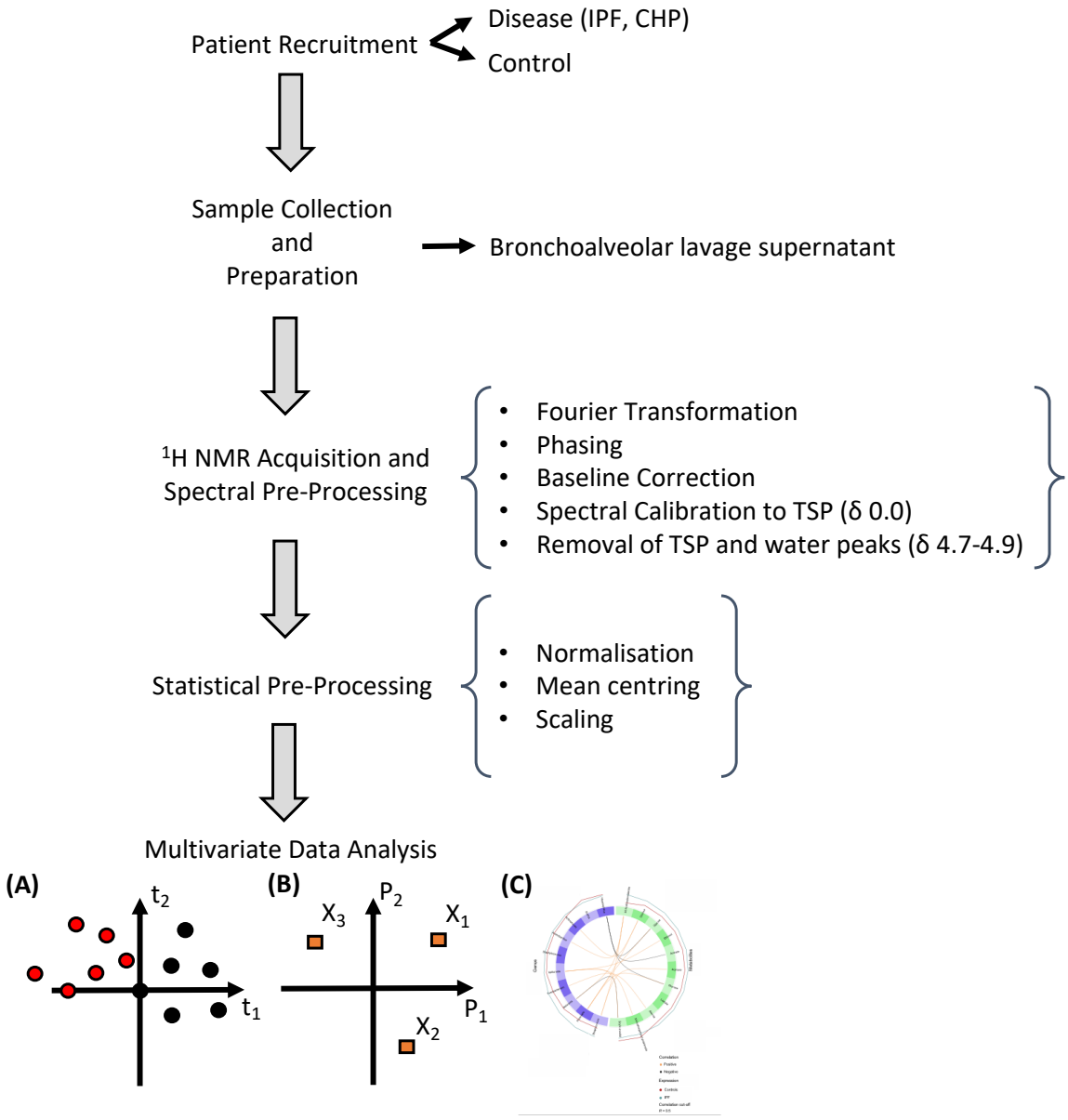


Figure 2.9 Overall strategy for metabonomic analysis.

Patients were recruited, and BAL samples were collected and analysed by conducting ^{1}H NMR. The datasets were pre-processed prior to multivariate statistical analysis. **(A)** PCA elucidated the biochemical variation in the dataset, the **(B)** loading plots were used to reveal the underlying metabolic differences driving the separation between the groups, and **(C)** the microbiome and metabolome datasets were integrated to elucidate potential associations.

2.5 Triple Quadrupole Gas Chromatography-Mass (GC-MS) Spectrometry

2.5.1 Chemicals and reagents

Methyl *tert*-butyl ether (MTBE, Cat #: 306975), *N*-*tert*-Butyldimethylsilyl-*N*-methylfluoroacetamide with 1% *tert*-Butyldimethylchlorosilane (MTBSTFA + 1% t-BDMSCI, Cat #: 00942), water for HPLC (H_2O, Cat #: 34877), hydrochloric acid (HCl, Cat #: 435570) and the internal standard methyl stearate (Cat #: 85769) were purchased from Sigma-Aldrich (UK). Methanol (CH_3OH) CHROMASOLV™ LC-MS Ultra (Cat #: 34966) was obtained from Honeywell (Germany). The following short-chain fatty acids (SCFAs) were all purchased from Sigma-Aldrich (UK): acetic acid (Cat #: A6283), propionic acid (Cat #: 94425), butyric acid (Cat #: BI03500), valeric acid (Cat #: 240370), isovaleric acid (Cat #: 129542), hexanoic acid (Cat #: 153745), isobutyric acid (Cat #: 1745), lactic acid (Cat #: L1750), 2-methyl-butyric acid (Cat #: 193070), and 2-hydroxybutyric acid (Cat #: 220116). All experiments were conducted in a fume hood.

2.5.2 Methyl *tert*-butyl ether preparation

A total of 0.070 g of internal standard (methyl stearate) was weighed on an electronic analytical balance (Sartorius, UK) and dissolved in 70 ml of MTBE.

2.5.3 Calibration curve preparation

A calibration curve was set up for the following SCFAs: acetic acid, propionic acid, butyric acid, valeric acid, isovaleric acid, hexanoic acid, isobutyric acid, lactic acid, 2-methyl-butyric acid, and 2-hydroxybutyric acid. Each compound was prepared in a total volume of 1 ml. The quantity added for each compound is listed in **Table 2.4**. Once each compound was carefully diluted to the specified volume a mother stock containing all of the compounds was prepared in one Eppendorf tube as described in **Table 2.5**. From the mother stock, the following dilutions were prepared in fresh Eppendorf tubes **(Table 2.6)**. From each series, 12 new Eppendorf tubes were labelled (1-12) and 5 µl of each stock and 95 µl of H_2O were aliquoted. A water sample was included as a blank. An aliquot of 1000 µl MTBE containing the internal standard (methyl stearate) prepared in section 2.6.2 and 2 µl HCl were added into each Eppendorf then vortexed and incubated for 1 h at 4°C. Following incubation, the samples were centrifuged at 1000 *x g* for 15 min at 4°C. The polar phase (90 µl) of each sample (1-12) was transferred into silanised vials into which 120 µl of MTBSTFA + 1% t-BDMSCI was also added. The samples were vortexed and incubated for 45 min at 60°C. Samples were allowed to cool in the fume hood for 20 min, following which 70 µl were transferred into non-silanised vials containing glass inserts ready for acquisition on the triple quadrupole mass spectrometer. The silanised vials were stored overnight in the fridge at 4°C.

Table 2.4 Illustrating the quantity added for each compound to set up the calibration curve.

Compound	mM	Volume (L)	MW	mmol/ ml	Density (g/ml)	Compound (mg)	(µl)	H_2O (µl)
Acetic acid	320	0.001	60.05	0.32	1.049	NA	18.3	981.7
Propionic acid	80	0.001	74.08	0.08	0.993	NA	6.0	994.0
Butyric acid	20	0.001	88.11	0.02	0.964	NA	1.8	998.2
Valeric acid	16	0.001	102.13	0.016	0.939	NA	1.7	998.3
Isovaleric acid	16	0.001	102.13	0.016	0.925	NA	1.8	998.2
Hexanoic acid	10	0.001	116.16	0.01	0.927	NA	1.3	998.7
Isobutyric acid	20	0.001	88.11	0.02	0.95	NA	1.9	998.1
Lactic acid	8000	0.00025	90.08	8	NA	180.2	NA	69.8
2-methyl-butyric acid	20	0.001	102.13	0.02	0.936	NA	2.2	997.8
2-hydroxybutyric acid	800	0.00025	104.1	0.8	NA	20.8	NA	229.2

Table 2.5 Illustrating the mother stock containing specific volumes of the respective compounds.

Compound	Stock (µM)	Dilution	Mother stock volume pipetted (µl)
Acetic acid	16000	20	50
Propionic acid	1600	50	20
Butyric acid	400	50	20
Valeric acid	80	200	5
Isovaleric acid	800	20	50
Hexanoic acid	200	50	20
Isobutyric acid	400	50	20
Lactic acid	800000	10	100
2-methyl-butyric acid	400	50	20
2-hydroxybutyric acid	40000	20	50
Total volume from each compound dilution (µl)			**355**
Total volume H_2O (µl)			**645**
Total volume of mother stock (µl)			**1000**

Table 2.6 Illustrating the dilution series performed to obtain the calibration curve.

Dilution series				Mother stock (µl)	H_2O (µl)
Series 1	Used **mother stock** as series 1				
Series 2	1.6	1.6		100	60
Series 3	1.25	2		100	100
Series 4	1.6	3.2	From Series 1	100	220
Series 5	1.5625	5		100	400
Series 6	2	10		100	900
Series 7	2	2		100	100
Series 8	2	4	From Series 6	100	300
Series 9	2.5	10		100	900
Series 10	2	2.5		100	150
Series 11	2	5	From Series 9	100	400
Series 12	2	10		100	900

2.5.4 Triple quadrupole mass spectrometer setup

GC-MS analysis was performed using the 7000C Agilent Triple Quadrupole GC/MS system (Agilent Technologies, Santa Clara, CA, USA) with a HB-5 ms capillary column (30 m x 0.25 mm x 0.25 µm). The injector, ion source, quadrupole, and the GC-MS interface temperature were 280°C, 150°C, 150°C, and 280°C, respectively. Each chromatogram from the calibration curve points were checked according to its retention time **(Table 2.7)**, transitions and qualifiers ions **(Table 2.8)**.

Table 2.7 Retention time for the respective short-chain fatty acids (SCFAs).

SCFA	RT (min)
Acetic acid	2.5
Isobutyric acid	3.2
2-methyl-butyric acid	3.7
Valeric acid	4.2
Lactic acid	6.14
Methyl stearate	7.88 (IS)
Propionic acid	3.07
Butyric acid	3.4
Isovaleric acid	3.8
Hexanoic acid	4.9
2-hydroxybutyric acid	6.4

RT: retention time; IS: internal standard

Table 2.8 Illustrating the quantifiers and qualifier ions for each compound.

Compound	Quantifier		Qualifier	
	Precursor Ion	Product Ion	Precursor Ion	Product Ion
Acetic acid	117	75	117	47
Propionic acid	131	75	131	47
Isobutyric acid	145	75	145	43
Butyric acid	145	75	145	43
2-methyl-butyric acid	159	75	159	57
Isovaleric acid	159	75	159	57
Valeric acid	159	75	159	57
Hexanoic acid	173	75	173	81
Lactic acid	147	73	147	45
2-hydroxybutyric acid	147	73	147	45
Methyl stearate (IS)	87	59	87	55

IS: internal standard.

2.5.5 BAL supernatant sample preparation

The BAL supernatant samples were thawed at 4°C and vortexed. A total of 100 μl of each sample was aliquoted into 2 ml Eppendorf tubes. Quality controls (QCs) were prepared by pooling all of the BAL supernatant samples into a single Eppendorf. Blanks containing 100 μl H_2O were prepared. The BAL supernatant samples, QCs and blanks were processed and run following the steps detailed in sections 2.6.3-2.6.4 starting with the addition of 1000 μl MTBE containing the internal standard (methyl stearate) and 2 μl HCl.

2.5.6 Statistical analysis

Data analysis was performed using the Agilent MassHunter Quantitative Analysis software version B.08. All data were expressed as mean ± SD (standard deviation). Differences with *P*-values ≤ 0.05 between groups was considered statistically significant.

2.6 Enzyme-linked immunosorbent assay

2.6.1 General ELISA protocol

The concentrations of several cytokines and mediators were assessed by commercially available enzyme-linked immunosorbent assay (ELISA) kits. The concentrations of the capture, standard and detection antibodies used are detailed in **Table 2.9.**

Table 2.9 Illustrating commercially available enzyme-linked immunosorbent assay (ELISA) kits utilised.

ELISA	Kit	Antibody type	Standard curve range (pg/ml)	Manufacturer (Catalogue #)
IL-6	Human IL-6	Capture (1:120) Standard (1:300) Detection (1:60)	9.38-600	R & D Systems (DY206-05)
TNF-α	Human TNF-α	Capture (1:120) Standard (1:130) Detection (1:60)	15.6-1000	R & D Systems (DY210-05)
IL-8	Human IL-8	Capture (1:120) Standard (1:100) Detection (1:60)	31.2-2000	R & D Systems (DY208-05)
Collagen	Human pro-collagen I α1	Capture (1:120) Standard (1:100) Detection (1:60)	31.2-2000	R & D Systems (DY6220-05)
Fibronectin	Human Fibronectin	Capture (1:120) Standard (1:100) Detection (1:60)	31.2-2000	R & D Systems (DY1918-05)
TGF-β	Human TGF-β	Capture (1:120) Standard (1:100) Detection (1:60)	31.2-2000	R & D Systems (DY240-05)

ELISA: enzyme-linked immunosorbent assay; IL-6: interleukin 6; TNF-α: tumour necrosis factor-alpha; IL-8: interleukin 8; TGF-β: transforming growth factor-beta.

High binding affinity 96-well half area plates were coated with respective purified capture antibodies in 1X PBS (VWR International, UK) and left overnight at 4°C. The following day, the plates were washed 3 times with 0.05% Tween-20 (Sigma-Aldrich, UK) in PBS wash buffer, after which they were blocked with 1% Bovine serum albumin (Sigma-Aldrich, UK) in PBS for 1 hour at room temperature. The plates were washed 3 times, and the standards and samples were loaded into the respective wells of the plate and incubated overnight at 4°C. The standards consisted of 8 serially diluted standard curve points which were loaded in duplicates. The plates were washed 3 times, and the respective biotinylated detection antibodies added into each well and incubated at room temperature for 2 hours. The plates were washed a further 3 times prior to the addition of streptavidin-horseradish peroxidase diluted 1:40 in 1% BSA in PBS and incubated at room temperature for 20 minutes. The plates were subsequently washed 3 a further three times before the addition of 3,3',5,5'-Tetramethybenzidine (Sigma-Aldrich, Cat #: T4444) to each well. The reaction was allowed to develop for 20 minutes at room temperature and was subsequently stopped by the addition of 20 μl of H_2SO_4. The absorbance values were recorded on a SpectraMax® i3x plate reader (UK) at a wavelength of 450 nm (reference wavelength: 540 nm), and cytokine concentrations were determined using SoftmaxPro software (version 6.5).

2.7 Human Cell Culture

2.7.1 Primary human bronchial epithelial cell (PBEC) culture

The bronchial brushes were obtained from patients undergoing bronchoscopy at the Royal Brompton Hospital, as described in **section 2.1.4**. These were agitated in media to detach the cells and centrifuged at 315 *x g* for 5 minutes to pellet the cells. The cell pellet was resuspended in Airway Epithelial Growth Medium (PromoCell, Cat #: 505-B) containing all components of the PromoCell SupplementPack (PromoCell, Cat #: C-39160): Bovine Pituitary Extract (BPE 0.004 ml/ml, human recombinant Epidermal Growth Factor (hrEGF 10 ng/ml), human recombinant insulin (hrinsulin 5 μg/ml), hydrocortisone (0.5 μg/ml), epinephrine (0.5 μg/ml), triiodo-L-thyronine (6.7 ng/ml), transferrin 10 (μg/ml), retinoic acid (ng/ml). Primocin 100 μg/ml (InvivoGen, Cat #: ant-pm-1) was added to the PromoCell media as an antimicrobial agent. Cells were cultured for expansion in a 25 cm^3 flask (Corning®, Cat #: 3055) and kept in an incubator (37°C, 5% CO_2). Media was refreshed every 48 hours, and cells were passaged at 80% confluence (passage 1) using TrypLE™ Express Enzyme (ThermoFisher Scientific, Cat #: 12604013) into two 75 cm^3 flasks (Corning®, Cat #: 3276). Cells were cryopreserved in liquid nitrogen at passage 2 in a solution containing 80% PromoCell media, 10% dimethyl sulfoxide (DMSO) (Sigma-Aldrich, Cat #: D2650) and 10% fetal bovine serum

(FBS) (Sigma-Aldrich, Cat #: 12103C). Cells were used later for experimentation at passage 3. Cell counting was performed with trypan blue (Sigma-Aldrich, Cat #: T8154).

2.7.2 Submerged culture of PBECs

PBECs obtained from IPF patients, and healthy controls (passage 2) were cultured for expansion in a 75 cm^3 flask (Corning, Cat #: 3276) using PromoCell media (supplemented as previously described in **section 2.7.1**). Once 80% confluent, the cells were passaged using TrypLE™ Express Enzyme (ThermoFisher Scientific, Cat #: 12604013) and seeded in a collagen-coated 96-well G (Corning® Costar®, Cat #: CLS3595) at a density of 10,000 cells per well. The cells were either treated with media or treated with the following doses (1 mM, 5 mM, 10 mM, or 20 mM) of the SCFAs acetate (Sigma-Aldrich, Cat #: S2889), propionate (Sigma-Aldrich, Cat #: P1880) and butyrate (Sigma-Aldrich, Cat #: 303410) and left in an incubator (37°C, 5% CO_2) until the respective time points (4h, 8h, 24h, 48h). At each time-point, the cell supernatant was collected, and IL-6 and TNF-α were quantified by ELISA as described in **section 2.6.1**. Following the collection of the cell supernatant, the cells were stained with crystal violet (Sigma-Aldrich, Cat #: C0775) to assess cell viability **(section 2.7.2.1)** and proliferation was assessed over 72 hours using the JuLI Stage Real-Time Cell History Recorder (NanoEnTek Inc.) **(section 2.7.2.2)**.

2.7.2.1 Crystal violet assay

Cytotoxic effects were investigated using the crystal violet assay (Sigma-Aldrich, Cat #: C0775) and expressed as % release of the positive controls treated with 0.01% (v/v) TRITON-X100 (Sigma-Aldrich, Cat #: 2315025).

2.7.2.2 Proliferation assay

PBECs from IPF and healthy control subjects were cultured for expansion in a 75-cm^3 flask (Corning), coated with type 1 collagen (30 µg/ml; Sigma-Aldrich), in PromoCell Airway Epithelial Growth Medium containing all components of the PromoCell SupplementPack. Once more than 80% confluence was achieved, the cells were removed with trypsin (TrypLE Express Enzyme, Gibco) and plated onto 96-well plates at a density of 5000 cells per well. The cells were supplemented with 200 µl PromoCell or PromoCell containing the SCFAs of interest (acetate, propionate or butyrate) at the respective doses (1 mM and 5 mM). The cells were placed in an incubator for 3 hours for acclimatisation and to avoid potential condensation from forming on the lid. Snapshots of each well were recorded at 30-min intervals over 72 hours using the JuLI Stage Real-Time Cell History Recorder (NanoEnTek Inc.). Percentage cell confluence was automatically recorded for every snapshot of each respective well. At the end of the experiment, the cells were fixed in 10% neutral buffered formalin

(Sigma-Aldrich) for subsequent analysis. Ki-67 staining was performed using a rat anti-mouse Ki-67 primary antibody diluted 1:500 (eBioscience). A Cy3-conjugated goat anti-rat antibody (Jackson ImmunoResearch) diluted 1:500 was subsequently added to enumerate Ki-67 positively stained cells. The cells were then counterstained with a DAPI dye (Sigma-Aldrich) diluted 1:1000. The number of DAPI-positive cells and the number of DAPI and Ki-67 double-positive cells were counted to determine the percentage of Ki-67 cells per well.

2.7.2.3 Wound healing assay

PBECs from IPF and healthy control subjects were cultured for expansion in a 75-cm^3 flask (Corning), coated with type 1 collagen (30 μg/ml; Sigma-Aldrich), in PromoCell Airway Epithelial Growth Medium containing all components of the PromoCell SupplementPack. Once more than 80% confluence was achieved, the cells were removed with trypsin (TrypLE Express Enzyme, Gibco) and plated onto collagen-coated 24-well plates at a density of 10,000 cells per well. The cells were grown to confluence and were supplemented with 500 μl PromoCell or PromoCell containing the SCFAs of interest (acetate, propionate or butyrate) at the respective doses (1 mM and 5 mM). A plastic pipette tip (P1000) was used to deliver a single, uniform, linear wound in each well and the cells were placed in an incubator for 3 hours for acclimatisation and to avoid potential condensation from forming on the lid. A JuLI Stage Real-Time Cell History Recorder (NanoEnTek Inc.) was used to capture time-lapse brightfield images of the wound every 30 minutes for 48-hours. Wound closure data was generated by the JuLi Stage software from these images.

2.7.3 Air-liquid interface culture

PBECs were cultured for expansion in a 75 cm^3 flask (Corning, Cat #: 3276) using PromoCell Airway Epithelial Cell Growth Medium (StemCell, Cat #: 505-B). Once 80% confluent, the cells were passaged using TrypLE™ Express Enzyme (ThermoFisher Scientific, Cat #: 12604013) and seeded into inserts (12 inserts/plate) into a 24-well plate at a density of 30,000 cells per insert. The cells were fed apically (50 μl) and basolaterally (500 μl) for 2 weeks, every two days with PneumaCult-air-liquid interface (ALI) Medium (StemCell, Cat #: 05021). After two weeks, the cells were air exposed by removing the apical media, and the cells were continued to be fed basolaterally as specified above. Experiments were performed 2 weeks after cilia were first observed. A total of 50 μl of the SCFAs (acetate, propionate and butyrate) were added apically at a concentration of 1 mM, 5 mM, 10 mM, 20 mM or media control. Apical supernatant was collected at each time-point (0 h, 24 h, 48 h, 96 h, and 144 h) by washing the apical side of the transwell with 100 μl PBS and stored at -80°C until subsequent analysis. Following the collection of the apical supernatant at each

respective time-point; the SCFAs were added again as specified above. Brightfield images of the cells were taken at each time-point with a LeicaDM2500 microscope (Leica Microsystems, UK). A summary of the steps involved in the culturing of the PBECs to differentiate at ALI is illustrated in **Figure 2.10**.

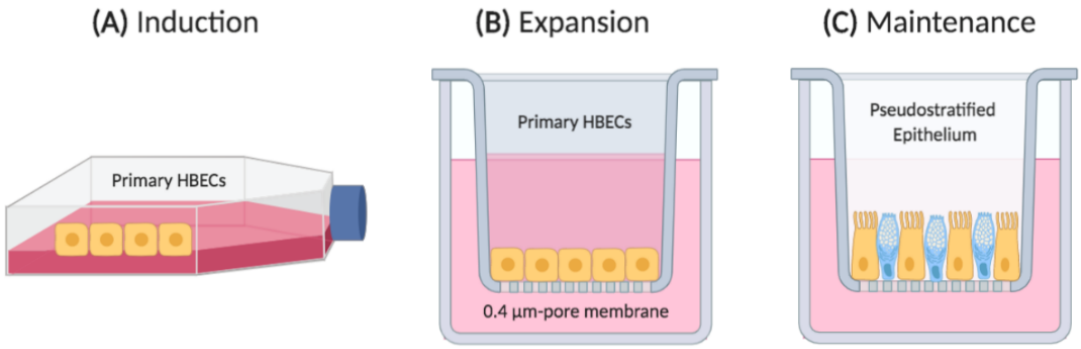


Figure 2.10 Air-liquid interface (ALI) cell culture schematic.

Illustrating the **(A)** induction, **(B)** expansion, and **(C)** maintenance steps involved in the culture of primary human bronchial epithelial cells (PBECs).

2.7.3.1 Lactate dehydrogenase assay

Cytotoxic effects were investigated using the lactate dehydrogenase (LDH) detection kit (LDH detection kit, Roche, ver. 10) by assessment of LDH in basal media and expressed as % release of the positive controls treated with 0.01% (v/v) TRITON-X100 (Sigma-Aldrich, Cat #: 2315025.

2.7.3.2 Trans-epithelial electrical resistance

TEER was measured using an electrometer EVOM2 (World Precision Instruments, Sarasota, FL). Ohm values were subsequently multiplied by the surface of the Transwell inserts (0.33 cm^2) to obtain the unit area resistance which was expressed as Ohm*cm2.

2.7.4 Isolation and culture of primary human lung fibroblasts

Lung tissue specimens were obtained from the Royal Brompton Hospital (REC 15/SC/0101) during cancer resections or lung-volume resection surgery for COPD. HLF were extracted from the tissue using an established protocol [307]. In a 6 well-plate four linear, parallel scratches were made with a scalpel on the surface of each well to encourage fibroblast outgrowth. Lung tissue was cut into 1mm^3 pieces and evenly distributed into each well (5 pieces/well). Explant cultures were maintained in DMEM (supplemented with 20% FBS and Primocin) in an incubator (37°C, 5% CO_2). Media was refreshed every 48 hours. Primary HLFs were passaged at 80% confluence and cryopreserved in DMEM with 10% DMSO and 10% FBS (passage 2).

2.7.5 Submerged culture of primary HLFs

HLF (passage 2) from IPF patients were cultured for expansion in a 75 cm^3 flask (Corning, Cat #: 3276) using DMEM media (supplemented with 20% FBS and Primocin). Once 80% confluent, the cells were passaged using TrypLE Express Enzyme (Gibco, Cat #: 12604013) and seeded in a 96-well plate (Corning, Cat #: CLS3595) at a density of 5,000 cells per well. The cells were either treated with 200 µl DMEM media or DMEM containing the SCFAs of interest (acetate, propionate or butyrate) at the respective doses (1 mM, 5 mM, 10 mM, 20 mM) and left in an incubator (37°C, 5% CO_2) until the respective timepoints (4h, 8h, 24h, 48h). At each time-point, the cell supernatant was collected and collagen I α1, fibronectin and TGF-β were quantified by ELISA as described in **section 2.6.1**. Following the collection of the cell supernatant the cells were stained with crystal violet (Sigma-Aldrich, Cat #: C0775) to assess cell viability **(section 2.7.5.1)** and proliferation was assessed over 72 hours using the JuLI Stage Real-Time Cell History Recorder (NanoEnTek Inc.) **(section 2.7.5.2)**.

2.7.5.1 Crystal violet assay

Cytotoxic effects were investigated using the crystal violet assay (Sigma-Aldrich, Cat #: C0775) and expressed as % release of the positive controls treated with 0.01% (v/v) TRITON-X100 (Sigma-Aldrich, Cat #: 2315025).

2.7.5.2 Proliferation assay

HLF from IPF subjects were cultured for expansion in a 75-cm^3 flask (Corning) in DMEM media (supplemented with Primocin). Once more than 80% confluence was achieved, the cells were removed with TrypLE Express Enzyme and plated onto 96-well plate at a density of 5000 cells per well. The cells were supplemented with 200 µl DMEM or DMEM containing the SCFAs of interest (acetate, propionate or butyrate) at the respective doses (1 mM, 5 mM, 10 mM, 20 mM). The cells were placed in an incubator for 3 hours for acclimatisation and to avoid potential condensation from forming on the lid. Snapshots of each well were recorded at 30-min intervals over 72 hours using the JuLI Stage Real-Time Cell History Recorder (NanoEnTek Inc.). Percentage cell confluence was automatically recorded for every snapshot of each respective well. At the end of the experiment, the cells were fixed in 10% neutral buffered formalin (Sigma-Aldrich) for subsequent analysis. Ki-67 staining was performed using a rat anti-mouse Ki-67 primary antibody diluted 1:500 (eBioscience). A Cy3-conjugated goat anti-rat antibody (ImmunoResearch) diluted 1:500 was subsequently added to enumerate Ki-67 positively stained cells. The cells were then counterstained with a DAPI dye (Sigma-Aldrich) diluted 1:1000. The number of DAPI-positive cells and the number of DAPI and Ki-67 double-positive cells were counted to determine the percentage of Ki-67 cells per well.

3 **Chapter 3:** The Respiratory Microbiome in Idiopathic Pulmonary Fibrosis

3.1 Introduction

As discussed previously in **Chapter 1**, IPF is a complex, multifactorial disease. Although the factors which initiate the fibrotic process remain elusive, susceptibility to infection may be contributing to disease progression [85]. Studies have identified a number of genetic variants, common and rare, which may increase the risk of both sporadic and familial forms of IPF, influence phenotype and predict disease outcome. These include SNPs in genes involved in host defence (e.g. *MUC5B*), telomere maintenance (e.g. *TERT*), and epithelial cell integrity (e.g. *DSP*) [41, 42]. These findings suggest that loss of normal pulmonary physiology and aberrant innate immunity due to genetic alterations may play an aetiological role in IPF, and support the hypothesis that infectious agents may act a source of injury to the alveolar epithelium, driving disease initiation and progression.

Occult viral infections, in particular, have long been suspected of playing a role in IPF pathogenesis by acting as persistent antigenic stimuli or co-factors, triggering aberrant activation of fibrotic processes in physiologically predisposed hosts. While a number of studies have highlighted a link between viruses and IPF initiation and progression as well as acute exacerbations, studies exploring the role of bacteria in this context are much scarcer [87, 88]. Indeed, there is growing evidence, albeit associative, to suggest that the respiratory microbiome may be linked to the development and progression of IPF [114, 115, 144]. Although the lungs have conventionally been considered sterile, the recent revolution in culture-independent technologies has prompted a reconsideration of IPF pathogenesis [110, 308]. Studies of the respiratory microbiome in IPF have reported associations between disturbances in the microbial communities, disease progression and mortality [114, 115, 309, 310]. Specifically, a higher bacterial burden has been consistently associated with an increased risk of mortality and differences in bacterial composition and diversity have also been correlated with disease outcomes [114, 115, 149]. Recently, animal models have been utilised to better understand the mechanisms by which perturbations in the microbial communities in the lower airways can drive fibrosis [149]. It has been shown that even at diagnosis and in the absence of infection there is a dramatic increase in bacterial load with an altered composition of the respiratory microbiome in the lower airways of patients with IPF [144]. However, whether bacterial burden directly influences the progression of fibrosis or if bacterial numbers instead reflect other underlying disease processes, remains poorly understood. Previous studies have accounted for disease severity using physiological parameters such FVC and DL_{CO}. To date, no study has yet assessed the relationship between bacterial burden and radiological markers of fibrotic distortion

of the lung or disease extent or severity. The overarching aim of this Chapter was to characterise the microbiome and quantify the bacterial load in the BAL of subjects diagnosed with IPF and healthy controls. The hypotheses examined were the following: (i) the bacterial burden and respiratory microbiome of IPF subjects are distinct compared to healthy controls; (ii) in IPF, bacterial burden and specific bacterial genera correlate with disease progression; (iii) in IPF, if the bacterial burden is elevated secondary to the underlying architectural destruction of the lung parenchyma, then there might be a clear association between radiographic findings such as honeycombing or traction bronchiectasis and bacterial burden in the lower airways of individuals with IPF.

3.2 Aims

- To quantify bacterial burden in the BAL of healthy controls and IPF subjects.
- To investigate how bacterial burden relates to baseline disease severity/progression, BAL differential cell counts, *MUC5B* genotype, and radiological disease extent.
- To characterise the taxonomic changes, differentially abundant bacteria, alpha- and beta-diversity differences in the BAL of healthy controls and IPF subjects.
- To investigate whether specific changes in bacterial genera in IPF correlate with disease progression.
- To assess whether there are changes in the respiratory microbiome in IPF according to baseline lung function, smoking history and *MUC5B* genotype.

3.3 Materials and Methods

3.3.1 Study design and sampling

Patients with IPF undergoing bronchoscopy with BAL were recruited at the Royal Brompton Hospital (London, UK). Healthy controls were also recruited at this time and included smokers and non-smokers with normal lung function. Further details on subject recruitment, exclusion and sample collection are detailed in **Chapter 2 (sections 2.1.1-2.1.4)**.

3.3.2 Measurement of BAL cells

Post collection, unfiltered and unprocessed BAL samples were immediately placed on ice and cells were counted. BAL differentials were performed as described in **Chapter 2 (section 2.1.5)** to assess counts of macrophages, lymphocytes, and neutrophils.

3.3.3 *MUC5B* genotyping

Genotypes of the *MUC5B* SNP *rs35705950* were determined using the Taqman SNP genotyping assay (ThermoFisher, UK) as described in **Chapter 2 (section 2.1.6)**.

3.3.4 Radiology

The CT variable definitions were based on the Fleischner Glossary of terms and current ATS/ERA/JRS/ALAT guideline definitions [13, 311]. CT data were scored by two independent radiologists using the following scales: to the nearest 5% per lobe (fibrosis, honeycombing), on a binary basis [UIP], on a lobar extent with the lingua defined as the 6th lobe [PPFE] and on an 18-point scale [traction bronchiectasis] [312, 313].

3.3.5 Bacterial DNA isolation, quantification and 16S rRNA gene sequencing

BAL samples (2 ml) were centrifuged at 21,000 x *g* for 30 minutes to pellet cell debris and bacteria, and genomic DNA was then isolated. High-throughput sequencing of bacterial 16S rRNA gene amplicons encoding the V4 region (150 bp read length, paired-end protocol) was performed using an Illumina MiSeq Sequencer (Illumina, San Diego, CA). Further details are provided in **Chapter 2 (sections 2.2.1-2.2.5)**.

3.3.6 Bioinformatics analysis

The 16S rRNA gene sequences were analysed using the QIIME2 pipeline for analysis of microbiome data. Details on bioinformatics analysis are provided in **Chapter 2 (section 2.2.6)**.

3.3.7 16S rRNA gene qPCR

Bacterial burden was measured by 16S rRNA gene qPCR as previously described in **Chapter 2 (section 2.3)**.

3.3.8 Statistical analysis

Continuous variables are presented as means (± standard deviation) and categorical variables as proportions. Data analysis was performed using *R* (version 3.4.3) [294]. Details of statistical analysis performed are described in **Chapter 2 (section 2.2.10)**.

3.4 Results

3.4.1 Study subjects and sampling

Two hundred and one subjects with IPF and 40 healthy control subjects were enrolled in the study and underwent bronchoscopy **(Table 3.1)**.

Table 3.1 Clinical characteristics and demographics of the study subjects.

Parameter	Controls (*N* = 40)	IPF (*N* = 201)	*P*-value*
Age, years (± SD)	60 (14)	74 (9)	0.0001
Male, *n* (%)	27 (67)	142 (71)	NS
Baseline PFT			
FVC, % predicted (± SD)	92 (30)	78 (18)	0.0001
FEV_1 % predicted (± SD)	81 (23)	82 (18)	NS
DL_{CO} % predicted (± SD)	78 (35)	42 (20)	0.0001
Smoking status			
Non-smoker, *n* (%)	17 (42)	74 (37)	-
Ex-smoker, *n* (%)	21 (52)	117 (58)	NS
Current-smoker, *n* (%)	2 (5)	10 (5)	-
BAL cell differentials			
Macrophages % (± SD)	74 (16)	71 (12)	NS
Neutrophils % (± SD)	7 (7)	12 (8)	<0.05
Lymphocytes % (± SD)	8 (7)	11 (6)	<0.01
Radiology			
UIP probable/definite (%)	-	58	-
Honeycombing (%)	-	30	-
PPFE (%)	-	21	-

Data are presented as percentage value or mean ± standard deviation (SD) as appropriate. IPF: idiopathic pulmonary fibrosis; PFT: pulmonary function test; FEV_1: forced expiratory volume in 1 second; FVC: forced vital capacity; DL_{CO}: diffusing capacity of the lung for carbon monoxide; BAL: bronchoalveolar lavage; UIP: usual interstitial pneumonia; honeycombing (presence/absence); PPFE: pleuroparenchymal fibroelastosis (presence/absence); NS: not significant. Smoking status: statistical testing by non-smoker versus ex/current-smoker. UIP defined as 'definite' or 'probable' UIP. **P*-values are calculated by Mann-Whitney U test or Fisher's exact test between controls and IPF subjects.

Individuals with IPF were predominantly male (71%), with a mean age of 74 years ± 9, and had moderately severe disease at enrolment (FVC, 78% predicted; DL_{CO}, 42% predicted). The 40 control subjects were also predominantly men (67%), had better lung function compared to the IPF cohort (DL_{CO}, 78% predicted; FVC, 92% predicted) and had a mean age of 60 years ± 14. Subjects were matched for sex and smoking status. Differential cell counts were performed in the control and IPF groups with identification of alveolar macrophages, lymphocytes and neutrophils. Alveolar macrophages predominated the BAL of both groups (controls 74% ± 16; IPF 71% ± 12). Mixed alveolitis with predominant neutrophils count was observed in IPF (neutrophils 12% ± 8, lymphocytes 11% ± 6). Out of the 201 individuals receiving a diagnosis of IPF, a total of 141 subjects had CT imaging performed within 4 months of bronchoscopy. Over half of the cohort had probable or definite UIP (57.7%) according to the current ATS/ERA/JRS/ALAT guidelines [13]. Twenty-one per cent (30/141) of subjects had radiological evidence of PPFE, and 30% had honeycombing (43/141) **(Table 3.1)**.

3.4.2 Survival analysis

3.4.2.1 Disease progression

Lung function parameters are used to define disease progression, assessing change in the FVC or DL_{CO} at 6 or 12 months [17]. The decrease in FVC% has been consistently shown to be the parameter of lung function that best predicts mortality [17]. A decline greater than or equal to 10% of FVC or greater than or equal to 15% of DL_{CO} at 12 months has been extensively used as cut-off points to define disease progression and used as established clinical trial endpoints [8]. Lung function data were available for 165 subjects out of the 201 IPF subjects recruited. Thirty subjects with IPF had a decrease in FVC of more than 10% in 12 months and were at a higher risk of mortality compared with subjects with no decline in lung function (hazard ratio: 2.47, confidence interval: 1.29-4.7, P = 0.006) **(Figure 3.1A)**. Sixty-three subjects with IPF had a decrease in DL_{CO} of more than 15% in 12 months and were at an increased risk of mortality compared with subjects with no decline in lung function (hazard ratio: 2.08, confidence interval: 1.11-3.9, P = 0.02) **(Figure 3.1B)**.

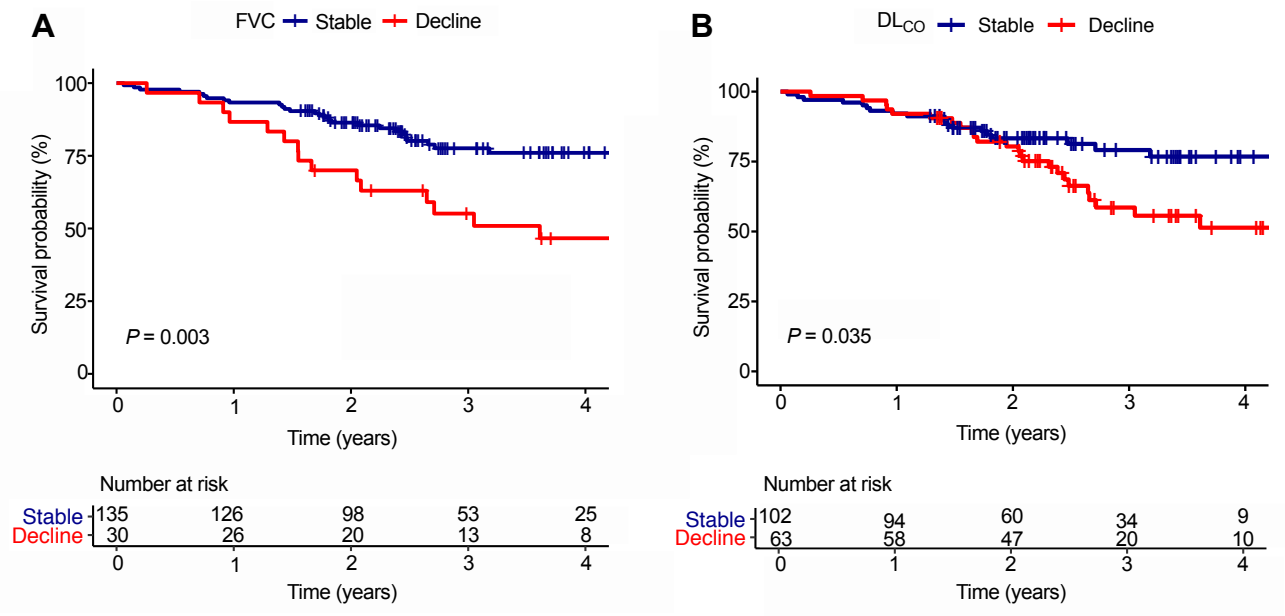


Figure 3.1 Kaplan-Meier curve demonstrating survival probability in idiopathic pulmonary fibrosis (IPF) subjects with stable or progressive disease.

Kaplan-Meier curve generated by Cox proportional-hazards model displaying the survival probability (%) of IPF subjects with **(A)** stable disease or a 10 % decline in forced vital capacity (FVC) or with **(B)** stable disease or a 15% decline in carbon monoxide diffusing capacity (DL_{CO}). Log-rank *P* test value reported.

3.4.2.2 MUC5B genotype

A common polymorphism in the promoter of the *MUC5B* gene has been consistently found to be associated with an increased risk of developing both familial and sporadic IPF in an allele dose-dependent manner [41, 84]. Specifically, the minor allele of this gene, *rs35705950*, confers an increased risk for IPF. Out of the 201 subjects with IPF, genotyping of the *MUC5B* gene was possible for 175 patients. The number of IPF patients in the GG, GT, and TT genotype groups were 90 (51.4%), 67 (38.2%), and 18 (10.2%), respectively. The median follow-up period was 907 (± 902 days) days. During follow-up, there were 36 deaths (12 GG, 22 GT, and 2 TT). In a multivariable Cox proportional hazards model adjusted for age, sex, smoking status and baseline % predicted FVC the TT and GT genotypes (risk for IPF) were associated with improved survival compared with GG (TT: hazard ratio: 2.23, confidence interval: 0.008-0.594, *P* = 0.02; GT: hazard ratio: 1.05, confidence interval: 0.010-0.481, *P* = 0.04) **(Figure 3.2)**. No differences in *MUC5B* genotype were observed when looking at age, FVC, sex or smoking history **(Figure 3.3)**.

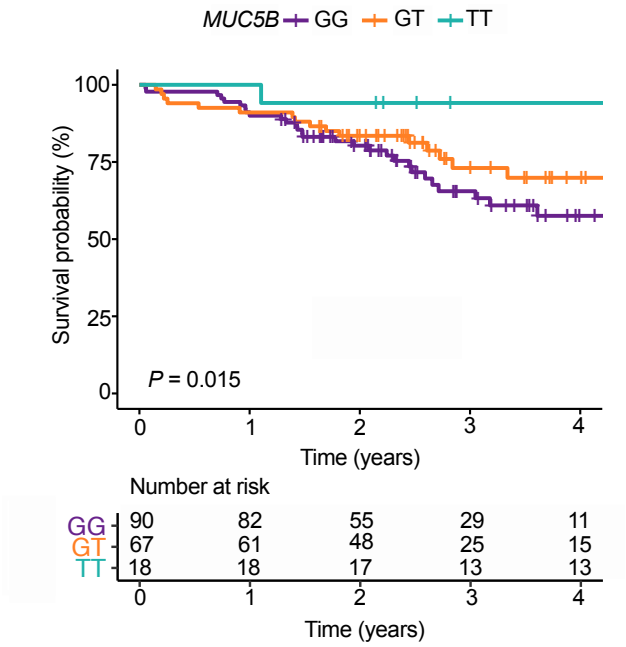


Figure 3.2 Survival probability of subjects with idiopathic pulmonary fibrosis (IPF) stratified according to *MUC5B* genotype.

Kaplan Meier curve generated by Cox proportional-hazards model adjusted for age, sex, smoking status and baseline % predicted forced vital capacity for IPF subjects stratified according to *MUC5B* genotype (GG, GT and TT). Log-rank *P* test value reported.

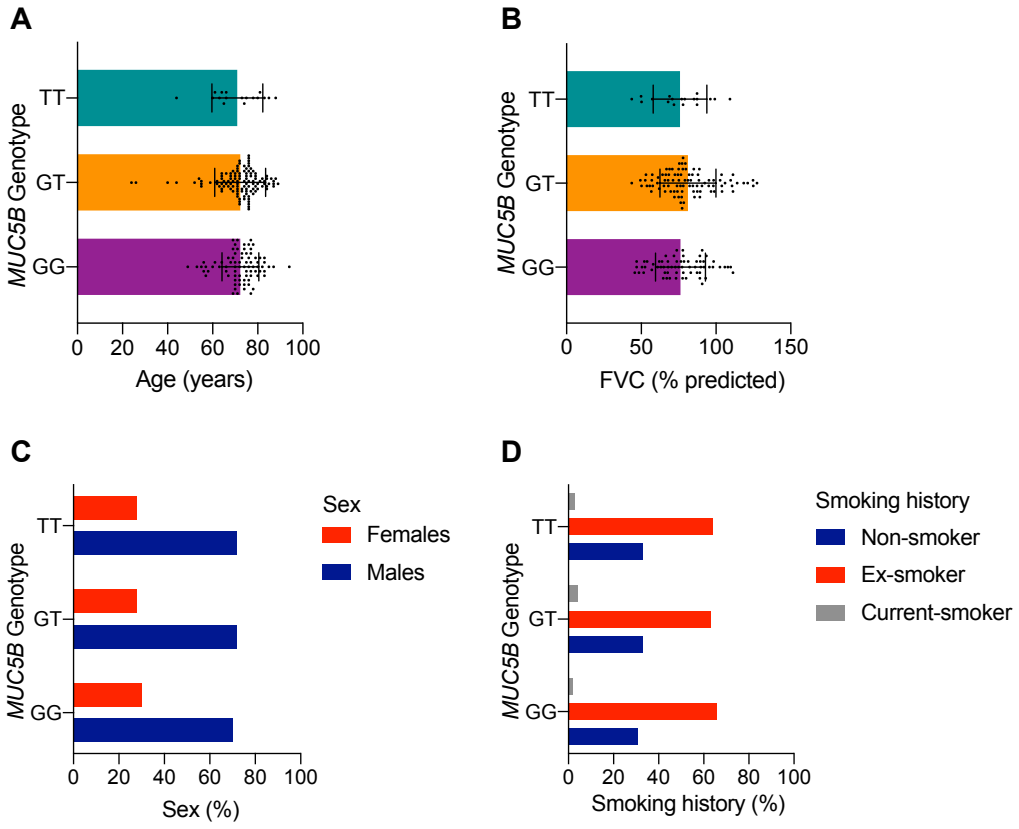


Figure 3.3 Clinical characteristics stratified according to *MUC5B* genotype.

MUC5B genotype distribution according to **(A)** age, **(B)** FVC, (**C**) sex, and **(D)** smoking history.

3.4.3 Sampling

3.4.3.1 DNA yield and sequencing quality

All 241 samples yielded genomic DNA and underwent amplification of the V4 region of the 16S rRNA gene using the barcoded universal primer pair 515F/806R on an Illumina MiSeq sequencer (New York, USA). The quantity (ng/µl) and quality (absorbance ratios: A260/280, A260/230) of the isolated DNA was assessed. Pure nucleic acids typically yield absorbance ratios of ~1.8 (A260/280) and ~2.00 (A260/230) [278]. The average yield from the control samples was 50.12 ng/µl ± 16 and from the IPF patients 53.12 ng/µl ± 18 (P = 0.34), with an A260/A280 ratio of 1.81 ± 0.23 and an A260/230 of 1.98 ± 0.21 indicating good quality DNA. Following paired-end Illumina sequencing, a total of 15 million sequences were obtained from the 258 samples. The lowest number of reads per sample was observed in the negative controls (kit controls and bronchoscopy controls) **(Figure 3.4)**.

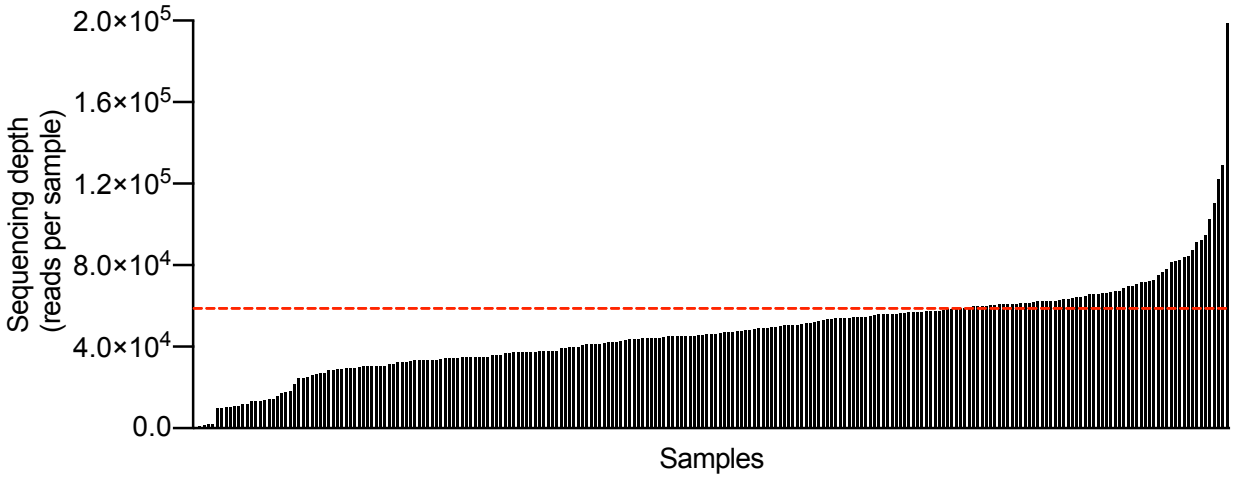


Figure 3.4 Distribution of samples sequence depth.

Illustrating the total number of reads present in each sample. The red dashed line represents the mean number of reads.

Following demultiplexing and denoising with DADA2 a total of 13,674,508 high-quality 16S rRNA reads remained, with on average 53,836 ASVs per sample **(Table 3.2)**. After singleton removal and filtering at 0.005% of total reads to remove low abundance ASVs likely representing sequencing artefacts, 12,523,345 sequences were retained, representing 553 unique ASVs across 241 samples. This final dataset was used in all subsequent analyses.

Table 3.2 Pre-processing of 16S rRNA sequence reads from bronchoalveolar lavage samples.

Metric	16S rRNA amplicon reads generated*	Post-Dada2 and clustering into ASVs	Post-filtering at 0.005% of total reads
Number of samples	258	258	241
Number of ASVs	6,799	4,791	553
Total number of sequence reads	15,001,358	13,674,508	12,523,345
Mean reads per sample ± SD	47,356 ± 66,868	53,836 ± 11,841	39,652 ± 10,114
Minimum reads per sample	248	248	1297
Maximum reads per sample	1,899,469	1,899,469	1,899,469

ASVs: amplicon sequence variants; DADA2: divisive amplicon denoising algorithm. *Including negative control samples which were included in sequencing run but filtered out from the final dataset.

3.4.3.2 Mock community

A mock community was included as a positive control in each run to enable accurate comparison across sequencing runs **(Figure 3.5)**. The number of sequencing reads and the composition of the mock community was similar across the three runs, confirming that all genera were represented in similar proportions.

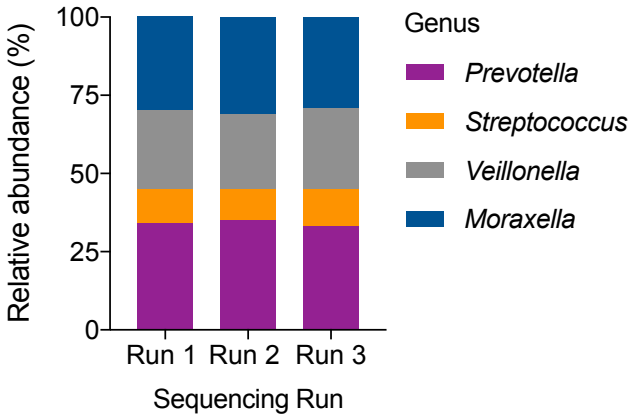


Figure 3.5 Mock community composition at the genus level from the three sequencing runs.

3.4.3.3 Contamination

To identify potential sources of contamination negative bronchoscopy controls and reagent controls were included in the sequencing run. By adding these controls, reads originating from contamination were identified and distinguished from those derived from actual constituent taxa. Despite no sample template being added at the DNA extraction step, several contaminant genera, which had been independently reported as contaminants in previous studies, were detected in the reagents and negative bronchoscopy controls [314-317] **(Table 3.3)**.

Table 3.3 Contaminant genera detected in the sequenced reagent and bronchoscopy negative controls.

Phylum	Constituent contaminant genera
Proteobacteria	
Alpha-Proteobacteria	*Bradyrhizobium*[a], *Caulobacter*, *Mesorhizobium*[b], *Novosphingobium*, *Sphingomonas*[a,c]
Beta-Proteobacteria	*Janthinobacterium*, *Pelomonas*, *Ralstonia*[a,b,c,d]
Gamma-Proteobacteria	*Acinetobacter*[a,c,d], *Enhydrobacter*, *Escherichia*[a,c,d], *Pseudomonas*[a,b]
Firmicutes	*Streptococcus*
Bacteroidetes	*Flavobacterium*[a]
Actinobacteria	*Aeromicrobium*, *Corynebacterium*, *Micrococcus*
Acidobacteria	Unclassified *Acidobacteria* Gp2 organisms
Spirochaetes	*Turneriella*

The listed genera were all detected in sequenced negative controls that were processed alongside human-derived samples (BAL samples). Genus names followed by a superscript letter indicate those that have previously been independently reported as contaminants. References: [a][314]; [b][315]; [c][316]; [d][317].

A PCA plot was constructed to compare the microbial profile of the BAL samples and the pooled negative bronchoscopy controls and reagent controls **(Figure 3.6A)**.

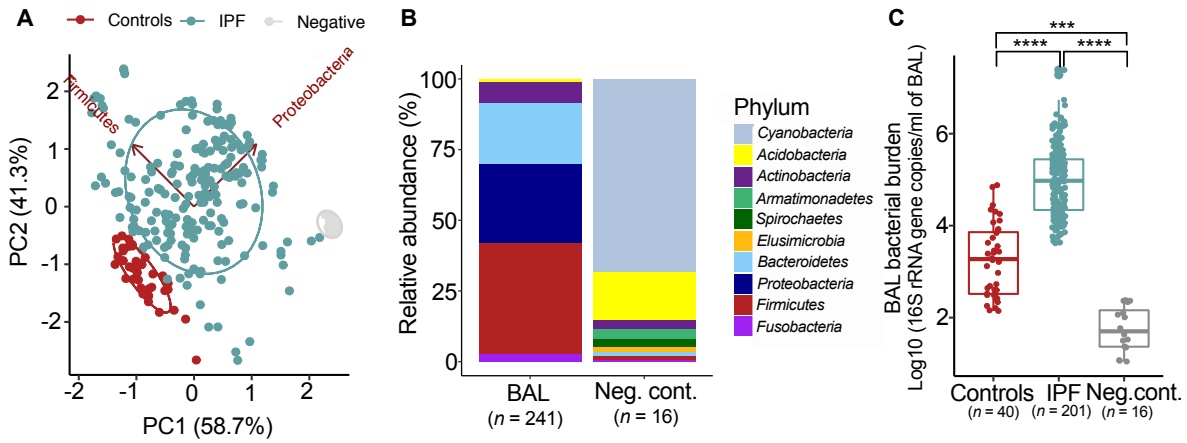


Figure 3.6 Comparison between all bronchoalveolar lavage (BAL) samples and negative controls.

(A) Principal component analysis (PCA) comparing the microbial composition of all of the BAL samples to that of the negative controls. **(B)** Taxonomic differences at the phylum level comparing all of the BAL samples (*n* = 241) to the negative control samples (*n* = 16) regardless of diagnosis. The number refers to the relative abundance (%) obtained prior to conducting any filtering step. **(C)** Bacterial burden calculated by qPCR comparing the bronchoalveolar lavage samples (controls = 40, IPF = 201) to the negative controls (*n* = 8 negative bronchoscopy controls, *n* = 8 reagent controls). Sample groups were compared by Kruskal-Wallis test with Dunn's multiple comparisons test adjusted with the Bonferroni method ($P < 2.2x10^{-16}$).

The representative points of the samples are mapped in space spanned by the first two principal components, PC1 and PC2. The PCA plot illustrates a clustering amongst the microbial profile according to sample type (58.7% of the variance was captured by PC1; 41.3% of the variance was captured by PC2). The PCA analysis revealed that the negative control samples had distinct microbial communities from the rest of the BAL samples. Furthermore, the negative control samples had a distinct taxonomic profile compared to the BAL samples analysed in the study and yielded a bacterial burden close to or below the lower limit of qPCR quantification (100 copies/ml of BAL) **(Figure 3.6B-C)**. The *decontam* package in R was used to additionally identify 30 ASV as contaminants by evaluating the prevalence (presence/absence across samples) of each ASV in the BAL samples compared to the prevalence in the negative controls. These ASV, together with the negative control samples, were removed from the dataset prior to conducting subsequent analyses.

3.4.4 Bacterial burden

3.4.4.1 Bacterial burden in health and disease

An increased bacterial burden has been consistently associated with fibrotic lung disease [114, 115, 149]. To confirm this finding, the differences in BAL bacterial burden between controls and subjects with IPF were tested. On average, IPF subjects had 5.11 x 10^5 copies of the 16S rRNA gene/ml of BAL, which was higher than the average copy number in control subjects (8.02 x 10^3, $P < 2.2 \times 10^{-16}$) **(Figure 3.7)**.

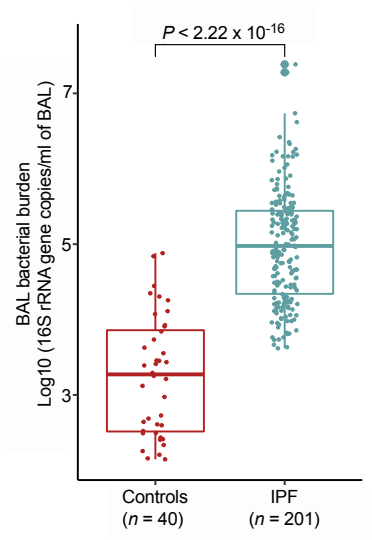


Figure 3.7 Bronchoalveolar lavage bacterial burden in healthy controls and idiopathic pulmonary fibrosis (IPF) samples.

BAL bacterial burden quantified using quantitative polymerase chain reaction and is expressed as log10 (16S rRNA gene copies/ml of BAL). Data are presented as median and interquartile range. Statistical significance tested by Mann-Whitney U test.

3.4.4.2 Bacterial burden and composite physiologic index

Disease progression was defined as a composite physiologic index (CPI) of either a relative decline in FVC of greater than 10% or death before 12 months. Individuals whose disease had progressed at 12 months (n = 53) demonstrated a significantly higher BAL bacterial burden when compared to subjects with stable disease (stable: $2.10 \times 10^5 \pm 3.63 \times 10^5$ 16S rRNA copies/ml of BAL, decline: $8.46 \times 10^5 \pm 2.74 \times 10^6$ 16S rRNA copies/ml of BAL, P = 0.02) **(Figure 3.8)**.

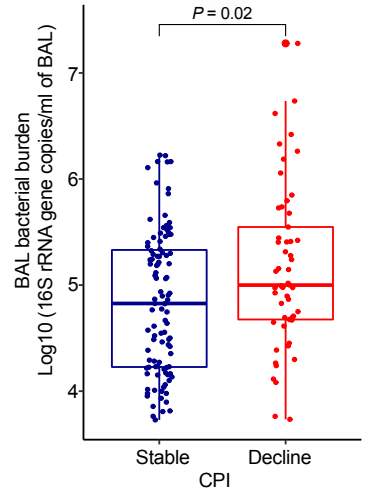


Figure 3.8 Bronchoalveolar lavage bacterial burden in idiopathic pulmonary fibrosis (IPF) subjects with either stable or progressive disease.

BAL bacterial burden of IPF subjects dichotomised into stable or progressive according to their composite physiologic index (CPI) at 12 months. BAL bacterial burden quantified using quantitative polymerase chain reaction and is expressed as log10 (16S rRNA gene copies/ml of BAL). Data are presented as median and interquartile range. Statistical significance tested by Mann-Whitney U test.

3.4.4.3 Bacterial burden and disease progression

An increased bacterial load in IPF has been consistently linked to an increased risk of mortality [114, 115, 149]. In an unadjusted Cox proportional hazards model, IPF patients with the highest bacterial burden had a significantly increased hazard compared to those with a lower bacterial burden (hazard ratio: 1.92, 95% confidence interval: 1.32-2.79, $P < 0.001$). To confirm these findings, the subjects with IPF were stratified into tertiles based on bacterial load and examined progression-free survival **(Figure 3.9)**. This association remained significant even following adjustment for age, sex, smoking status and baseline % predicted FVC in a multivariable Cox proportional hazards model (hazard ratio: 11.76, 95% confidence interval: 1.45-95.5, $P = 0.02$). Results of the multivariate Cox proportional hazards model are reported in **Table 3.4**.

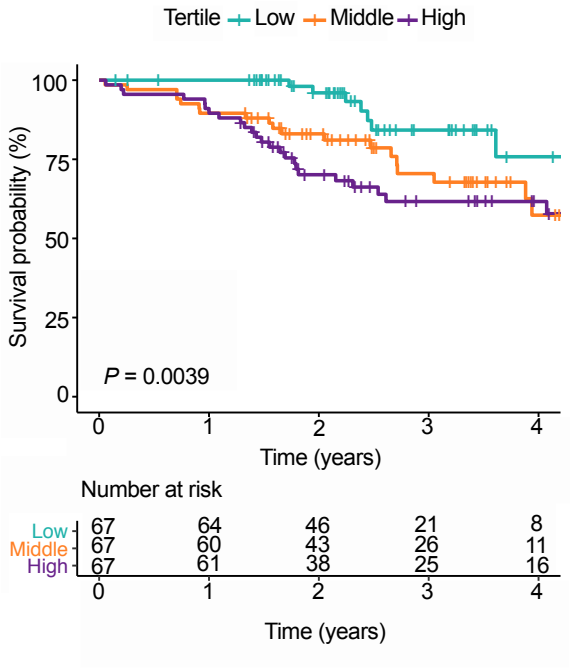


Figure 3.9 Survival probability of subjects with idiopathic pulmonary fibrosis (IPF) stratified into tertiles by bacterial burden.

Kaplan Meier curve generated by Cox proportional-hazards model for IPF subjects stratified into tertiles by bacterial burden (low: 4.15 x 10^3 – 3.76 x 10^4 16S rRNA gene copies/ml of BAL; middle: 3.78 x 10^4 – 2.11 x 10^5 16S rRNA gene copies/ml of BAL; high: 2.12 x 10^5 – 2.42 x 10^7 16S rRNA gene copies/ml of BAL). Log-rank *P* test value reported.

Table 3.4 Cox proportional hazards model of disease progression in idiopathic pulmonary fibrosis (IPF).

Variable	Group	*N*	HR	95% CI Lower	95% CI Upper	*P*-value*
Burden tertiles	low	67	reference	-	-	-
	middle	67	5.39	0.617	47.1	0.128
	high	67	11.76	1.449	95.5	**0.021**
Age	-	201	1.01	0.943	1.1	0.81
Sex	Females	59	reference	-	-	-
	Males	142	1.44	0.455	4.5	0.535
Smoking status	Current/Ex-smoker	10	reference	-	-	-
	Non-smoker	74	0.39	0.037	4.0	0.424
FVC (% predicted)	-	201	1.02	0.994	1.0	0.152

Multivariate model adjusted for age, sex, smoking status and baseline % predicted forced vital capacity (FVC).
*Log *P*-rank test value reported.
Events: 18; Global *P*-value (Log-rank): 0.047695. AIC: 139.87; Concordance Index: 0.78.

3.4.4.4 Bacterial burden and BAL differential cell counts in IPF

The relationship between BAL differential cell count and bacterial burden was assessed in the IPF cohort. No correlation was found between BAL bacterial load and the percentage of either alveolar macrophages, lymphocytes or neutrophils **(Figure 3.10)**.

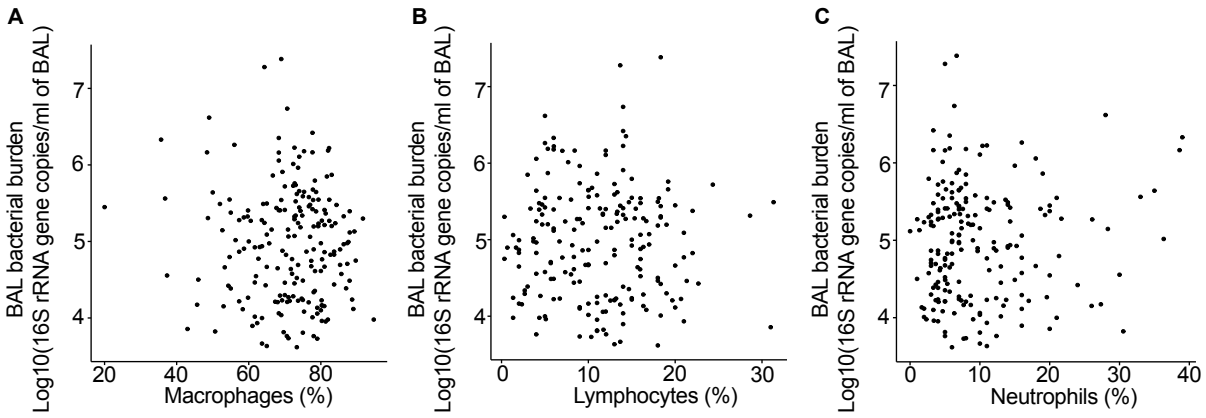


Figure 3.10 Correlations between bronchoalveolar lavage (BAL) differential cell counts and bacterial burden of subjects with idiopathic pulmonary fibrosis (IPF).

Percentage of total **(A)** macrophages **(B)** lymphocytes and **(C)** neutrophils. BAL bacterial burden in subjects diagnosed with IPF was calculated by qPCR and expressed as log10 16S rRNA gene copies/ml of BAL. Data was log-transformed to approximate normal distribution.

3.4.4.5 Bacterial burden and MUC5B genotype

Within the IPF cohort, the bacterial burden was found to associate independently with the *rs35705950* polymorphism in the promoter region of the mucin *MUC5B* genotype. IPF patients with the GG and GT *MUC5B* genotypes had on average, 3.99 x 10^5 and 2.31 x 10^5 16S rRNA gene copies/ml of BAL, respectively. Patients with IPF possessing the minor allele (TT) were found to have a lower bacterial burden (3.68 x 10^4 16S rRNA gene copies/ml of BAL) compared to patients possessing the GG ($P < 0.005$) and GT ($P < 0.012$) alleles **(Figure 3.11)**.

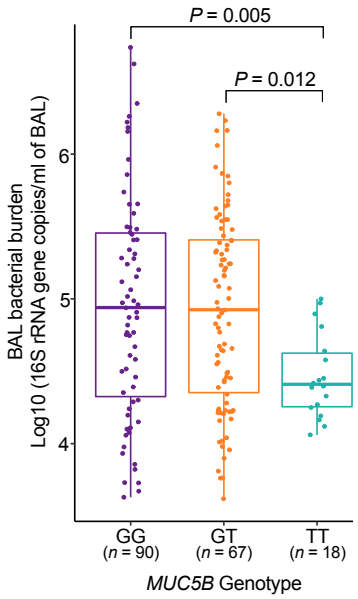


Figure 3.11 Bronchoalveolar lavage (BAL) bacterial burden in idiopathic pulmonary fibrosis (IPF) subjects stratified according to *MUC5B* genotype.

BAL bacterial burden quantified using quantitative polymerase chain reaction and is expressed as log10 (16S rRNA gene copies/ml of BAL). Data are presented as median and interquartile range. Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

3.4.4.6 Bacterial burden and radiological disease extent

Multivariate stepwise regression was used to explore the relationship between radiological markers of disease severity and bacterial burden. Strong collinearity was observed between CT fibrosis extent and CT traction bronchiectasis (Spearman's *rho* = 0.57) and were therefore included in separate models. No associations were found between bacterial burden and radiological extent of disease or extent of either traction bronchiectasis or honeycombing **(Table 3.5)**. No difference in bacterial burden was observed when specifically considering patients with definite or probable UIP, the presence or extent of PPFE, and the presence or absence of honeycombing **(Figure 3.12A-C)**. To assess whether bacterial burden predicted survival independent of radiographic features, radiographic features of disease severity were incorporated into the survival model. Bacterial burden remained an independent predictor of survival even when incorporating radiographic

features into the survival models and following adjustment for age, sex, baseline % predicted FVC; baseline % predicted DL_{CO} and smoking status. Furthermore, given that CT fibrosis extent and CT traction bronchiectasis score were strongly co-linear (Spearman's *rho* = 0.57), they were included in separate models. In both models, bacterial burden remained predictive of disease progression (hazard ratio: 1.16, 95% confidence interval: 1.01-1.35, *P* = 0.04).

Table 3.5 Radiological features and their univariate prediction of the bacterial burden in bronchoalveolar lavage.

Radiological feature	Mean (±SD)	β Estimate	SE	Adjusted *P*-value
Traction bronchiectasis extent score	6.7 (3.0)	0.691	1.750	0.492
Fibrosis extent (%)	22.0 (10.7)	1.050	0.921	0.297
Honeycombing extent (%)	2.6 (6.8)	-0.028	1.830	0.978
Definite or probable UIP pattern (*n* (%))	88.0 (62.4)	-0.597	0.278	0.552
PPFE lobar extent (lobes)	0.4 (0.9)	0.414	1.140	0.680

Binary CT variables expressed as absolute number and percentage. Variables expressed as mean and standard deviations (SD) of per cent of the whole lung proportion, with the exception of UIP pattern, which is expressed as an absolute number and percentage. Traction bronchiectasis scored on an 18-point scale. PPFE scored by number of lobes involved. Univariate regression performed against log-transformed 16S rRNA gene copy number. SE: standard error; UIP: usual interstitial pneumonia; PPFE: pleuroparenchymal fibroelastosis.

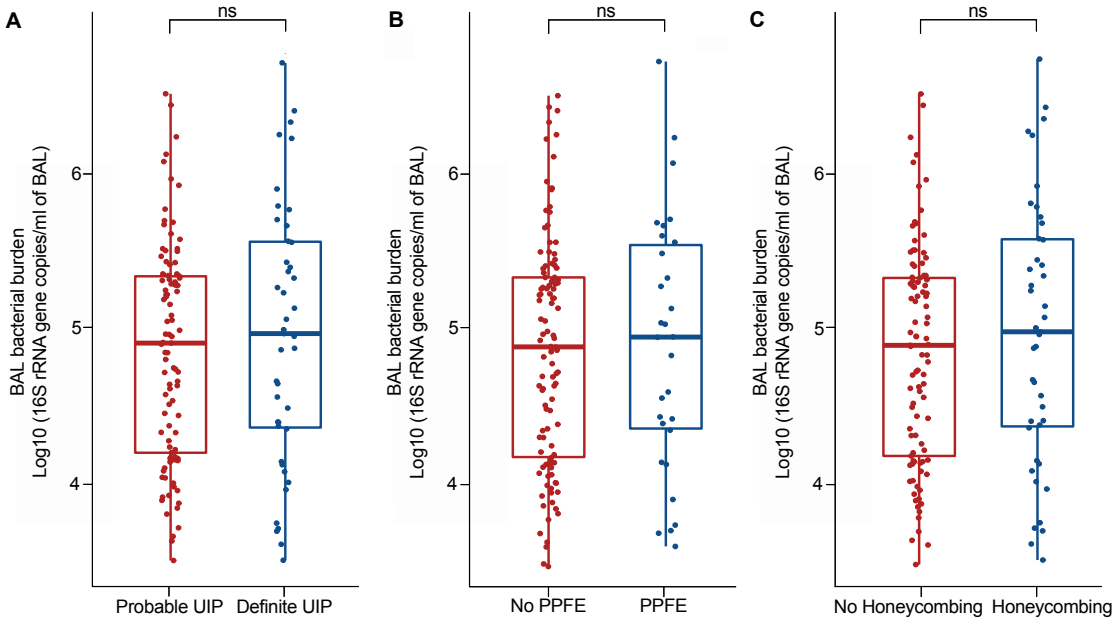


Figure 3.12 Relationship between radiological features and bronchoalveolar lavage (BAL) bacterial burden.

No difference in bacterial burden log10 (16S rRNA gene copies/ml of BAL) based on **(A)** a definite or probable pattern of usual interstitial pneumonia (UIP), **(B)** the presence or absence of radiological pleuroparenchymal fibroelastosis (PPFE) or **(C)** the presence or absence of honeycombing. Statistical significance tested by Mann-Whitney U test. Data are presented as median and interquartile range.

3.4.5 Microbiota composition in health and disease

3.4.5.1 Relative abundance

The composition of the microbial communities in the two groups was dominated by the following five phyla: *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Fusobacteria* **(Figure 3.13)**. At the phylum level, the microbiota of control subjects was dominated by *Firmicutes* (46%), *Bacteroidetes* (31%), *Proteobacteria* (15%), *Actinobacteria* (5%) and *Fusobacteria* (3%) **(Figure 3.13A)**. Although the same phyla predominated the BAL of IPF subjects, the *Firmicutes* and the *Proteobacteria* phyla accounted for 50% and 16% of the total reads **(Figure 3.13B)**. Comparison at the genus level identified that the microbial profile of the controls and IPF subjects was dominated by *Prevotella* (controls 29%; IPF 24%), *Streptococcus* (controls 22%; IPF 27%) and *Veillonella* (controls 16%; IPF 21%). Of these genera, *Streptococcus* (P = 0.01) and *Veillonella* (P = 0.04) were found to be more abundant in the BAL of IPF subjects compared to controls. Furthermore, *Staphylococcus* (controls 0.47%; IPF 2%) was also elevated in the BAL of IPF subjects compared to controls (P = 0.04), whereas *Actinobacillus* (controls 3%; IPF 2%) was found to be significantly lower in the BAL of IPF subjects (P = 0.007).

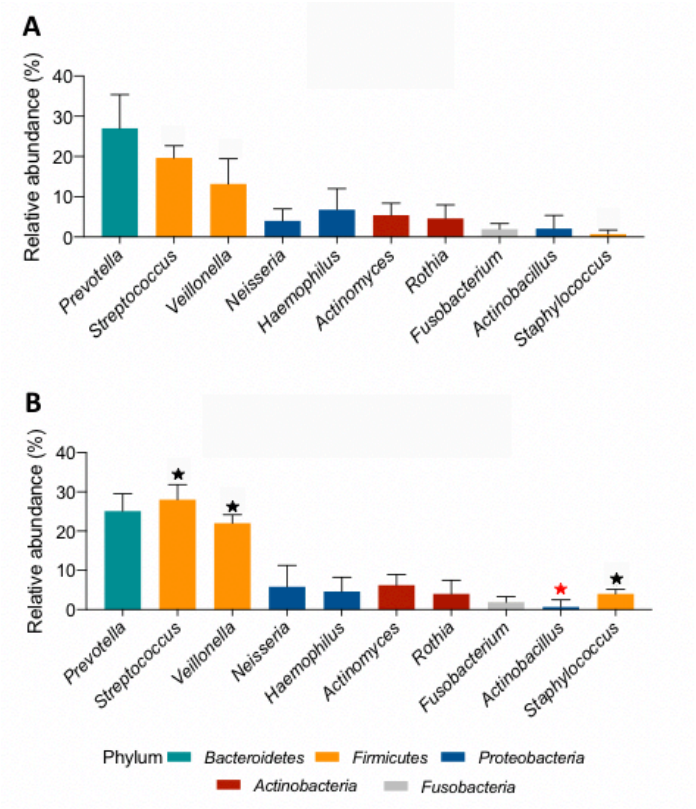


Figure 3.13 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) is altered compared to healthy controls.

Taxonomic changes at the phylum and genus levels in the bronchoalveolar lavage (BAL) of **(A)** healthy controls and **(B)** IPF subjects. Statistical significance tested by Mann-Whitney U test. Statistical differences marked as follows: ⋆ (increase in IPF), and ⋆ (decrease in IPF). Data are presented as mean ± standard deviation.

No correlation was found between bacterial genera and BAL differential cell counts **(Table 3.6)**.

Table 3.6 Spearman's rank correlations (*rho*) between respiratory microbiota (genus level) of idiopathic pulmonary fibrosis (IPF) subjects and bronchoalveolar lavage cell differentials.

Comparison	Spearman's *rho*	*P* value
Macrophages vs *Prevotella*	0.045	0.52
Macrophages vs *Streptococcus*	0.075	0.28
Macrophages vs *Veillonella*	-0.025	0.72
Macrophages vs *Neisseria*	0.100	0.13
Macrophages vs *Haemophilus*	0.030	0.67
Macrophages vs *Staphylococcus*	0.059	0.40
Lymphocytes vs *Prevotella*	-0.110	0.11
Lymphocytes vs *Streptococcus*	-0.085	0.22
Lymphocytes vs *Veillonella*	-0.069	0.32
Lymphocytes vs *Neisseria*	-0.083	0.23
Lymphocytes vs *Haemophilus*	-0.084	0.23
Lymphocytes vs *Staphylococcus*	0.026	0.27
Neutrophils vs *Prevotella*	0.010	0.89
Neutrophils vs *Streptococcus*	0.002	0.98
Neutrophils vs *Veillonella*	0.018	0.80
Neutrophils vs *Neisseria*	-0.032	0.65
Neutrophils vs *Haemophilus*	0.110	0.12
Neutrophils vs *Staphylococcus*	-0.021	0.77

3.4.5.2 Differential abundance

To further explore the community composition changes between healthy controls and IPF, differentially abundant bacteria were identified using LEfSe. Features significantly discriminating among groups with an LDA score > 2.0 were represented as a cladogram, as produced by LEfSe with default parameters. Multiple taxonomic differences were found: the BAL of healthy controls was enriched with bacteria belonging to the *Bacteroidetes* and *Actinobacteria* phyla, and the IPF group was enriched with bacteria belonging to the *Firmicutes* and *Proteobacteria* phyla **(Figure 3.14)**.

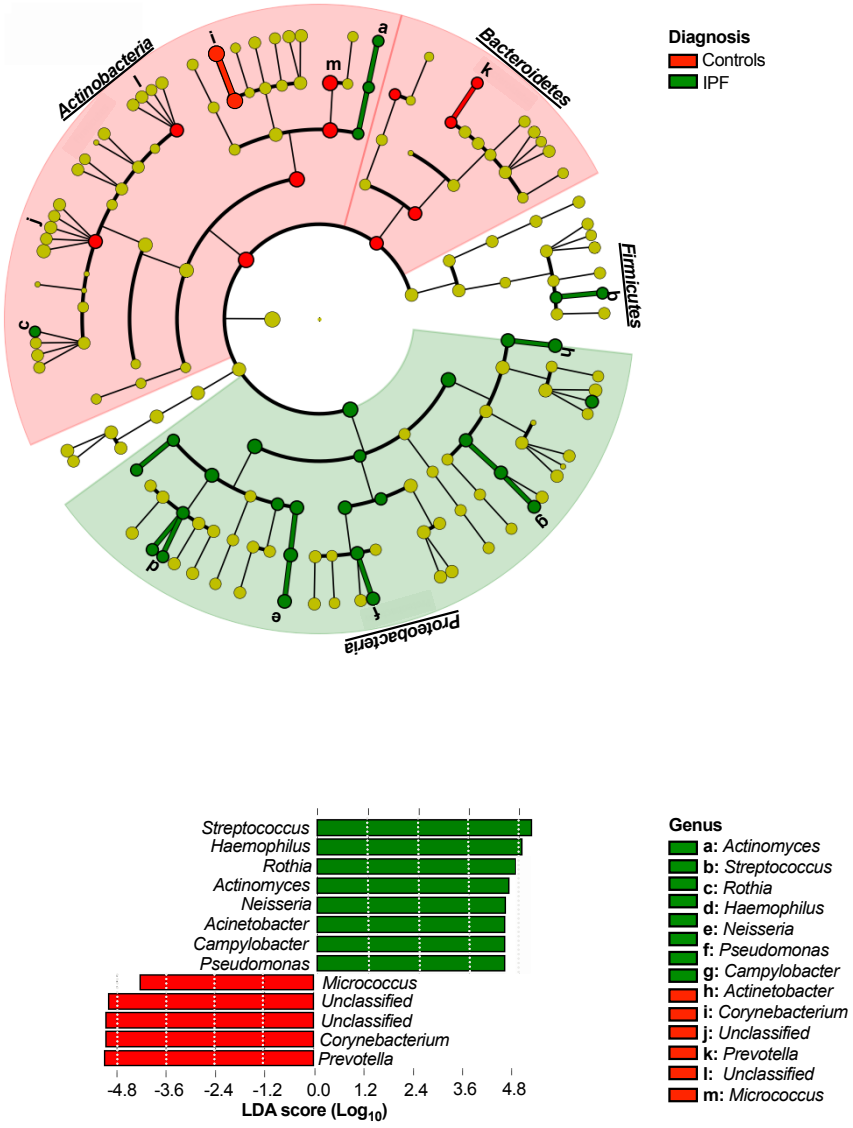


Figure 3.14 Differentially abundant bacteria within the healthy control and idiopathic pulmonary fibrosis (IPF) groups.

Cladogram representing results from calculated linear discriminant analysis (LDA) effect size (LEfSe) comparing the taxonomic composition of BAL samples from the two groups, $P < 0.05$. Multiple significant taxonomic differences were observed at different phylogenetic levels, with labels in the cladogram written for differences at the phylum level and indicated by letters for differences at the genus level. Only taxa meeting an LDA threshold > 2.0 are shown.

3.4.5.3 Diversity analyses

Alpha-diversity metrics were used to characterise the bacterial diversity within each sample **(Figure 3.15).** Alpha-diversity was assessed using the Shannon and Chao1 indexes, which estimate the evenness and richness, respectively, of the microbiome in each sample. Compared to the controls, subjects diagnosed with IPF were found to have a decreased evenness (P = 0.0039) **(Figure 3.15A)**. However, no significant difference (P = 0.074) was found when comparing microbial richness **(Figure 3.15B)**.

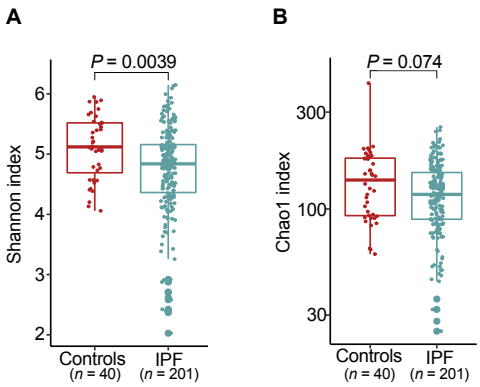


Figure 3.15 Alpha-diversity of microbial communities by diagnosis.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Data are presented as median and interquartile range. Statistical significance was tested by Mann-Whitney U test.

Changes in the community composition (β-diversity) between the two groups were investigated using PCA and tested by PERMANOVA which confirmed differences in the microbial profiles of the subjects based on diagnosis (P < 0.01) **(Figure 3.16)**. The PCA revealed that the variation amongst the groups was mostly driven by changes in the *Proteobacteria* and *Firmicutes* phyla, with a tendency toward an enrichment of *Firmicutes* and *Proteobacteria* in the BAL from IPF patients.

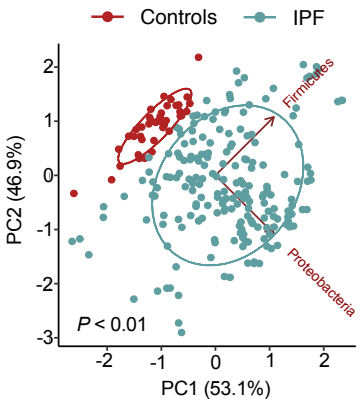


Figure 3.16 Beta-diversity of microbial communities by diagnosis.

Principal Component Analysis based on Euclidean distances statistically tested using Permutational Multivariate Analysis of Variance (PERMANOVA) adjusted for age, sex, smoking status, FVC (% predicted) and DL_{CO} (% predicted) (PERMANOVA: P < 0.01).

3.4.6 Microbiota composition and disease progression

It has previously been reported that an increased abundance of either *Streptococcus* or *Staphylococcus* in subjects with IPF is associated with a significant reduction in progression-free survival time [114]. Using a univariate analysis, both *Streptococcus* (hazard ratio: 2.05, confidence interval: 0.80-5.14, $P < 0.03$) and *Staphylococcus* (hazard ratio: 1.81, confidence interval: 0.78-4.07, $P < 0.03$) significantly associated with increased mortality in IPF. However, using a multivariable Cox proportional hazards model adjusted for age, sex, baseline % predicted FVC and smoking status, only an increased abundance of *Staphylococcus* (hazard ratio: 2.09, 95% confidence interval: 0.90-1.0, $P = 0.03$) but not of *Streptococcus* (hazard ratio: 1.54, 95% confidence interval: 0.84-4.19, $P = 0.05$) in the BAL of IPF subjects conferred an increased risk of mortality **(Figure 3.17).** The enrichment of these bacteria was observed in less than half of the cohort; therefore, it remains unlikely that these bacterial genera alone can explain the disease pathogenesis.

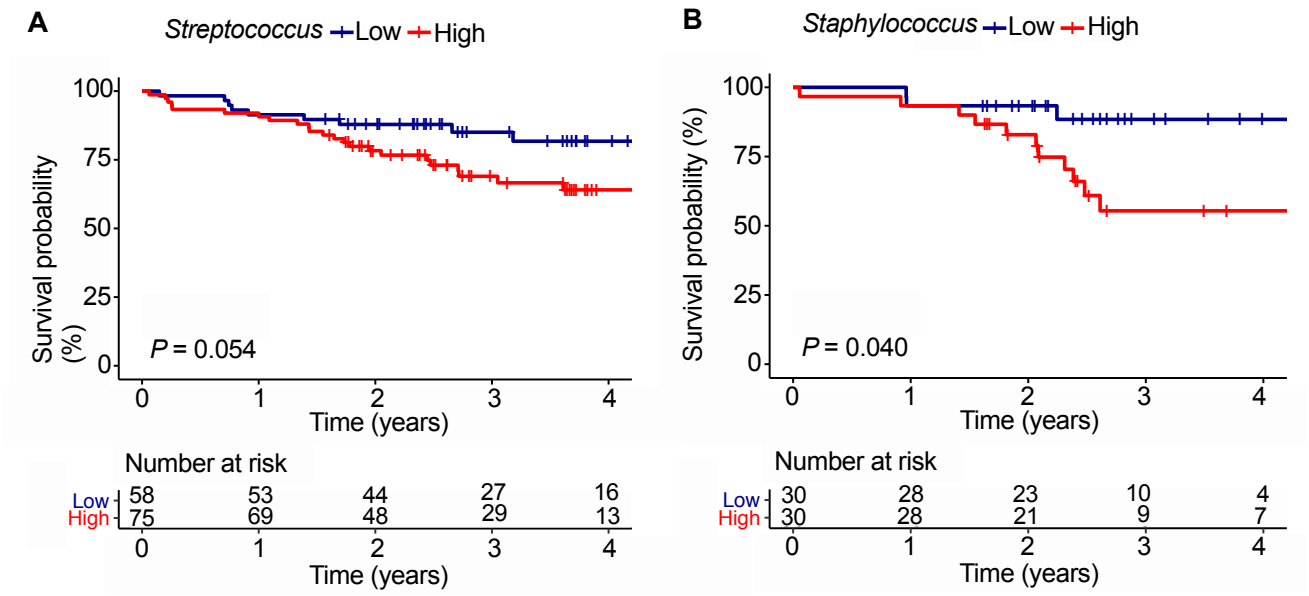


Figure 3.17 Increased relative abundance of *Staphylococcus* in the bronchoalveolar lavage (BAL) of idiopathic pulmonary fibrosis (IPF) subjects is associated with increased mortality.

Kaplan Meier curve generated by Cox proportional-hazards model stratified by relative abundance of **(A)** *Streptococcus* and **(B)** *Staphylococcus* in BAL of subjects diagnosed with IPF. Subjects were dichotomised using a threshold identified with the *cox.main* function from the R *AIM* package. Log-rank *P* test value reported.

Despite the significant differences in overall bacterial burden **(section 3.4.4.2)** when comparing IPF subjects with either stable or progressive disease, as determined by the CPI, no obvious changes were observed in microbial composition at the phylum or genus levels **(Figure 3.18)**.

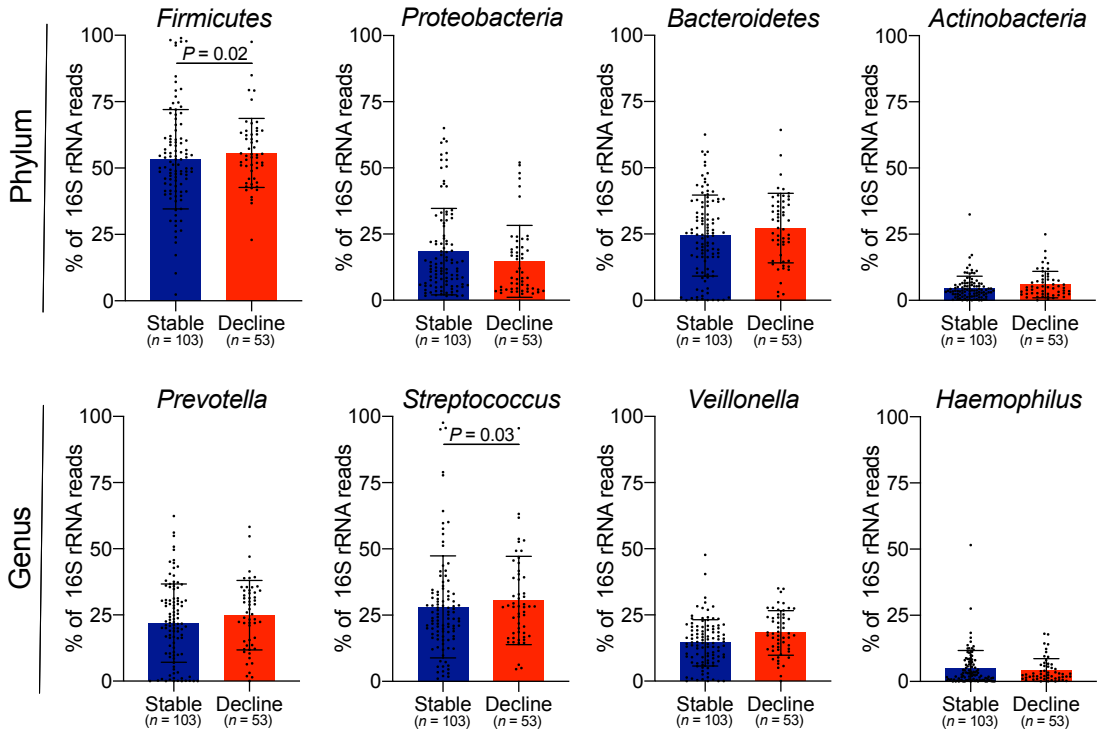


Figure 3.18 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects with stable or progressive disease.

Taxonomic composition (phylum and genus levels) of IPF subjects with stable disease or progressive disease. Statistical significance tested by Mann-Whitney U test.

Compared to subjects with stable disease, an increased abundance of *Firmicutes* and *Streptococcus* was observed in the BAL of patients with a decline in the CPI. However, individuals with progressive disease were not found to have a decreased microbial evenness or richness **(Figure 3.19AB)**.

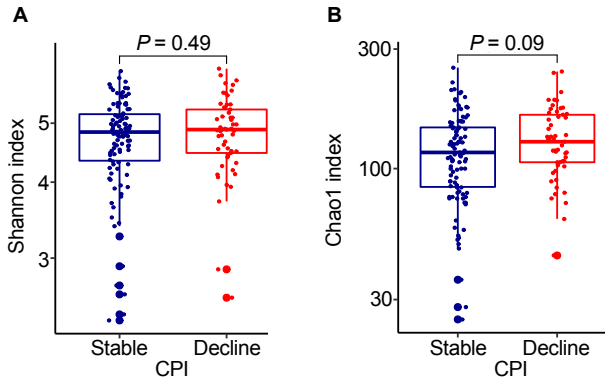


Figure 3.19 Alpha-diversity of idiopathic pulmonary fibrosis (IPF) subjects with stable or progressive disease.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Statistical significance tested by Mann-Whitney U test.

Similarly, when stratifying individuals diagnosed with IPF into tertiles based on their levels of bacterial burden (low, middle, high), no significant differences in the respiratory microbiota were observed at either the phylum or genus levels **(Figure 3.20)**.

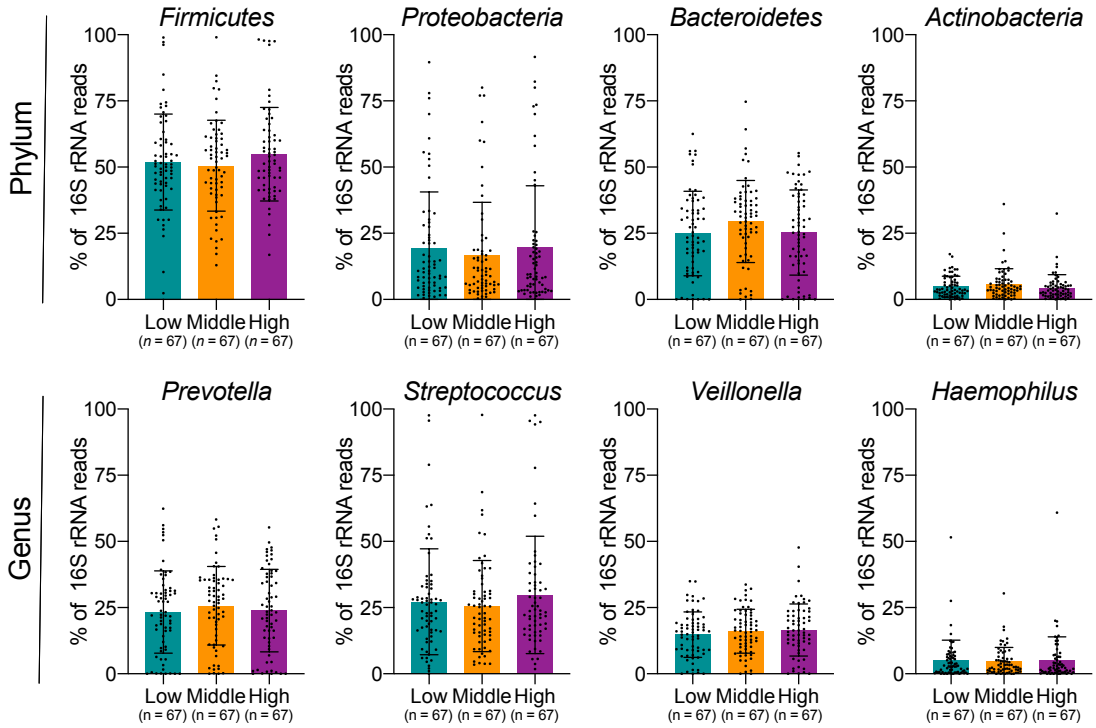


Figure 3.20 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects stratified according to bronchoalveolar lavage (BAL) bacterial burden.

Taxonomic composition (phylum and genus levels) for IPF subjects stratified into tertiles by bacterial burden (low: 4.15 x 10^3 – 3.76 x 10^4; middle: 3.78 x 10^4 – 2.11 x 10^5; high: 2.12 x 10^5 – 2.42 x 10^7 16S rRNA gene copies/ml of BAL). Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

No differences in microbial evenness or richness were observed according to bacterial burden tertile **(Figure 3.21AB)**.

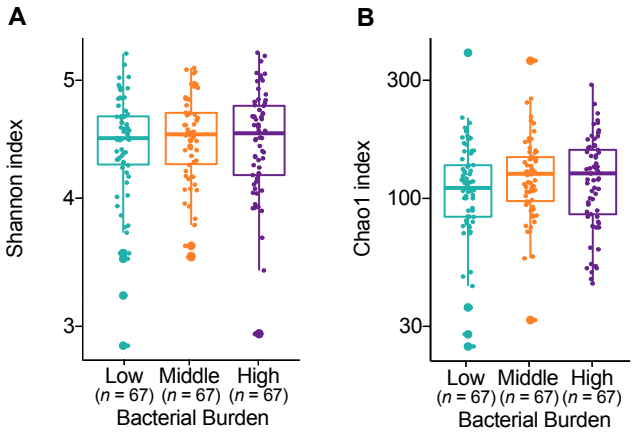


Figure 3.21 Alpha-diversity of idiopathic pulmonary fibrosis (IPF) subjects stratified according to the bacterial burden.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

3.4.7 Microbiota composition and smoking history

Of the 201 subjects diagnosed with IPF, 74 were classified as non-smokers, 117 as ex-smokers and 10 as current smokers. No demonstrable changes were observed at the phylum or genus levels according to smoking history except for an increased abundance of *Bacteroidetes*, particularly *Prevotella*, in the BAL of ex-smokers compared to non-smokers **(Figure 3.22)**. No differences in microbial evenness or richness were observed according to smoking history **(Figure 3.23AB)**.

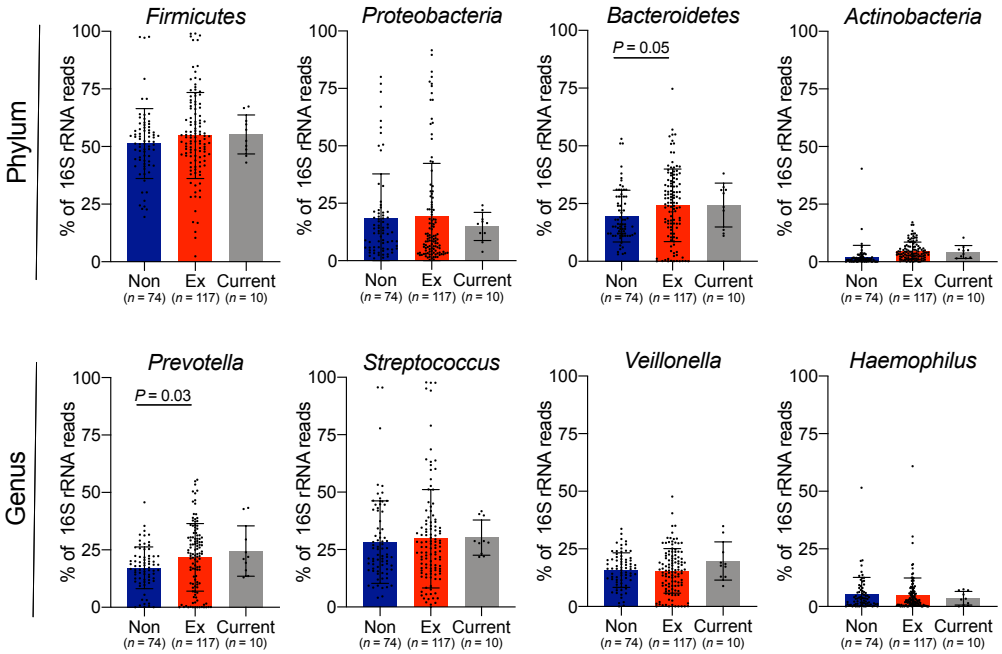


Figure 3.22 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects stratified according to smoking history.

Taxonomic composition (phylum and genus levels) for IPF subjects stratified according to smoking history (non-smoker, ex-smoker and current smoker). Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

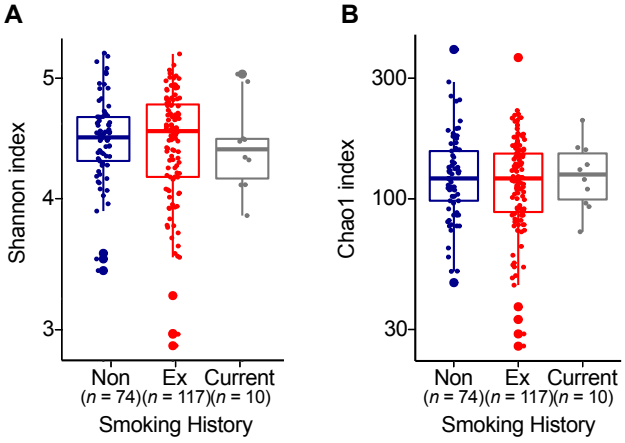


Figure 3.23 Alpha-diversity of idiopathic pulmonary fibrosis (IPF) individuals stratified according to smoking history.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

3.4.8 Microbiota composition and *MUC5B* genotype

Changes in the microbiota composition of BAL from IPF patients were evaluated according to *MUC5B* genotype. Although patients with IPF possessing the minor allele (TT) were found to have a lower bacterial load compared to patients possessing the GG and GT alleles **(section 3.4.4.5)**, no significant differences were observed between the microbiota, either at the phylum or genus levels, based on genotype **(Figure 3.24)**. No differences in microbial evenness or richness were observed according to genotype **(Figure 3.25AB)**.

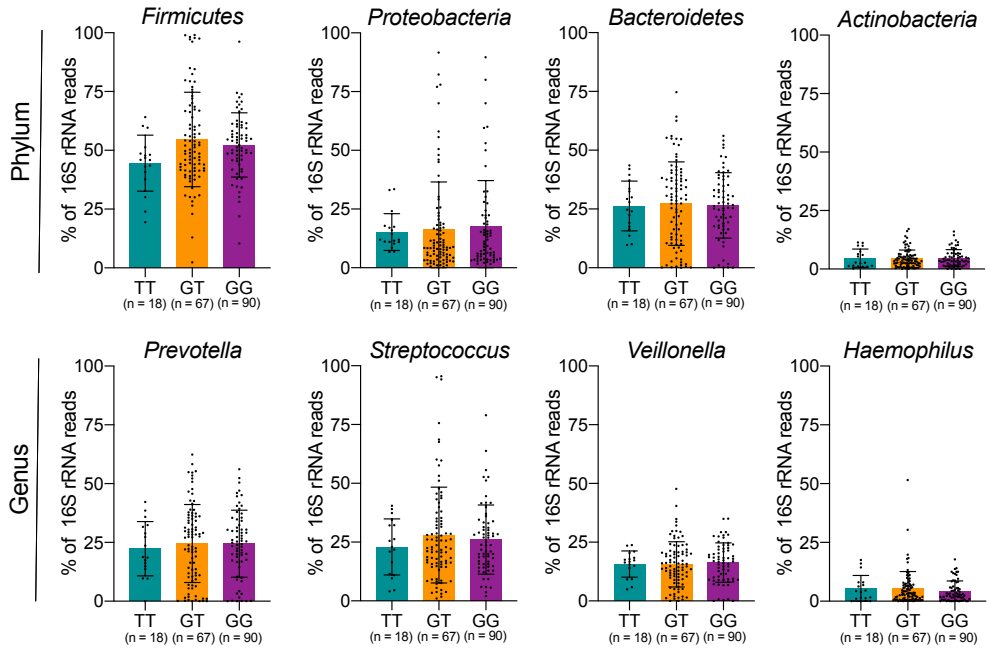


Figure 3.24 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects based on *MUC5B* genotype.

Taxonomy at the phylum and genus levels based on *MUC5B* genotype (TT, GT, and GG). Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

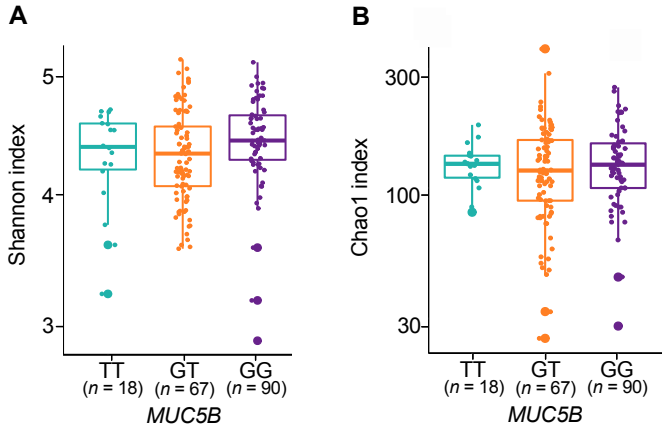


Figure 3.25 Alpha-diversity of idiopathic pulmonary fibrosis (IPF) individuals stratified according to *MUC5B* genotype.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Statistical significance was tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

3.5 Discussion

IPF is a progressive, irreversible and inevitably fatal scarring condition of the parenchyma [9]. The triggers that initiate the fibrotic response remain unknown. In order to investigate the role of bacteria in the pathogenesis and progression of IPF, the airway microbiome of patients with IPF was characterised using 16S rRNA sequencing and compared to that of healthy controls.

This Chapter validates previous findings showing there are significant differences in the composition and diversity of the respiratory microbiome in subjects with IPF compared to healthy controls. Furthermore, it confirms the finding that patients with IPF have an increased bacterial burden at the time of diagnosis, which confers a worse survival. For the first time, it was demonstrated that this increased bacterial burden is independent of radiological markers of disease morphology and severity. This study is the largest prospective study of bacteria in the lower airways of individuals with IPF performed to date, and confirms the importance of the respiratory microbiota and airway bacterial burden in IPF pathology. While previous studies have investigated bacterial burden and disease severity using physiological factors, such as FVC, none have looked at the relationship between bacterial burden and radiological markers of lung deterioration or disease severity. In the present Chapter, bacterial burden was found to be an independent predictor of survival, even when CT scan results were incorporated into the survival model, and adjustments for demographics and physiological parameters were made. The lack of association seen with measures of disease severity and fibrotic destruction of the lung suggests that changes in the bacterial burden are disease-relevant and unlikely simply secondary to the architectural distortion of the lungs [277].

The existence of a microbiome in the healthy lung is now generally accepted; at the phylum level, the healthy lung is characterised by the presence of a 'core' pulmonary microbiome, typified by the presence of *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Fusobacteria* and *Actinobacteria* [318]. At the genus level *Prevotella*, *Veillonella* and *Streptococcus* genera have been shown to dominate [132, 318, 319]. In this work, there were marked differences in the distribution of phyla between the healthy controls and subjects with IPF. Specifically, an increased abundance of *Firmicutes* and *Proteobacteria*, which harbour potentially pathogenic microbial species, was noted in the fibrotic group. In contrast, *Bacteroidetes* and *Actinobacteria* were found to be less abundant in disease than in control subjects. These findings are consistent with the literature which shows that during respiratory disease there is an increased prevalence of *Firmicutes* and *Proteobacteria* and a reduction in non-pathogenic commensals, such as *Bacteroidetes*, when compared to healthy

controls [320]. This observation suggests that a dysbiotic state of the lung microbiome could be directly associated with the pathophysiology of IPF.

Despite infection being a potential cause of alveolar epithelial injury, the role of microbial colonisation in the initiation and progression of fibrosis in patients with IPF has only been recently addressed [321]. Using BAL from the lower airways, Molyneaux and colleagues have reported that, compared to age-matched healthy controls, patients with IPF displayed significantly less bacterial communities and were found to harbour a greater abundance of *Haemophilus*, *Neisseria*, and *Streptococcus* bacterial genera. Furthermore, individuals with IPF were found to have an increased bacterial burden which was associated with increased mortality [115]. In a separate study, Han and colleagues prospectively recruited patients in the COMET study to evaluate the lung microbiome in patients with IPF. An increased abundance of *Streptococcus* or *Staphylococcus* genera was associated with reduced progression-free survival time [114]. Concordant with these findings, in this Chapter, differentially abundant bacterial genera were identified with LEfSe and, compared to healthy controls, the BAL of IPF subjects was found to be enriched with *Streptococcus*, *Haemophilus*, and *Neisseria*. In further agreement with the literature, the baseline bacterial communities of the healthy control subjects were enriched with *Prevotella*, which is known to dominate the airways of healthy subjects [156]. Moreover, in IPF, an increased abundance of *Staphylococcus* in the BAL of IPF subjects was found to confer an increased mortality risk. Collectively, these findings suggest that alterations in the microbiome characterised by an increased abundance of potentially pathogenic bacteria may play a potential role in the pathogenesis of IPF.

Usually, microbial disturbances are associated with the outgrowth of dominant and typically pathogenic bacterial genera over community diversity [156]. This pattern of decreased diversity has been reported in patients with CF and COPD, where a shift towards a microbial community predominated by pathogenic species has been shown [322, 323]. Indeed, previous studies have demonstrated that a loss of diversity of the lung microbiome correlated with the progression of IPF. Specifically, research by Takahashi and colleagues found an increase in *Firmicutes* and *Bacteroidetes* and a decrease in *Proteobacteria*; this reduction in microbial diversity was associated with disease activities in IPF [320]. These findings may indicate that a shift in the respiratory microbiome toward more densely populated and less diverse bacterial communities of the lower airways in IPF may provide persistent stimuli for repetitive alveolar injury [252]. In the present study, alpha diversity metrics were used to characterise the number (richness) and distribution (evenness) of taxa expected within each BAL sample. Compared to the controls, subjects diagnosed with IPF were

found to have a decreased evenness. A signal toward a lower bacterial richness was also observed, but this was not statistically significant. In contrast, beta diversity analyses revealed significant differences in whole community composition between the two groups. These results validate the finding that the airway microbiota in IPF is different from that of healthy controls. Nevertheless, causal conclusions cannot be drawn about this altered respiratory microbiome in IPF ahead of functional and longitudinal studies which are urgently needed to better understand pathogenic host-microbial interactions.

Several factors have been proposed as potential sources of injury to the airway epithelium, including occupational dust exposure and cigarette smoke. Specifically, smoking may influence the microbiome via changes to immune homeostasis, biofilm formation and oxygen tension which may be involved in a range of diseases [324]. Indeed, smoking has been previously reported to contribute to modifications of the oral, lung and gut microbiome in health and disease [324]. For example, Charlson and colleagues examined the microbial composition and effects of cigarette smoke by using culture-independent high-density sequencing to analyse the microbiota from the right and left nasopharynx and oropharynx of 29 smoking and 33 healthy non-smoking adults. They found that the microbiota from smokers was significantly more diverse than non-smokers as well as reported enrichment of anaerobic bacterial genera associated with periodontal disease in the oropharynx [325]. Similarly, in a study of patients with asthma, ex-smokers were found to have a higher prevalence of the *Bacteroidetes* and *Fusobacteria* phyla in sputum compared with non-smokers [326]. In this Chapter, the BAL-associated microbial communities in IPF patients was compared between non-smokers, ex-smokers and current smokers. In agreement with other studies, the BAL of non-smokers had an increased abundance of *Bacteroidetes* compared to ex-smokers. At the genus level, *Prevotella* and *Haemophilus* were found to be increased in the BAL of non-smokers compared to ex-smokers. However, no differences were reported when comparing the changes in the microbial communities between non-smokers and current smokers. This lack of difference may be explained by the small sample size in the current smokers' group. However, in IPF, the exact relationship between smoking and microbiome alterations remains to be further explored. More studies examining the effect of smoking on airway microbiome in individuals with IPF are warranted.

It is of vital importance to identify patients with increased risk of mortality in order to allow prompt and timely referral for lung transplant assessment. Whilst there have been conflicting reports regarding age and smoking history and their association with mortality, lung function parameters

have been consistently used to define the progression, stability or improvement of the disease [17]. Specifically, the percentage value on the predicted FVC and DL_{CO} are used in the monitoring and evaluation of the disease. The worsening, stability or improvement of the disease is defined by changes between 6 or 12 months in the percentage of FVC or DL_{CO} [8]. A decline greater than or equal to 10% of FVC or greater than or equal to 15% of DL_{CO} have been extensively used as cut-off points to define disease progression. In this Chapter, a twelve-month decline of greater than 10% in FVC or a decrease in DL_{CO} of more than 15% were both found to associate with an increased risk of mortality in IPF [17]. Furthermore, it has consistently been shown that the bacterial burden in the BAL of subjects with IPF is increased compared to healthy controls and that the bacterial load at the time of diagnosis predicts rapidly progressive IPF and an increased risk of mortality [114, 115, 149]. Similar to previous studies, bacterial load was increased in the BAL of subjects with IPF and was found to predict subsequent progression of IPF. However, when stratifying individuals diagnosed with IPF into tertiles based on their levels of bacterial burden (low, middle, high), no significant differences in alpha-diversity or the respiratory microbiota were observed at either the phylum or genus levels.

Although there was an overall increase in bacterial burden in IPF subjects, when examining BAL differential cell counts, there was no correlation with a specific cellular response. Conflicting evidence exists regarding the prognostic implications of BAL neutrophilia in IPF. In a study by Kinder and colleagues, an increased BAL neutrophil count led to an increased risk of mortality at 12 months. The authors hypothesised that they might be identifying a subgroup of individuals with more active disease [327]. In the present study, an increased proportion of neutrophils was observed in the airways of patients with IPF, however, there was no association with their counts and overall bacterial burden. Interestingly, there was no association between BAL cellular profiles and the bacterial burden in the airways, emphasising the difference between the stable microbiome and infection, where a cellular response is seen. Future studies should consider integrating methods of detecting response, such as the characterisation of the epithelial and airway transcriptome [277].

The single most reliable genetic risk factor for the development of IPF is the *MUC5B* promoter variant *rs35705950* [41]. Paradoxically, the presence of the minor allele of the *MUC5B* promoter polymorphism has been associated with a lower bacterial burden and better prognosis in patients with IPF [54, 115]. In the present study, the bacterial burden associated independently with *rs35705950* genotype; patients with IPF possessing the minor allele (TT) were found to have a lower bacterial burden compared to patients possessing the GG and GT alleles and were at a significantly

lower risk of mortality. This protective effect may result from improved bacterial clearance due to enhanced mucin production, suggesting that increased bacterial colonisation of the airways may play a role in IPF progression. This is further supported by the fact that conditions such as gastro-oesophageal reflux disease, which may lead to increased immigration of potentially pathogenic gastric bacteria to the lungs, is a common comorbidity in patients with IPF [328]. However, based on genotype, no significant differences were observed in alpha-diversity or taxonomic composition at the phylum or genus levels.

While the minor (T) allele is the risk allele for development of IPF, it also confers a survival advantage among patients with IPF. Taken together, these results suggest that the IPF phenotype is heterogeneous and consists of at least two clinical subtypes which are partly distinguishable by the *MUC5B* genotype [41]. However, the mechanism by which variants in *MUC5B* confer an increased risk of lung fibrosis is yet to be elucidated. Immune dysregulation has been proposed as a potential mechanism by which increased mucin expression contributes to the pathophysiology of IPF. Alternatively, IPF may be a disease of mucociliary clearance, whereby overexpression of *MUC5B* leads to impaired ciliary function, thus allowing the retention of particles and consequently recurrent lung injury [56].

Although traction bronchiectasis and honeycombing are defined as distinct entities, they radiologically reflect two aspects of the pathological remodelling which occurs in IPF. In addition to the extent of overall fibrosis, both bronchiectasis and honeycombing have been associated with disease behaviour and clinical outcome [329]. However, no correlation was found between honeycombing extent or traction bronchiectasis and bacterial burden, which remained a predictor of progression even after accounting for both of these entities. The relative proportions of CT UIP morphologies (definite, probable, indeterminate, inconsistent) in this cohort of IPF subjects are consistent to proportions which have been previously reported [330]. No associations were found with an extent of fibrosis or traction bronchiectasis, neither were there any differences between the radiological diagnostic cohorts of probable versus definite UIP. Despite the clinical association between PPFE and infection, there was no relationship between either the presence or extent of PPFE and bacterial burden [331]. Although this is an interesting finding, only one-third of the subjects in the cohort had evidence of PPFE. Thus, this lack of association will have to be investigated in larger cohorts where it would also be interesting to characterise the microbiome in subjects with isolated PPFE.

Dickson and colleagues recently set out to investigate the relationship between the microbiome and the presence or absence of honeycombing in a subgroup of 68 subjects with IPF [332]. The authors found no association with the lower airways bacterial burden when dichotomising the cohort based on the presence or absence of honeycombing. The prospective nature and quantitative radiological analysis undertaken in this Chapter expands upon their work. Furthermore, the study design employed here enabled me to answer questions regarding the graded relationships between radiological change and bacterial burden, strengthening the conclusion that increased bacterial burden is not simply related to the extent of radiological disease.

However, this work raises several issues regarding the potential for antimicrobial therapies in IPF, some of which are being currently trialled (NCT02759120 and 17464641) [83, 333]. Although it is clear that the lower airways of patients with IPF harbour a distinct microbiome compared to healthy controls, the only consistent finding has been that of an increased bacterial burden. Currently, none of the therapies which are being trialled have demonstrated an effective reduction of bacterial burden. In a *post hoc* analysis of multiple trials of inhaled antibiotics, it was recently revealed that patients with the higher bacterial burden demonstrated a clinical response and improved quality of life [334]. Therefore, rather than treating indiscriminately, a more targeted approach may be required for patients with IPF.

The advent of next-generation sequencing technologies has allowed direct and culture-independent sequence-based interrogation of microbial communities, both in homeostatic and perturbed states [335]. The most commonly used approach to the study of bacterial communities exploits high-throughput sequencing of amplicons of the 16S rRNA gene, a highly conserved prokaryotic gene. 16S rRNA gene sequencing is a highly accurate and versatile technology for bacterial classification and identification in complex biological mixtures [318]. Although 16S rRNA gene sequencing is a powerful tool for understanding the role of the microbial community in health and disease, there are limitations. 16S rRNA sequencing results only permit interrogation of the relative, rather than the actual, quantity of a particular bacterium [335]. Chimeras are hybrid reads between multiple parent sequences that can be incorrectly classified or falsely identified as novel organisms, thus confounding the real biological diversity [336]. However, in the present study, these were removed following demultiplexing and denoising of reads with DADA2. Another significant concern when using 16S rRNA sequencing to characterise microbial communities is the potential of contamination during sample acquisition. Contamination presents a critical problem when evaluating samples with low microbial biomass, such as the lungs [337]. The low amount of starting material may be

swamped by the amplification of reagent-derived contaminants and results in erroneous representations of the microbial communities [337, 338]. In this study, to minimise the risk of contamination, negative bronchoscopy controls and reagent controls were included in the sequencing run. The negative control samples yielded a bacterial burden close to or below the lower limit of qPCR quantification (100 copies/ml of BAL) and had a taxonomic composition which was distinct from the actual BAL samples. This was further confirmed by the PCA analysis, which showed that BAL samples clustered separately from the negative controls, supporting the hypothesis that the qPCR results are not simply an artefact from a sampling or processing error. Another significant limitation is that 16S rRNA gene sequencing provides a snapshot of the microbial community at a particular moment in time, and it is therefore challenging to determine whether the altered microbial community is a cause or a consequence of the disease [335]. Thus, functional and longitudinal studies examining host-microbial interactions are necessary. Evidence also indicates the involvement of viruses and fungi, which play an important role in health and disease. Another caveat of 16S rRNA sequencing is that it does not assess other potential nonbacterial lung pathogens present in the lower airways, which may play a role in the disease. The lung virome and mycobiome have been shown to have a significant impact on clinical outcome of chronic respiratory diseases such as asthma, COPD, CF and IPF [145, 339, 340]. Furthermore, respiratory viruses are a major cause of acute exacerbations of chronic pulmonary diseases, and it has been demonstrated that the host response to viral infection is dysregulated in these diseases. Studies have also supported a mechanistic role for viruses in the initiation and progression of IPF. In particular, cytomegalovirus, Epstein-Barr virus, HHV-7 and HHV-8 of the HHV family have been considerably studied in the context of acute exacerbations of IPF [88]. Moreover, infection by respiratory viruses is a major trigger of wheezing in infants as well as exacerbations of asthma in older children. Respiratory syncytial virus, human rhinovirus, coronavirus, parainfluenza, as well as other viruses have been associated with these wheezing events [341]. Although very little is known about the lung mycobiome compared to the bacterial microbiome, fungi have recently been identified in respiratory diseases [342, 343]. Commensal fungi affect the immune system and modulate bacterial communities and as a result are involved in the restoration of the bacterial microbiota following antibiotic treatment [344, 345]. Although the importance of lung viral and fungal communities is increasingly recognised, mechanistic research is urgently needed to further explore their role in the pathogenesis of chronic lung diseases.

3.6 Conclusions

The microbial profile of subjects with IPF is distinct from that of healthy controls. Compared to controls, these individuals have an increased bacterial burden at the time of diagnosis, which confers a worse survival. This is the first study to show that this increased bacterial burden is independent of radiological markers of disease morphology and severity. The lack of association seen with measures of disease severity and fibrotic destruction of the lung suggests that changes in bacterial burden are disease-relevant and unlikely simply secondary to the architectural distortion. Future work should aim to understand whether the total bacterial load and/or the relative abundances of the altered bacterial genera may be acting synergistically to advance the fibrotic process. Collectively, these findings provide a strong rationale for trials of targeted antimicrobial therapy which will help elucidate the aetiological role of these microbial organisms and at the same time aid in the search for a novel potential treatment.

4 **Chapter 4**: The Respiratory Metabolome in Idiopathic Pulmonary Fibrosis

4.1 Introduction

The work in **Chapter 3** has demonstrated that the lungs of subjects with IPF are characterised by an increased bacterial load, by changes in bacterial community composition, and decreased bacterial diversity compared to those of healthy individuals. Additionally, increased numbers of bacterial communities in the airways were found to have detrimental effects on survival in patients with IPF. Although these observational and association studies support the growing hypothesis that host-environment interaction is crucial in the pathogenesis of IPF, they do not provide any functional or mechanistic insights.

It is now well recognised that the intestinal microbiota is an important contributor to host health and immune homeostasis in a range of inflammatory diseases, including type 2 diabetes, asthma, obesity, and COPD [346]. The respiratory microbiome develops in tandem with that of the gut [347]. Indeed, a bidirectional dialogue occurs between these two compartments, which has been termed the 'gut-lung axis'. This cross-talk is driven by the metabolic state of the microbiota and is pivotal for maintaining health and homeostasis in the individual host. However, dysregulation of this dynamic has been shown to be a key driver for many diseases [168, 348]. Manipulation of the gut microbiome, by changes in diet, can alter the respiratory microbiome providing beneficial effects in asthma and protection against respiratory viruses [167, 349]. The gut microbiota influences the microbial composition within the lung either by direct seeding of the respiratory tract with bacteria or by the distribution of bacterial microbial metabolites, including the SCFAs [163]. The prevailing hypothesis in the literature is that SCFAs, produced from the bacterial fermentation of non-digestible carbohydrates by intestinal bacteria in the gut lumen, are integral for tight junction formation and are directly anti-inflammatory [350]. SCFAs are straight-chain saturated fatty acids with fewer than six carbon atoms, which have widespread effects on the gut but have also been shown to exert their effects in peripheral tissues such as the lung [351]. The most common SCFAs produced by the microbiota are acetate, a two-carbon SCFA; propionate, a three-carbon SCFA; and butyrate, a four-carbon SCFA. SCFAs act as a link between the microbiota and the immune system by interacting with immune cells and modulating their recruitment, differentiation, activation and survival within different tissues [157]. The concentrations at which SCFAs occur in the lumen of the large intestine range from 30 to 150 mM [352]. They are absorbed into the portal circulation and reach the bloodstream, where they potentially exert their effects [163]. However, SCFAs can also be detected in sputum at lower concentrations (0.1-5 mM), demonstrating that they reach the lungs and airways [353]. SCFAs elicit their effects by binding to GPRCs, including GPCR43, GPCR41,

GPR109A and olfactory receptor 78. These GPCRs are expressed on several cell types such as immune cells and bronchial epithelial cells where they are upregulated in disease [354, 355]. For example, once activated, GPCR43 has been shown to mediate the expression of several antimicrobial peptides and defensins in an mTOR and STAT3-dependent manner [356].

The pro- and anti-inflammatory effects of these dietary fibre fermentation products are now well documented in a number of chronic inflammatory diseases [161, 353, 355, 357]. However, while the role of SCFAs and the gut-lung axis in pulmonary inflammation has been investigated, little is known about the impact of microbial metabolites on the development of fibrosis. In recent years, there has been a clear increase in the incidence of IPF in developed countries, concurrent with alterations in diet, including reduced consumption of fibre [358]. Interestingly, in chronic lung diseases such as IPF, the airway microenvironment has been shown to become increasingly hypoxic, providing a niche in which these anaerobic bacteria can thrive in [134]. Previous studies in the respiratory field suggest that SCFAs may play a key role in the regulation of allergic airway inflammation [163]. Therefore, investigating which metabolites are involved in IPF may aid in the understanding of the pathomechanisms underlying this disease and ultimately provide an opportunity for the development of more targeted therapeutic interventions. Currently, there is limited data available on the lung metabolome of IPF patients. To our knowledge, there has been no study on the impact of the lung metabolome, and in particular the SCFAs, on disease progression in IPF. The overarching aim of this Chapter was to characterise the metabolites present in the BAL supernatant of subjects with IPF and healthy controls. The hypothesis examined was that the respiratory metabolome of IPF subjects is distinct compared to healthy controls and might be contributing to disease progression.

4.2 Aims

- To phenotype the airway metabolic profile in patients with IPF and healthy controls by untargeted ^{1}H NMR spectroscopy and by targeted GC-MS.
- To relate specific changes in the BAL metabolic profile of individuals with IPF to clinical outcomes and disease features.
- To test for associations between the respiratory microbiome, bacterial burden and metabolites.

4.3 Materials and Methods

4.3.1 Study design

Patients with IPF undergoing bronchoscopy with BAL and healthy controls were recruited at the Royal Brompton Hospital (London, UK). Further details on subject recruitment, exclusion and sample collection are detailed in **Chapter 2 (sections 2.1.1-2.1.4)**.

4.3.2 Sample collection and preparation

For metabolomics analysis, BAL supernatants (2 ml) were processed for ^{1}H NMR spectroscopy and for Triple Quadrupole GC-MS. See **Chapter 2 (sections 2.4-2.5)** for further details.

4.3.3 ^{1}H NMR spectroscopy

^{1}H NMR spectroscopy was performed at 300 K (26.85 C) on a Bruker 600 MHz spectrometer (Bruker Biospin, Karlsruhe, Germany). All samples were acquired using the 1D NOESY-presat sequence. BAL supernatant spectra were referenced to TSP at δ 0 ppm. Typical acquisition parameters were: spectral width: 20 ppm; time-domain points: 65536; acquisition time: 2.7 s; number of scans: 32; number of dummy scans: 4 and line broadening: 0.3 Hz. See **Chapter 2 (section 2.4)** for further details.

4.3.4 Triple Quadrupole GC-MS spectrometry

GC-MS of the SCFAs was conducted on an Agilent 7890A (Agilent Technologies, CA, USA), equipped with an automatic liquid sampler (Gerstel, Germany) and coupled to an Agilent 5975C mass selective detector. The acquisition was made using Chemstation software (Hewlett-Packard, CA, USA). The GC was fitted with high polarity, polyethylene glycol (PEG), fused silica capillary column DB-WAXetr (30 m, 0.25 mm id, 0.25 μm film thickness) and helium was used as the carrier gas at 1 mL/min. See **Chapter 2 (section 2.5)** for further details.

4.3.5 Statistical analysis

Continuous variables are presented as means (± standard deviation) and categorical variables as proportions. Data analysis was performed using *R* (version 3.4.3) [359]. Details of statistical analysis performed are described in **Chapter 2 (section 2.4-2.5)**.

4.4 Results

4.4.1 Study subjects

Two hundred and one subjects with IPF and 40 healthy control subjects were enrolled in the study and underwent bronchoscopy. These subjects were the same ones described in **Chapter 3** (refer to **section 3.4.1** for further details).

4.4.2 Metabolome composition in health and disease

4.4.2.1 BAL spectral profiles of healthy controls and IPF subjects

First, an untargeted and hypothesis-driven approach was used to look for changes in the metabolic profiles of the control and IPF groups. BAL supernatant metabolic profiles were measured using ^{1}H NMR spectroscopy, and differences in the metabolic landscape based on diagnosis were identified using univariate and multivariate data analysis techniques. Typical ^{1}H NMR spectra of BAL collected from control and subjects diagnosed with IPF are illustrated in **Appendix 2**. Following pre-processing of the NMR spectra, a total of 23 metabolites were detected of which only 16 could be identified **(Table 4.1)**.

Table 4.1 ^{1}H NMR chemical shift data for metabolites identified in bronchoalveolar lavage (BAL) supernatant of controls and idiopathic pulmonary fibrosis (IPF) subjects.

Illustrating compound name, chemical group, chemical shift in part per million (ppm) and multiplicity.

Compound	Assignment	δ^{1} H	Multiplicity
Butyrate	CH3	0.9	triplet
Propionate	βCH3	1.06	triplet
Ethanol	CH3	1.18	triplet
Lactate	CH3	1.33	doublet
Acetate	CH3	1.92	singlet
Acetone	CH3	2.23	singlet
Sarcosine	CH3	2.74	singlet
Dimethyl sulfone	-	3.16	singlet
Glucose	-	3.25	doublets of doublets
Scyllo-inositol	6xCH	3.36	singlet
3-ketobutyrate	-	3.45	singlet
Glycine	CH	3.57	singlet
Alanine	αCH	3.79	quartet
Oxidised glutathione	αCH2	4.75	multiplet
Formate	CH	8.46	singlet
Choline	CH20H	4.07	multiplet

The peaks arising from the 16 identified metabolites were integrated in order to determine the relative concentration of each metabolite in each sample and to compare these across the two groups. The compound name, chemical group, chemical shift and multiplicity for each metabolite are detailed in **Table 4.1**.

4.4.2.2 Integration of metabolites detected in BAL

To tease out the differences between the groups, the identified metabolites were integrated to compare the concentration (arbitrary units) between controls and subjects with IPF **(Figure 4.1)**.

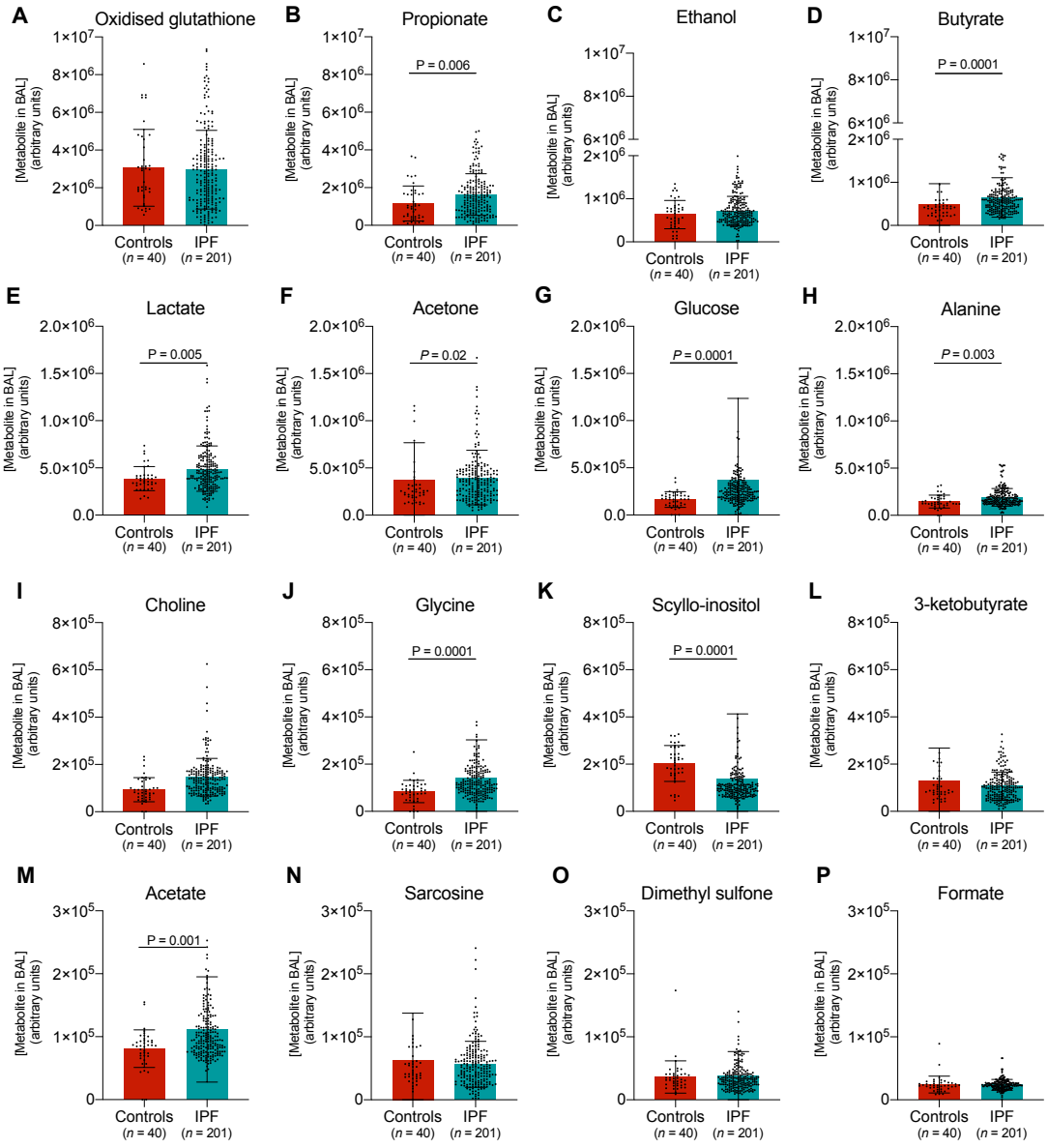


Figure 4.1 Integration of the metabolites detected by ^{1}H nuclear magnetic resonance (NMR) spectroscopy in the bronchoalveolar lavage (BAL) supernatant of healthy controls and subjects diagnosed with idiopathic pulmonary fibrosis (IPF).

Metabolites from healthy controls (*n* = 40) and subjects with IPF (*n* = 201) were integrated. Integrated metabolites: **(A)** oxidised glutathione; **(B)** propionate; **(C)** ethanol; **(D)** butyrate; **(E)** lactate; **(F)** acetone; **(G)** glucose; **(H)** alanine; **(I)** choline; **(J)** glycine; **(K)** scyllo-inositol; **(L)** 3-ketobutyrate; **(M)** acetate; **(N)** sarcosine; **(O)** dimethyl sulfone; **(P)** formate. Data are presented as mean ± standard deviation. Statistical significance tested by Mann-Whitney U test.

The metabolites belonged to the following 7 classes: SCFA, alcohol, acid, ketone, amino acid, amine, carbohydrate. The following metabolites were found to be significantly increased in IPF compared to the control group: propionate, butyrate, lactate, acetone, glucose, alanine, glycine, and acetate. Significantly increased levels of glucose in IPF suggests altered glycolysis in IPF lungs **(Figure 4.1G)**. Furthermore, an increased presence of lactate in IPF lungs supports this hypothesis and may suggest that the products of glycolysis are being shuttled towards lactate production **(Figure 4.1E)**. The carnitine synthesis pathway produces carnitine and glycine [360]. Of these, glycine was found to be increased in IPF compared to controls **(Figure 4.1J)**. Glycine is a nonessential amino acid that can be produced via the metabolism of glucose through the *de novo* serine/glycine biosynthetic pathway and has been recently shown to be required for transforming growth factor-β-induced (TGF-β) collagen production in human lung fibroblasts [361]. Given that the respiratory microbiome is dominated by SCFA-producing bacteria, their increased proportion in IPF compared to controls may suggest a key physiological role.

Multivariate statistical analysis facilitates data identification and analysis by converting the peaks of a spectrum into more easily interpretable coordinates in multidimensional plots of metabolic activity in order to differentiate between classes within the data. Changes in the metabolic profiles between the two groups were further investigated using a PCA biplot **(Figure 4.2)**.

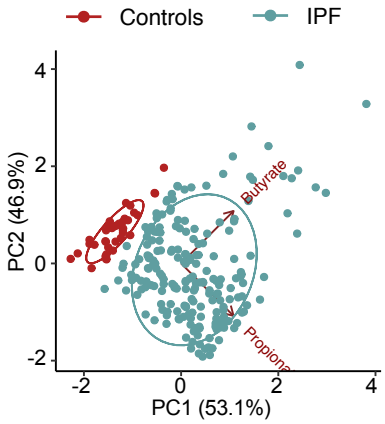


Figure 4.2 Multivariate statistical analysis of the bronchoalveolar lavage (BAL) ^{1}H nuclear magnetic resonance (NMR) spectroscopy spectra.

Principal component analysis (PCA) biplot comparing the BAL metabolic profile of controls (n = 40) and idiopathic pulmonary fibrosis (IPF) subjects (n = 201). PCA elucidated the biochemical variation in the dataset, and the loadings were used to reveal the underlying metabolic differences between the two groups.

The scores plot illustrate a clustering amongst pulmonary metabolic profiles according to diagnosis (53.1% of the variance was explained by PC1; 46.9% of the variance was explained by PC2). The loadings illustrate that butyrate and propionate were the metabolites that were most influential for

the separation in the BAL samples between the groups. The arrows illustrate the direction and indicate with which axis the variable is correlated. Butyrate and propionate were found to be most strongly correlated with the IPF subjects.

4.4.3 Associations between metabolites and clinical variables in IPF

To address whether the metabolic differences observed in the BAL of subjects diagnosed with IPF could be attributed to a clinical phenotype characteristic of IPF, the relationship between clinical parameters and individual metabolites was explored. Firstly, the association of each of the identified metabolites with the BAL differential cell counts was assessed. No correlation was found between the metabolites detected in the BAL of IPF subjects and BAL differential cell counts **(Figure 4.3)**. Unsupervised hierarchical clustering revealed areas of highly correlated metabolites. Interestingly, several of these areas correspond to the biochemical pathway to which these metabolites belong. Since some of these metabolites reside in overlapping pathways, a high correlation between them is expected. The summary of the Spearman rank correlation test is illustrated in **Appendix 3**.

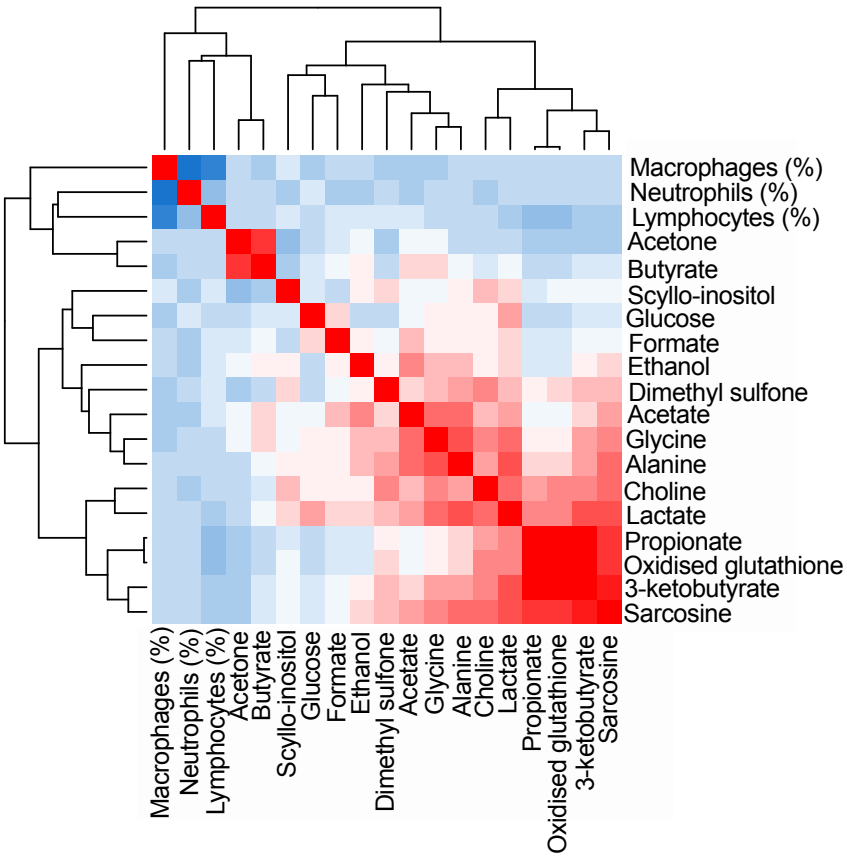


Figure 4.3 Correlation heatmap between metabolites identified in the bronchoalveolar lavage (BAL) of idiopathic pulmonary fibrosis (IPF) subjects and BAL differential cell counts.

Each square represents the Spearman's correlation coefficient (*rho*) between the variable in the row with the variable in the column. Boxes are colour coded according to magnitude and correlation between variables as measured by Spearman's correlation coefficient (dark red: $r^2 = 1.0$; red: $r^2 \geq 0.5 \leq 0.9$; white: $r^2 = 0$; blue: $r^2 \geq -0.1 \leq -0.4$; dark blue: $r^2 \leq -0.5$). Variable order is determined as in hierarchical clustering using the complete linkage method.

Next, subjects with IPF were stratified according to those with stable or progressive disease, and metabolic changes were assessed. No statistically significant metabolic differences were observed when comparing the BAL metabolic profile of subjects with IPF with either stable or progressive disease, as determined by a decline in FVC greater than 10% **(Figure 4.4)**.

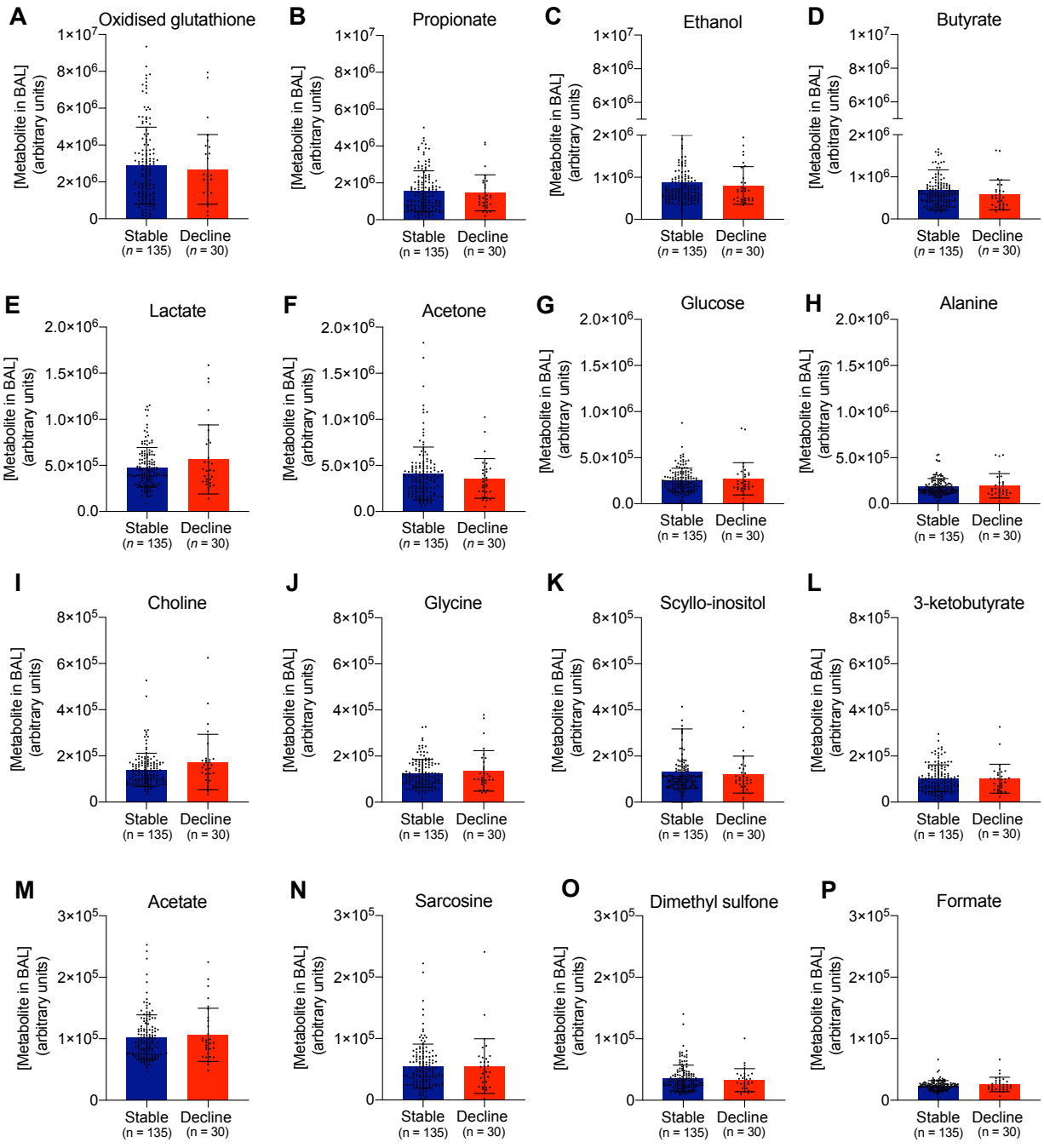


Figure 4.4 The respiratory metabolome in idiopathic pulmonary fibrosis (IPF) subjects with stable or progressive disease.

Comparison of the bronchoalveolar lavage metabolic profile of IPF subjects with stable disease (*n* = 135) or a 10% decline (*n* = 30) in forced vital capacity (FVC). Metabolites compared: **(A)** oxidised glutathione; **(B)** propionate; **(C)** ethanol; **(D)** butyrate; **(E)** lactate; **(F)** acetone; **(G)** glucose; **(H)** alanine; **(I)** choline; **(J)** glycine; **(K)** scyllo-inositol; **(L)** 3-ketobutyrate; **(M)** acetate; **(N)** sarcosine; **(O)** dimethyl sulfone; **(P)** formate. Data are presented as mean ± standard deviation. Statistical significance tested by Mann-Whitney U test.

Similarly, of the 201 subjects diagnosed with IPF, 74 were classified as non-smokers, 117 as ex-smokers and 10 as current smokers. No statistically significant metabolic differences were found when comparing the BAL metabolic profile of subjects with IPF based on smoking history **(Figure 4.5)**.

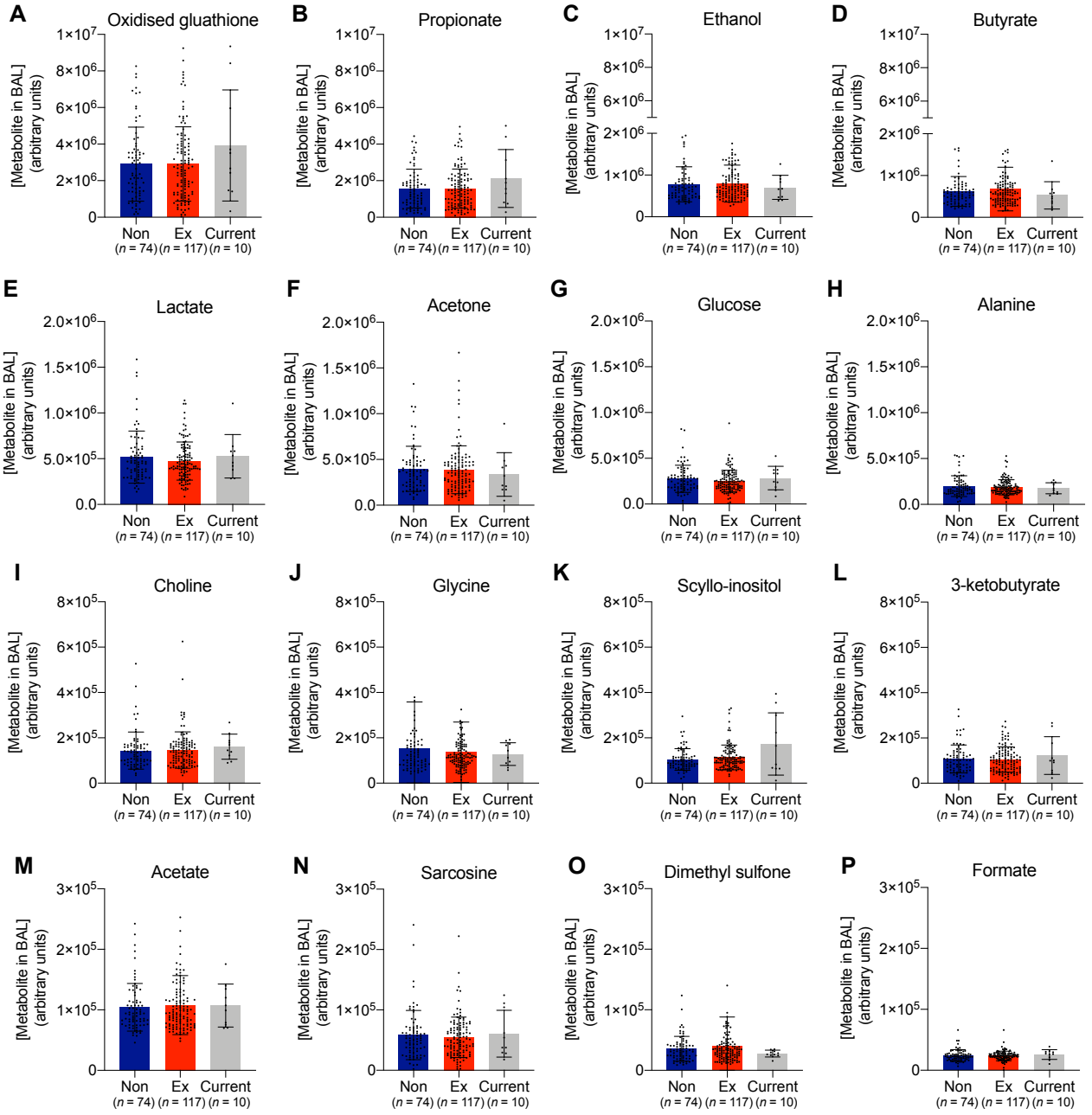


Figure 4.5 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects stratified according to smoking history.

Comparison of the bronchoalveolar lavage metabolic profile of IPF subjects stratified according to smoking history: non-smoker (*n* = 74), ex-smoker (*n* = 117), and current smoker (*n* = 10). Metabolites compared: **(A)** oxidised glutathione; **(B)** propionate; **(C)** ethanol; **(D)** butyrate; **(E)** lactate; **(F)** acetone; **(G)** glucose; **(H)** alanine; **(I)** choline; **(J)** glycine; **(K)** scyllo-inositol; **(L)** 3-ketobutyrate; **(M)** acetate; **(N)** sarcosine; **(O)** dimethyl sulfone; **(P)** formate. Data are presented as mean ± standard deviation. Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test.

To further confirm this, a PCA was constructed to compare the BAL metabolic profile of IPF subjects stratified according to stable or progressive disease **(Figure 4.6A)** or smoking history **(Figure 4.6B)**. No separation was found in the metabolic profile according to FVC or smoking history **(Figure 4.6)**.

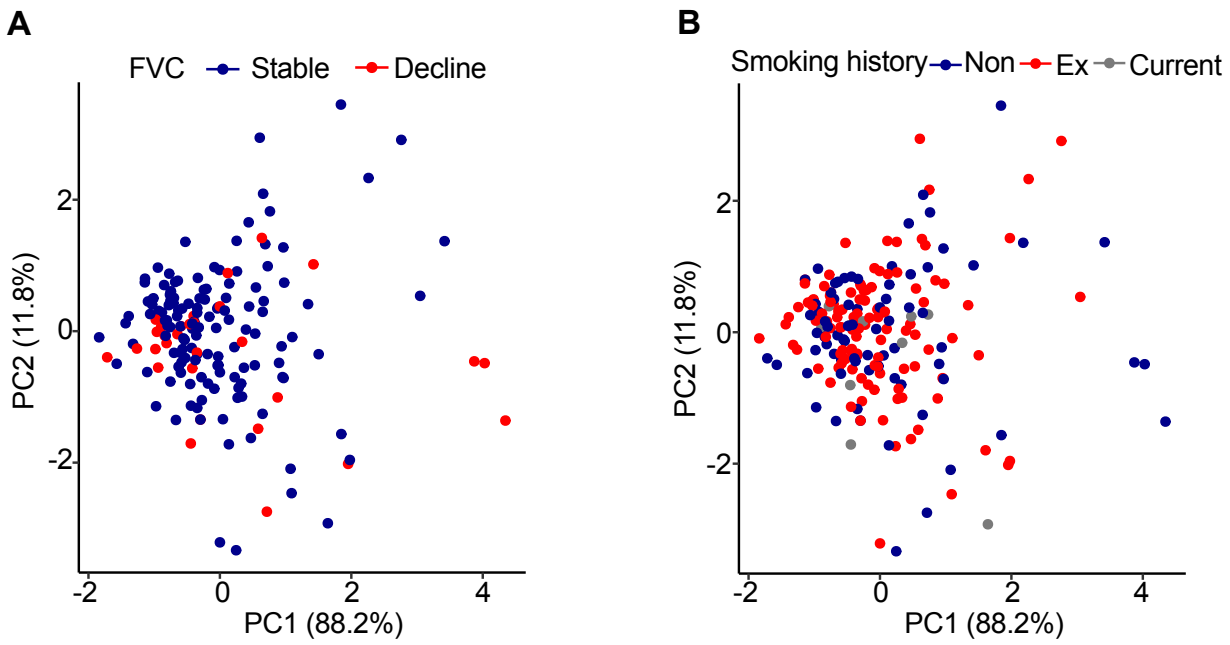


Figure 4.6 Principal component analysis (PCA) of the respiratory metabolome in idiopathic pulmonary fibrosis (IPF) subjects stratified according to forced vital capacity (FVC) and smoking history.

(A) FVC defined as stable or 10% decline. **(B)** Smoking history defined as either non-smoker, ex-smoker or current smoker. PCA on Euclidean distance with statistical significance tested by permutational multivariate analysis of variance (PERMANOVA). PERMANOVA adjusted for age and sex.

4.4.4 Associations between the respiratory microbiome and metabolites in IPF

To study the potential associations between the microbiome and metabolome, a correlation analysis between the two datasets were performed. Firstly, Spearman's correlation coefficients were computed for the full set of normalised individual metabolic features identified by ^{1}H NMR and the relative abundances of microbial genera in the BAL of healthy controls and subjects with IPF **(Figure 4.7)**. In the BAL of controls, pairwise correlations between the bacterial genera and metabolites revealed weak associations **(Figure 4.7A)**. In contrast, in the BAL of IPF subjects, acetate was found to positively correlate with *Prevotella* (*rho* = 0.60, *P* = 0.04) and propionate with *Veillonella* (*rho* = 0.72, *P* = 0.03) **(Figure 4.7B)**.

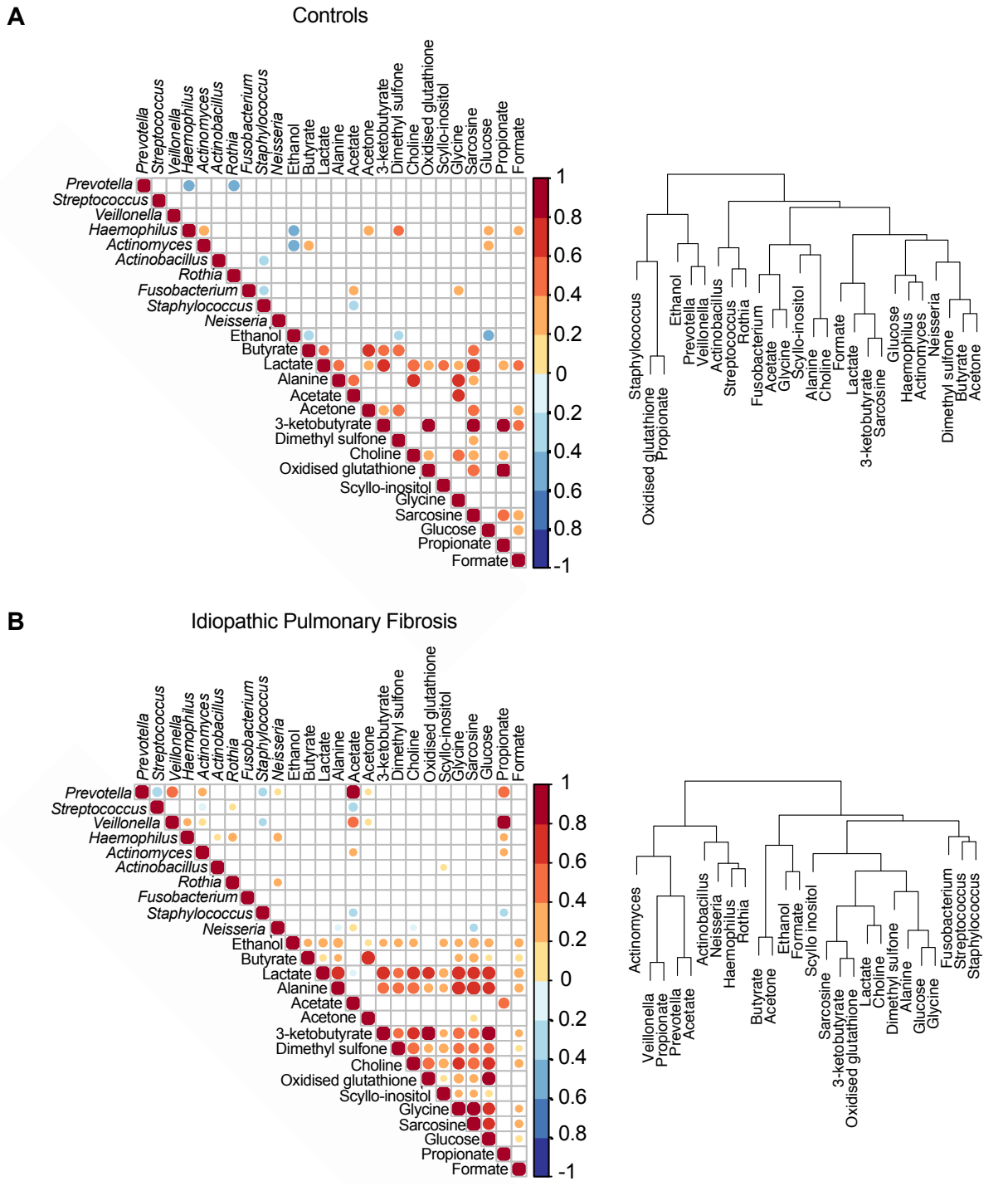


Figure 4.7 Associations between microbial genera and metabolites.

Spearman rank correlation coefficients were computed between the microbiome (at the genus level) and the metabolites in the bronchoalveolar lavage of **(A)** healthy controls and **(B)** subjects with idiopathic pulmonary fibrosis (IPF). Hierarchical clustering dendrogram of bacterial genera and metabolites in BAL was constructed for healthy controls and subjects with IPF.

To further explore the associations between the relative abundance of dominant bacterial genera in BAL samples and the intensity of derived BAL metabolites, sparse partial least-squares-discriminant analysis (sPLS-DA) in canonical mode was performed using MixOmics [362, 363]. Significant correlations were found between the bacterial genera and the metabolites. The circos plot illustrates a significant correlation between the bacterial genera and metabolites regardless of diagnosis **(Figure 4.8)**. The SCFAs acetate, propionate and lactate were found to positively correlate with several bacterial genera including *Prevotella*, *Streptococcus*, and *Veillonella*. The scores plot illustrate a separation in both datasets according to diagnosis **(Figure 4.9A),** and the loadings plot of each metabolic feature and bacterial genera explain the features which are driving the separation between the two groups **(Figure 4.9B)**.

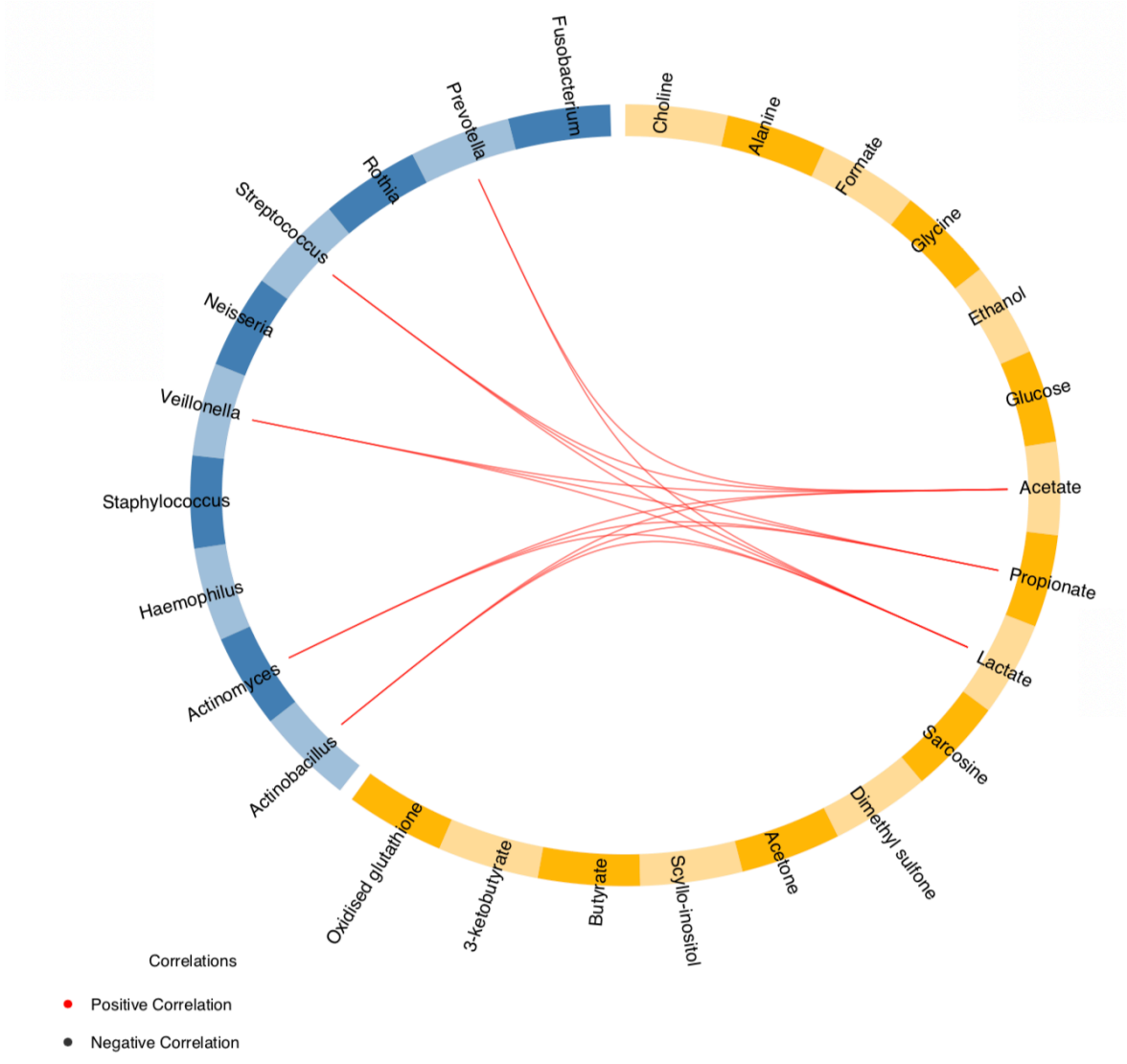


Figure 4.8 Circos plot of horizontal sparse partial least-squares-discriminant analysis displaying correlations between the respiratory microbiome and metabolome datasets.

Depicting correlations between bacterial genera (*n* = 10) and metabolites (*n* = 16). Positive correlations (r > 0.7) are displayed in red. Relative abundance of bacterial genera ≥ 0.01%.

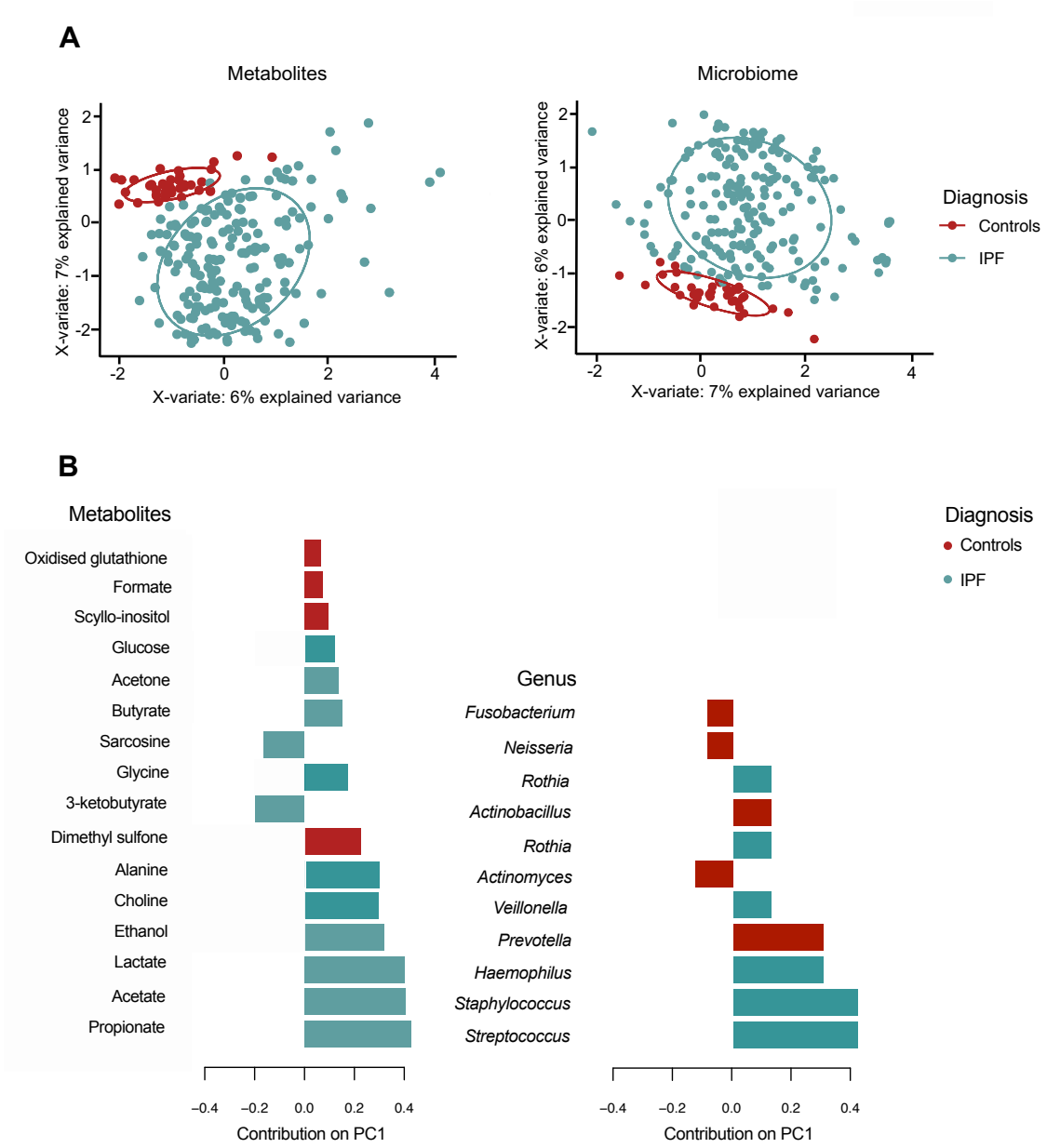


Figure 4.9 Sparse partial least-squares-discriminant analysis depicting the contribution of the selected features in driving the separation between healthy controls and individuals with idiopathic pulmonary fibrosis (IPF).

Illustrating the **(A)** Scores plot for the metabolome and microbiome datasets. **(B)** Loading plot of each feature selected on the first component in each dataset. The respective colours indicate the class with the maximal expression value for each metabolite and bacterial genera.

4.4.5 Associations between bacterial burden and metabolites in IPF

Subjects diagnosed with IPF were stratified into tertiles based on their levels of bacterial burden (low, middle, high). The following metabolites were found to be significantly increased in subjects with a higher bacterial load compared to patients with the lowest bacterial burden: oxidised glutathione, propionate, butyrate, lactate, and acetone. Conversely, scyllo-inositol was increased in those patients possessing a lower bacterial load compared to those with the highest bacterial burden **(Figure 4.10)**.

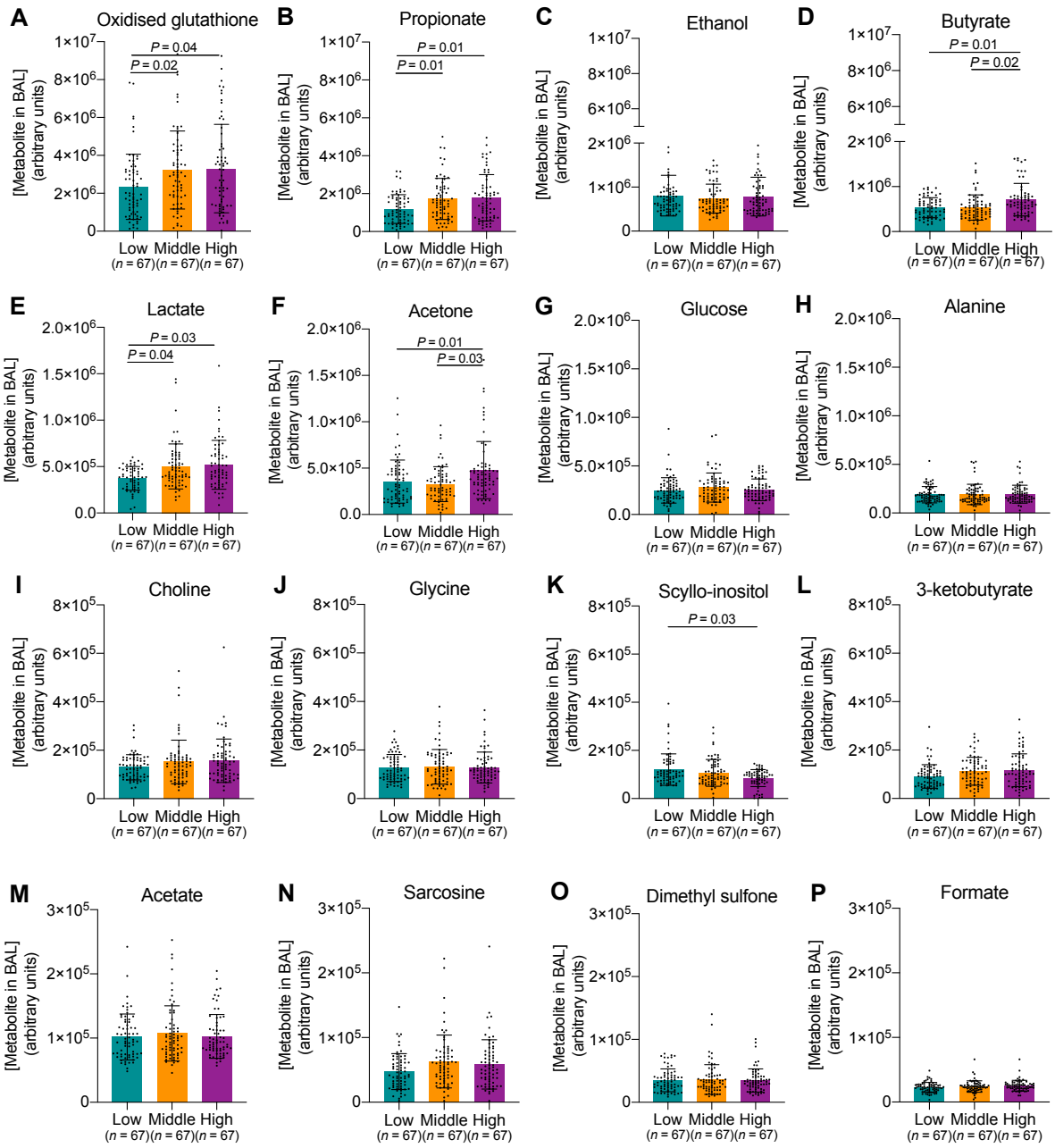


Figure 4.10 The respiratory microbiome in idiopathic pulmonary fibrosis (IPF) subjects stratified according to bronchoalveolar lavage (BAL) bacterial burden tertiles.

Comparison of the BAL metabolic profile of IPF subjects stratified according to bacterial burden tertiles (low: 4.15 x 10^3 – 3.76 x 10^4; middle: 3.78 x 10^4 – 2.11 x 10^5; high: 2.12 x 10^5 – 2.42 x 10^7 16S rRNA gene copies/ml of BAL): low (*n* = 67), middle (*n* = 67), and high (*n* = 67). BAL bacterial burden quantified using quantitative polymerase chain reaction and is expressed as log10 (16S rRNA gene copies/ml of BAL). Data are presented as mean ± standard deviation. The following metabolites were investigated: **(A)** oxidised glutathione, **(B)** propionate, **(C)** ethanol, **(D)** butyrate, **(E)** lactate, **(F)** acetone, **(G)** glucose, **(H)** alanine, **(I)** choline, **(J)** glycine, **(K)** scyllo-inositol, **(L)** 3-ketobutyrate, **(M)** acetate, **(N)** sarcosine, **(O)** dimethyl sulfone, and **(P)** formate. Statistical significance tested by Kruskal-Wallis test with Dunn's Multiple Comparison test adjusted with the Bonferroni method.

Based upon the previously confirmed a relationship between bacterial burden and disease progression in IPF **(Chapter 3, section 3.4.4.3)**, a nonparametric Spearman rank correlation was used to look for associations between the BAL bacterial load and the metabolites **(Table 4.2)**. Weak

correlations were observed between oxidised glutathione, lactate, acetone, scyllo-inositol, 3-ketobutyrate, formate and bacterial burden. The only two metabolites showing a strong positive correlation with bacterial burden were the SCFAs propionate and butyrate.

Table 4.2 Spearman's rank correlations (*rho*) between the metabolites detected in the bronchoalveolar lavage (BAL) of idiopathic pulmonary fibrosis (IPF) subjects and BAL bacterial burden.

Comparison	Spearman's *rho*	*P*-value
Bacterial burden vs oxidised glutathione	0.17	<0.05
Bacterial burden vs propionate	0.47	<0.0001
Bacterial burden vs ethanol	-0.04	NS
Bacterial burden vs butyrate	0.38	<0.01
Bacterial burden vs lactate	0.24	<0.05
Bacterial burden vs acetone	0.28	<0.05
Bacterial burden vs glucose	0.06	NS
Bacterial burden vs alanine	0.03	NS
Bacterial burden vs choline	0.13	NS
Bacterial burden vs glycine	0.02	NS
Bacterial burden vs scyllo-inositol	-0.30	<0.05
Bacterial burden vs 3-ketobutyrate	0.16	<0.05
Bacterial burden vs acetate	0.42	NS
Bacterial burden vs sarcosine	0.13	NS
Bacterial burden vs dimethyl sulfone	0.03	NS
Bacterial burden vs formate	0.15	<0.05

BAL bacterial burden quantified using quantitative polymerase chain reaction and log-transformed to approximate normal distribution (log10 16S rRNA gene copies/ml of BAL). NS: not significant.

Given the known association between increased bacterial load and increased risk of mortality in IPF, the next question which was addressed was whether having an increased abundance of the SCFAs in the BAL, particularly propionate and butyrate associated with survival independent of burden. Subjects with IPF were stratified into two groups based on their levels of propionate and butyrate, and a multivariable Cox proportional hazards model adjusted for age, sex, burden, smoking status and baseline % predicted FVC was used to examine progression-free survival. No associations with survival were found for either propionate (hazard ratio: 1.28, 95% confidence interval: 0.49-3.30, *P* = 0.609) or butyrate (hazard ratio: 1.28, 95% confidence interval: 0.49-3.30, *P* = 0.613).

4.4.6 Triple Quadrupole GC-MS targeted analysis of the SCFAs in BAL samples

While the effects of the SCFAs, such as acetate, propionate and butyrate, are now well documented in the gut and in a number of chronic inflammatory diseases [161, 357], little is known about their impact on the development of fibrosis. To address this, a targeted GC-MS approach was used to quantify the SCFAs in the BAL supernatant of a subset of controls and subjects diagnosed with IPF. These subjects were the same ones included in the ^{1}H NMR study **(section 4.4.2)** for which matched BAL supernatant had been collected. GC-MS was utilised because of its greater sensitivity compared to ^{1}H NMR. The clinical characteristics and demographics are outlined in **Table 4.3**. Healthy controls and IPF subjects were matched for age, sex and smoking history. Individuals with IPF were predominantly men (80%), with a mean age of 72 years ± 8, and had moderately severe disease at enrolment (DL_{CO}, 41% predicted; FVC, 72% predicted). The control subjects were also predominantly male (90%), had better lung function compared to the IPF cohort (DL_{CO}, 94% predicted; FVC, 98% predicted) and had a mean age of 66 years ± 6.

Table 4.3 Clinical characteristics and demographics of the study subjects.

Parameter	Controls (*N* = 10)	IPF (*N* = 10)	*P*-value*
Age, yrs (± SD)	66 (6)	72 (8)	NS
Male, *n* (%)	9 (90)	8 (80)	NS
Baseline PFT			
FVC, % predicted (± SD)	98 (5)	72 (15)	0.0001
FEV_1 % predicted (± SD)	83 (26)	78 (17)	NS
DL_{CO} % predicted (± SD)	94 (10)	41 (17)	0.0001
Smoking status			
Ex-/current-smoker, *n* (%)	6 (60)	6 (60)	NS
BAL cell differentials			
Macrophages % (± SD)	73 (9)	69 (9)	NS
Neutrophils % (± SD)	6 (3)	10 (4)	<0.05
Lymphocytes % (± SD)	8 (6)	13 (5)	<0.05

Data are presented as percentage value or mean ± standard deviation (SD) as appropriate. PFT: pulmonary function test; FEV_1: forced expiratory volume in 1 second; FVC: forced vital capacity; DL_{CO}: diffusing capacity of the lung for carbon monoxide; BAL: bronchoalveolar lavage. **P*-values are calculated by Mann-Whitney U test or Fisher's exact test between controls and idiopathic pulmonary fibrosis (IPF) subjects.

Differential cell counts were performed in the control and IPF groups with identification of alveolar macrophages, lymphocytes and neutrophils. Alveolar macrophages predominated the BAL of both groups (controls 73% ± 9; IPF 69% ± 9). Mixed alveolitis with predominant neutrophils count was observed in IPF (neutrophils 10% ± 4, lymphocytes 13% ± 5)**.**

The following SCFA were quantified: acetate, propionate, isobutyrate, butyrate, isovalerate, 2-methylbutyrate, valerate, caproate, lactate and α-hydroxybutyrate. Amongst these, isobutyrate, isovalerate, 2-methylbutyrate, caproate, and α-hydroxybutyrate were undetectable in the BAL of healthy controls and subjects with IPF. In contrast, the SCFAs acetate, propionate, lactate and butyrate were found to be increased in the BAL of IPF subjects compared to controls **(Figure 4.11)**. To validate these findings, the SCFAs detected in the BAL supernatant of the controls (*n* = 10) and subjects diagnosed with IPF (*n* = 10) were correlated with the ^{1}H NMR data of the same individuals, which confirmed strong correlations between the metabolites **(Appendix 4)**.

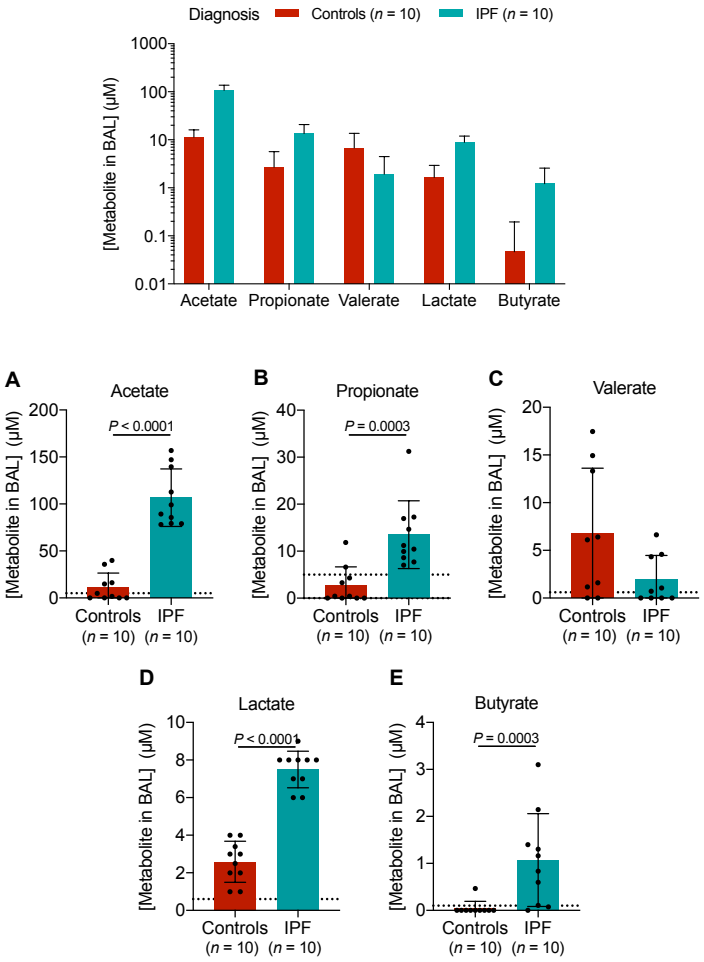


Figure 4.11 Gas chromatography-mass spectrometry (GC-MS) of the short-chain fatty acids (SCFAs) in the bronchoalveolar lavage (BAL) supernatant of healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Metabolites that were quantified with GC-MS: **(A)** acetate, **(B)** propionate, **(C)** valerate, **(D)** lactate, and **(E)** butyrate. Data are presented as mean ± standard deviation. Statistical significance tested by Mann Whitney U test.

The inter-individual variability in the BAL metabolic profile of the healthy controls and subjects with IPF was further explored by comparing the SCFA composition in the BAL of the subjects **(Figure 4.12)**. Significant interindividual variability in the BAL metabolic profile of both groups was observed, which reflects the fact that, although as a whole there are differences between the groups, the metabolic signature of each individual is unique. For example, in the control group, the BAL supernatant of subject 1 was dominated by acetate (61%) and valerate (20%), whilst in the BAL supernatant of subject 2 there was a higher proportion of valerate (39%), and lactate (25%). In comparison, in the IPF group, subject 1 and subject 2 had an overall higher proportion of acetate (76% and 62%) and propionate (16% and 30%).

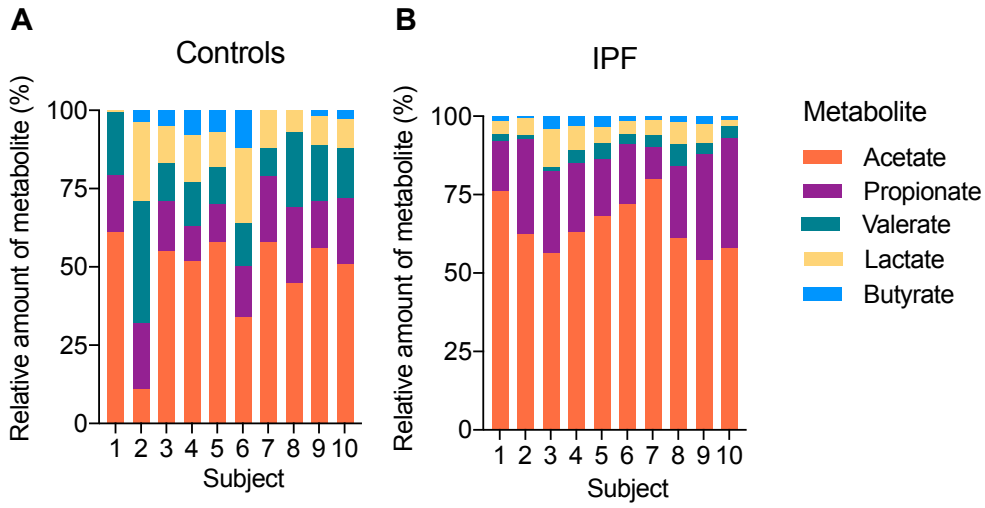


Figure 4.12 Interindividual variability in the short-chain fatty acid (SCFA) landscape of the bronchoalveolar lavage (BAL) from healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Comparing ten **(A)** healthy controls and ten **(B)** IPF patients. Relative proportions of each metabolite are represented as a function of the total metabolite concentration (μM)/patient.

Hierarchical clustering of the BAL microbiome in the ten healthy controls and ten IPF subjects was conducted in order to address whether some of these differences in SCFA-production could be due to distinct changes in the BAL microbiome. To elucidate this, heatmaps representing the relative abundance of the top bacterial phyla and bacterial genera in the BAL of these subjects were constructed **(Figure 4.13)**. At the phylum level, the microbiota of control subjects was dominated by *Bacteroidetes* and *Actinobacteria* **(Figure 4.13A)**. In contrast, there was an increased proportion of *Firmicutes* in the BAL of IPF subjects. This was also reflected at the genus level, where *Prevotella* was found to be increased in the BAL of controls and *Streptococcus* and *Veillonella*, known producers of lactate and propionate, respectively, in the BAL of subjects with IPF **(Figure 4.13B)**. In IPF, a positive correlation was observed between *Streptococcus* and propionate (*rho* = 0.62, *P* = 0.02), as well as *Veillonella* and propionate (*rho* = 0.60, *P* = 0.03).

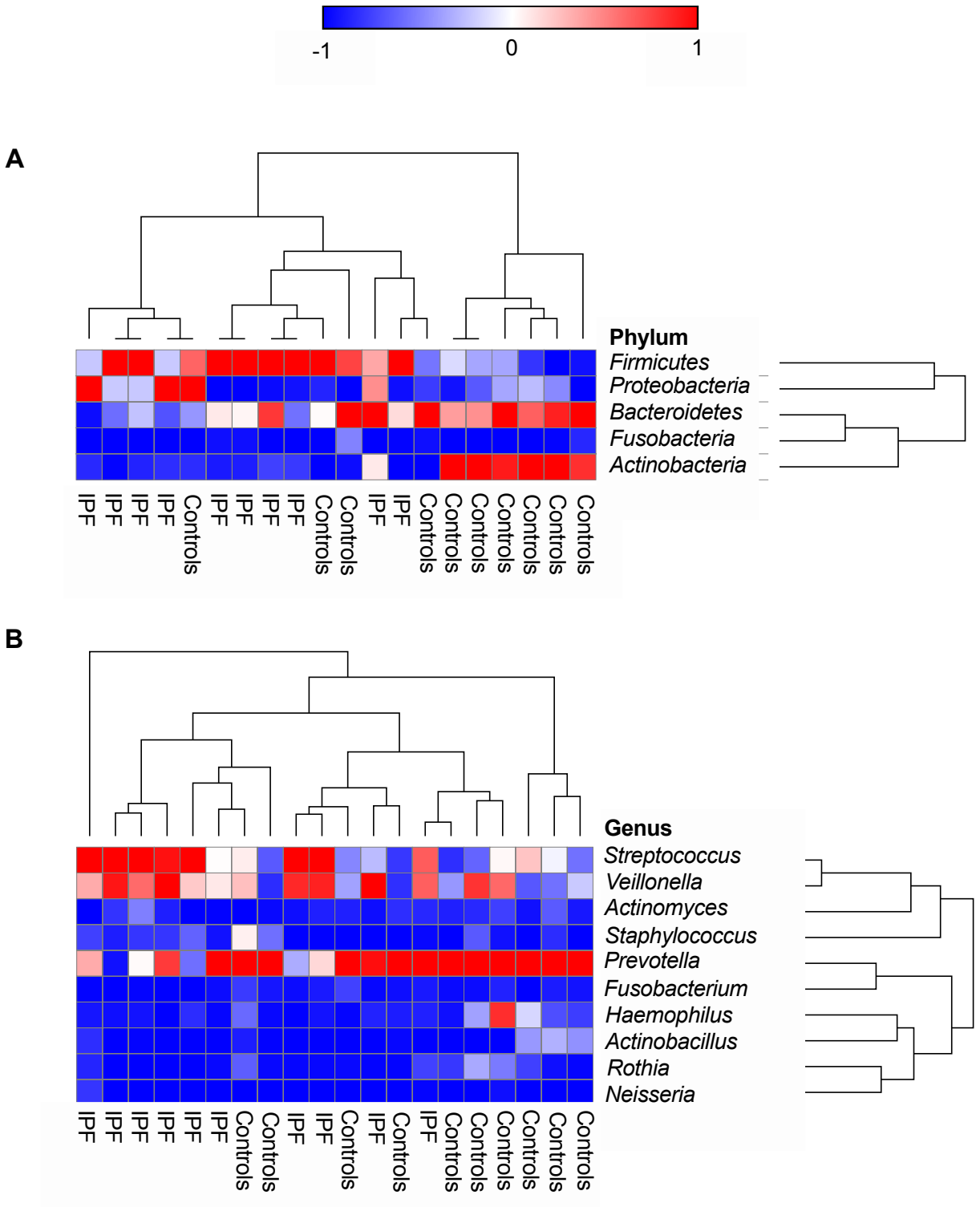


Figure 4.13 Hierarchical clustering of the bronchoalveolar lavage (BAL) microbiome in healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Heatmaps representing the relative abundance of **(A)** bacterial phyla and **(B)** bacterial genera across the matched BAL samples from the controls (*n* = 10) and IPF (*n* = 10) subjects included in the gas chromatography-mass spectrometry targeted analysis. Rows (microbial taxa at each level) and columns (samples) were ordered by hierarchical clustering. Heatmap colour (dark blue to dark red) displays the row scaled relative abundance of each taxon across all samples.

4.5 Discussion

As of yet, no study has investigated the global BAL metabolic profile from IPF patients using both untargeted and targeted approaches. The overarching goal of this Chapter was to characterise the metabolites present in the BAL supernatant of healthy controls and subjects with IPF using untargeted ^{1}H NMR spectroscopy as well as to quantify the SCFAs using targeted GC-MS. Specifically, it was hypothesised that there would be differences in the metabolic profiles of healthy controls and subjects with IPF and that certain metabolites increased in IPF may be contributing to disease progression. Using ^{1}H NMR spectroscopy, it was demonstrated that the overall metabolic profile in BAL supernatant was different between healthy controls and subjects with IPF. Specifically, in IPF, acetate, butyrate, propionate, lactate, acetone, glucose, glycine, and alanine were found to be increased. In IPF, no differences in metabolites were observed based on BAL differential cell counts, smoking history and FVC (stable or progressive). However, specific associations between the respiratory microbiome and bacterial metabolites were observed in IPF; the BAL supernatant was characterised by a unique SCFA signature which was associated with the bacterial burden. The results of this Chapter support a greater understanding of the metabolic basis of pathogenesis of IPF formation.

The current paradigm for IPF pathogenesis suggests that prolonged exposure to environmental stimuli causes recurrent AEC injury or AEC death. These abnormally activated AECs, in turn, secrete pro-inflammatory mediators, which stimulate fibroblast activation and proliferation, ultimately resulting in the excessive deposition of extracellular matrix within the lungs [15]. Although a diagnosis of IPF typically requires the exclusion of infectious causes, the idea that alterations in the airway microbiota may play a role in the pathogenesis and progression of the disease has recently gained significant attention. In agreement with the literature, results from **Chapter 3** revealed that the lungs of patients with IPF are characterised by increased bacterial loads, changes in bacterial community composition, and decreased bacterial diversity, compared to those of healthy individuals [114, 115, 149]. Furthermore, previous work has demonstrated that the expression of genes involved in host defence against microbes, positively correlate with BAL bacterial burden, suggesting that perturbations in the airway microenvironment may be acting as a persistent antigenic stimulus for repetitive injury to AEC in IPF [309]. However, there is limited data looking at the products of the airway microbiota, specifically the metabolites, and their involvement in the development and progression of IPF.

It is now well recognised that the airways consist of diverse habitats which shape the composition of the lung microbiota. The microbial communities that reside in the lung differ in metabolite production. Strikingly, even bacteria of the same species, but belonging to distinct communities, produce different metabolites [153]. Metabolic profiling is increasingly being used in the investigation of molecular processes underlying several diseases. Metabolomics is defined as a comprehensive and quantitative analysis of all metabolites in a system [299]. It provides a multifactorial phenotypic signature of the physiology of an organism and can give an insight into the mechanisms of disease. Differences in the metabolic profiles of individuals may be used to discriminate amongst clinical phenotypes and provide an insight into the biological processes underlying the pathophysiology of the disease. Following exclusion of drug metabolites (e.g. lidocaine), 16 metabolic features were isolated in the BAL from IPF patients by non-targeted ^{1}H NMR spectroscopy-based metabolomics. The metabolites belonged to the following 7 classes: SCFA, alcohol, acid, ketone, amino acid, amine, carbohydrate. Out of the 16 identified metabolites, the following were found to be increased in IPF: glucose, lactate, glycine, alanine, acetone, acetate, propionate and butyrate.

The catabolism of major nutrient substrates, including sugars, amino acids and fatty acids, provides acetyl-CoA for the tricarboxylic acid (TCA) cycle to generate energy and metabolites for the biosynthesis of nucleotides, proteins and lipids [364]. Cells require adenosine triphosphate (ATP) in order to drive various cellular processes and maintain homeostasis [365]. Metabolic homeostasis is crucial for regulating cell proliferation, and differentiation and dysregulation has been shown to contribute to several inflammatory responses [366]. These cellular processes have been implicated in the pathogenesis of chronic lung disorders, including IPF. However, the role of these metabolic dysregulations in IPF are not fully understood. Glucose is one of the main sources of ATP production for most tissues [367]. Upon transport into the cell, glucose can be converted to energy as follows: by the activation of glycolysis under anaerobic conditions, by complete oxidation under anaerobic conditions, and by the pentose phosphate pathway [365]. Several fibrotic diseases have described a perturbation in the glycolytic pathway. In the skin, upregulation and downregulation of extracellular matrix deposition have been shown to be regulated by changes in glycolysis and fatty acid oxidation [368]. In the lungs of IPF patients, a study conducted by Yin's group found that the glycolytic enzyme hexokinase 2 was highly abundant in lung fibroblasts from patients with IPF and that TGF-β induced the accumulation of hexokinase 2 and stimulated glycolysis in human lung fibroblasts [369]. Zhao and colleagues found altered metabolite production in IPF lung samples and disrupted glycolysis compared with healthy controls [360]. In agreement with these findings,

glucose was found to be significantly increased in the BAL of IPF subjects compared to healthy controls.

During oxygen deprivation, pyruvate is metabolised to lactate in the cytosol [370]. Lactate is a microbiota-derived SCFA, and several studies have supported its role in health and disease [371]. Kottomann and colleagues investigated the role of lactic acid in IPF [372]. The authors found that lactic acid is elevated in the lung tissue of patients with IPF and suggested that it induced myofibroblast differentiation via the activation of TGF-β and subsequently increased the expression of hypoxia-inducible factor-1-α and LDH-5 [372]. Lactate is produced by LDH5 and is then exported by monocarboxylate transporter (MCT)-4 and taken up by adjacent cells expressing MCT-1. It can then be oxidised to pyruvate to provide additional energy through the TCA cycle. In this Chapter, lactate was found to be increased in the BAL supernatant of subjects with IPF compared to healthy controls. Detection of lactate was not unexpected, given that the lungs are known to produce lactate, and this metabolite has been formerly detected in airway secretions [373, 374]. Such finding may suggest that lactate is elevated in the presence of inflammation, which supports previous studies that have shown that lung inflammation results in an increased production of this metabolite [375]. To further confirm this, future research should aim to quantify inflammatory cytokines in the BAL of subjects with IPF and determine whether there are any associations with lactate. Elevated levels of lactate in the BAL supernatant may also reflect the low levels of oxygen concentrations present in the lower airways of patients with IPF. As the cells of the lower airways reside in this increasingly hypoxic environment, they may have to rely on anaerobic metabolism for energy, which could culminate in the elevated production of the lactate observed in our samples. Collectively, significantly increased levels of glucose and an increased presence of lactate in IPF, may suggest altered glycolysis in IPF lungs and supports the hypothesis that the products of glycolysis are being shuttled towards lactate production.

Although the glycolytic pathway can lead to lactate or pyruvate production, glycolysis can be diverted for the *de novo* synthesis of serine and glycine. This process is mediated by the phosphoglycerate dehydrogenase (PHGDH) and serine hydroxymethyltransferase 2 (SHMT2) enzymes, which have been shown to be transcriptional targets of TGF-β1 in lung fibroblasts [361, 376]. Of these, glycine was found to be increased in the BAL supernatant of subjects with IPF compared to healthy controls. Glycine is a nonessential amino acid and is the most abundant amino acid residue in collagen, accounting for approximately 33% of its content. Fibroblasts rely on the *de novo* serine/glycine synthesis pathway for collagen production [377]. This was recently shown in a

study by Nigdelioglu and colleagues, which demonstrated that the *de novo* serine/glycine synthesis is required for TGF-β-induced collagen production in human lung fibroblasts [361]. Furthermore, the authors observed an increased expression of PHGDH and SHMT2 in the lungs of patients with IPF, further validating the involvement of this pathway in the pathogenesis of this chronic lung disease [361].

Remodelling of the extracellular matrix is a key feature of IPF. Elastin is the main component of elastic fibres in matrix tissue which provides tissue resilience and elasticity, and has been shown to be dysregulated in IPF lung tissue [378]. Elastin fibres are degraded through a process called elastinolysis, which has been reported to occur in the lungs of patients with IPF [378]. The amino acid alanine was also found to be increased in the BAL supernatant of subjects with IPF. Given the high abundance of alanine in elastin, an increased presence of alanine in the BAL could reflect an altered turnover of elastin occurring in the lungs of subjects with IPF. This finding is further supported by a recent study of Gaugg and colleagues where they found significantly elevated levels of alanine in the breath of patients with IPF. This further coincided with the increased abundance of the amino acid in the pulmonary tissue, which could reflect progressive fibrosis [379].

Acetone is the simplest and smallest ketone compound found naturally in the human body [380]. Although there is contrasting evidence as to whether acetone is metabolically inert or not, some studies have shown this metabolite can be metabolised either by cytochrome P450 2E1 via methylglyoxal to D-lactate and pyruvate and ultimately to glucose, or by another pathway via propylene glycol to pyruvate, lactate and acetate [381-383]. Using ^{1}H NMR spectroscopy metabolic profiling of exhaled breath condensate in patients with CF, acetone was found to be increased in CF compared to healthy controls and was amongst the most discriminant metabolites between the groups [384]. Furthermore, acetone has been previously detected by ^{1}H NMR spectroscopy in the BAL of children with CF with varying degrees of inflammation [385]. The presence of acetone in the BAL supernatant is expected given that this compound is known to be present at relatively high and stable concentrations within the human airways [386]. The two pathways that contribute to the conversion of acetone to glucose are the methylglyoxal and the propanediol pathways [387]. In fasting humans, Reichard and colleagues have shown that up to 11% of plasma glucose production may be derived from acetone [388]. Nonetheless, the presence of this ketone body within the BAL supernatant of IPF patients is still poorly understood. Airway acetone levels can be affected by diabetic ketoacidosis [386]; however, none of the subjects included in this study had diabetes.

The SCFAs are straight-chain saturated fatty acids with fewer than six carbon atoms, which have widespread effects on the gut but have also been shown to exert their effects in peripheral tissues such as the lung [351]. Using untargeted ^{1}H NMR, the SCFAs acetate, butyrate, propionate and lactate were found to be elevated in the BAL supernatant of subjects with IPF compared to healthy controls. This was further confirmed by conducting targeted GC-MS on the same samples, which facilitated quantification of the levels of these SCFAs in health and disease. In health, the pulmonary microenvironment does not generally permit excessive bacterial growth. However, alterations that occur in the context of inflammation, such as increased pH, vascular permeability, temperature changes, mucus production, cytokines, free ATP, and growth factors may create selectively favourable local growth conditions for bacterial colonisation of the lower airways [129]. Given that the respiratory microbiome is dominated by SCFA-producing bacteria, their increased proportion in IPF compared to controls may suggest a key physiological role. In the BAL of controls, pairwise correlations between the bacterial genera and metabolites revealed weak associations. In contrast, in the BAL of IPF subjects, acetate was found to positively correlate with *Prevotella* and propionate with *Veillonella.* This is of interest because studies in the gut have shown that acetate is produced by most of the enteric bacteria, including *Prevotella* species via the Wood-Ljungdhal and acetyl-CoA pathways [389, 390]. Propionate is produced by *Veillonella* via the succinate, acrylate and propanediol pathways [153]. In contrast, distinct families within the *Firmicutes* are capable of producing butyrate from two molecules of acetyl-CoA [391]. Therefore, elevated BAL SCFA concentrations in subjects with IPF could reflect the metabolism of lung resident bacteria such as *Streptococcus* and *Veillonella*, which degrades pyruvate into end products of metabolism, including lactate and propionate [392]. However, significant inter-individual variability in the BAL metabolic profile of both groups was observed, which reflects the fact that, although on a whole there are differences between the groups, the metabolic signature of each individual is unique. Overall, compared to healthy controls, the BAL supernatant of IPF subjects was characterised by a unique SCFA signature. Furthermore, the SCFAs propionate and butyrate positively correlated with increased bacterial burden although there were no associations with survival. Instead, in IPF, no differences were found in the metabolites when stratified according to clinical features including BAL differential cell counts, smoking history and FVC. However, the strong correlation between all significantly altered compounds further suggests a shared metabolic process.

A clear link between the gut and lungs termed the ‘gut-lung axis’ has been recently established, adding complexity in that besides disturbances in the airway microbiome, imbalances in the intestinal microbiota may also play a role in the development of chronic lung diseases, including

asthma, COPD, and CF [393-395]. Several extrinsic and intrinsic factors, such as drugs, diet, age, genetics, sex and body mass index, can alter the composition of the gut microbiota and promote intestinal inflammation [163]. Subsequently, this may lead to disruptions in intestinal barrier integrity, culminating in the translocation of bacterial antigens from the intestine, into the bloodstream, which may promote the initiation of a systemic inflammatory response [396]. For example, dietary fibre intake can alter gut and lung microbial composition by shifting the bacterial ratio [167, 172]. Recent evidence indicates an association between imbalances in the intestinal microbiome and the development of lung disorders and respiratory infections in mice [167]. Consumption of a high fibre diet leads to increased production of SCFAs in the circulation and the gut, which is protective against allergic asthma [167]. Chronic helminth infection has been shown to alter the intestinal habitat, increasing SCFA production, leading to an attenuation of house dust mite-induced allergic airway inflammation [397]. However, our current understanding of the gut-lung axis rests on mouse work, and this cross-talk has not yet been studied in the context of IPF [398].

Indeed, the pro- and anti-inflammatory effects of these dietary fibre fermentation products are now well documented in a number of chronic inflammatory diseases [161, 357]. However, studies on the pro-inflammatory roles of these SCFAs in the lungs are just emerging. Indeed, in addition to anti-inflammatory effects, pro-inflammatory roles of SCFAs have also been reported. Mirković and colleagues found that anaerobic bacteria identified in the lungs of CF patients secrete a distinct set of SCFAs, which trigger the release of IL-8 from CF bronchial epithelial cells. This effect was reduced in normal bronchial epithelium compared to CF [355]. Similarly, Ghorbani and colleagues provided evidence that SCFAs stimulate the release of proinflammatory cytokines, including IL-8 and IL-6 from CF bronchial epithelial cells. The secretion of SCFAs was associated with increased neutrophil recruitment, suggesting that SCFAs may drive innate immunity in CF airways [353]. Recently, it was found that treatment of primary HLFs and ASM cells with supraphysiological concentrations of butyrate or propionate in combination with TNF-α, resulted in a significantly greater release of IL-6 and IL-8 from HLFs and ASM cells, compared to treatment with TNF-α alone [399]. However, while the role of SCFAs in CF has been investigated, very few studies have addressed the impact of SCFAs and other microbial metabolites on the development of fibrosis. Specifically, SCFA receptors are expressed on bronchial epithelial cells, upregulated in disease and could be driving pulmonary fibrosis via the activation of a number of mechanistic downstream pathways implicated in fibrinogenesis. The *in vitro* work conducted in **Chapter 5** will address the functional impact that these metabolites have on the development of fibrosis.

4.6 Conclusions

Metabolic changes are increasingly being recognised as a hallmark of pulmonary fibrosis, and studies suggest that metabolic dysregulation may contribute to the pathogenesis of the disease. In this Chapter, it was demonstrated that there are clear differences in the metabolic profile of subjects with IPF compared to healthy controls. Specific metabolic pathways, such as the glycolytic pathway, which may contribute to disease pathogenesis, are altered in IPF patients compared to healthy controls. The BAL supernatant of IPF subjects was characterised by a unique SCFA signature. Specifically, the SCFAs propionate and butyrate were found to associate with the bacterial burden. Measurement of these metabolic changes may serve as potential biomarkers for the diagnosis of the disease and to better understand how specific metabolic changes may affect immune response and inflammation. However, the nature of this study was purely descriptive, and future work needs to move from correlative studies to mechanistic studies, which will be key to better understand the role of the microbiota and its metabolites in the development of pulmonary fibrosis. Clarifying the mechanistic role of SCFAs in the context of this disease will be necessary to gain a better understanding of the host-microbe interactions that underpin IPF.

5 **Chapter 5**: Investigating the Influence of the Short-Chain Fatty Acids on Primary Human Bronchial Epithelial Cells and Primary Human Lung Fibroblasts from Healthy Individuals and Subjects with Idiopathic Pulmonary Fibrosis

5.1 Introduction

IPF is characterised by progressive and irreversible scarring within the pulmonary interstitium and is associated with architectural destruction, impaired gas exchange and a median survival of 3 years post-diagnosis [16, 400]. This loss of epithelial cells and abnormal tissue repair results in the accumulation of fibroblasts and myofibroblasts, and excessive deposition of extracellular matrix components which over time leads to interstitial remodelling and respiratory failure. Current evidence suggests that IPF is caused by an abnormal wound healing driven by persistent or repeated micro injuries to the alveolar epithelium in individuals rendered susceptible by ageing or genetic predisposition [15]. Exposure to cigarette smoke, viruses and bacteria have all been implicated in IPF pathogenesis, but a definitive causal link with the disease is yet to be established [9, 75, 115, 401, 402].

Lending support to the theory that host-environment interactions drive fibrosis, the microbiome of the lower airways is altered in IPF with changes in microbial composition, diversity and a greater bacterial burden associated with disease progression [115]. The host peripheral blood transcriptome in IPF also provides evidence for a response to the presence of these bacterial communities [309]. Bacteria such as *Streptococcus* and *Staphylococcus* have been recognised as key pathogens in the IPF airways; an increased abundance of either bacteria is associated with a significant reduction in progression-free survival time [114]. Through the fermentation of carbohydrates and amino acids, these facultative anaerobic bacteria produce vast amounts of SCFAs in the gut, including acetate, propionate and butyrate [357]. These SCFAs can also be detected in sputum at concentrations ranging from 0.1 to 5 mM, demonstrating that they reach the lungs and the airways [353]. By binding to G protein-coupled receptors (GPCRs), including GPCR41 and GPCR43, these metabolites regulate a multitude of functions including cell proliferation, cell differentiation, and immune and inflammatory responses [403-405].

In **Chapter 4**, these SCFAs were found to be elevated in the BAL of subjects with IPF compared to healthy controls. As of yet, no study has investigated the impact of these microbial metabolites in the context of fibrosis. The overarching aim of this Chapter was to investigate the effect of microbial metabolites, specifically the SCFAs acetate, propionate and butyrate on PBECs and HLF. Given that the airways of subjects with IPF are dominated by SCFA-producing bacteria and that these microbial

metabolites are increased in the BAL compared to healthy controls, the hypothesis examined here was that these SCFAs might be stimulating the secretion of epithelial-cell derived cytokines and contribute to the induction of the inflammatory and fibrotic response characteristic of this chronic lung disease. To address this, submerged PBEC cultures from healthy controls and subjects with IPF were either left untreated or exposed to increasing concentrations of the SCFAs and cell viability, cell proliferation, wound healing and pro-inflammatory response assessed. To recapitulate the *in vivo* phenotypic and functional characteristics of the human airway, these experiments were also conducted in PBECs from IPF patients and healthy controls differentiated at ALI. Given the crucial role of fibroblasts in mediating fibrosis, the effects of these SCFAs were also assessed on primary cultures of HLF isolated from subjects with IPF.

5.2 Aims

- To establish the target pharmacological range of the SCFAs on HLF and PBECs (submerged and at ALI).
- To investigate the pro-inflammatory response in IPF PBECs (submerged and at ALI) post-SCFA treatment.
- To assess the production of collagen I α1, fibronectin and TGF-β in HLF post-SCFA treatment.
- To quantify the TEER in cultures of healthy and IPF PBECs differentiated at ALI.
- To compare the proliferation in PBECs (submerged) from healthy controls and subjects with IPF.
- To compare the proliferation in IPF HLF and PBECs (submerged) post-SCFA treatment.
- To compare the wound healing capacity in PBECs (submerged) from healthy controls and subjects with IPF.
- To compare the wound healing capacity in IPF PBECs (submerged) post-SCFA treatment.

5.3 Materials and Methods

5.3.1 PBECs culture

PBECs were obtained from bronchial brushes from healthy controls and IPF patients undergoing bronchoscopy and cultured as described in **Chapter 2 section 2.7.1**.

5.3.2 Submerged culture of PBECs

PBECs (passage 2) from healthy controls and subjects with IPF were cultured in PromoCell or PromoCell containing the SCFAs (acetate, propionate or butyrate) at the respective doses (1 mM, 5 mM, 10 mM, or 20 mM) as described in **Chapter 2 section 2.7.2**. Experiments were performed as described in **Chapter 2 sections 2.7.2.1-2.7.2.3**.

5.3.3 ALI culture of PBECs

PBECs (passage 2) from IPF and healthy control subjects were cultured in PromoCell Airway Epithelial Cell Growth Medium or PromoCell containing the SCFAs (acetate, propionate or butyrate) at the respective doses (1 mM and 5 mM) as described in **Chapter 2 section 2.7.3**. Experiments were performed as described in **Chapter 2 sections 2.7.3.1-2.7.3.2**.

5.3.4 Isolation of primary HLF

Lung tissue specimens were obtained from the Royal Brompton Hospital (REC 15/SC/0101) during cancer resections or lung-volume resection surgery for COPD. Human lung fibroblasts were extracted from the tissue using an established protocol [307] as described in **Chapter 2 section 2.7.4**.

5.3.5 Submerged culture of primary HLF

HLF (passage 2) from IPF patients were cultured in DMEM media or DMEM containing the SCFAs (acetate, propionate or butyrate) at the respective doses (1 mM, 5 mM, 10 mM, or 20 mM) as described in **Chapter 2 section 2.7.5**. Experiments were performed as detailed in **Chapter 2 sections 2.7.5.1-2.7.5.2**.

5.3.6 Statistical analysis

Data were analysed with GraphPad Prism 8.2 software (GraphPad Software, San Diego, CA). Unless stated otherwise, all data are presented as mean ± SEM and are representative of at least three independent experiments. Data were compared using a Kruskal-Wallis test with Dunn's multiple comparisons test. Differences were considered statistically significant at $P < 0.05$.

5.4 Results

5.4.1 Patient characteristics

A summary of the characteristics of the subjects recruited is illustrated in **Table 5.1**. PBECs were cultured from 6 subjects with IPF and 6 healthy controls. Individuals with IPF were predominantly men (83%), with a mean age of 73 years ± 8, and had moderately severe disease at enrolment (FVC, 76% predicted; DL_{CO}, 44% predicted). The 6 control subjects were also predominantly men (67%) and had a mean age of 56 years ± 15. Subjects were matched for sex and smoking status.

Table 5.1 Donor characterisation.

Parameter	Controls (*N* = 6)	IPF (*N* = 6)	*P*-value*
Age, years (± SD)	56 (15)	73 (8)	0.0001
Male, *n* (%)	4 (67)	5 (83)	NS
Baseline PFT			
FVC, % predicted (± SD)	90 (25)	76 (16)	0.0001
FEV_1 % predicted (± SD)	82 (22)	81 (18)	NS
DL_{CO} % predicted (± SD)	77 (32)	44 (22)	0.0001
Smoking status			
Non-smoker, *n* (%)	2 (33)	0 (0)	-
Ex-smoker, *n* (%)	3 (50)	4 (67)	NS
Current-smoker, *n* (%)	1 (17)	2 (33)	-

Data are presented as percentage value or mean ± standard deviation (SD) as appropriate. IPF: idiopathic pulmonary fibrosis; PFT: pulmonary function test; FEV_1: forced expiratory volume in 1 second; FVC: forced vital capacity; DL_{CO}: diffusing capacity of the lung for carbon monoxide; BAL: bronchoalveolar lavage; NS: not significant. Smoking status: statistical testing by non-smoker versus ex/current-smoker. **P*-values are calculated by Mann-Whitney U test or Fisher's exact test between controls and IPF subjects.

5.4.2 Submerged culture of PBECs

5.4.2.1 Assessing the cytotoxicity post-SCFA treatment

Firstly, the effects of the SCFAs acetate, propionate and butyrate on cell viability of healthy and IPF PBECs was investigated to exclude any SCFA-induced cytotoxicity **(Figure 5.1)**. The cells were either left untreated (0 mM) or treated with increasing doses of the SCFAs (1mM, 5mM, 10 mM and 20 mM) over a time-course of 48 hours. In both healthy and IPF PBECs the lower doses (1 mM and 5 mM) of the SCFAs were found to be non-toxic to the cells at all time-points. In contrast, treatment of healthy and IPF PBECs with the higher doses (10 mM and 20 mM) of the SCFAs was found to significantly reduce cell viability at 48 hours. Butyrate was the only SCFA which was found to be non-toxic to the healthy PBECs even at the higher doses over the 48-hour time-course.

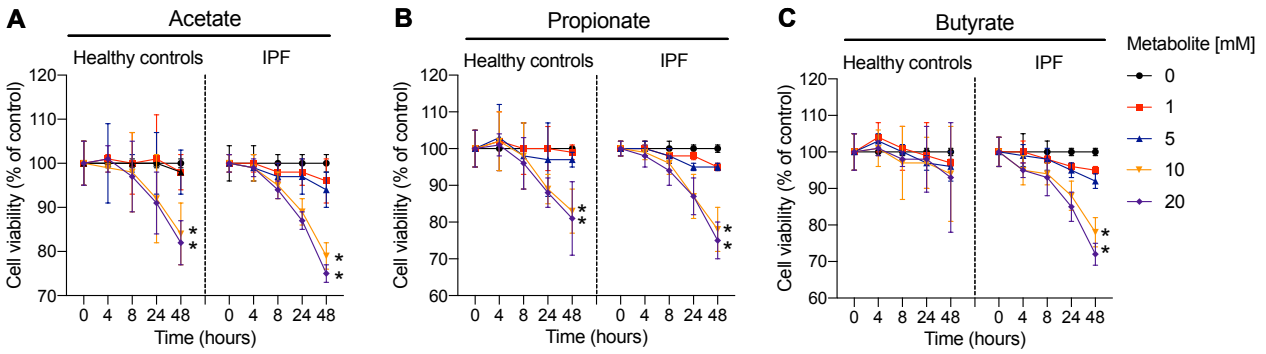


Figure 5.1 Cytotoxic effects of the short-chain fatty acids (SCFAs) on submerged cultures of human primary bronchial epithelial cells (PBEC) from healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Cell viability of healthy (*n* = 4) and IPF (*n* = 4) PBECs post **(A)** acetate, **(B)** propionate, and **(C)** butyrate treatment (0 mM, 1 mM, 5 mM, 10 mM and 20 mM). All data are presented as mean ± SEM of four independent experiments. **P* < 0.05. All comparisons are made with the untreated control (0 mM).

5.4.2.2 Assessing the pro-inflammatory response post-SCFA treatment

SCFAs have been shown to exert pro-inflammatory effects on bronchial epithelial cells from cystic fibrosis patients [175]. The effects of SCFAs on inflammatory responses of IPF PBECs were investigated **(Figure 5.2)**.

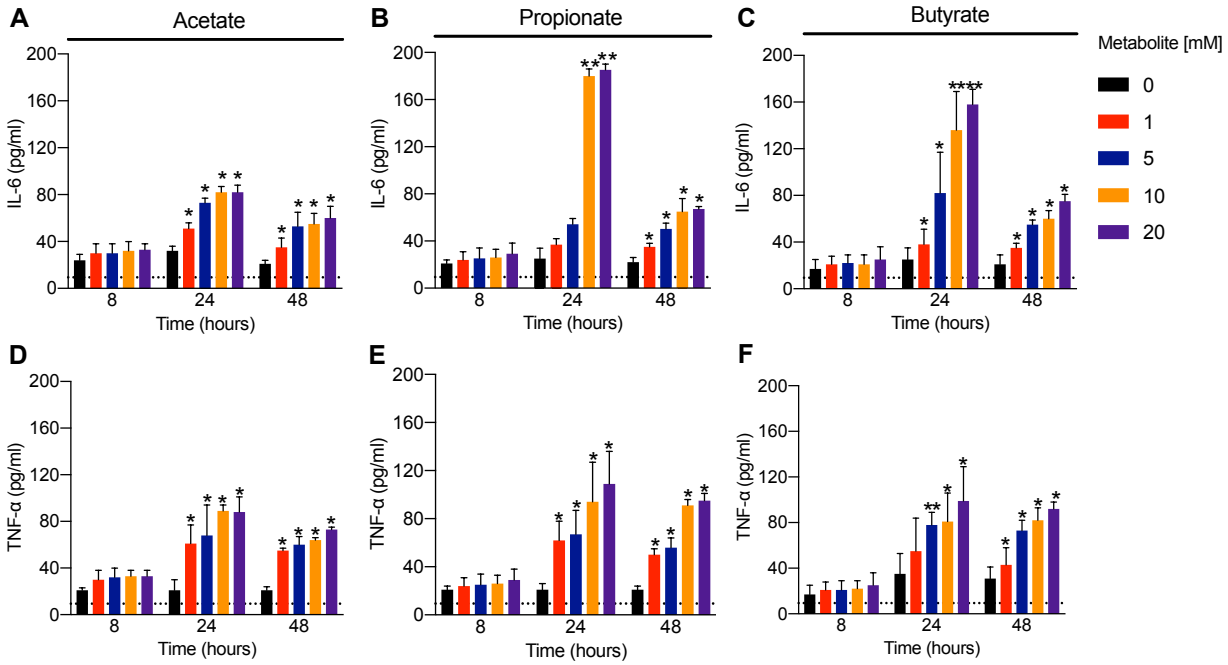


Figure 5.2 Short-chain fatty acids induce a dose-dependent IL-6 and TNF-α secretion in submerged cultures of human primary bronchial epithelial cells (PBECs) from idiopathic pulmonary fibrosis (IPF) subjects.

Secretion of IL-6 and TNF-α in cell culture supernatant of IPF (n = 4) PBECs post **(A, D)** acetate, **(B, E)** propionate, and **(C, F)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM, 5 mM, 10 mM and 20 mM. All data are presented as mean ± SEM of four independent experiments. **P* < 0.05, ***P* < 0.01. All comparisons are made with the untreated control.

PBECs from four IPF patients were cultured and either left untreated (0 mM) or exposed to increasing doses (1 mM, 5 mM, 10 mM and 20 mM) of the SCFAs acetate, propionate and butyrate for 4-, 8-, 24- and 48 hours. Both IL-6 and TNF-α were found to be undetectable at 4 hours. At 8 hours, treatment of the cells with the SCFAs induced no response. All SCFAs induced a dose-dependent release of IL-6 and TNF-α at 24- and 48-hours post-treatment **(Figure 5.2)**.

5.4.2.3 Assessing the proliferation: pre-SCFA treatment

To assess the proliferative capacity of IPF PBECs in submerged culture the confluence was assessed over 72 hours using a real-time cell-history recorder. Although cell seeding density was constant, differences were noted in the starting confluence given the distribution of the seeded cells. The data were normalised to a starting confluence of 15% to control for this difference. No differences were found in cell proliferation when comparing PBECs from healthy controls and from subjects with IPF **(Figure 5.3)**.

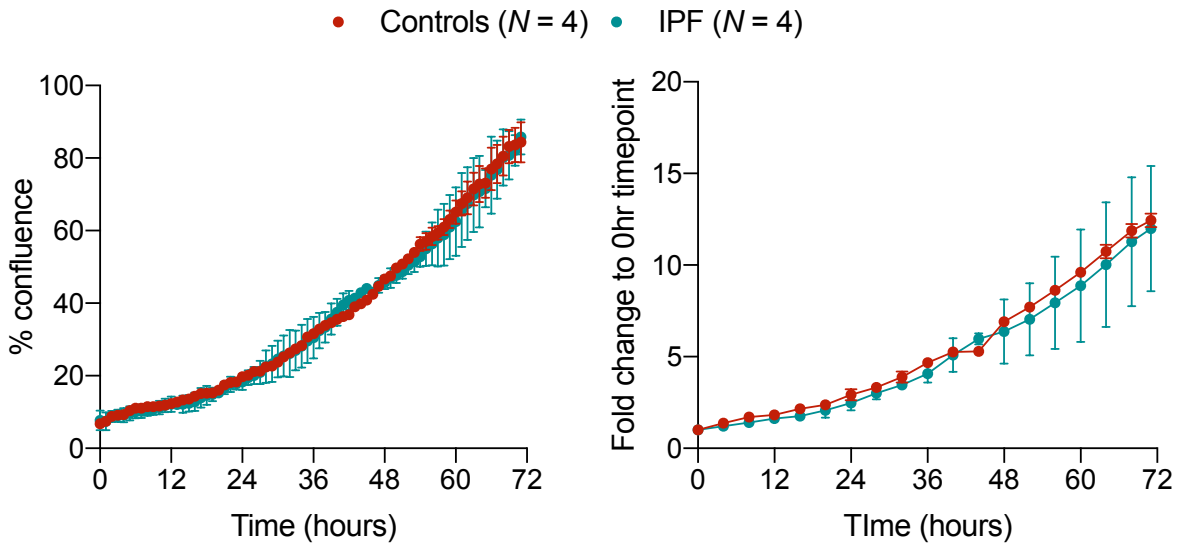


Figure 5.3 Proliferation assay of untreated primary bronchial epithelial cells (PBECs).

PBECs from healthy controls (*n* = 4) and idiopathic pulmonary fibrosis (*n* = 4) subjects were seeded (passage 3) into a 96-well plate pre-coated with type 1 bovine collagen. To correct for initial differences in % confluence at the start of the assay, data is presented as confluence normalised to a starting confluence of 30%. Snapshots of each well were recorded at 30-min intervals over 72 hours using the JuLI Stage Real-Time Cell History Recorder. Data plotted represents the mean confluence at 3 separate positions. All data are presented as mean ± SEM.

5.4.2.4 Assessing the proliferation: post-SCFA treatment

The effects of the SCFAs on PBEC proliferation were assessed in PBECs from three patients with IPF over a time-course of 72 hours **(Figure 5.4)**. PBECs were either left untreated or exposed to 1 mM or 5mM of acetate, propionate and butyrate. Following exposure of the PBECs to the SCFAs, a trend towards a reduction in proliferation was observed over time, although not statistically significant. Ki67 staining was performed at 72 hours, and the number of DAPI-positive cells and the number of DAPI and Ki-67 double-positive cells were counted to determine the percentage of Ki-67 cells per well. No significant differences were observed in PBEC proliferation when comparing cells cultured in media (untreated) or treated with 1 mM or 5 mM of the respective SCFAs **(Figure 5.4AB)**.

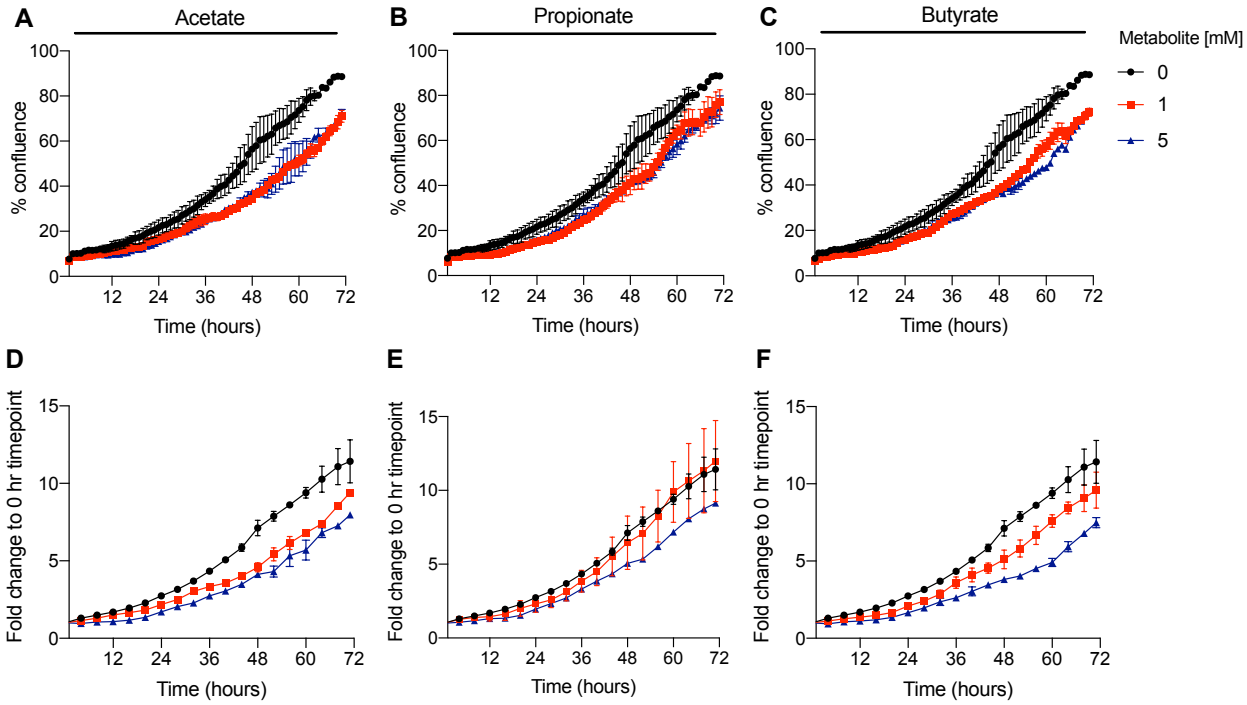


Figure 5.4 Proliferation of human primary bronchial epithelial cells (PBECs) from idiopathic pulmonary fibrosis (IPF) subjects following exposure to the short-chain fatty acids (SCFAs).

IPF PBECs were cultured in media (untreated), or media supplemented with different concentrations of the SCFAs, propionate and butyrate and visualised over a period of 72 hours using a JuLI Stage automated cell imaging system. **(A-C)** Illustrating % cell confluence over time and **(D-F)** intra-group gold change in % cell confluence of cells cultured in media (untreated) or 1 mM or 5 mM of the SCFAs at 72 hours relative to 0 hours. Snapshots of each well were recorded at 30-min intervals over 72 hours using the JuLI Stage Real-Time Cell History Recorder. Data plotted represents the mean confluence at 3 separate positions. All data are presented as mean ± SEM.

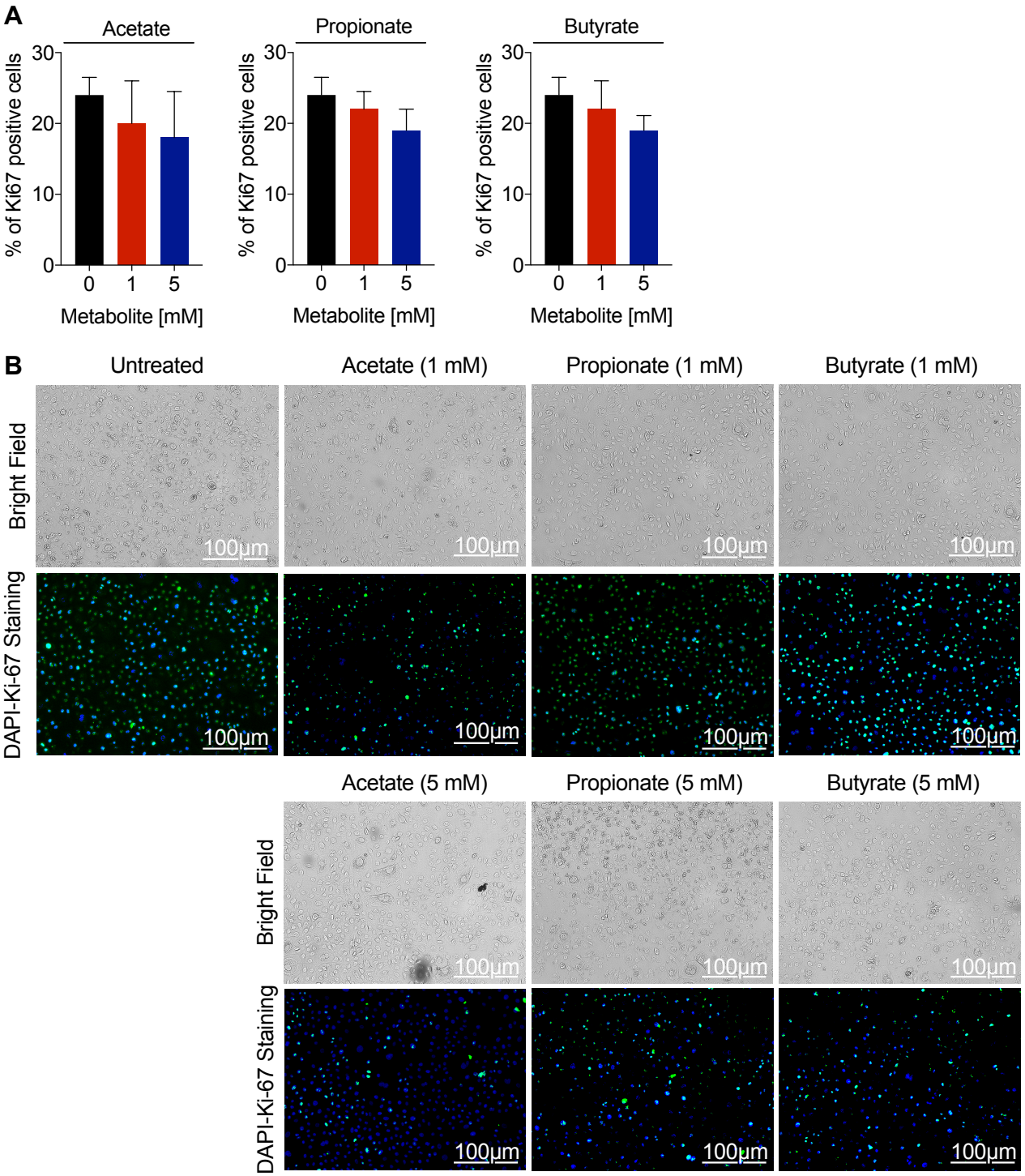


Figure 5.5 Ki67 staining of human primary bronchial epithelial cells (PBECs) from idiopathic pulmonary fibrosis (IPF) subjects following exposure to the short-chain fatty acids (SCFAs).

PBECs were cultured in media (untreated), or media supplemented with different concentrations of the SCFAs acetate, propionate and butyrate and visualized over a period of 72 hours using a JuLI Stage automated cell imaging system. **(A)** After 72 hours, the cells were fixed and stained for Ki67 and Ki67 positive cells expressed as a percentage of DAPI positive cells. **(B)** Representative Bright Field images and images of cells stained for DAPI and Ki67 after 72 hours of culture in media (untreated) or 1 mM or 5 mM of acetate, propionate or butyrate. Scale bar = 100 µm. Results depicted as mean ± SEM. **P* < 0.05 using ANOVA with Bonferroni correction.

5.4.2.5 Assessing the wound healing capacity: pre-SCFA treatment

Given that epithelial injury is a key driver of pulmonary fibrosis, the wound healing capacity of IPF PBECs was assessed. A linear wound was created in the confluent cell monolayer, and the % wound closure was calculated by the JuLI Stage software. No significant difference in wound healing capacity was observed when comparing PBECs from healthy controls to IPF PBECs **(Figure 5.6)**.

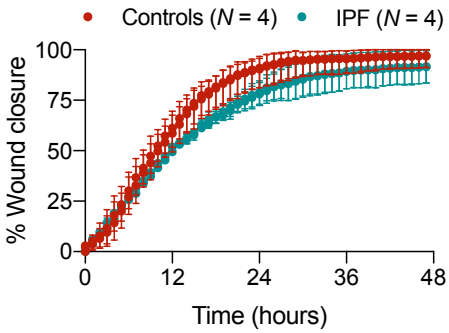


Figure 5.6 Wound healing assay of untreated primary bronchial epithelial cells (PBECs).

PBECs from healthy controls (*n* = 4) and idiopathic pulmonary fibrosis (*n* = 4) subjects were seeded (passage 3) into a 24-well plate pre-coated with type 1 bovine collagen. Cells were grown to a 100% confluent monolayer, and a single linear wound was generated with a P1000 pipette tip in the wells. Snapshots of each well were recorded at 30-min intervals over 48 hours using the JuLI Stage Real-Time Cell History Recorder. Data plotted represents the mean confluence at 3 separate positions. All data are presented as mean ± standard deviation.

5.4.2.6 Assessing the wound healing capacity: post-SCFA treatment

The effects of the SCFAs on PBEC wound healing were assessed in PBECs from three patients with IPF over a time-course of 48 hours **(Figure 5.7)**. PBECs were either left untreated or exposed to 1 mM or 5mM of acetate, propionate and butyrate. Compared to cells treated with media, exposure to either 1 mM of 5 mM of the respective SCFAs led to a decrease in wound closure at 48 hours.

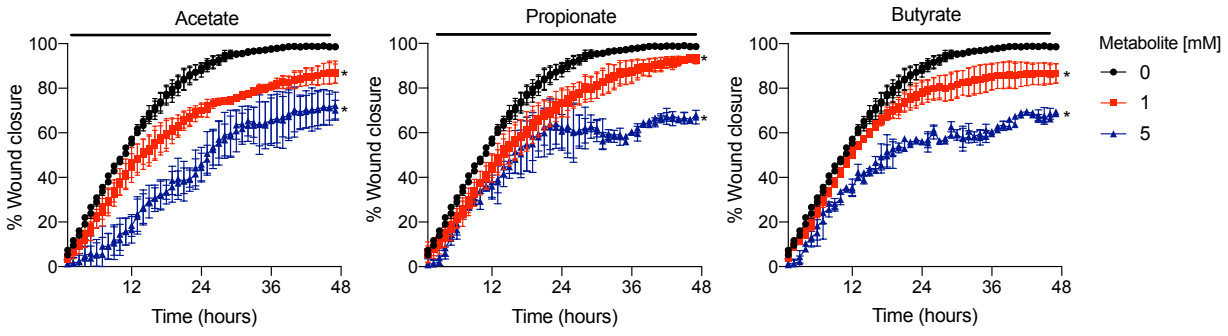


Figure 5.7 Wound healing assay of human primary bronchial epithelial cells (PBECs) from idiopathic pulmonary fibrosis (IPF) subjects following exposure to the short-chain fatty acids (SCFAs).

IPF PBECs were cultured in media (untreated), or media supplemented with different concentrations of the SCFAs, propionate and butyrate and visualised over a period of 48 hours using a JuLI Stage automated cell imaging system. PBECs were seeded (passage 3) into a 24-well plate pre-coated with type 1 bovine collagen. Cells were grown to a 100% confluent monolayer, and a single linear wound was generated with a P1000 pipette tip in the wells. Snapshots of each well were recorded at 30-min intervals over 48 hours using the JuLI Stage Real-Time Cell History Recorder. Data plotted represents the mean confluence at 3 separate positions. All data are presented as mean ± standard deviation.

5.4.3 ALI culture of PBECs

PBECs from six healthy controls and six subjects with IPF were differentiated at ALI to recapitulate the *in vivo* phenotypic and functional characteristics of the human airway. Experiments were performed 2 weeks after cilia were first observed. The SCFAs acetate, propionate and butyrate were added apicolaterally at a concentration of 1 mM and 5 mM or media control at the following time points: 24h, 48h, 72h, 96h and 144h. As illustrated in **sections 5.4.3.1-5.4.3.4**, at each time-point, cell death, TEER, and pro-inflammatory cytokine release were measured. Furthermore, brightfield images were captured to observe the changes in cell phenotype over time.

5.4.3.1 Assessing the cytotoxicity: post-SCFA treatment

The effects of the SCFAs acetate, propionate and butyrate on cell viability of healthy and IPF PBECs differentiated at ALI were investigated **(Figure 5.8)**. The cells were either left untreated (0 mM) or treated with 1 mM or 5 mM of the SCFAs at 24h, 48h, 72h, 96h, and 144h; cell death was measured using the LDH assay. In the healthy PBECs, compared to the untreated cells, there were no differences observed in healthy PBEC survival after exposure to either dose of acetate, propionate, or butyrate. Similarly, IPF PBECs treated with 1 mM of either acetate, propionate or butyrate had no effect on cell viability. In contrast, IPF PBECs exposed to 5 mM of the SCFAs was found to significantly reduce cell viability at 48, 96 and 144h hours.

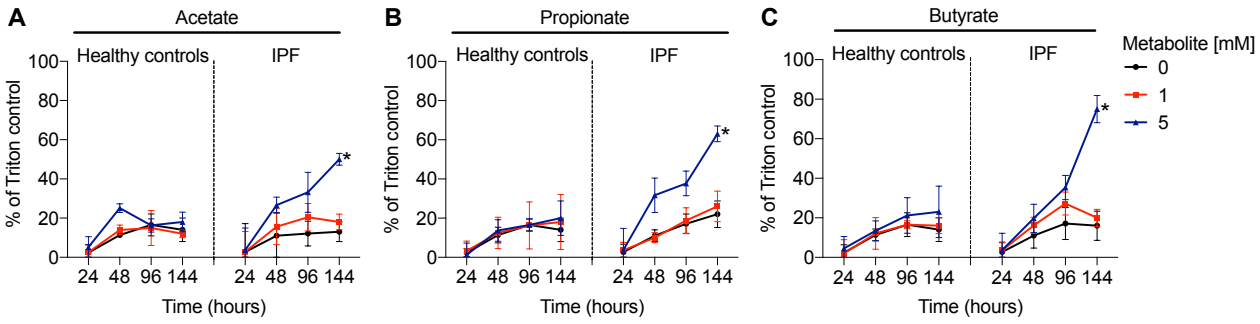


Figure 5.8 Cytotoxic effects of the short-chain fatty acids (SCFAs) on air-liquid interface (ALI) cultures of human primary bronchial epithelial cells (PBEC) from healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Cell viability of healthy (n = 6) and IPF (n = 6) PBECs post **(A)** acetate, **(B)** propionate, and **(C)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM and 5 mM. All data are presented as mean ± SEM of four independent experiments. *P < 0.05. All comparisons are made with untreated control.

5.4.3.2 Assessing the trans-epithelial electrical resistance: post-SCFA treatment

TEER is a validated, non-invasive technique for quantifying the integrity of tight junctions in cell culture models [406]. TEER, was measured pre- (0 h) and post-SCFA treatment (24 h, 48 h, 96, and 144 h) in cells differentiated at ALI from six healthy controls and six subjects with IPF **(Figure 5.9)**. Interestingly, compared to the cells cultured in media alone (untreated), the TEER of the PBECs from the healthy controls remained consistent over time following exposure to 1 mM and 5 mM of the SCFAs. In contrast, compared to the untreated cells, IPF PBECs treated with 1 mM and 5 mM of acetate, propionate and butyrate showed a significant decrease in TEER at 48, 96 and 144 hours **(Figure 5.9 A-C)**.

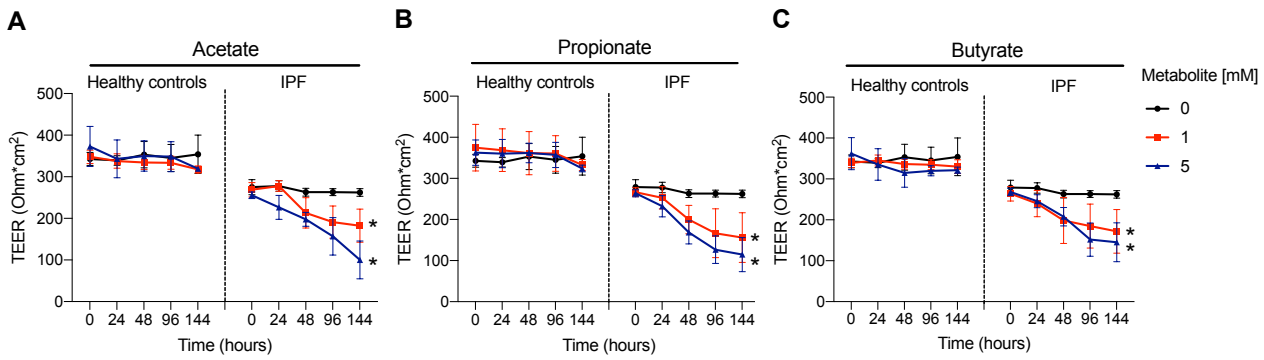


Figure 5.9 Effects of the short-chain fatty acids (SCFAs) on the trans-epithelial electrical resistance (TEER) of human primary bronchial epithelial cells (PBECs) from healthy controls and subjects with idiopathic pulmonary fibrosis (IPF) differentiated at air-liquid interface (ALI).

TEER of healthy (n = 6) and IPF (n = 6) PBECs post **(A)** acetate, **(B)** propionate, and **(C)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM and 5 mM. All data are presented as mean ± SEM of four independent experiments. *P < 0.05. All comparisons are made with untreated control.

5.4.3.3 Assessing the changes in morphology: post-SCFA treatment

The morphological changes in the PBECs from healthy controls and subjects with IPF were assessed at each time-point following the addition of the SCFAs. As illustrated in **Figure 5.10**, at 24 hours, 48 hours, 96 hours and 144 hours the PBECs from the healthy controls exposed to acetate, propionate and butyrate at either 1 mM or 5 mM showed no significant alterations in morphology compared to the PBECs cultured in media (untreated). In contrast, exposure of the IPF PBECs to 5 mM but not 1 mM of the SCFAs appeared to induce substantial cell death over the time-course compared to the untreated control **(Figure 5.11)**. At the 144h time-point, PBECs from IPF subjects treated with 1 mM of either acetate, propionate or butyrate appeared to be morphologically similar compared to the untreated cells. In contrast, PBECs from IPF subjects treated with 5 mM of the respective SCFAs appeared to be mostly dead.

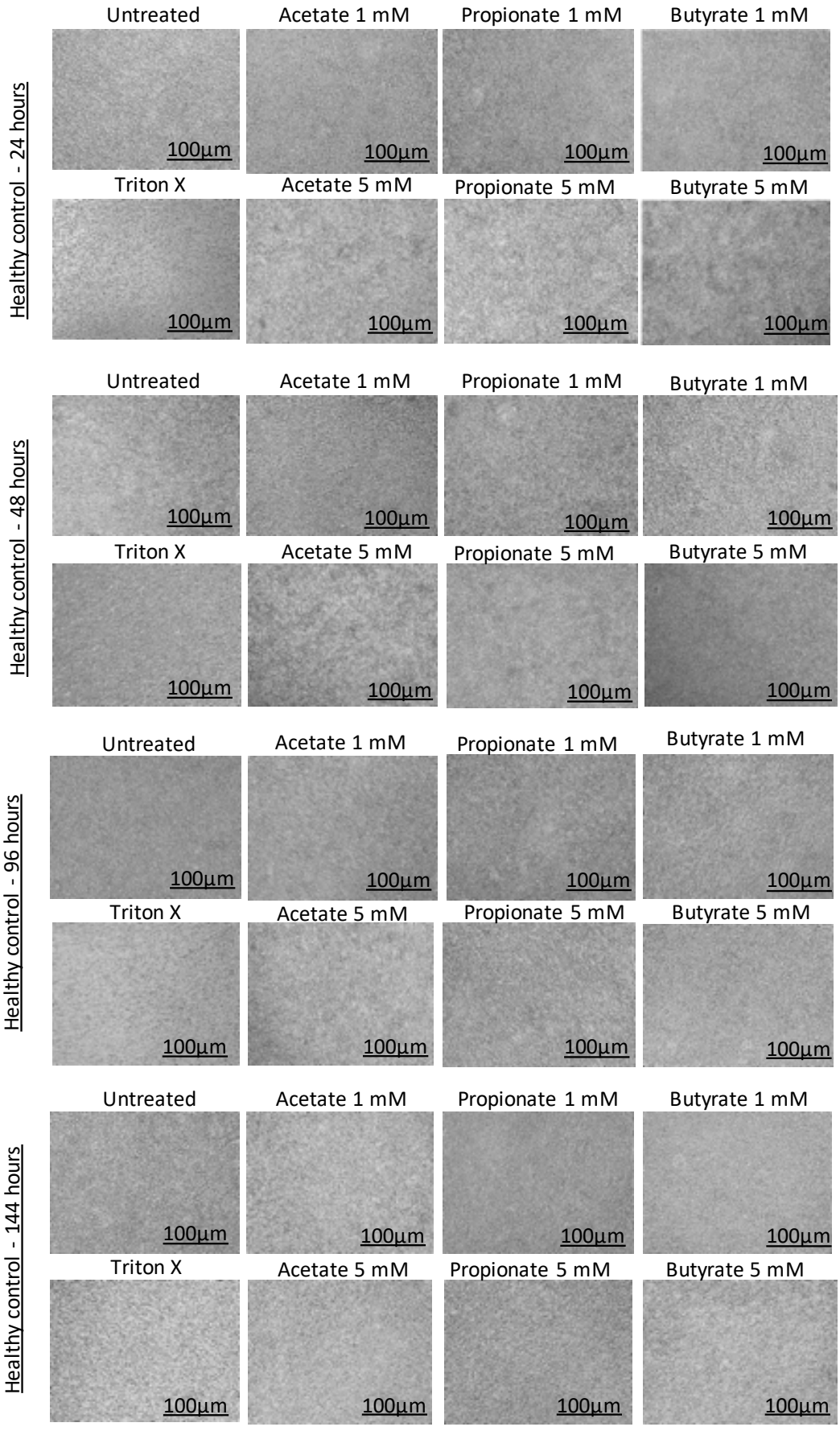


Figure 5.10 Morphology of primary bronchial epithelial cells from a healthy control differentiated at air-liquid interface in different culture conditions.

Brightfield images of the cells were taken at each time-point with a LeicaDM2500 microscope (Leica Microsystems, UK).

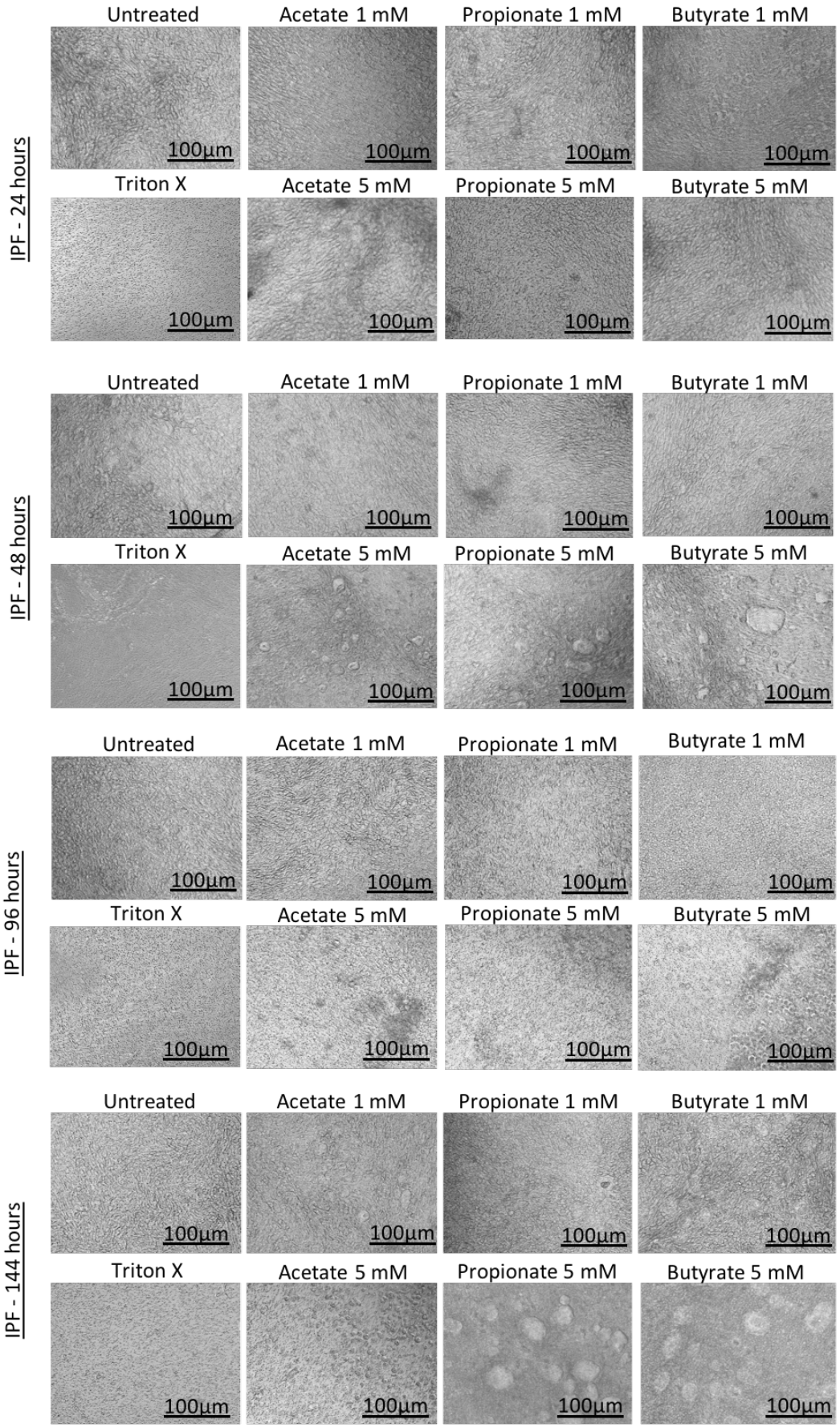


Figure 5.11 Morphology of primary bronchial epithelial cells from a subject with idiopathic pulmonary fibrosis differentiated at air-liquid interface in different culture conditions.

Brightfield images of the cells were taken at each time-point with a LeicaDM2500 microscope (Leica Microsystems, UK).

5.4.3.4 Assessing the pro-inflammatory response: post-SCFA treatment

The pro-inflammatory effects of the SCFAs were assessed by quantifying IL-6 in the apical cell supernatant from the PBECs of healthy controls and subjects with IPF differentiated at ALI **(Figure 5.12)**. Acetate, propionate and butyrate induced a dose-dependent release of IL-6 in both healthy and IPF cultures over the time-course. The secretion of IL-6 was found to be significantly increased in IPF compared to healthy controls at 24-, 48-, 96-, and 144 hours post-exposure to all concentrations of the SCFAs ($P < 0.05$, Mann-Whitney U test). However, compared to PBECs treated with media alone, a significant difference was only observed at the 24-hour time-point in the PBECs from IPF patients following treatment with either 1 mM or 5 mM of propionate or with 5 mM of acetate or butyrate **(Figure 5.12 D-F)**.

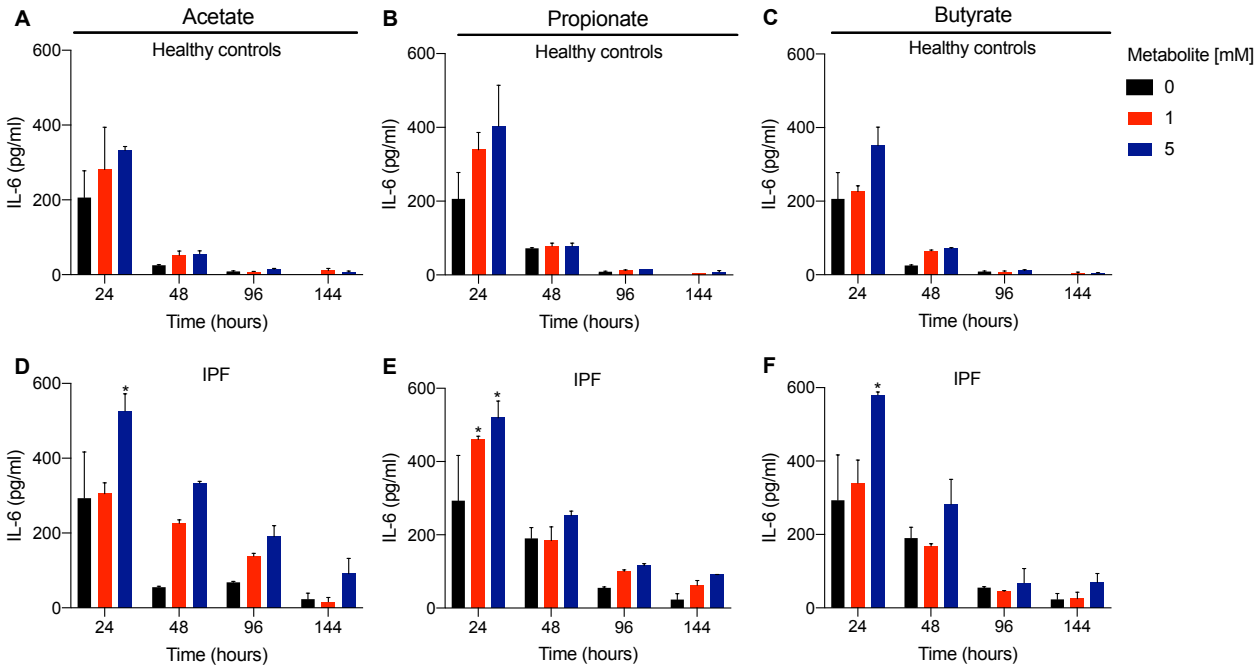


Figure 5.12 Secretion of IL-6 in the apical supernatant of human primary bronchial epithelial cells (PBECs) from healthy controls and idiopathic pulmonary fibrosis (IPF) subjects differentiated at air-liquid interface (ALI).

Secretion of IL-6 in apical cell culture supernatant of healthy control ($n = 6$) and IPF ($n = 6$) PBECs post **(A, D)** acetate, **(B, E)** propionate, and **(C, F)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM and 5 mM. All data are presented as mean ± standard deviation of six independent experiments. *$P < 0.05$. All comparisons are made with untreated control.

5.4.4 Submerged culture of primary HLFs

5.4.4.1 Assessing the cytotoxicity: post-SCFA treatment

The effects of the SCFAs acetate, propionate and butyrate on cell viability of IPF HLF was then investigated to exclude any SCFA-induced cytotoxicity **(Figure 5.13)**. The fibroblasts were either left untreated (0 mM) or treated with increasing doses of the SCFAs (1 mM, 5 mM, 10 mM and 20 mM) over a time-course of 48 hours. Exposure of the HLF to the higher doses of the SCFAs (10 mM and 20 mM) induced a significant decrease in cell viability over time. In contrast, HLF treated with the lower doses (1 mM and 5 mM) of the respective SCFAs had no effects on cell viability **(Figure 5.13)**.

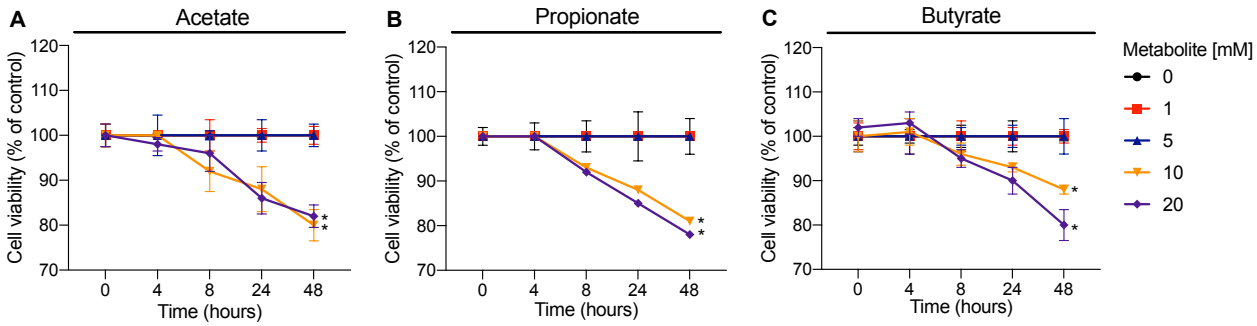


Figure 5.13 Cytotoxic effects of the short-chain fatty acids (SCFAs) on submerged cultures of human lung fibroblasts (HLF) from subjects with idiopathic pulmonary fibrosis (IPF).

Cell viability of IPF (n = 4) HLF post **(A)** acetate, **(B)** propionate, and **(C)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM, 5 mM, 10 mM and 20 mM. All data are presented as mean ± SEM of four independent experiments. *P < 0.05. All comparisons are made with untreated control.

5.4.4.2 Assessing the production of collagen I α1, fibronectin and TGF-β: post-SCFA treatment

To assess whether the SCFAs acetate, propionate and butyrate had an effect on HLF from IPF subjects, the secretion of TGF-β, collagen I α1 and fibronectin were quantified in the cell supernatants of untreated HLF and in HLF treated with increasing doses of the SCFAs **(Figure 5.14)**. At 24 hours, compared to the untreated control (0 mM), HLF treated with acetate at 10 mM and 20 mM were found to have a greater concentration of collagen I α1 in the cell culture supernatant **(Figure 5.14A)**. Similarly, fibronectin was increased in the cell culture supernatant following treatment with 20 mM acetate **(Figure 5.14D)**. No changes were observed in collagen I α1 and fibronectin when treating the cells with the lower doses of acetate. Compared to the untreated control TGF-β was found to be unchanged following the addition of acetate even at the highest concentration **(Figure 5.14G)**. No changes were observed in collagen I α1 **(Figure 5.14BC)**, fibronectin **(Figure 5.14EF)** and TGF-β **(Figure 5.13HI)** in the cell culture supernatants of the IPF HLF following treatment with propionate and butyrate at 1 mM, 5 mM, 10 mM or 20 mM.

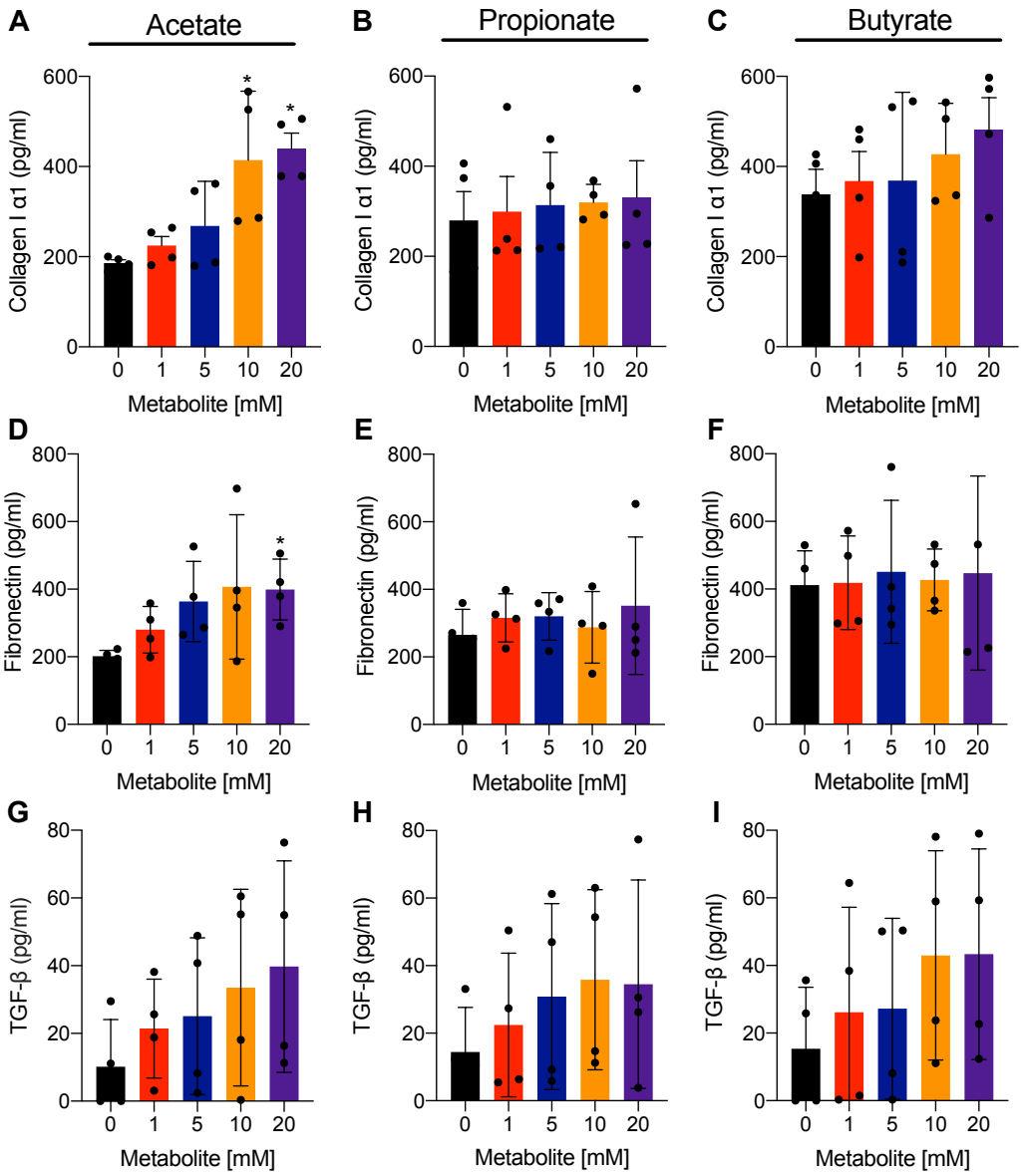


Figure 5.14 Collagen I α1, fibronectin and TGF-β production post-short-chain fatty acid (SCFA) treatment in submerged cultures of human lung fibroblasts (HLF) from idiopathic pulmonary fibrosis (IPF) subjects.

Secretion of collagen I α1, fibronectin and TGF-β at 24 hours in the cell culture supernatant of IPF (*n* = 4) HLF post **(A, D, G)** acetate, **(B, E, H)** propionate, and **(C, F, I)** butyrate treatment at the following doses: 0 mM (untreated), 1 mM, 5 mM, 10 mM and 20 mM. All data are presented as mean ± SEM of four independent experiments. **P* < 0.05. All comparisons are made with untreated control.

5.4.4.3 Assessing the proliferation: post-SCFA treatment

The effects of the SCFAs on HLF proliferation were assessed in HLF from three patients with IPF over a time-course of 72 hours **(Figure 5.15)**. HLF were either left untreated or exposed to 1 mM or 5mM of either acetate, propionate or butyrate. Addition of acetate was found to have no effect on HLF proliferation compared to the untreated control (0 mM) **(Figure 5.15A)**. In contrast, following exposure of the HLF to propionate and butyrate, a trend towards a dose-dependent increase in proliferation was observed compared to cells treated with media alone **(Figure 5.15BC)**. Ki67 staining was performed at 72 hours, and the number of DAPI-positive cells and the number of DAPI

and Ki-67 double-positive cells were counted to determine the percentage of Ki-67 cells per well **(Figure 5.16AB)**. No significant differences were observed in HLF proliferation when comparing cells cultured in media (untreated) or treated with 1 mM or 5 mM of acetate. In contrast, compared to the untreated cells, HLF treated with 5 mM of either propionate or butyrate were found to have proliferated more at 72 hours.

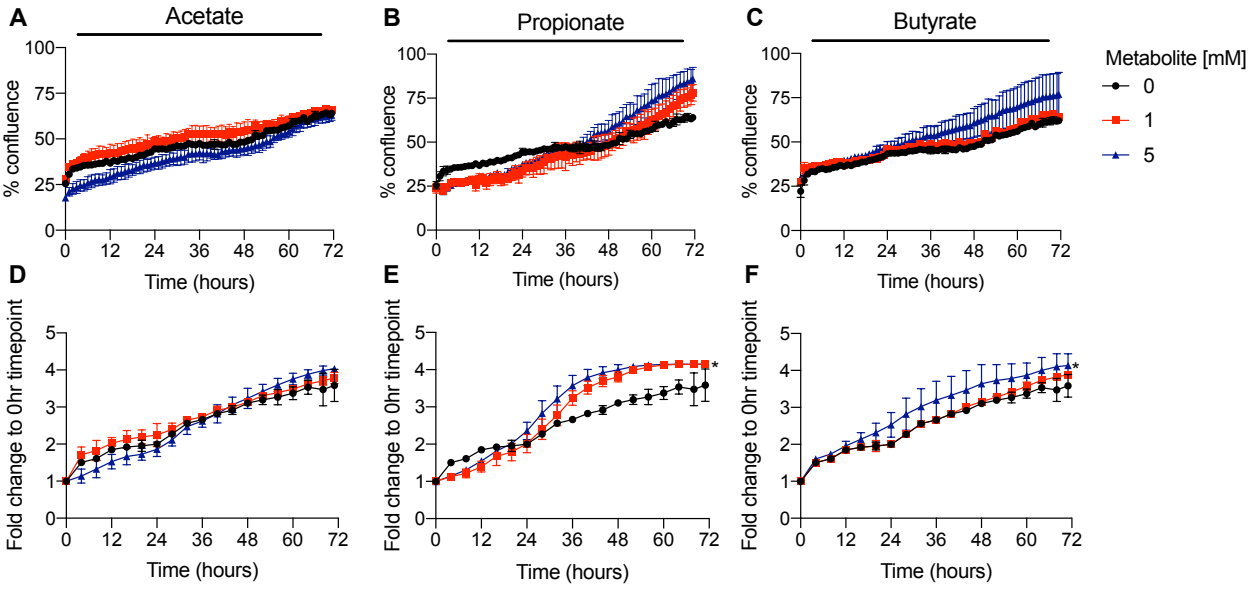


Figure 5.15 Proliferation of human lung fibroblasts (HLF) from idiopathic pulmonary fibrosis (IPF) subjects following exposure to the short-chain fatty acids (SCFAs).

IPF HLF (*n* = 4) were cultured in media (untreated), or media supplemented with different concentrations of the SCFAs (1 mM, 5 mM, 10 mM, and 20 mM), propionate and butyrate and visualised over a period of 72 hours using a JuLI Stage automated cell imaging system. **(A-C)** Illustrating % cell confluence over time and **(D-F)** intra-group gold change in % cell confluence of cells cultured in media (untreated) or 1 mM or 5 mM of the SCFAs at 72 hours relative to 0 hours. Snapshots of each well were recorded at 30-min intervals over 72 hours using the JuLI Stage Real-Time Cell History Recorder. Data plotted represents the mean confluence at 3 separate positions. All data are presented as mean ± standard deviation.

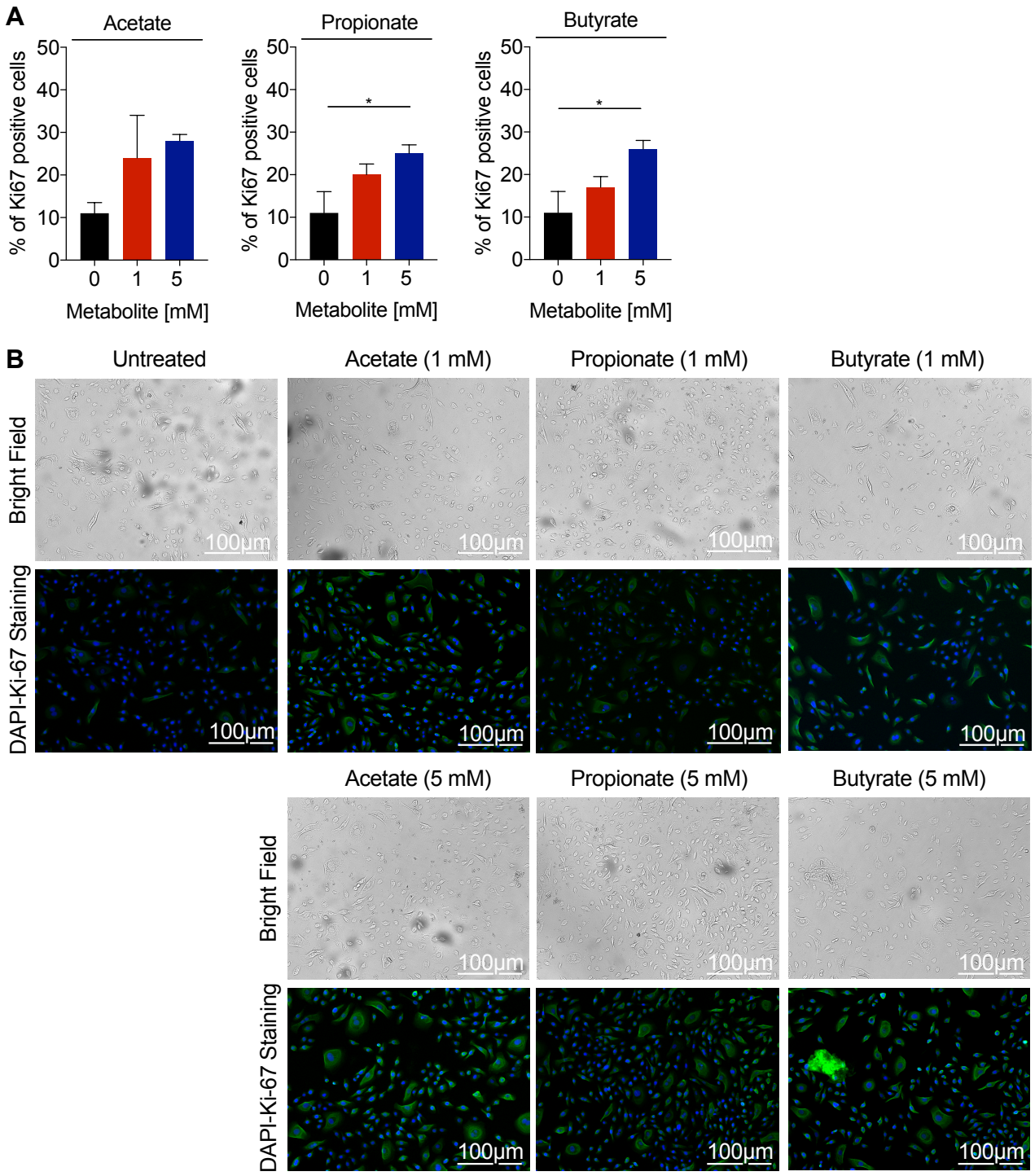


Figure 5.16 Ki67 staining of human lung fibroblasts (HLF) from idiopathic pulmonary fibrosis (IPF) subjects following exposure to the short-chain fatty acids (SCFAs).

IPF HLF (n = 4) were cultured in media (untreated) or media supplemented with different concentrations of the SCFAs acetate, propionate and butyrate and visualized over a period of 72 hours using a JuLI Stage automated cell imaging system. **(A)** After 72 hours the cells were fixed and stained for Ki67 and Ki67 positive cells expressed as a percentage of DAPI positive cells. **(B)** Representative Bright Field images and images of cells stained for DAPI and Ki67 after 72 hours of culture in media (untreated) or 1 mM or 5 mM of acetate, propionate or butyrate. Scale bar = 100 µm. Results depicted as mean ± SEM. *P < 0.05 using ANOVA with Bonferroni correction.

5.5 Discussion

IPF is characterised by a recurrent injury to the alveolar epithelium, loss of lung epithelial cells, and uncontrolled proliferation of fibroblasts and myofibroblasts. Over time, this process results in an exaggerated deposition of extracellular matrix components and abnormal tissue repair, which leads to respiratory failure [407]. The hypoxic environment of IPF airways allows the persistence of facultative anaerobic bacteria, which are responsible for the production of SCFAs. However, the role that SCFAs play in IPF is yet to be elucidated. For the first time, the effects of the SCFAs, specifically acetate, propionate and butyrate were characterised on submerged and ALI cultures of healthy and IPF PBECs, as well as on HLF isolated from subjects with IPF. In this Chapter it was demonstrated that supplementation of IPF PBECs with increasing concentrations of acetate, propionate or butyrate have effects on cell viability, tight junctions, cell proliferation, wound healing and inflammation.

Previous studies have only been observational and functional studies in IPF are currently lacking. The use of primary cells allows the direct comparison of cells from different donors in controlled *in vitro* conditions. Bronchial epithelial cells are the first anatomical barrier to external environmental stimuli and are thought to contribute to the initiation of airway remodelling through the production of proinflammatory mediators, ECM, growth factors and matrix metalloproteinases [408]. To explore the potential effects of the SCFAs on the airway epithelium and how it may promote IPF initiation and progression, submerged cultures of PBECs from healthy individuals and patients with IPF were either left untreated or exposed to increasing concentrations of acetate, propionate and butyrate over a time-course of 48 hours. Firstly, the effects of the SCFAs on cell viability of healthy and IPF PBECs was investigated to exclude any SCFA-induced cytotoxicity. In both healthy and IPF PBECs the lower doses (1 mM and 5 mM) of the SCFAs were found to be non-toxic to the cells at all time-points. In contrast, treatment of healthy and IPF PBECs with the higher doses (10 mM and 20 mM) of the SCFAs was found to significantly reduce cell viability at 48 hours. Butyrate was the only SCFA which was found to be non-toxic to the healthy PBECs even at the higher doses over the 48-hour time-course.

The airway epithelium participates in the inflammatory response by producing both pro- and anti-inflammatory factors [409]. Cytokines are small secreted proteins produced by nearly every cell, which regulate and influence the immune response [208]. The release of pro-inflammatory cytokines leads to the activation of immune cells and a release of further cytokines which contribute to the inflammatory process [410]. In intestinal epithelial cells, SCFAs can either exert pro- or anti-

inflammatory effects depending on the concentration [411-413]. Therefore, it was hypothesised that the SCFAs might be playing a role in inflammatory responses in the IPF airways. To test this, the inflammatory cytokines IL-6 and TNF-α were quantified in IPF PBECs which were either left untreated or treated with increasing doses of the SCFAs acetate, propionate and butyrate. Both cytokines were undetectable in the cell culture supernatant at 4 hours. Similarly, at 8 hours, treatment of IPF PBECs with the SCFAs induced no response. At 24 hours and 48 hours, all SCFAs induced a dose-dependent release of IL-6 and TNF-α. This provides evidence that SCFAs at concentrations measured in IPF BAL affect the inflammatory responses of airway epithelial cells. Since the cell viability experiment confirmed that the higher doses of the SCFAs (10 mM and 20 mM) were toxic to the cells, exposure of the cells to these concentrations might indicate that the cytokines are being released due to increased cell death. Therefore, in the following experiments, the cells were treated with nontoxic concentrations of the SCFAs (1 mM and 5 mM). It could be speculated that intrinsic abnormalities, including genetic and epigenetic changes, which characterise IPF airway epithelial cells, may be an underlying cause for the continuous secretion of these proinflammatory mediators, which may, in turn, be involved in the fibrotic process resulting in the excessive deposition of fibrotic tissue. It may also be that PBECs from patients with IPF were exposed to high concentrations of stimulating factors such as acetate, propionate and butyrate *in vivo*, resulting in the continued production of these mediators *in vitro*. Alternatively, cytokines released by the PBECs themselves, such as IL-6 and TNF-α, may act synergistically with the SCFAs in an autocrine manner to further stimulate their production.

To assess the proliferative capacity of IPF airway epithelial cells in submerged culture, the confluence of the cells over a 72-hour time-course was examined using a real-time cell-history recorder. No differences in cell proliferation were found when comparing PBECs from healthy controls and subjects with IPF. Airway epithelial cell proliferation represents a key event in IPF pathogenesis and progression. Cell proliferation was further assessed in IPF PBECs post-exposure to 1 mM or 5 mM of either acetate, propionate or butyrate. Compared to the untreated cells, there was a trend towards a reduction in PBEC proliferation following exposure to the SCFAs at both concentrations. Although not statistically significant, this reduction in proliferation is likely due to the fact that the SCFAs, including acetate, and propionate have been previously shown to induce cell death by acting on mitochondria [414]. Nevertheless, it is challenging to conclude whether the results observed were due to increased cell death or whether the SCFAs were inhibiting cell proliferation. Similar to studies investigating the role of SCFAs in inflammatory and immune responses, studies examining the impact of SCFA on epithelial cell proliferation have yielded

contradictory results, suggesting that the effects observed on cell growth may be context-dependent.

Persistent epithelial injury is considered a key driver of pulmonary fibrosis; therefore, the wound healing capacity of IPF PBECs was assessed after creating a linear wound in the confluent cell monolayer. No significant difference in wound healing capacity was observed when comparing PBECs from healthy controls to IPF PBECs. Wound healing was further assessed in IPF PBECs following exposure to 1 mM or 5 mM of either acetate, propionate or butyrate. Compared to the untreated cells, cells exposed to either 1 mM or 5 mM of the respective SCFAs led to a decrease in wound closure at 48 hours. This might indicate that the SCFAs are inhibiting wound closure, although the mechanism behind this is unclear. It is also important to note that the molecular events by which PBECs proliferate and migrate for lung wound healing are yet to be elucidated.

In contrast to the use of submerged cultures of epithelial cells, ALI cultures provide a differentiated airway epithelium more representative of the *in vivo* pseudostratified epithelium with ciliated, secretory and basal cells [415]. These systems have been developed to mimic the native conditions of the airway epithelium *ex vivo*. Several mechanistic studies have been conducted with these systems to study the function of epithelial cells as drug permeation barriers as well as to investigate host-pathogen interactions [416, 417]. Compared to submerged cultures, ALI cultures offer several advantages including the *in vitro* reconstitution of a pseudostratified epithelium, containing visible cilia and mucus, which are comparable to the features observed *in vivo* [416, 418-420]. Indeed, the transcriptome of the ALI epithelium correlates with the airway epithelium *in vivo* [421, 422]. Importantly, cells at passage 2-3 retain the ability to differentiate at ALI [423].

The effects of the SCFAs acetate, propionate and butyrate on cell viability of healthy and IPF PBECs differentiated at ALI were investigated by measuring cell death using the LDH assay. Cells were either cultured with media alone or treated with 1 mM or 5 mM of the SCFAs at 24h, 48h, 72h, 96h, and 144h. In the healthy PBECs, compared to the untreated cells, there were no differences observed in cell death after exposure to either dose of acetate, propionate, or butyrate. Similarly, IPF PBECs treated with 1 mM of either acetate, propionate or butyrate had no effect on cell viability. In contrast, exposure of IPF PBECs to 5 mM of the SCFAs was found to significantly increase cell death at 48, 96 and 144h hours. One potential explanation for the heightened response of the IPF PBECs to the SCFAs may be due to their cell-intrinsic defects, particularly mitochondrial dysfunction, which may increase the risk of mitochondria-mediated cell death following exposure to the SCFAs. Several studies have shown that SCFAs retain the ability to induce apoptosis and cell cycle arrest. In

a study by Jan and colleagues, propionate and acetate were found to induce apoptosis in two human colorectal carcinoma cell lines and were further shown to promote mitochondrial swelling when added to purified mitochondria *in vitro*, indicating that SCFAs may be inducing cell death by acting on mitochondria [414]. In a separate study, treatment with propionate was found to inhibit lung cancer cell proliferation and induce cell cycle arrest [424].

At each time-point, epithelial integrity was confirmed by measurement of TEER. TEER is a validated, non-invasive technique used for the evaluation of the cell permeability in transwell inserts [406]. Cell layers in insert systems under different culture conditions have been previously shown to differ in TEER measurements [419, 425, 426]. Furthermore, TEER can vary between cells from different donors and may decrease with the continued passaging of cells [419, 426, 427]. TEER, was measured pre- (0 h) and post-SCFA treatment (24 h, 48 h, 96, and 144 h) in PBECs differentiated at ALI from six healthy controls and six subjects with IPF. TEER analysis of ALI-cultured healthy PBECs showed values of approximately 342 Ωcm^2 at baseline. In contrast, IPF PBECs showed a reduction in TEER at baseline with a TEER of approximately 279 Ωcm^2. This finding might imply that the health status of donors can affect the development of TEER. Interestingly, compared to the cells cultured in media alone, the TEER of the PBECs from the healthy controls remained consistent over time following exposure to 1 mM and 5 mM of the SCFAs. In contrast, compared to the untreated cells, IPF PBECs treated with 1 mM and 5 mM of acetate, propionate and butyrate showed a significant decrease in TEER at 48, 96 and 144 hours. These findings provide evidence that the SCFAs acetate, propionate and butyrate can all alter tight junction permeability in PBECs from subjects with IPF but not from healthy controls. Immunofluorescence for expression of molecules that form tight junctions (ZO-1, cadherin) pre- and post-SCFA exposure could be conducted as a future experiment to validate this finding.

An increase in the number of epithelial cells characterised by a large, elongated and flattened phenotype is a distinctive feature of the epithelial cells observed in the lungs of subjects with IPF [428]. In accordance with what has been observed in the literature, the PBECs from the subjects with IPF differed morphologically to the PBECs from the healthy controls. Morphologically, the PBECs from the IPF subjects differentiated at ALI became larger and more flattened in appearance following exposure to acetate (1mM), propionate (1mM) and butyrate (1mM) following 24 hours. Elevation of the exposure concentration to 5 mM induced more prominent changes in the morphology of the IPF PBECs over time. In contrast, no morphological changes were observed when exposing the healthy PBECs differentiated at ALI at either dose of the respective SCFAs.

In addition to the morphological changes observed, the pro-inflammatory effects of the SCFAs were assessed by quantifying IL-6 in the apical cell supernatant from the PBECs of healthy controls and subjects with IPF differentiated at ALI. Compared to healthy PBECs, in the absence of treatment, IPF PBECs released significantly greater amounts of IL-6, suggesting that the cells in the lungs of patients with IPF may be in a state of persistent activation and are spontaneously producing proinflammatory mediators regardless of the presence of any stimulating factors. Acetate, propionate and butyrate induced a dose-dependent release of IL-6 in both healthy and IPF cultures over the time-course. The secretion of IL-6 was found to be significantly increased in IPF compared to healthy controls at 24-, 48-, 96-, and 144 hours post-exposure to all concentrations of the SCFAs. Compared to IPF PBECs treated with media alone, a significant difference was only observed at the 24-hour time-point in the PBECs from IPF patients following treatment with either 1 mM or 5 mM of propionate or with 5 mM of acetate or butyrate.

Fibroblasts are pivotal in the maintenance of normal tissue function as well as in the wound healing response. They are mainly located in the subepithelial layer of the conducting airways and the interstitium of the lung parenchyma, where they interact with epithelial and endothelial cells [180]. Through the generation of ECM fibroblasts provide a scaffold for cells and several mediators which allow the lungs to respond and adapt to varying degrees of stress. The ECM is composed of extracellular macromolecules such as collagen, enzymes and glycoproteins, which provide structural and biochemical support to surrounding cells [429]. Collagen is the most abundant ECM component in most tissues; collagen I and III are the principal collagens elevated in the parenchyma of patients with IPF [430]. The exaggerated deposition of ECM contributes to the development of the clinical symptoms associated with IPF.

Given the crucial role of fibroblasts in mediating fibrosis, the effects of the SCFAs were also assessed on primary cultures of HLF isolated from subjects with IPF. The effects of the SCFAs acetate, propionate and butyrate on cell viability of IPF HLF was investigated to exclude any SCFA-induced cytotoxicity. HLF were either left untreated or exposed to increasing doses of the SCFAs (1 mM, 5 mM, 10 mM and 20 mM) over a time-course of 48 hours. Exposure of the HLF to the higher doses of the SCFAs (10 mM and 20 mM) induced a significant decrease in cell viability over time. In contrast, HLF treated with the lower doses (1 mM and 5 mM) of the respective SCFAs had no effects on cell viability.

The role of exogenous TGF-β in fibroblast proliferation and collagen production is well known. TGF-β is a key cytokine for the transdifferentiation of fibroblasts to myofibroblasts which exhibit a

contractile phenotype with increased expression of alpha-smooth muscle actin [431]. Lung fibroblasts secrete collagen during normal tissue repair as well as in the development of fibrosis [432]. Fibronectin is a glycoprotein which mediates cell-matrix adhesion by binding proteins, including collagen and facilitates cell migration and differentiation during wound healing [433]. To assess whether the SCFAs acetate, propionate and butyrate had an effect on HLF from IPF subjects, the secretion of TGF-β, collagen I α1 and fibronectin were quantified in the cell supernatants of either untreated or treated HLF with increasing doses of the SCFAs. At 24 hours, compared to cells treated with media alone, HLF exposed to 10 mM, and 20 mM of acetate were found to have a greater concentration of collagen I α1 in the cell culture supernatant. Similarly, fibronectin was increased in the cell culture supernatant following treatment with 20 mM acetate. No changes were observed in collagen I α1 and fibronectin when treating the cells with the lower doses of acetate. Compared to the untreated control TGF-β was found to be unchanged following the addition of acetate even at the highest concentration. No changes were observed in collagen I α1, fibronectin and TGF-β in the cell culture supernatants of the IPF HLF following treatment with propionate and butyrate at 1 mM, 5 mM, 10 mM or 20 mM. It should be noted, however, that only supraphysiologic concentrations of acetate, which have been shown to decrease cell viability, were found to induce a response and therefore might not be biologically relevant.

The effects of the SCFAs on HLF proliferation were assessed in HLF from three patients with IPF over a time-course of 72 hours. HLF were either left untreated or exposed to increasing doses of acetate, propionate or butyrate. Addition of acetate was found to have no effect on HLF proliferation compared to the cells treated with media alone. In contrast, following exposure of the HLF to 5 mM of propionate and butyrate, an increase in proliferation was observed compared to the untreated control. This finding may suggest that acetate may be acting through a different receptor compared to propionate and butyrate. The molecular mechanisms responsible for the effects of SCFA on PBECs and HLF in IPF are still lacking. Future research should characterise the expression of the GPCRs on these cell types to gain a better understanding of the host-microbe interactions that underpin IPF.

There are several limitations that must be considered. Although primary cells are more physiologically relevant to *in vivo* organs, contamination with microorganisms can occur during the isolation process. Several factors, including sample type, sample handling, cell isolation protocol, type of media and culturing techniques may affect the primary airway epithelial cultures [434]. The number of passages also needs to be considered, as cells will dedifferentiate and lose their phenotype at later stages. For this reason, experiments using PBECs and HLF were all performed at

earlier stages (passage 3). Another limitation which must be considered is the high degree of variability between donors [435]. Although SCFAs have been studied in the context of allergic airway inflammation [436] and cystic fibrosis [175, 353], the present study is the first to investigate the effects of the SCFAs in the lower airways of subjects with IPF. It is therefore challenging to compare the results observed to the literature as SCFA have been shown to elicit pro and anti-inflammatory effects in other lung diseases such as asthma and cystic fibrosis.

Having established the direct effects of acetate, propionate and butyrate on both epithelial cells and fibroblasts future research should aim at using co-culture systems to examine the influence of these microbial metabolites on the interactions between PBECs and HLF that drive fibrosis (via direct cell-cell contact or soluble mediators) [400]. A co-culture model would facilitate a better understanding of the effects of the SCFAs on these cell types and would be critical to recapitulate the epithelial-fibroblast cross-talk, which is thought to be key in the development of IPF. Fibroblast response in co-culture could be assessed by proliferation, collagen synthesis and secretion of growth factors while cellular permeability, mucous production, wound healing, and key cytokine secretion could be used to assess the epithelial response. Epithelial repair responses were interrogated using standard wound healing assays in real-time using the JuLI Stage Real-Time Cell History Recorder. Future experiments should assess the impact of SCFAs on key features of epithelial repair/remodelling, including cell spreading, actin cytoskeletal re-organisation, adherens/tight junctions and focal adhesion formation. The impact of SCFA on pro-fibrotic genes and the secretion of specific growth factors including epidermal growth factor and platelet-derived growth factor should also be investigated.

5.6 Conclusions

In summary, these findings show that SCFAs in concentrations found in the BAL from IPF patients can influence the inflammatory response of airway epithelium and may contribute to differences between normal and IPF airway inflammation. In submerged cultures of PBECs, the SCFAs acetate, propionate and butyrate were shown to have an effect cell viability. Supplementation of the IPF PBECs with non-toxic concentrations of these SCFAs was found to induce an increased release of IL-6 and TNF-α in the cell culture supernatant at 24 and 48 hours. In IPF, no differences in cell proliferation were observed pre- and post-exposure to the SCFAs. However, supplementation of the IPF PBECs with 5 mM of either acetate, propionate or butyrate led to a significant decrease in wound closure over time. These results were also validated in PBECs differentiated at ALI, where non-toxic doses of the SCFAs were shown to have no effects on the healthy PBECs but were found to decrease

the TEER of the IPF PBECs and impact their morphology. Exposure of the IPF HLF to propionate and butyrate was found to increase cell proliferation at 72 hours. Future work is needed to better understand the nature of the dialogue occurring between the metabolites and epithelial cells and fibroblasts in the lungs to further elucidate the function and mechanisms through which the respiratory microbiota shapes the development of fibrosis in IPF.

6 **Chapter 6**: Changes in the Respiratory Microbiome and Metabolome during Acute Exacerbations of Idiopathic Pulmonary Fibrosis

6.1 Introduction

A diagnosis of IPF can be made based on the exclusion of other known causes of ILD and the presence of characteristic histopathological and radiological patterns referred to as UIP [437]. Despite recent guideline revisions, correctly diagnosing IPF remains a key issue for health professionals as the clinical manifestations and the pathological patterns of UIP that characterise the disease may also be found in others forms of ILD [9]. Further adding to this diagnostic complexity is the significant heterogeneity among individual patients, as the clinical course of the disease is both variable and unpredictable [16]. While some patients with IPF experience a steady decline in lung function, others suffer sudden acute respiratory deterioration, following a period of relative stability. These episodes termed AE-IPF are an important clinical feature of IPF disease progression and have been defined as clinically significant respiratory deteriorations of unidentifiable cause characterised by new, widespread alveolar abnormalities [25]. The exact incidence of AEs is unclear but is known to confer substantial morbidity and mortality (often > 50%) and it is estimated that each year between 4-20% of individuals experience an AE [27, 28]. Prior to the latest revised diagnostic criteria [25], the exclusion of any infective trigger was required for the diagnosis of an AE-IPF [438]. Nonetheless, our current understanding of the pathophysiology of AE-IPF is limited and it remains unclear whether AE-IPF represent an accelerated phase of the underlying fibrotic condition or a response to occult infection. Although there is a paucity of information regarding the aetiology, pathophysiology and treatment of AE, recent observational studies in IPF patients indicate that infection, both viral and bacterial, might play a role in some cases of AE-IPF [145, 321]. While past studies have mainly focused on the role of viruses in AE-IPF [90, 439, 440], the long-held, but incorrect, theory of sterility outside of clinical infection, means historically little work has evaluated the role of bacteria in IPF pathogenesis [145]. More recently, thanks to advances in high-throughput molecular sequencing technologies, the role of the respiratory microbiome in chronic lung diseases, including IPF, is beginning to be elucidated [114, 149, 277, 321]. The changes occurring in the respiratory microbiome of patients experiencing an AE-IPF were recently characterised [144]. However, this study had a number of limitations, including lack of paired sampling prior to and during an exacerbation in addition to limited cohort size. In this, the largest prospective study of the respiratory microbiome in AE-IPF, the potential changes which occur in the respiratory microbiome and metabolites during AE were examined in a cohort of IPF subjects (*n* =

40) who underwent bronchoscopy at baseline and during an AE. The hypothesis examined was that, compared to baseline, subjects with AE-IPF have an altered bacterial burden, respiratory microbiome, and respiratory metabolome, which contributes to the increased morbidity and mortality.

6.2 Aims

- To quantify the bacterial burden in paired samples of BAL supernatants of patients with IPF at baseline and during AE.
- To characterise the taxonomic changes, differentially abundant bacteria, alpha- and beta-diversity differences at baseline and during AE-IPF
- To evaluate the metabolomic and cytokine profiles in the BAL supernatants of patients with IPF at baseline and during AE.

6.3 Materials and Methods

6.3.1 Study design and sampling

Subjects were prospectively recruited at the Tosei General Hospital (Seto, Japan), between January 2008 and July 2015. Forty subjects with IPF were included in this study. All cases were subjected to MDD to establish a diagnosis, using the patients' clinical history, physical examination, laboratory results, chest HRCT scans, BAL fluid and histopathological findings. Following MDD, the final diagnoses were assigned as IPF using established criteria [2, 13, 441-443]. Diagnoses of AE-IPF were made according to the latest diagnostic criteria [25]. Pulmonary function testing was performed at baseline, and clinical measurements were recorded. Matched samples were obtained for all patients at baseline and during an AE. Total cell count was done manually using a haemocytometer, and BAL differentials were performed following cytospin centrifugation and Wright-Giemsa stain to assess cell counts. All patients underwent fibreoptic bronchoscopy with BAL as part of their routine clinical evaluation. In subjects experiencing an AE, this procedure was undertaken within 48 h of hospital admission. BAL was performed in the most affected lobe, and three 50 ml aliquots of saline were instilled. BAL was centrifuged to remove mucous and cells and supernatant stored at -80°C until further analysis.

6.3.2 Measurement of clinical parameters

Lung function parameters, including FVC and DL_{CO} were measured according to the American Thoracic Society (ATS)/European Respiratory Society (ERS) recommendations [444, 445]. Serum measurements of C-reactive protein (CRP), Krebs von den Lungen-6 (KL-6), and surfactant protein D (SP-D) were performed at the clinic at baseline and during an AE-IPF.

6.3.3 ELISA

BAL supernatants were assessed for concentrations of IL-6, TNF-α, TGF-β, IL-8 and IL-17 by ELISA as per manufacturer's instructions (R & D Systems).

6.3.4 Bacterial DNA isolation, quantification and 16S rRNA gene sequencing

BAL samples (2 ml) were centrifuged at 21,000 x *g* for 30 minutes to pellet cell debris and bacteria, and genomic DNA was then isolated. High-throughput sequencing of bacterial 16S rRNA gene amplicons encoding the V4 region (150 bp read length, paired-end protocol) was performed using an Illumina MiSeq Sequencer (Illumina, San Diego, CA). Further details are provided in **Chapter 2 (sections 2.2.1-2.2.5)**.

6.3.5 Bioinformatics analysis

The 16S rRNA gene sequences were analysed using the QIIME2 pipeline for analysis of microbiome data. Details on bioinformatics analysis are provided in **Chapter 2 (section 2.2.6)**.

6.3.6 16S rRNA gene qPCR

Bacterial burden was measured by 16S rRNA gene qPCR as previously described in **Chapter 2 (section 2.3)**.

6.3.7 Triple Quadrupole Gas Chromatography-Mass spectrometry

Targeted GC-MS of the SCFA was conducted on an Agilent 7890A (Agilent Technologies, CA, USA). See **Chapter 2 (section 2.5)** for further details.

6.3.8 Statistical analysis

Continuous variables are presented as means (± standard deviation) and categorical variables as proportions. Data analysis was performed using *R* (version 3.4.3) [294]. Details of statistical analysis performed are described in **Chapter 2 (section 2.2.10)**.

6.4 Results

6.4.1 Study subjects and sampling

Forty patients with IPF enrolled at the Tosei General Hospital (Seto, Japan) and underwent bronchoscopy **(Table 6.1)**. The subjects were predominantly male (87%), with a mean age of 68 years ± 9, and had moderately severe restrictive lung disease at enrolment (FVC, 78% predicted; DL_{CO}, 58% predicted). Over half of the patients were ex-smokers (53%), whilst 37% were classified as non-smokers and only 10% as current-smokers. Seventeen out of the forty patients received antibiotics prior to the second bronchoscopy (i.e. at the time of an acute exacerbation). Seven of these patients also received corticosteroids.

Table 6.1 Baseline clinical characteristics of study subjects.

Parameter	IPF (*N* = 40)
Age, years (± SD)	68 (9)
Male, *n* (%)	35 (87)
BMI (± SD)	23 (3)
Baseline PFT	
FVC % predicted (± SD)	78 (18)
DL_{CO} % predicted (± SD)	58 (19)
PaO_2, torr (± SD)	77 (12)
Smoking status	
Non-smoker, *n* (%)	15 (37)
Ex-smoker, *n* (%)	21 (53)
Current-smoker, *n* (%)	4 (10)

Data are presented as percentage value or mean ± standard deviation (SD) as appropriate. IPF: idiopathic pulmonary fibrosis; BMI: body mass index; PFT: pulmonary function test; FVC: forced vital capacity; DL_{CO}: diffusing capacity of the lung for carbon monoxide; PaO_2: partial pressure of oxygen.

Differential cell counts demonstrated differences in the airway populations and serum markers at baseline and during an acute exacerbation **(Table 6.2)**. Neutrophils ($P < 0.0001$), lymphocytes ($P < 0.05$) and eosinophils ($P < 0.01$) were higher during an AE compared to baseline. Concentrations of CRP, KL-6 and SP-D were measured in the serum at baseline and during an AE-IPF. CRP ($P < 0.0001$) and KL-6 ($P < 0.0001$) were elevated during an AE compared to baseline whilst SP-D was unchanged **(Table 6.2)**. All AE-IPF events were deemed to be non-infective using clinical diagnostic tools.

Table 6.2 Changes in bronchoalveolar lavage (BAL) cell differentials and serum marker concentrations.

Parameter	Baseline-IPF (N = 40)	AE-IPF (N = 40)	*P*-value*
BAL cell differentials			
Neutrophils % (± SD)	5 (11)	24 (24)	<0.0001
Lymphocytes % (± SD)	8 (10)	13 (12)	<0.05
Eosinophils % (± SD)	1 (2)	4 (9)	<0.01
Serum markers			
CRP, mg/l (± SD)	0 (1)	8 (6)	< 0.0001
KL-6, U/ml (± SD)	1478 (1094)	1699 (1133)	< 0.0001
SP-D, ng/ml (± SD)	364 (307)	429 (353)	NS

Data are presented as percentage value or mean ± standard deviation (SD) as appropriate. IPF: idiopathic pulmonary fibrosis; AE-IPF: acute exacerbation of idiopathic pulmonary fibrosis; BAL: bronchoalveolar lavage; CRP: C-reactive protein; KL-6: Krebs von den Lungen-6; SP-D: surfactant protein D. **P*-values are calculated by Wilcoxon matched-pairs signed-rank test between baseline-IPF and AE-IPF subjects.

6.4.2 Sampling

6.4.2.1 DNA yield and sequencing quality

All 86 samples yielded genomic DNA and underwent amplification of the V4 region of the 16S rRNA gene using the barcoded universal primer pair 515F/806R on an Illumina MiSeq sequencer (New York, USA). The quantity (ng/µl) and quality (absorbance ratios: A260/280, A260/230) of the isolated DNA was assessed. Pure nucleic acids typically yield absorbance ratios of ~1.8 (A260/280) and ~2.00 (A260/230) [278]. The average yield from the baseline-IPF samples was 19.61 ng/µl ± 9.8 and from the AE-IPF samples 19.28 ng/µl ± 7.34 (P = 0.29), with an A260/A280 ratio of 1.77 ± 0.08 and an A260/230 of 1.86 ± 0.17 indicating good quality DNA. Following paired-end Illumina sequencing, a total of 7 million sequences were obtained from the 86 samples. The lowest number of reads per sample was observed in the negative controls (kit controls and bronchoscopy controls) **(Figure 6.1)**. Following demultiplexing and denoising with DADA2 a total of 1,014 high-quality 16S rRNA reads remained, with on average 50,336 ASVs per sample **(Table 6.3)**. After singleton removal and filtering at 0.005% of total reads to remove low abundance ASVs likely representing sequencing artefacts, 3,962,856 sequences were retained, representing 135 unique ASVs across 80 samples. This final dataset was used in all subsequent analyses.

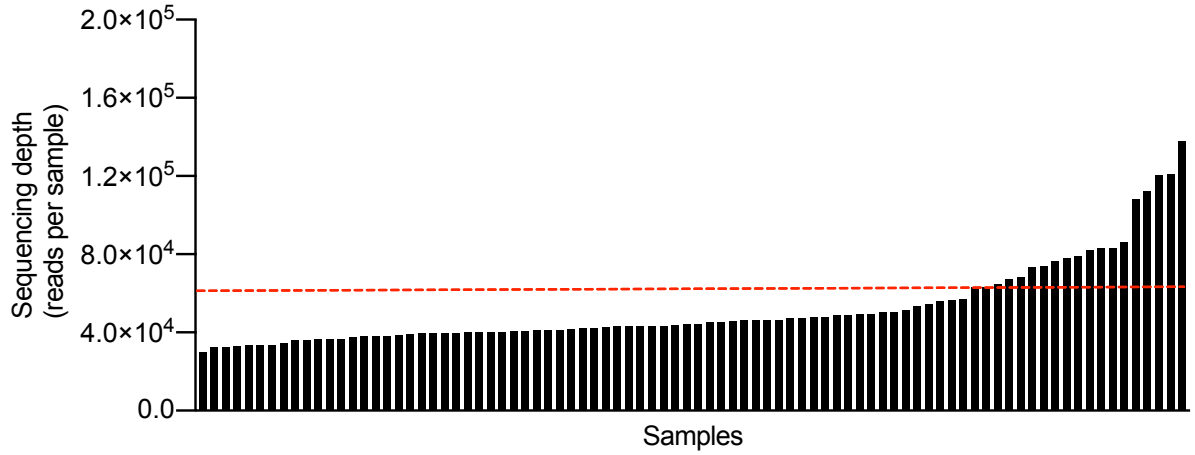


Figure 6.1 Distribution of samples sequence depth. Illustrating the total number of reads present in each sample.

The red dashed line represents the mean number of reads.

Table 6.3 Pre-processing of 16S rRNA sequence reads from bronchoalveolar lavage samples.

Metric	16S rRNA amplicon reads generated	Post-Dada2 and clustering into ASVs	Post-filtering at 0.005% of total reads
Number of samples	86	86	80
Number of ASVs	3,199	1,014	135
Total number of sequence reads	7,099,795	6,896,156	3,962,856
Mean reads per sample ± SD	51,823 ± 19,749	50,336 ± 19,885	30,483 ± 19,870
Minimum reads per sample	246	246	3,504
Maximum reads per sample	138,051	136,518	109,523

ASVs: amplicon sequence variants; DADA2: divisive amplicon denoising algorithm. *Including negative control samples which were included in sequencing run but filtered out from the final dataset.

6.4.2.2 Contamination

To identify potential sources of contamination several negative bronchoscopy controls and reagent controls were included in the sequencing run. By adding these controls, reads originating from contamination were identified and distinguished from those derived from actual constituent taxa. Negative control samples had a distinct taxonomic profile compared to the BAL samples and yielded a bacterial burden close to or below the lower limit of qPCR quantification (100 copies/ml of BAL) **(Figure 6.2)**.

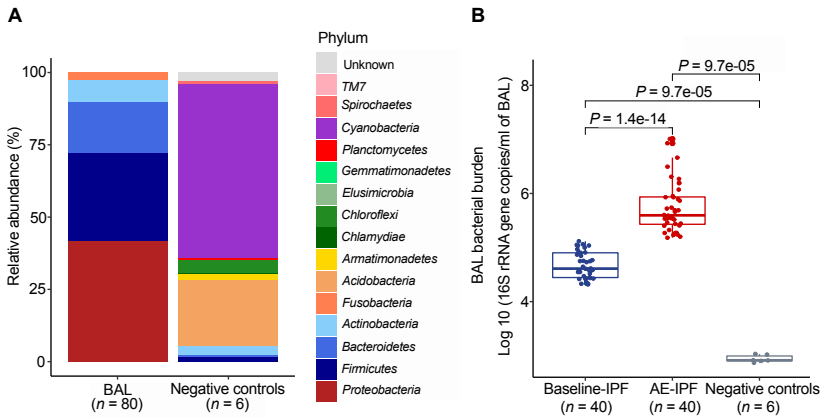


Figure 6.2 Comparison between all bronchoalveolar lavage (BAL) samples and negative controls.

(A) Taxonomic differences at the phylum level comparing all of the BAL samples (*n* = 80) to the negative control samples (*n* = 6) regardless of diagnosis. The number refers to the relative abundance (%) obtained prior to conducting any filtering step. **(B)** Bacterial burden calculated by qPCR comparing the bronchoalveolar lavage samples (Baseline-IPF = 40, AE-IPF = 40) to the negative controls (*n* = 3 negative bronchoscopy controls, *n* = 3 reagent controls). Sample groups were compared by Wilcoxon matched-pairs signed-rank test between Baseline-IPF and AE-IPF subjects. Overall *P*-value: $P < 1.8 \times 10^{-15}$.

6.4.2.3 Mock Community

A mock community was included as a positive control in each run to enable accurate comparison across sequencing runs **(Figure 6.3)**. The number of sequencing reads and the composition of the mock community was similar in both runs, confirming that all genera were represented in similar proportions.

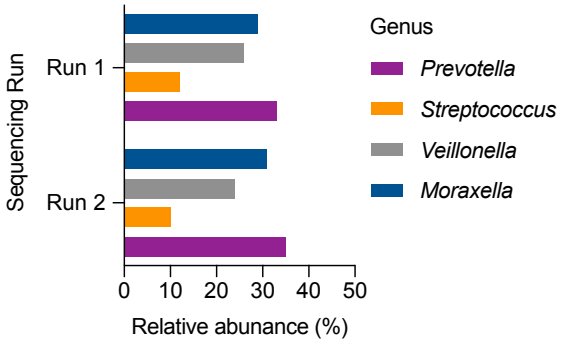


Figure 6.3 Mock community composition at the genus level from the two sequencing runs.

6.4.3 Bacterial burden

6.4.3.1 Bacterial burden at baseline and during an AE-IPF

It has previously been reported that AE-IPF subjects have a higher bacterial load compared to subjects with IPF at baseline [144]. This finding was validated in this cohort of patients. On average, AE-IPF subjects had 1.17×10^{6} copies of the 16S rRNA gene/ml of BAL, which was significantly higher ($P < 1.4 \times 10^{-14}$) than the average copy number at baseline (5.56×10^{4} copies of the 16S rRNA gene/ml of BAL) **(Figure 6.4)**.

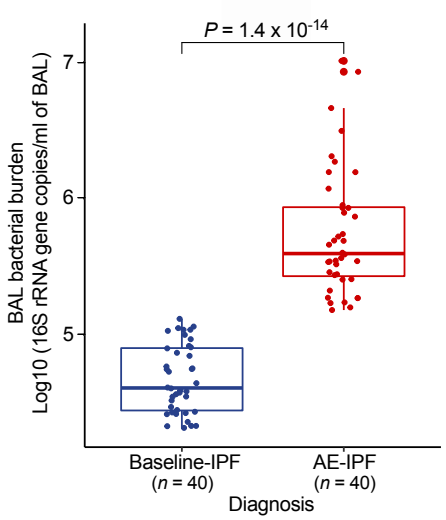


Figure 6.4 Bacterial burden in bronchoalveolar lavage (BAL) of stable and acute exacerbation subjects.

Comparison of bacterial load of idiopathic pulmonary fibrosis subjects (IPF) at baseline and during an acute exacerbation (AE). Bacterial burden calculated by qPCR and expressed as log10 16S rRNA gene copies/ml of BAL. Data are presented as median and interquartile range. Statistical significance tested by Wilcoxon matched-pairs signed-rank test.

6.4.3.2 Bacterial burden and time to exacerbation

To test for a relationship between the bacterial burden at the time of diagnosis and the time taken to develop an AE-IPF, patients were stratified into tertiles according to bacterial load. In an adjusted multivariable Cox proportional hazards model (adjusted for age, sex, smoking status and baseline % predicted FVC) bacterial burden at baseline did not predict the rate of disease progression (i.e. the time to an AE-IPF) (hazard ratio: 0.90, 95% confidence interval: 0.37-2.20, $P < 0.824$) **(Figure 6.5)**.

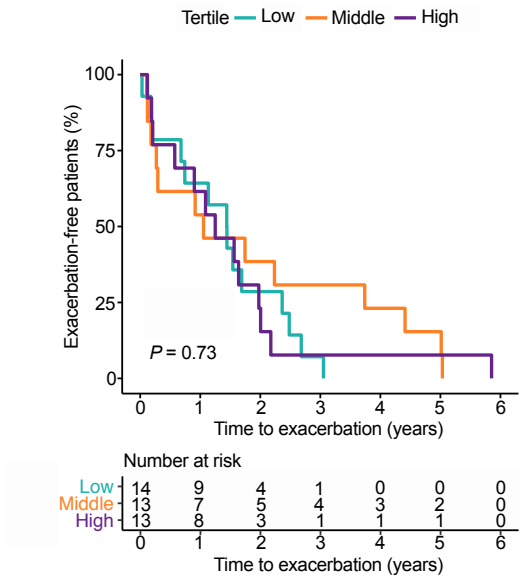


Figure 6.5 Kaplan-Meier curve comparing the time from diagnosis to an acute exacerbation of IPF (AE-IPF).

Kaplan-Meier curve generated by Cox proportional-hazards model for IPF subjects stratified into tertiles by bacterial burden (low: 2.08 x 10^4 – 3.52 x 10^4 16S rRNA gene copies/ml of BAL; middle: 3.53 x 10^4 – 5.86 x 10^4 16S rRNA gene copies/ml of BAL; high: 5.87 x 10^4 – 1.31 x 10^5 16S rRNA gene copies/ml of BAL). Log-rank *P* test value reported.

6.4.3.3 Bacterial burden and BAL differential cell counts

Differential cell counts were performed in the baseline and AE-IPF groups with identification of neutrophils, lymphocytes and eosinophils. The relationship between BAL differential cell count and bacterial burden was examined. No correlation was found between BAL bacterial load and the BAL differential cell counts at baseline or during an AE **(Figure 6.6)**.

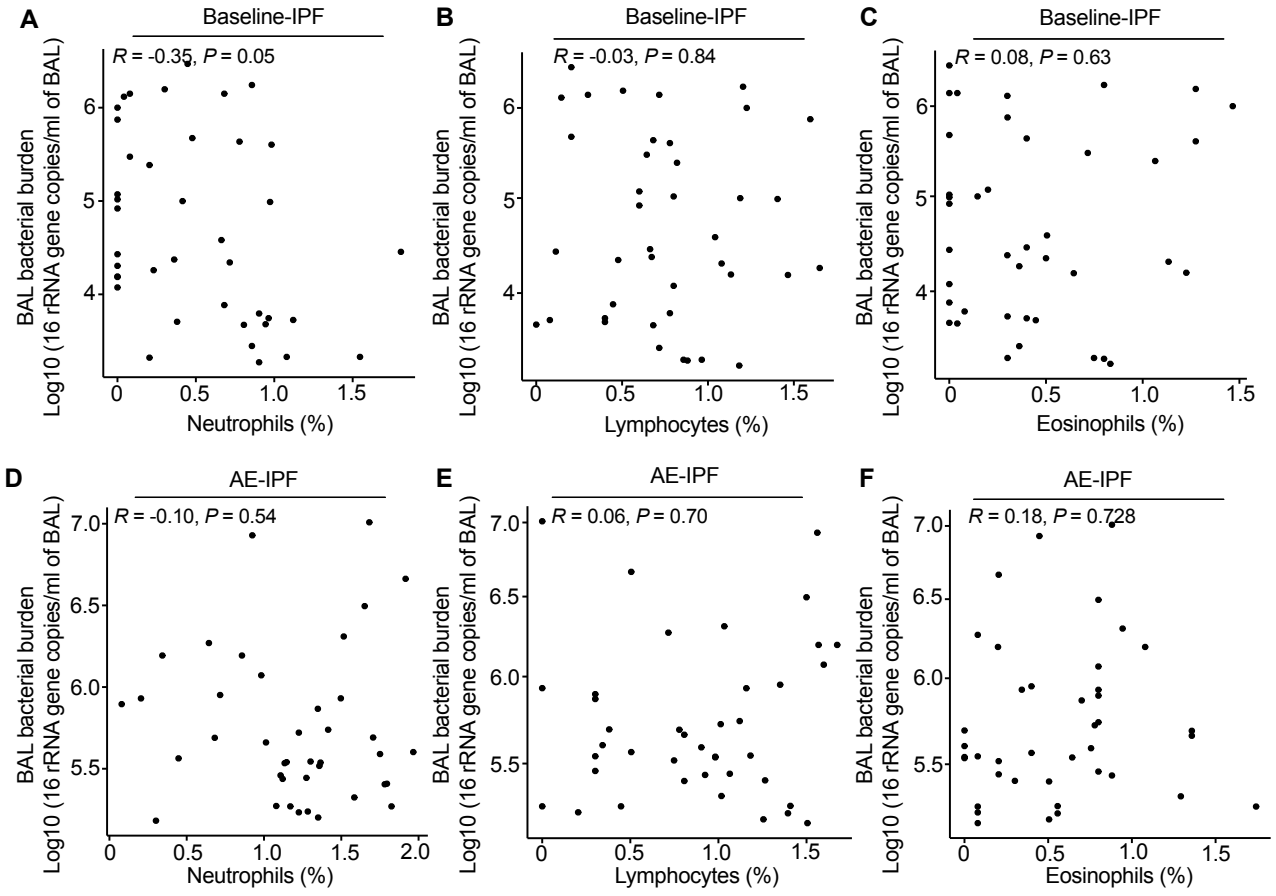


Figure 6.6 No correlation between bronchoalveolar lavage (BAL) differential cell counts and bacterial burden of subjects with idiopathic pulmonary fibrosis (IPF).

No correlation between BAL bacterial burden and percentage of neutrophils, lymphocytes, and eosinophils at **(A-C)** baseline and during an **(D-F)** AE-IPF. BAL bacterial burden was calculated by qPCR and expressed as log10 16S rRNA gene copies/ml of BAL. Correlations between cytokines and propionate were determined by Spearman's rank test.

6.4.3.4 Bacterial burden and antibiotic treatment

Out of the 40 patients included in the study, a total of 17 patients were administered antibiotics prior to the second bronchoscopy. In order to determine the potential impact of antibiotic intake on BAL bacterial burden, subjects were dichotomised into two groups based on whether they had received antibiotics or not. On average, subjects who had received antibiotics had 1.17×10^6 copies of the 16S rRNA gene/ml of BAL, which was not significantly lower ($P = 0.83$) than the average copy number of those who had not received antibiotics (7.50×10^5 copies of the 16S rRNA gene/ml of BAL) **(Figure 6.7)**.

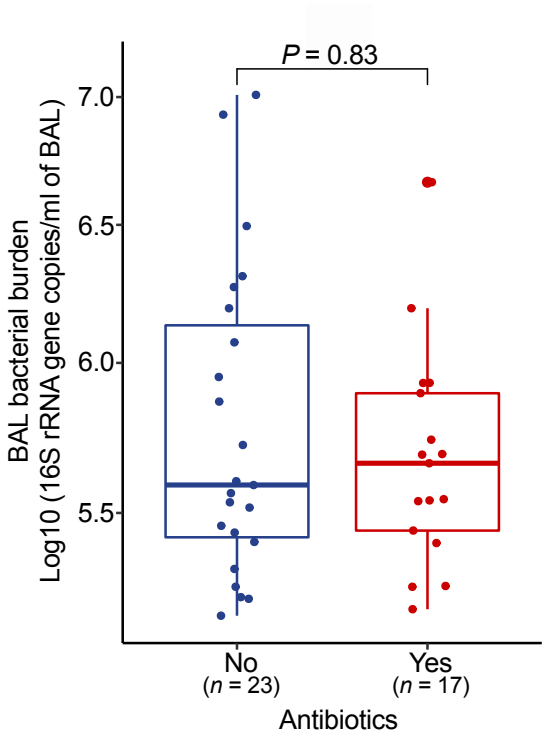


Figure 6.7 Bacterial burden in bronchoalveolar lavage (BAL) of subjects with idiopathic pulmonary fibrosis (IPF) stratified according to exposure to antibiotics.

BAL bacterial burden quantified using quantitative polymerase chain reaction and is expressed as log10 (16S rRNA gene copies/ml of BAL). Data are presented as median and interquartile range. Statistical significance tested by Mann-Whitney U test.

6.4.4 Microbiota composition at baseline and during an AE-IPF

6.4.4.1 Relative abundance

The composition of the microbial communities at baseline was dominated by the following phyla: *Proteobacteria* (45%), *Firmicutes* (30%), *Bacteroidetes* (17%), *Actinobacteria* (5%), and *Fusobacteria* (2%). The same phyla predominated the BAL during an AE-IPF, and the differences in the overall relative abundance at the phylum level between baseline and AE were not statistically significant **(Figure 6.8)**. Comparison at the genus level identified that the microbial landscape of the subjects with IPF was dominated by *Streptococcus*, *Prevotella* and *Veillonella*. Compared to baseline-IPF, *Streptococcus* ($P < 0.05$) and *Campylobacter* ($P < 0.05$) were found to be increased in the AE-IPF samples. In contrast, *Porphyromonas* ($P < 0.01$) was found to be decreased in the AE-IPF samples compared to baseline **(Figure 6.9)**.

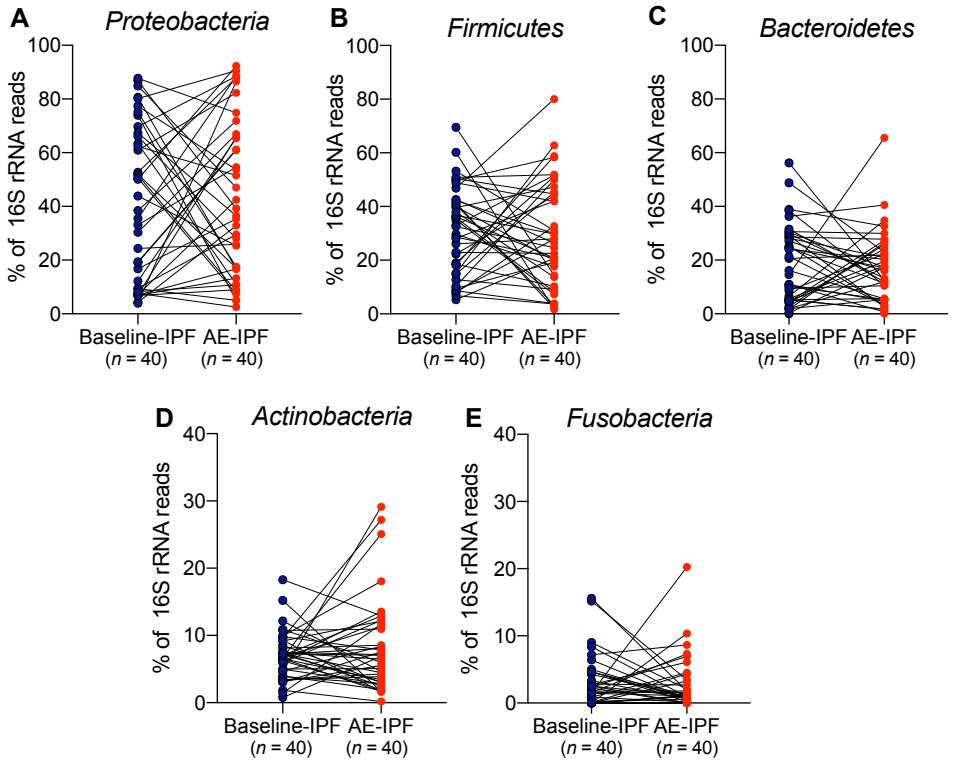


Figure 6.8 Mean taxonomic distribution at the phylum level.

Percentage of total 16S rRNA reads of **(A)** *Proteobacteria*, **(B)** *Firmicutes*, **(C)** *Bacteroidetes*, **(D)** *Actinobacteria*, and **(E)** *Fusobacteria* at baseline and during an AE-IPF. Statistical significance tested by Wilcoxon matched-pairs signed-rank test.

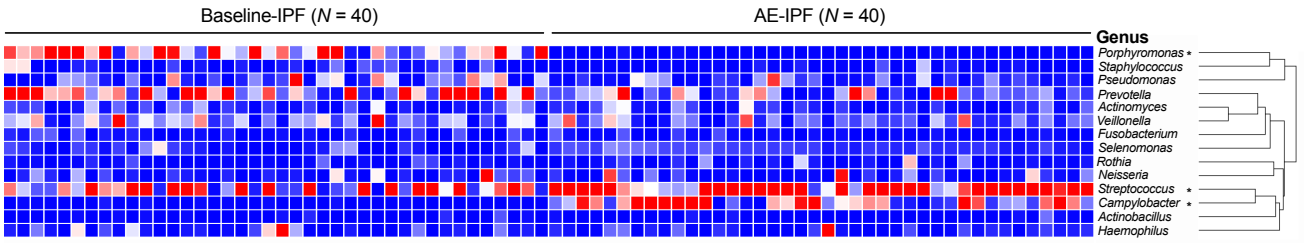


Figure 6.9 Heatmap of bacterial genera in the bronchoalveolar lavage (BAL) of subjects with idiopathic pulmonary fibrosis (IPF).

Hierarchical clustering of the top 14 bacterial genera comparing baseline-IPF and AE-IPF. The variation in colour from blue (decrease) to red (increase) illustrates the changes in the BAL of each bacterial genera.

6.4.4.2 Differential abundance

To further investigate the community composition changes at baseline and during an AE-IPF, differentially abundant bacteria were identified using LEfSe. Features significantly discriminating among groups with an LDA score > 2.0 were represented as a cladogram, as produced by LEfSe with default parameters. Multiple taxonomic differences were found: the BAL supernatant of subjects at baseline was enriched with *Prevotella* and *Porphyromonas*, with the BAL supernatant of IPF subjects experiencing an acute exacerbation enriched with bacteria belonging to the *Campylobacteraceae* and *Pseudomonadaceae* families **(Figure 6.10)**.

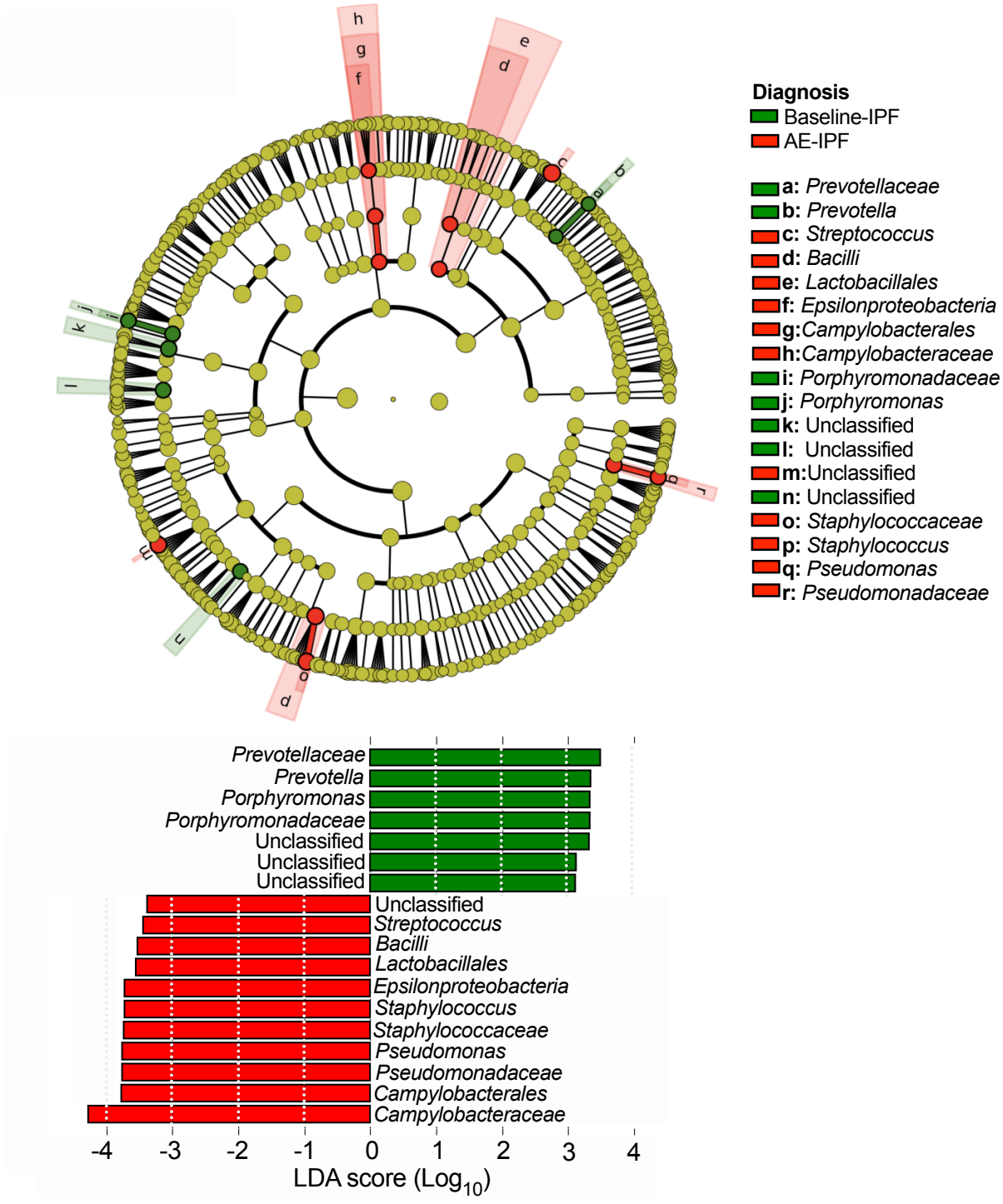


Figure 6.10 Differentially abundant bacteria within the baseline- idiopathic pulmonary fibrosis (IPF) and acute exacerbation of IPF (AE-IPF) groups.

Cladogram representing results from calculated linear discriminant analysis (LDA) effect size (LEfSe) comparing the taxonomic composition of BAL samples from the two groups, $P < 0.05$. Multiple significant taxonomic differences were observed at different phylogenetic levels. Letters in the cladogram indicate differences at the class, order and family levels. Only taxa meeting an LDA threshold > 2.0 are shown.

6.4.4.3 Diversity analyses

Next, the Shannon index and Chao1 diversity index alpha-diversity metrics were used to determine whether the subjects with IPF at baseline and during an AE differed in the distribution (evenness) and number (richness) of taxa. Compared to baseline-IPF, during an exacerbation, there were no changes in microbial richness or evenness **(Figure 6.11).**

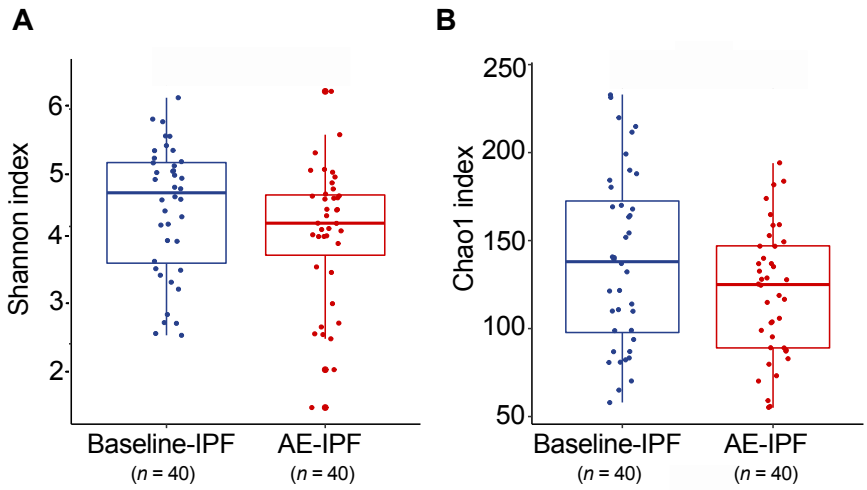


Figure 6.11 Alpha-diversity of microbial communities at baseline and during an acute exacerbation.

Alpha-diversity assessed using the **(A)** Shannon index and **(B)** Chao1 index. Data are presented as median and interquartile range. Statistical significance was tested by Wilcoxon matched-pairs signed-rank test.

Changes in the community composition (β-diversity) at baseline and during an acute exacerbation were investigated using PCA and tested by PERMANOVA which confirmed that there were no significant changes in the microbial profiles between the groups ($P = 0.82$) **(Figure 6.12)**.

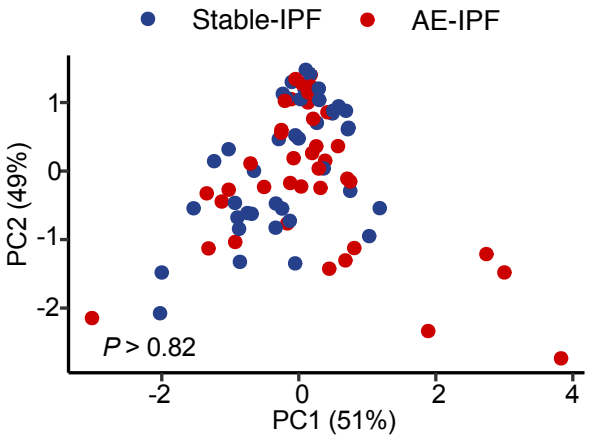


Figure 6.12 Beta-diversity of microbial communities at baseline and during an acute exacerbation.

Principal Component Analysis based on Euclidean distances statistically tested using Permutational Multivariate Analysis of Variance (PERMANOVA) adjusted for age, sex, smoking status, FVC (% predicted) and DL_{CO} (% predicted) (PERMANOVA: $P < 0.82$).

6.4.4.4 Paired samples

The variability in the microbial profile of the subjects with IPF at baseline and experiencing an AE was investigated. The taxonomic profile at the phylum level was characterised for each individual at baseline and during an AE. Comparing the relative abundance of the bacterial phyla of the two patients at baseline revealed significant differences in terms of the proportions of the phyla in each group. Clear changes were observed during an AE in all patients with a tendency towards an increased relative abundance of *Proteobacteria* and *Firmicutes* compared to baseline. For example, patient 1 had a 36% rise in the abundance of *Proteobacteria* during an exacerbation with an associated drop of *Firmicutes* (12%), *Bacteroidetes* (13%), *Actinobacteria* (4%) and *Fusobacteria* (7%). In contrast, in the case of the second patient, the microbiome at baseline was dominated by *Bacteroidetes* (56%) and *Firmicutes* (28%), and during an AE there was an evident drop in *Bacteroidetes* (44%) with an associated increase in *Firmicutes* (14%) and *Proteobacteria* (29%) **(Figure 6.13)**.

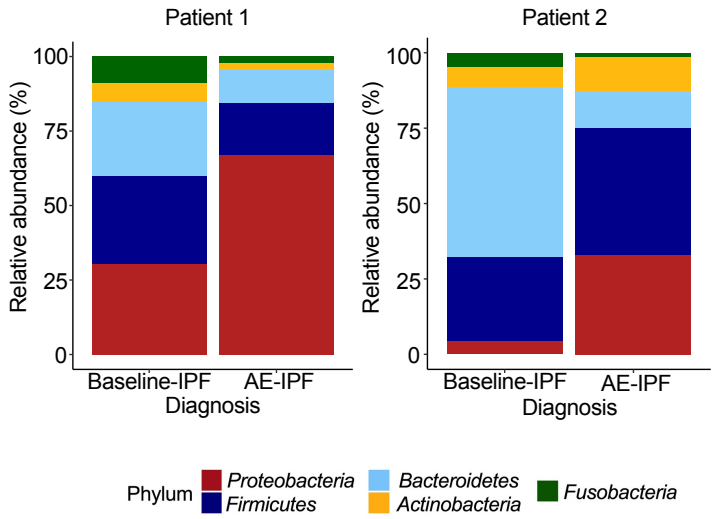


Figure 6.13 Taxonomic distribution at the phylum level in paired samples from two subjects diagnosed with idiopathic pulmonary fibrosis (IPF).

Illustrating the changes occurring at baseline and during an acute-exacerbation.

6.4.5 Cytokine profiles at baseline and during an AE-IPF

BAL supernatant concentrations of IL-6, IL-8, IL-17, TGF-β, TNF-α and IL-1β at baseline and during an AE-IPF were quantified by ELISA. IL-8, IL-6, and TNF-α were found to be significantly increased during AE-IPF compared to baseline ($P < 0.0001$) **(Figure 6.14)**. In contrast, IL-17, TGF-β and IL-1β were undetectable in the BAL supernatant. At baseline, there were no significant correlations between BAL bacterial burden and levels of IL-6, IL-8 and TNF-α **(Figure 6.15A-C)**. During an AE-IPF, there was a significant weak, positive correlation between levels of IL-6 and bacterial burden (R =

0.33, *P* = 0.04) and a significant, weak, negative correlation between levels of TNF-α and bacterial burden (R = -0.35, *P* = 0.03) (**Figure 6.15D,F**). No correlation was found between BAL bacterial burden and IL-8 during an AE-IPF **(Figure 6.15E)**.

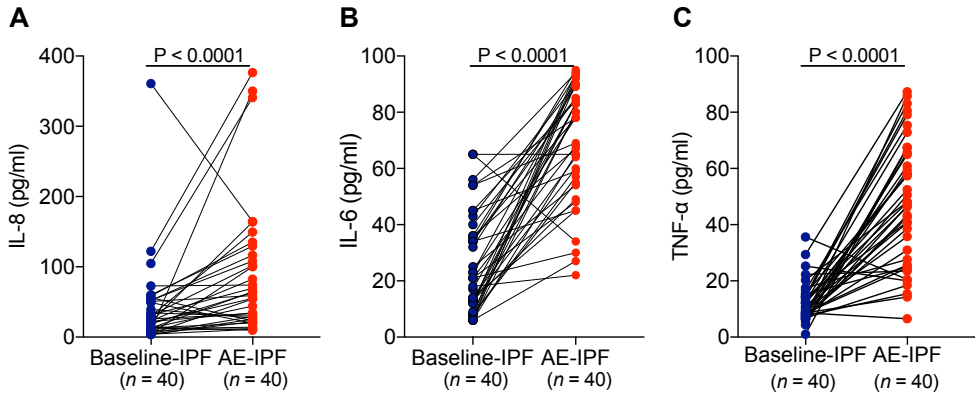


Figure 6.14 Bronchoalveolar lavage (BAL) supernatant levels of inflammatory cytokines at baseline and during an acute exacerbation of idiopathic pulmonary fibrosis (AE-IPF).

Concentrations (pg/ml) in BAL supernatant of **(A)** IL-8, **(B)** IL-6, and **(C)** TNF-α at baseline and during an AE-IPF. Statistical significance tested by Wilcoxon matched-pairs signed-rank test.

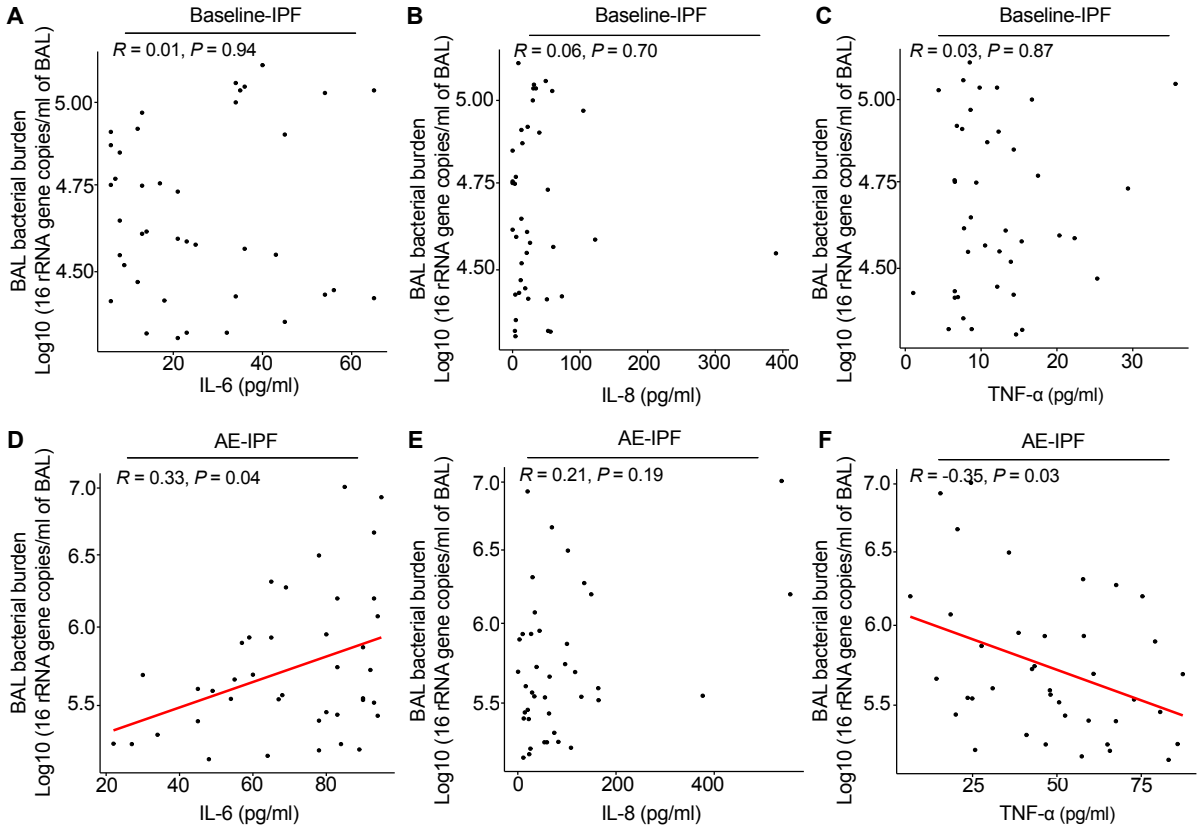


Figure 6.15 Correlation between cytokines detected in bronchoalveolar lavage (BAL) and BAL bacterial burden of subjects with idiopathic pulmonary fibrosis (IPF) at baseline and during an acute exacerbation (AE).

Correlation between BAL bacterial burden and concentrations of IL-6, IL-8, and TNF-α in BAL at **(A-C)** baseline and during an **(D-F)** AE-IPF. BAL bacterial burden was calculated by qPCR and expressed as log10 16S rRNA gene copies/ml of BAL. Correlations between cytokines and propionate were determined by Spearman's rank test.

The percentage of neutrophils in BAL were significantly increased during an AE-IPF compared to baseline (baseline 5% ± 11; AE-IPF 24% ± 24; $P < 0.0001$) **(Table 6.2)**. Since IL-8 is known to mediate the recruitment of neutrophils to the site of injury, the relationship between levels of IL-8 in the BAL and the percentage of neutrophils was assessed [446]. No correlation was found between BAL bacterial load and the BAL differential cell counts at baseline or during an AE **(Figure 6.6)**. However, a weak but significant positive correlation was found between IL-8 and BAL neutrophil percentage (R = 0.34, $P < 0.05$) in the BAL of subjects with AE-IPF, suggesting that IL-8 may promote the recruitment of neutrophils during these events **(Figure 6.16)**. In contrast, no correlations were found between IL-6 and TNF-α and neutrophils or between IL-6, IL-8, TNF-α and percentage of lymphocytes and eosinophils in BAL.

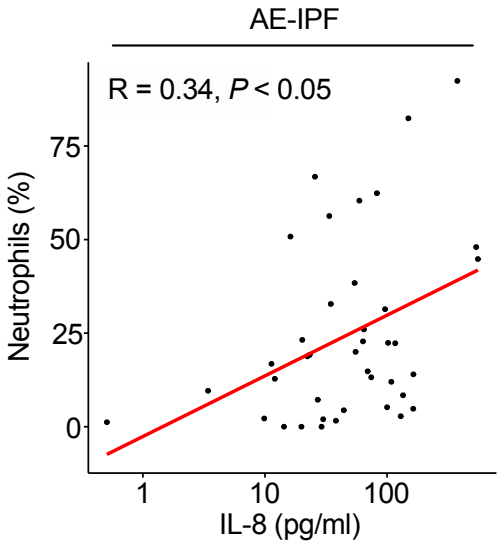


Figure 6.16 Correlation between IL-8 detected in bronchoalveolar lavage (BAL) and percentage of neutrophils in BAL of subjects with an acute exacerbation of idiopathic pulmonary fibrosis (AE-IPF).

Correlations between percentage of neutrophils and IL-8 were determined by Spearman's rank test.

6.4.6 Metabolic changes at baseline and during an AE-IPF

Targeted GC-MS was conducted in order to determine the concentrations of the SCFAs at baseline and during an AE-IPF. The following SCFAs were quantified: acetate, propionate, isobutyrate, butyrate, isovalerate, 2-methylbutyrate, valerate, caproate, lactate and α-hydroxybutyrate **(Figure 6.17)**. Amongst these, 2-methylbutyrate, caproate, butyrate, and valerate were undetectable in the BAL supernatant of subjects with IPF at baseline and during an AE. Compared to baseline, the SCFAs α-hydroxybutyrate ($P < 0.001$), acetate ($P < 0.0001$), and propionate ($P < 0.001$) were found to be increased in the subjects experiencing an AE **(Figure 6.17B-D)**.

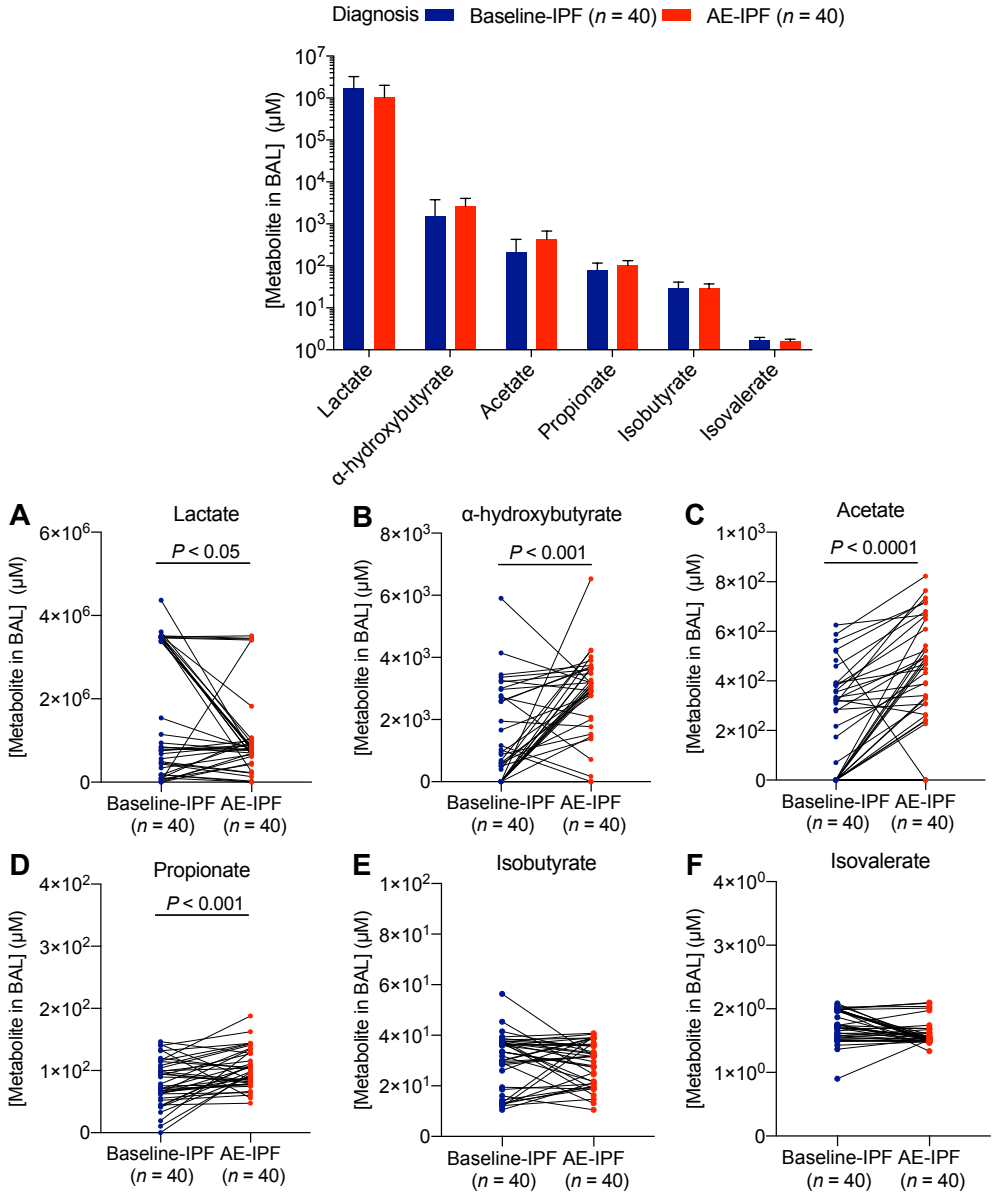


Figure 6.17 Gas chromatography-mass spectrometry (GC-MS) of the short-chain fatty acids (SCFAs) in the bronchoalveolar lavage (BAL) supernatant of subjects with idiopathic pulmonary fibrosis (IPF) at baseline and during an acute exacerbation (AE).

Metabolites that were quantified with GC-MS: **(A)** lactate, **(B)** α-hydroxybutyrate, **(C)** acetate, **(D)** propionate, **(E)** isobutyrate, and **(F)** isovalerate. Data are presented as mean ± standard deviation. Statistical significance tested by Wilcoxon matched-pairs signed-rank test.

In contrast, lactate was found to be increased ($P < 0.05$) at baseline compared to subjects experiencing an AE **(Figure 6.17A)**. No significant differences in the metabolic profiles of these subjects were found when stratified according to smoking status or between patients who had received antibiotics before the second BAL and those who did not. Furthermore, propionate was the only SCFA to be positively correlated with BAL bacterial burden both at baseline (Spearman's *rho* = 0.47, $P < 0.01$) and particularly during an AE (Spearman's *rho* = 0.84, $P < 0.0001$) **(Figure 6.18)**.

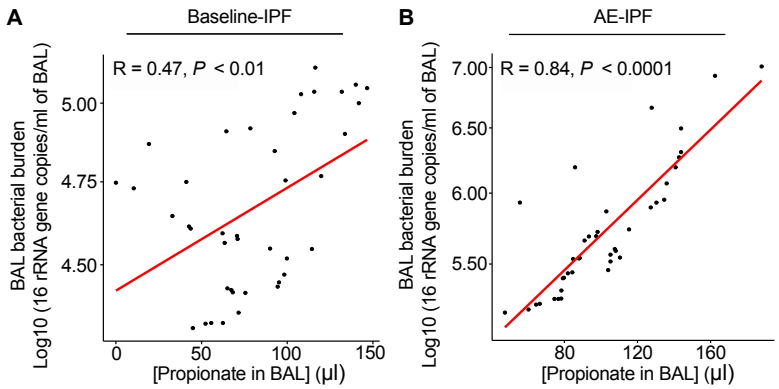


Figure 6.18 Correlation between propionate detected in bronchoalveolar lavage (BAL) and BAL bacterial burden of subjects with idiopathic pulmonary fibrosis (IPF) at baseline and during an acute exacerbation.

Correlation between BAL bacterial burden and propionate in BAL at **(A)** baseline and during an **(B)** AE-IPF. BAL bacterial burden was calculated by qPCR and expressed as log10 16S rRNA gene copies/ml of BAL. Correlations between cytokines and propionate were determined by Spearman's rank test.

Furthermore, a positive, albeit weak, correlation was found between levels of IL-6 and propionate at baseline (R = 0.36, $P < 0.05$) and during an AE-IPF (R = 0.32, $P < 0.05$) **(Figure 6.19AD)**. During an AE-IPF, a weak negative correlation was also found between levels of TNF-α and propionate in the BAL supernatant (R = -0.42, $P < 0.01$) **(Figure 6.19F).** No correlations were found between BAL cytokine levels and the other SCFAs.

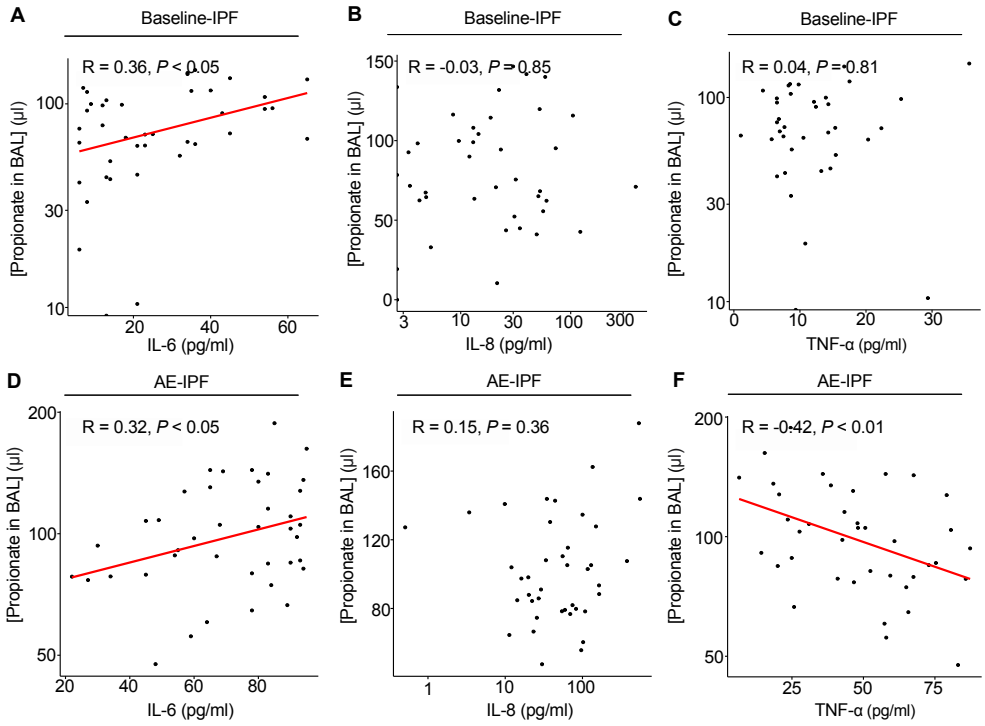


Figure 6.19 Correlations between cytokines detected in bronchoalveolar lavage (BAL) and propionate detected in BAL of subjects with idiopathic pulmonary fibrosis (IPF) at baseline and during an acute exacerbation (AE).

Correlation between BAL propionate concentration and concentrations of IL-6, IL-8, and TNF-α in BAL at **(A-C)** baseline and during an **(D-F)** AE-IPF. Correlations between cytokines and propionate were determined by Spearman's rank test.

To test for the relationship between BAL propionate levels at baseline and the time taken to experience an AE-IPF, subjects with IPF were dichotomised into two groups based on their BAL concentration of propionate at baseline **(Figure 6.20)**. Subjects in the group with higher levels of propionate were found to experience significantly faster rates of disease progression compared to patients in the lower group ($P < 0.05$), with a mean timeframe from diagnosis to an AE-IPF being 10.8 (± 10.2) *versus* 18 (± 21.6) months in the high and low propionate groups, respectively. No relationship was found between BAL concentrations of either α-hydroxybutyrate or acetate at baseline and time to an AE-IPF. Furthermore, higher levels of propionate during an AE-IPF were associated with an increased risk of death within a 90-day period following an AE-IPF. Fifteen out of twenty patients (75%) were alive after 90 days following a diagnosis of AE-IPF in the low propionate group compared to nine out of twenty (55%) in the high propionate group. In addition, there was a significant weak correlation between levels of propionate and the relative abundance of *Streptococcus* during AE-IPF (Spearman's *rho* = 0.4; $P < 0.05$).

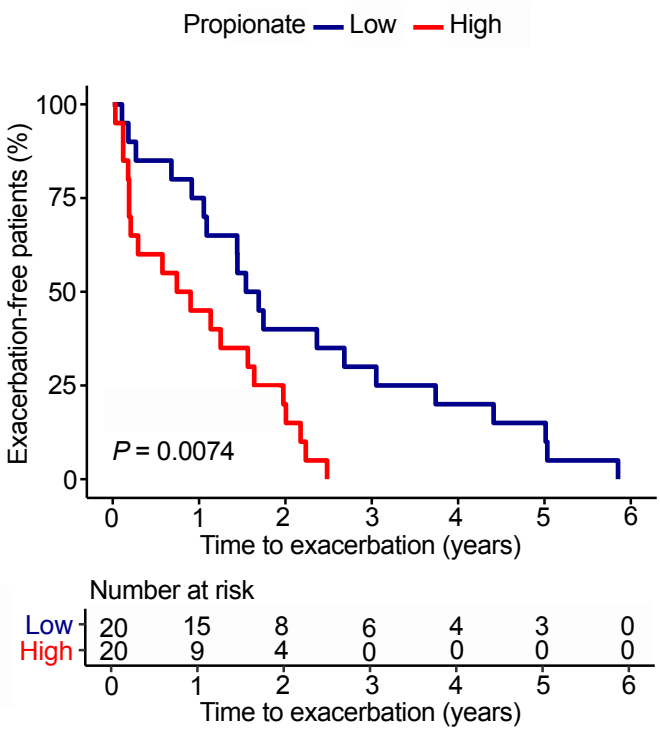


Figure 6.20 Kaplan-Meier curve comparing the time from diagnosis to an acute exacerbation of IPF (AE-IPF) according to the concentration of propionate in the bronchoalveolar lavage (BAL) supernatant at baseline.

IPF subjects at baseline were stratified into two groups (low and high) based on their concentration of propionate in the BAL supernatant. Log-rank *P* test value reported.

6.5 Discussion

The objective of this Chapter was to examine the potential changes which occur in the bacterial burden, respiratory microbiome and cytokines during an AE-IPF. This Chapter also aimed to dissect for the first time the changes which occur in the microbial metabolites, particularly the SCFAs, at baseline and during an AE-IPF. Using paired baseline and AE-IPF BAL supernatant samples from the same individual, it was demonstrated that, compared to baseline, the bacterial burden was consistently and significantly increased during an AE-IPF. Furthermore, compared to baseline, changes in microbial community composition, SCFAs and the pro-inflammatory mediators IL-6, IL-8 and TNF-α were also observed during an AE-IPF.

The first aim of this study was to quantify the bacterial burden in paired samples of BAL supernatants of patients with IPF at baseline and during an AE. In a smaller cohort of IPF subjects, it has previously been shown that bacterial load is significantly increased in the BAL supernatant of subjects experiencing an AE-IPF compared to baseline [144]. This finding was validated in this Chapter in a larger cohort of paired IPF subjects. Furthermore, the increased bacterial load has been consistently shown to predict the rate of functional decline and risk of death in subjects with IPF [115, 149]. It was, therefore, speculated that higher bacterial loads at baseline might be associated with a shorter time taken for patients to experience an AE-IPF from the time of IPF diagnosis. However, the bacterial burden at baseline did not predict the rate of disease progression, suggesting that increased bacterial burden alone may not fully account for these events. In addition, dichotomising the subjects with IPF based on whether they had received antibiotics or not did not alter the BAL bacterial burden, suggesting that the overuse and misuse of antibiotics could be a significant driver in the emergence of antibiotic resistance.

An increasing number of studies have also highlighted changes in the pulmonary microbiota composition in subjects with progressive IPF compared to subjects with stable disease. For example, Han and colleagues demonstrated that the presence of specific operational taxonomic units (OTUs) within the *Streptococcus* and *Staphylococcus* genera, above a certain relative abundance threshold, significantly associated with reduced progression-free survival in patients with IPF [114]. Compared to individuals with stable disease, subjects experiencing an AE-IPF have also been shown to have higher levels of potentially pathogenic bacteria, including *Campylobacter* and *Stenotrophomonas* and reduced levels of *Veillonella* [144]. In this study, compared to baseline, the BAL supernatant of subjects with an AE-IPF was characterised by an increase in *Streptococcus* and *Campylobacter* and a reduction in the relative abundance of *Porphyromonas*. This was further confirmed by exploring

the differentially abundant bacteria between the two groups using LEfSe, which revealed that the BAL supernatant of subjects at baseline was enriched with bacteria in the *Bacteroidetes* phylum, whilst the BAL supernatant of subjects experiencing an AE-IPF was enriched with bacteria belonging to the *Proteobacteria* phylum. An increased abundance of *Campylobacter*, a known gastrointestinal pathogen, reinforces the concept that gastrointestinal reflux and aspirations may drive acute exacerbations in some individuals. Overall, the increased relative abundance of these bacteria in the BAL supernatants of subjects experiencing an AE-IPF may suggest that bacterial infections may potentially be involved in a proportion of AE-IPF cases. Furthermore, these microbes may be acting synergistically together with increased BAL bacterial burden to promote the initiation and development of an AE-IPF. Despite differences in bacterial loads and changes in bacterial community composition, studies have also highlighted that the respiratory microbiome of subjects with IPF is characterised by a reduced bacterial diversity compared to healthy controls [115, 447]. Furthermore, impaired diversity, including lower community diversity, has been shown to be associated with disease activities of IPF [320]. However, it has previously been shown that this difference does not appear to exist when comparing the microbiota diversity of patients at baseline compared to those experiencing an AE-IPF [144]. In this study, compared to baseline-IPF, no changes in microbial richness or evenness were observed during an acute exacerbation. Moreover, there were no significant differences in the microbial community structure, marked by the lack of clustering between the two groups. Comparison of the taxonomic composition between samples of paired individuals may partially explain this. Although clear changes were observed during an AE in the BAL supernatant of both individuals, significant differences in terms of the proportions of bacterial phyla in each individual were noted. This may indicate that if an exacerbation is a result of changes in the microbiome, the bacteria involved may be specific to the individual, which might explain why some types of community-acquired pneumonias in healthy individuals are not always caused by the same bacteria [448]. Collectively, these findings suggest that the development of an AE-IPF may not rely entirely on an acceleration of the underlying fibroproliferative processes occurring in IPF, but rather necessitates a combination of factors, including an occult infectious trigger, to promote its initiation.

The last aim of this study was to investigate whether shifts in microbiota composition may be accompanied by changes in BAL microbial metabolites and distinct pro-inflammatory and pro-fibrotic cytokine profiles. Compared to baseline, the BAL supernatant of patients experiencing an AE-IPF showed significantly increased levels of the SCFAs acetate, propionate and α-hydroxybutyrate. Furthermore, these subjects revealed higher levels of the pro-inflammatory

cytokines IL-6, IL-8 and TNF-α. In contrast, TGF-β and IL-1β, two fundamental profibrogenic cytokines, were undetectable in most BAL samples both at baseline and during an AE-IPF. Specifically, the choice of cytokines was based on previous studies regarding their role in exacerbations of various pulmonary disorders, where an upregulation of pro-inflammatory cytokines is often observed, including cystic fibrosis, chronic obstructive pulmonary disease, asthma and IPF [149, 175, 449, 450]. Furthermore, compared to baseline, patients experiencing an AE-IPF had a significantly increased percentage of neutrophils in their BAL, suggesting that infectious agents including bacterial pathogens, may be involved in the development of AE-IPF in a subset of patients. Given that recruitment of neutrophils to the site of injury is an important hallmark of the acute phase response, and that IL-8 is known to mediate the recruitment of neutrophils to the site of injury [446], the positive correlation observed during an AE-IPF between the percentage of neutrophils in BAL, and IL-8 concentration strengthens the concept that AE-IPF initiation may have occurred in response to an occult external trigger. In addition, the correlation observed between BAL levels of propionate, but not acetate or α-hydroxybutyrate, at baseline and during an AE-IPF, indicates that bacteria present in the lungs, rather than at distal sites, may have been involved in the production of the majority of propionate detected in the BAL supernatant. Nonetheless, it is challenging to assess whether a proportion of these SCFAs may have originated in the gastrointestinal tract, given that the microbiota residing there is responsible for the production of large amounts of SCFAs [352]. Furthermore, given the presence of *Campylobacter*, a bacterium normally confined to the gastrointestinal tract, in the BAL supernatant of these patients, may call into question the origin of these metabolites. However, given the correlation between propionate and the relative abundance of *Streptococcus* during an AE-IPF, may suggest that *Streptococcus* is a potential source of propionate production in the airways during these events. This finding reflects the observations made in **Chapter 4**; propionate was elevated in the BAL of subjects with IPF compared to healthy controls. Increased BAL propionate concentrations in individuals with IPF could indeed reflect the metabolism of lung resident bacteria such as *Streptococcus*, which degrades pyruvate into end products of metabolism, including lactate and propionate. Interestingly, propionate positively correlated with BAL bacterial burden and an increased concentration of propionate in the BAL supernatant at baseline was also predictive of a shorter time taken to experience an AE-IPF. Patients with higher levels of propionate in the BAL supernatant at the time of AE-IPF diagnosis were at an increased risk of death within a 90-day period. Collectively, these findings may suggest that propionate may be a precipitating factor in IPF initiation and progression.

This work has a number of limitations. Firstly, undertaking bronchoscopy in subjects with IPF and particularly in those individuals experiencing an AE-IPF is not recommended, as an AE after BAL is increasingly recognised as a complication for patients with IPF [451]. These sampling difficulties have, therefore hindered research in the field and may explain the limited understanding of the nature of these events. Additionally, given the insidious nature of these events, which may develop over a period of days or weeks, the BAL samples may not have been collected at the same stage of disease for each individual. Consequently, this may have influenced the comparisons made in this study. Although antibiotic intake was not found to significantly alter the BAL bacterial burden, it is plausible that exposure of the individuals to antibiotics may have affected the accuracy of the results obtained. Another limitation to take into consideration is that it is challenging to understand the origin of the SCFAs quantified in the BAL supernatant. A proportion of the SCFAs measured in the BAL supernatant may have originated at distal sites such as the gut. Given that intestinal disturbances have been associated with several diseases, including chronic lung diseases, future work should not only aim to characterise the metabolites present in the BAL supernatant but also aim to phenotype the metabolites present in the gut and systemically, in order to better understand the cross-talk which occurs between these two sites and its role in the development and progression of IPF and AE-IPF. Furthermore, analysis of the microbiome and metabolites in the BAL does not consider changes that are occurring over time and may not fully provide an accurate representation of the regional differences in bacterial burden, microbial and metabolite composition in fibrotic lung disease. Therefore, future studies should try and assess the changes which are occurring longitudinally and, in parallel to BAL analyses, should aim to measure the molecular factors within the lung tissue itself. These factors might be undetectable in the BAL but be locally present at the site of fibrosis and thus allow a more accurate reflection of the molecular and cellular composition of the lung tissue itself, where the fibrotic processes are taking place. The subjects recruited in this study were from Japan. Therefore, compared to the literature, the variability in genetic, environmental and lifestyle factors may have contributed to the differences observed in the results. Furthermore, although subjects were excluded if they had a history of self-reported upper or lower respiratory tract infection, it is impossible to entirely exclude the possibility that occult viral infections may have been partially responsible for the changes observed and/or may have directly contributed to the epithelial damage leading to these events. Lastly, it is important to note that this was an observational study. Future work should assess the role of the SCFAs, particularly propionate, on epithelial/fibroblast cell cultures and co-cultures. These *in vitro* studies may offer

potential mechanistic insights into the pathogenesis of IPF and may provide a better understanding of the effects of these microbial metabolites on the airway epithelium and their role in AE-IPF.

6.6 Conclusions

Collectively, the observations from this Chapter suggest that alterations in the pulmonary microbiome may be implicated in a proportion of AE-IPF cases. Specifically, the increased bacterial load and relative abundances of potentially pathogenic bacteria reinforces the hypothesis that specific bacteria may play a causative role in these events. The findings of this Chapter also shed light on a potential link between pulmonary disturbances, microbial metabolites, inflammation and AE-IPF. Overall, the findings suggest that AE-IPF may represent an acute response to an occult event, occurring in the context of the underlying fibroproliferative processes driving IPF. These events may be accelerated by various factors, including the SCFA propionate. Elucidating the mechanistic role of the SCFAs in the context of IPF will be pivotal to better understand the host-microbe interactions that underpin AE-IPF pathogenesis.

7 **Chapter 7**: Discussion

7.1 General discussion

IPF is a progressive and fatal condition characterised by progressive and irreversible scarring within the pulmonary interstitium associated with architectural destruction, impaired gas exchange and a median survival of 3 years from diagnosis [321]. Despite several environmental factors having been shown to confer an increased risk for developing IPF, the exact triggers which initiate the fibrotic process remain unknown [75]. Multiple exposures including cigarette smoke, environmental pollutants, viruses and bacteria have been implicated in IPF pathogenesis. There is clear evidence that individual genetic background plays a role in IPF pathogenesis, with common genetic variants, including *MUC5B* involved in host defence, responsible for approximately one-third of the risk of developing IPF [40]. In addition, there is growing evidence suggesting that subclinical or occult viral infections may play a role in exacerbating the disease [452]. Although the lungs have conventionally been considered sterile, the recent revolution in culture-independent technologies has prompted a reconsideration of IPF pathogenesis [110, 308]. Recent, albeit associative, evidence suggests that the respiratory microbiome may be linked to the development and progression of IPF [114, 115, 144, 277, 309]. Indeed, bacteria have been hypothesised to act as a continued driver of epithelial injury. Specifically, a higher bacterial burden has been consistently associated with an increased risk of mortality and differences in bacterial composition and diversity have also been correlated with disease outcomes [114, 115, 308].

These observational and association studies support the hypothesis that host-environment interaction is key in the pathogenesis of IPF, yet they do not provide any functional or mechanistic insight. At present, no conclusion can be drawn on whether the changes observed in the microbiome are a cause or a consequence of the disease. Is the perturbed microbial environment driving the disease, or is it the result of an altered lung environment? It is conceivable that, regardless of the initiating trigger of fibrosis, a self-amplifying cycle is established whereby the altered growth conditions result in a disordered microbiome; this, in turn, provokes further inflammation via several mechanisms, including microbial metabolite signalling to epithelial cells, thus propagating the lung damage and creating a further selective pressure [110]. With these considerations in mind, the overarching aim of this thesis was to phenotype the airway microbiome and metabolome in patients with IPF and healthy controls and to investigate the influence of microbial metabolites, specifically the SCFAs acetate, propionate and butyrate on airway epithelial cells and fibroblasts.

Two hundred and one patients with IPF and 40 healthy controls were recruited and followed up over a four-year period. Comparable to a typical IPF cohort, the subjects diagnosed with IPF were predominantly men, had a mean age of 74 years and had moderately severe disease at enrolment [453]. The clinical course of IPF is both variable and unpredictable, with significant heterogeneity among individual patients. Two predictors of poor prognosis in IPF, which were confirmed in **Chapter 3** are a baseline DL_{CO} of less than 40% predicted or a decline in FVC >10% predicted over 12 months. Furthermore, disease progression was defined as a composite of either a relative decline in FVC of greater than 10% of death before 12 months. Validating what has been previously observed in the literature, individuals whose disease had progressed at 12 months demonstrated a significantly higher BAL bacterial burden when compared to subjects with stable disease.

The epithelial surfaces of the human respiratory tract, previously thought to be sterile, have been shown using culture-independent techniques to harbour a dynamic community of microbes [137]. More recently, the advent of sequencing techniques has revolutionised the field of metagenomics, revealing a vast microbial diversity which had been previously missed by cultivation-based methods. Molecular sequencing of the variable regions of the bacterial 16S rRNA gene is the most commonly used approach to the study of bacterial communities. Clusters of bacteria sharing similar gene sequences are classified into OTUs, which can be compared to 16S rRNA gene reference databases [109]. Bioinformatics analysis of the data produced on platforms such as Illumina MiSeq allows quantifying the total and relative abundance of the bacterial groups [110, 111]. The application of these molecular techniques to phenotype the microbial communities of the respiratory tract in health and disease has allowed to uncover associations which indicate that host-microbiome interactions may play a key role in the aetiology and progression of lung disease [113, 144, 454, 455]. In **Chapter 3**, concordant with what has been consistently shown in the literature, clear differences in the respiratory microbiome of patients with IPF compared to healthy controls were demonstrated in a larger cohort of patients. Specifically, patients with IPF have a higher bacterial load and significant differences in the composition and diversity compared to healthy individuals **(Chapter 3, sections 3.4.4-3.4.6)**. The BAL of healthy controls and subjects with IPF was dominated by *Prevotella*, *Streptococcus*, *Veillonella*, *Neisseria* and *Haemophilus*. These bacterial communities have been consistently found in the airways of healthy subjects, but also of asthmatics and subjects with COPD [113, 117, 252, 325]. Specific differences were observed when comparing the microbiome in healthy subjects to that of IPF. Specifically, *Streptococcus*, *Veillonella*, and *Staphylococcus* were increased in IPF. These differences together with the finding that there is a significantly increased bacterial load in the BAL of IPF patients compared to health may support the

hypothesis that these bacterial communities may be acting synergistically to promote airway inflammation as well as drive repetitive alveolar injury. These findings may, therefore provide an opportunity for the development of targeted therapeutic interventions, which do not target the entire microbiome.

Differential abundance analysis confirmed these findings. Multiple taxonomic differences are evident when comparing the BAL of healthy controls and subjects with IPF. The BAL of healthy controls was enriched with bacteria belonging to the *Bacteroidetes* and *Actinobacteria* phyla, and the IPF group was enriched with bacteria belonging to the *Firmicutes* and *Proteobacteria* phyla. These findings are consistent with the literature which shows that during respiratory disease, there is an increased prevalence of *Firmicutes* and *Proteobacteria* and a reduction in non-pathogenic commensals, such as *Bacteroidetes* when compared to healthy controls. This observation suggests that a dysbiotic state of the lung microbiome could be directly associated with the pathophysiology of IPF [320].

A loss of microbial diversity has been previously associated with disease activities in IPF [320]. In **Chapter 3 (section 3.4.5.3)** diversity analyses were used to characterise the bacterial diversity within each sample. Alpha-diversity was assessed using the Shannon and Chao1 indexes, which estimate the evenness and richness, respectively, of the microbiome in each sample. Compared to the controls, subjects diagnosed with IPF were found to have a decreased evenness. Additionally, a signal toward a lower bacterial richness was also observed, but this was not statistically significant. These findings may indicate that a shift in the respiratory microbiome toward more densely populated and less diverse bacterial communities of the lower airways in IPF may provide persistent stimuli for repetitive alveolar injury [320]. Beta-diversity analyses revealed significant differences in whole community composition between the two groups. These results validate the finding that the airway microbiota in IPF is different from that of healthy controls. Nevertheless, causal conclusions cannot be drawn about this altered respiratory microbiome in IPF ahead of functional and longitudinal studies which are urgently needed to better understand pathogenic host-microbial interactions. In addition, an increased bacterial load at the time of diagnosis identified patients with more rapidly progressive IPF and an increased risk of mortality. This research has shown, for the first time, that bacterial burden is an independent predictor of survival even when incorporating radiographic features into the survival models and following adjustment for confounding factors **(Chapter 3, section 3.4.4.6)**. The lack of association observed with measures of disease severity, and

fibrotic destruction of the lung suggests that changes in bacterial burden are disease-relevant and unlikely simply secondary to the architectural distortion of the lungs [277].

Within the IPF cohort, the bacterial burden was found to associate independently with the *rs35705950* polymorphism in the promoter region of the mucin *MUC5B* genotype, which provides a link between host immunity and the respiratory microbiome **(Chapter 3, section 3.4.4.5)**. It has been shown that the minor allele predisposes to a milder form of IPF associated with a longer survival [54]. The results conducted in **Chapter 3** are supportive of this finding since IPF subjects possessing the minor allele were found to have a lower bacterial burden and a longer survival compared to patients possessing the GG and GT alleles. A plausible explanation is that the bacterial communities residing in the lower airways are acting as a persistent stimulus for repetitive alveolar injury and that a higher bacterial load could be accelerating this process. Regional differences in the respiratory microbiome could also explain the temporal and spatial heterogeneity observed in usual interstitial pneumonia, which is the histological correlate of IPF [117].

Having characterised the respiratory microbiome in health and disease, the second aim of this thesis was to phenotype the airway metabolome using untargeted ^{1}H NMR and using targeted GC-MS to specifically look at the SCFAs [159, 354]. Specifically, it was hypothesised that there would be differences in the metabolic profiles of healthy controls and subjects with IPF and that certain metabolites upregulated in IPF may be contributing to disease progression. In **Chapter 4**, it was demonstrated that there are clear differences in the metabolic profile of subjects with IPF compared to healthy controls **(Chapter 4, section 4.4.2)**. Specific metabolic pathways, such as the glycolytic pathway, which may contribute to disease pathogenesis, are altered in IPF patients compared to healthy controls. Indeed, significantly increased levels of glucose and an increased presence of lactate in IPF, have been previously reported to be increased in the BAL and lung tissue of subjects with IPF in the presence of inflammation [373, 375]. This finding is concordant with the increase in lactate and glucose which was observed in the BAL of IPF subjects compared to healthy controls and may suggest altered glycolysis in IPF lungs and supports the hypothesis that the products of glycolysis are being shuttled towards lactate production. Other metabolites, including acetate, butyrate, propionate, acetone, glycine and alanine were found to be upregulated in IPF. Furthermore, specific associations between the respiratory microbiome and metabolome were investigated in IPF **(Chapter 4, section 4.4.4)**. Firstly, Spearman's correlation coefficients were computed to assess potential relationships between the two datasets. Weak associations were observed between the respiratory microbiome and the metabolites in the BAL of healthy controls.

In contrast, in the BAL of IPF subjects, acetate was found to positively correlate with *Prevotella* and propionate with *Veillonella*. To further explore these associations, sPLS-DA was performed using MixOmics. In IPF, the SCFAs acetate, propionate and lactate were found to positively correlate with several bacterial genera including *Prevotella*, *Streptococcus*, and *Veillonella*. This is of interest because studies in the gut have shown that acetate is produced by most of the enteric bacteria including *Prevotella* species via the Wood-Ljungdhal and acetyl-CoA pathways [389, 390], whilst propionate is produced by *Veillonella* via the succinate, acrylate and propanediol pathways [153]. In contrast, distinct families within the *Firmicutes* are capable of producing butyrate from two molecules of acetyl-CoA [391]. Therefore, elevated BAL SCFA concentrations in subjects with IPF could reflect the metabolism of lung resident bacteria such as *Streptococcus* and *Veillonella*, which degrades pyruvate into end products of metabolism, including lactate and propionate [392].

Parallels and lessons can be drawn from studies of the gastrointestinal tract microbiome and its interactions with the host immune system. There, a mutually beneficial partnership between the microflora and the host has been established over thousands of years of co-evolution. The bacteria benefit from the stable habitat, which is rich in energy sources from the food we ingest. On the other hand, commensals assist in the metabolism of dietary constituents such as cellulose, produce essential cofactors such as vitamin K_1 and SCFAs which constitute a fundamental source of energy for enterocytes [163]. SCFAs improve gut health through several local effects, including the maintenance of intestinal barrier integrity, mucus production, and protection against inflammation [398].

Similar to the gut, the airway microbiota in healthy lungs is dominated by SCFA-producing bacteria: *Bacteroidetes*, *Firmicutes* and *Proteobacteria* [153]. These are metabolically active bacteria which produce SCFAs via the bacterial fermentation of non-digestible carbohydrates by the gut microbiota [351]. It is now well documented that the intestinal microbiota influences the microbial composition within the lung either by direct seeding of bacteria into the airways or by the distribution of SCFAs. These SCFAs, not only act as a fuel for intestinal epithelial cells but also exert distinct physiological effects, acting as a link between the microbiota and the immune system and playing a key role in maintaining mucosal immunity [456]. Nonetheless, very few studies have characterised the metabolites in the BAL of subjects with IPF, and no studies have yet addressed the impact of the SCFAs on the development of fibrosis. Interestingly, acetate, propionate, lactate and butyrate were all found to be increased in the BAL supernatant of IPF subjects compared to healthy controls **(Chapter 4, section 4.4.6)**. Particularly, the SCFAs acetate and propionate have been previously

linked to mucosal epithelial interaction in the gut and might, therefore be playing a similar role in the lungs [398]. A 'leaky gut' in humans refers to increased gut permeability, disturbed microbial balance and impaired mucosal immunity, and plays a role in the initiation of inflammatory diseases and autoimmunity [163]. Interestingly, subjects with IPF are often diagnosed with inflammatory bowel disease (IBD), which is an intestinal disease characterised by an impaired gut-blood barrier [95]. This might provide evidence for the increased gut-to-blood penetration of SCFAs. In order to validate the increased translocation of SCFAs into the circulation, their presence should be also quantified in serum samples from patients with IPF that present concomitant IBD [95].

The results of **Chapter 4** highlight the importance of metabolomics in exploratory studies. Searching for differences in the metabolic profiles may ultimately serve to identify novel pathways and potential biomarkers to aid in the prediction of the development of IPF. In the literature, the metabolic products from bacteria have been shown to stimulate host metabolism and could provide an insight into the pathogenic processes occurring during fibrosis. Chronic lung diseases, including IPF, are complex, progressive disorders with increasing incidence and mortality [457]. Managing these conditions can be improved using personalised approaches for more targeted treatments. In the literature, several studies highlight the role of metabolic dysfunction in the development of chronic lung diseases. Metabolomics is increasingly being used as a technique to aid in the diagnosis and prognosis of such disorders. The first diagnoses of these diseases were based on the integration of clinical, radiological and histopathological features. For example, histological assessment is an invasive technique which can be impractical for some patients as it typically requires surgical biopsy. Therefore, incorporating techniques in standard clinical practice which eliminate the need for invasive diagnostic methods such as bronchoscopy and surgical lung biopsy, may benefit patients with chronic lung diseases, including IPF. Metabolomics thus holds considerable potential to elucidate biomarkers which may aid in the diagnosis and treatment with better sensitivity, specificity and precision. ^{1}H NMR and MS are the two main analytical techniques utilised for metabolomics. Both methods have been extensively used to acquire information on a variety of metabolites in a single measurement. ^{1}H NMR spectroscopy has been shown to be a non-destructive technique useful for the detection of metabolic changes. On the other hand, GC-MS is efficacious for the detection of smaller-molecular weight metabolites that cannot be detected by ^{1}H NMR spectroscopy. The metabolomics data obtained from these two analytical platforms were combined to fully characterise the metabolites present in the BAL of the subjects with IPF as well as healthy controls. However, there are a number of limitations that must be taken into consideration. For example, the metabolic profile was only characterised in BAL supernatant samples. In the future,

concomitant analysis of faecal and serum profiles could provide information regarding the cross-talk that occurs between the gut and the lung and further support these findings. Furthermore, it is important to take into consideration the fact that BAL supernatant does not necessarily reflect the molecular and cellular composition of the lung tissue itself, where the fibrotic processes are taking place. Therefore, future studies should aim to characterise the metabolites within the lung tissue itself, where factors that are undetectable in BAL supernatant, may be locally present at the site of fibrosis. This may allow for a more accurate overview of the overall molecular changes occurring during these events. Furthermore, these techniques only provide a snapshot of the metabolome at a particular moment in time and is therefore challenging to draw conclusions regarding whether the altered metabolic profile is a cause or a consequence of the disease. Although these findings support a greater understanding of the metabolic basis of pathogenesis of IPF formation yet, they do not provide any functional insight. Therefore, to explore the influence of the microbial metabolites identified in aim 2 on the respiratory epithelium, PBECs and HLF were exposed to increasing concentrations of the SCFAs acetate, propionate and butyrate. Bronchial epithelial cells are the first anatomical barrier to external environmental stimuli and are responsible for the initiation of airway remodelling through the production of proinflammatory mediators, ECM, growth factors and matrix metalloproteinases [408]. In submerged cultures of PBECs, supplementation of healthy and IPF cells with 10 mM and 20 mM of the respective SCFAs was found to impact cell viability, whilst the lower doses (1 mM and 5 mM) were found to be non-toxic to the cells **(Chapter 5, section 5.4.2.1)**. IL-6 and TNF-α were quantified in the cell culture supernatants pre- and post-exposure to the SCFAs; a dose-dependent secretion of these cytokines was observed in the cell supernatant from the IPF PBECs, suggesting that the SCFAs at concentrations measured in IPF BAL affect the inflammatory responses of airway epithelial cells **(Chapter 5, section 5.4.2.2)**. Whilst no differences in IPF PBEC cell proliferation was observed **(Chapter 5, section 5.4.2.4)**, treatment of these cells with 5 mM acetate, propionate or butyrate led to a decrease in wound closure at 48 hours **(Chapter 5, section 5.4.2.6)**. This might indicate that the SCFAs are inhibiting wound closure, although the mechanism behind this is unclear. It is also important to note that the molecular events by which PBECs proliferate and migrate for lung wound healing are yet to be elucidated. To provide a differentiated airway epithelium more representative of the normal *in vivo* pseudostratified epithelium, healthy and IPF cells were cultured at ALI and exposed to the SCFAs **(Chapter 5, section 5.4.3)**. Non-toxic doses of the SCFAs were shown to have no effects on the healthy PBECs, whilst they were found to decrease the TEER of the IPF PBECs, impact their morphology and lead to an increased production of IL-6 in the apical cell supernatant at 24 hours. Given the crucial role of fibroblasts in mediating

fibrosis, the effects of the SCFAs were also assessed on primary cultures of HLF isolated from subjects with IPF. Similar to the PBECs, HLF treated with the lower doses (1 mM and 5 mM) of the respective SCFAs had no effects on cell viability **(Chapter 5, section 5.4.4.1)**. The effects of the SCFAs on HLF proliferation were assessed in HLF from three patients with IPF over a time-course of 72 hours **(Chapter 5, section 5.4.4.3)**. Addition of acetate was found to have no effect on HLF proliferation compared to the cells treated with media alone. In contrast, following exposure of the HLF to 5 mM of propionate and butyrate, an increase in proliferation was observed compared to the untreated control. This finding may suggest that acetate may be acting through a different receptor compared to propionate and butyrate. Additional experiments looking at the expression of GPCRs on these cells and their expression following treatment with SCFAs are urgently needed.

While some patients with IPF experience a steady decline in lung function, others suffer sudden acute respiratory worsening, following a period of relative stability. These acute exacerbations are an important clinical feature of IPF disease progression and have been defined as clinically significant respiratory deteriorations of unidentifiable cause characterised by new, widespread alveolar abnormalities. Given that the respiratory microbiome was characterised in individuals with IPF at baseline, the final aim of this thesis was to establish if and how the microbiome changed during an AE-IPF [25]. This was achieved this by establishing a collaboration with the Tosei General Hospital (Japan), where subjects were prospectively recruited between January 2008 and July 2015. Very few studies have explored the variation in the respiratory microbiome at baseline and during an AE-IPF. In a smaller cohort of IPF subjects, Molyneaux and colleagues showed that bacterial load is significantly increased in the BAL supernatant of subjects experiencing an AE-IPF compared to baseline and that there are specific changes in the respiratory microbiome [144]. This finding was validated in **Chapter 6 (section 6.4.3)** in a larger cohort of paired IPF subjects. An increased bacterial load has been consistently shown to predict the rate of functional decline and risk of death in subjects with IPF [115, 149]. It was hypothesised that higher bacterial loads at baseline may be associated with a shorter time taken for patients to experience an AE-IPF from the time of IPF diagnosis. However, the bacterial burden at baseline did not predict the rate of disease progression, suggesting that increased bacterial burden alone may not fully account for these events **(Chapter 6, section 6.4.3.2)**. *Proteobacteria*, *Firmicutes*, *Bacteroidetes*, and *Actinobacteria* dominated the BAL at baseline and during an AE. At the genus level, compared to baseline-IPF, *Streptococcus* and *Campylobacter* were found to be increased in the AE-IPF samples **(Chapter 6, section 6.4.4.1)**. This was further confirmed by exploring the differentially abundant bacteria between the two groups using LEfSe, which revealed that the BAL supernatant of subjects at baseline was enriched with

bacteria in the *Bacteroidetes* phylum, whilst the BAL supernatant of subjects experiencing an AE-IPF was enriched with bacteria belonging to the *Proteobacteria* phylum **(Chapter 6, section 6.4.4.2)**. An increased abundance of *Campylobacter*, a known gastrointestinal pathogen, reinforces the concept that gastrointestinal reflux and aspirations may drive acute exacerbations in some individuals. However, a limitation when using 16S rRNA sequencing to characterise microbial communities is the potential of incorrect taxonomic annotation. Therefore, to confirm the presence of *Campylobacter* in the BAL of patients with IPF, a BLAST-based method should be used to validate the assignments made by less precise sequence classification databases. Overall, the increased relative abundance of these bacteria in the BAL supernatants of subjects experiencing an AE-IPF suggests that bacterial infections may play a causative role in a proportion of AE-IPF cases. Furthermore, these microbes may be acting synergistically together with increased BAL bacterial burden to promote the initiation and development of an AE-IPF. Overall no significant differences in alpha-diversity and beta-diversity were observed when comparing IPF patients at baseline and during an AE **(Chapter 6, section 6.4.4.3)**. This may be partially explained by the fact that heterogeneous changes were observed across the cohort, as highlighted by the taxonomic analysis of paired samples **(Chapter 6, section 6.4.4.4)**. This may indicate that if an exacerbation is a result of changes in the microbiome, the bacteria involved may be specific to the individual. Overall, these findings suggest that the development of an AE-IPF may not rely entirely on an acceleration of the underlying fibroproliferative processes occurring in IPF, but rather necessitates a combination of factors, including an occult infectious trigger, to promote its initiation. Further analysis of bacterial burden and taxonomic composition in individuals with non-IPF acute lung injury as well as longitudinal sampling during an AE-IPF may be key to gaining a better understanding of these events. Lastly, for the first time, the SCFAs were quantified at baseline and during an AE-IPF **(Chapter 6, section 6.4.6)**. Compared to AE-IPF, the BAL supernatant of patients at baseline had significantly lower levels of the SCFAs acetate, propionate and α-hydroxybutyrate. Furthermore, these subjects revealed lower levels of the pro-inflammatory cytokines IL-6, IL-8 and TNF-α **(Chapter 6, section 6.4.5)**. The correlation between propionate and the relative abundance of *Streptococcus* during an AE-IPF, may indicate that *Streptococcus* is a potential source of propionate production in the airways during these events. Interestingly, propionate positively correlated with BAL bacterial burden and an increased concentration of propionate in the BAL supernatant at baseline was also predictive of a shorter time taken to experience an AE-IPF **(Chapter 6, section 6.4.6).**

7.2 Future work

The findings from this thesis provide a strong rationale for trials of antimicrobial therapy in IPF. Nevertheless, these findings require further validation in separate cohorts of IPF subjects from multiple centres. Although prospective trials of broad-spectrum antibiotics in IPF are currently underway, the findings described herein indicate that a more personalised approach may be necessary in order to target those patients with a higher BAL bacterial burden or where specific changes in the microbiome and metabolome are occurring. A significant concern when using 16S rRNA sequencing to characterise microbial communities is the potential of contamination during sample acquisition. Contamination presents a critical problem when evaluating samples with low microbial biomass, such as the lungs [337]. The low amount of starting material may be swamped by the amplification of reagent-derived contaminants and result in erroneous representations of the microbial community [337, 338]. In this work, to minimise the risk of contamination, negative bronchoscopy controls and reagent controls were included in the sequencing run. PCA analysis showed that BAL samples clustered separately from the negative controls, demonstrating that the samples were not influenced by contamination. This was further confirmed by comparing the bacterial burden amongst groups, assessed by using the total number of reads in each sample. Future research could make use of protected catheter brushes which could help minimise the risk of potential contamination. Another significant limitation is that 16S rRNA gene sequencing provides a snapshot of the microbial community at a particular moment in time, and is therefore challenging to draw conclusions regarding whether the altered microbial community is a cause or a consequence of the disease [335]. Thus, functional and longitudinal studies examining the host-microbial interactions are necessary. Regional differences in the microbiome throughout the lung have also been previously demonstrated [117]. Therefore, assessing the regional variation in the microbiome across the respiratory tract in IPF could allow to correlate with the extent of fibrosis. This could help uncover the role of specific bacterial microorganisms or bacterial burden in driving the fibrotic process.

To build upon the cell culture work which has been conducted in **Chapter 5**, the next step would be to determine whether the GPCRs are expressed on the PBECs and whether their expression is altered following the addition of the SCFAs. These SCFA receptors should be expressed on bronchial epithelial cells, upregulated in disease and could be driving pulmonary fibrosis via the activation of a number of mechanistic downstream pathways implicated in fibrinogenesis. Studies addressing the functional aspects of the microbiota will be critical to determine whether the lung microbiota has a causal role in driving the development and progression of IPF. Interactions between pulmonary

epithelial cells and fibroblasts appear to be central to the pathogenesis of IPF. Future work should also investigate how specific IPF-associated microbes/metabolites, identified in aims 1 and 2, impact these interactions leading to the activation of pro-fibrotic pathways. Specifically, the effects of the SCFAs in a co-culture system of PBECs and HLF will be critical in uncovering whether these microbial metabolites are driving the fibrotic process in IPF.

What still remains unclear is whether these shifts in the microbiome are a cause or a consequence of fibrosis; is the perturbed microbial environment causing the disease or are these microbial changes the result of an altered lung environment? Furthermore, is the microbial diversity, or the abundance of a particular species, responsible for the development and progression of IPF? It is clear that more information is needed to better understand the nature of the dialogues occurring between the respiratory microbiota and epithelial cells in the lungs to further elucidate the function and mechanisms through which the microbiota shapes the development of fibrosis in IPF. Further studies are urgently needed to validate these findings, address the crucial question of causality and potentially shape future interventional trials.

8 Bibliography

1. Antoniou KM, Margaritopoulos GA, Tomassetti S, Bonella F, Costabel U, Poletti V. Interstitial lung disease. *Eur Respir Rev* 2014;23:40-54.

2. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, Ryerson CJ, Ryu JH, Selman M, Wells AU, Behr J, Bouros D, Brown KK, Colby TV, Collard HR, Cordeiro CR, Cottin V, Crestani B, Drent M, Dudden RF, Egan J, Flaherty K, Hogaboam C, Inoue Y, Johkoh T, Kim DS, Kitaichi M, Loyd J, Martinez FJ, Myers J, Protzko S, Raghu G, Richeldi L, Sverzellati N, Swigris J, Valeyre D. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2013;188:733-748.

3. Zibrak JD, Price D. Interstitial lung disease: raising the index of suspicion in primary care. *NPJ Prim Care Respir Med* 2014;24:14054.

4. Azadeh N, Limper AH, Carmona EM, Ryu JH. The Role of Infection in Interstitial Lung Diseases: A Review. *Chest* 2017;152:842-852.

5. Bradley B, Branley HM, Egan JJ, Greaves MS, Hansell DM, Harrison NK, Hirani N, Hubbard R, Lake F, Millar AB, Wallace WA, Wells AU, Whyte MK, Wilsher ML. Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax* 2008;63 Suppl 5:v1-58.

6. Hutchinson J, Fogarty A, Hubbard R, McKeever T. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J* 2015;46:795-806.

7. Navaratnam V, Fleming KM, West J, Smith CJ, Jenkins RG, Fogarty A, Hubbard RB. The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax* 2011;66:462-467.

8. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med* 2000;161:646-664.

9. Maher TM, Wells AU, Laurent GJ. Idiopathic pulmonary fibrosis: multiple causes and multiple mechanisms? *Eur Respir J* 2007;30:835-839.

10. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. *Lancet* 2017;389:1941-1952.

11. Gibb C, Smith PJ, Miller R. Clubbing. *Br J Hosp Med (Lond)* 2013;74:C170-172.

12. Hodnett PA, Naidich DP. Fibrosing interstitial lung disease. A practical high-resolution computed tomography-based approach to diagnosis and management and a review of the literature. *Am J Respir Crit Care Med* 2013;188:141-149.

13. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J, Cottin V, Danoff SK, Morell F, Flaherty KR, Wells A, Martinez FJ, Azuma A, Bice TJ, Bouros D, Brown KK, Collard HR, Duggal A, Galvin L, Inoue Y, Jenkins RG, Johkoh T, Kazerooni EA, Kitaichi M, Knight SL, Mansour G, Nicholson AG, Pipavath SNJ, Buendia-Roldan I, Selman M, Travis WD, Walsh S, Wilson KC. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med* 2018;198:e44-e68.

14. Torres P, Rabahi MF, Moreira MAC, Meirelles GSP, Marchiori E. Usual interstitial pneumonia: typical, possible, and "inconsistent" patterns. *J Bras Pneumol* 2017;43:393-398.

15. King TE, Jr., Pardo A, Selman M. Idiopathic pulmonary fibrosis. *Lancet* 2011;378:1949-1961.

16. Ley B, Collard HR, King TE, Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011;183:431-440.

17. Collard HR, King TE, Jr., Bartelson BB, Vourlekis JS, Schwarz MI, Brown KK. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2003;168:538-542.

18. Hallstrand TS, Boitano LJ, Johnson WC, Spada CA, Hayes JG, Raghu G. The timed walk test as a measure of severity and survival in idiopathic pulmonary fibrosis. *Eur Respir J* 2005;25:96-103.

19. Sumikawa H, Johkoh T, Colby TV, Ichikado K, Suga M, Taniguchi H, Kondoh Y, Ogura T, Arakawa H, Fujimoto K, Inoue A, Mihara N, Honda O, Tomiyama N, Nakamura H, Muller NL. Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *Am J Respir Crit Care Med* 2008;177:433-439.

20. King TE, Jr., Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: scoring system and survival model. *Am J Respir Crit Care Med* 2001;164:1171-1181.

21. Mejia M, Carrillo G, Rojas-Serrano J, Estrada A, Suarez T, Alonso D, Barrientos E, Gaxiola M, Navarro C, Selman M. Idiopathic pulmonary fibrosis and emphysema: decreased survival associated with severe pulmonary arterial hypertension. *Chest* 2009;136:10-15.

22. Nicholson AG, Fulford LG, Colby TV, du Bois RM, Hansell DM, Wells AU. The relationship between individual histologic features and disease progression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2002;166:173-177.

23. Collard HR, Ward AJ, Lanes S, Cortney Hayflinger D, Rosenberg DM, Hunsche E. Burden of illness in idiopathic pulmonary fibrosis. *J Med Econ* 2012;15:829-835.

24. Raghu G, Amatto VC, Behr J, Stowasser S. Comorbidities in idiopathic pulmonary fibrosis patients: a systematic literature review. *Eur Respir J* 2015;46:1113-1130.

25. Collard HR, Ryerson CJ, Corte TJ, Jenkins G, Kondoh Y, Lederer DJ, Lee JS, Maher TM, Wells AU, Antoniou KM, Behr J, Brown KK, Cottin V, Flaherty KR, Fukuoka J, Hansell DM, Johkoh T, Kaminski N, Kim DS, Kolb M, Lynch DA, Myers JL, Raghu G, Richeldi L, Taniguchi H, Martinez FJ. Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report. *Am J Respir Crit Care Med* 2016;194:265-275.

26. Sakamoto S, Homma S, Mun M, Fujii T, Kurosaki A, Yoshimura K. Acute exacerbation of idiopathic interstitial pneumonia following lung surgery in 3 of 68 consecutive patients: a retrospective study. *Intern Med* 2011;50:77-85.

27. Moua T, Westerly BD, Dulohery MM, Daniels CE, Ryu JH, Lim KG. Patients With Fibrotic Interstitial Lung Disease Hospitalized for Acute Respiratory Worsening: A Large Cohort Analysis. *Chest* 2016;149:1205-1214.

28. Ryerson CJ, Collard HR. Acute exacerbations complicating interstitial lung disease. *Curr Opin Pulm Med* 2014;20:436-441.

29. Collard HR, Moore BB, Flaherty KR, Brown KK, Kaner RJ, King TE, Jr., Lasky JA, Loyd JE, Noth I, Olman MA, Raghu G, Roman J, Ryu JH, Zisman DA, Hunninghake GW, Colby TV, Egan JJ, Hansell DM, Johkoh T, Kaminski N, Kim DS, Kondoh Y, Lynch DA, Muller-Quernheim J, Myers JL, Nicholson AG, Selman M, Toews GB, Wells AU, Martinez FJ. Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2007;176:636-643.

30. Ahluwalia N, Shea BS, Tager AM. New therapeutic targets in idiopathic pulmonary fibrosis. Aiming to rein in runaway wound-healing responses. *Am J Respir Crit Care Med* 2014;190:867-878.

31. Daccord C, Maher TM. Recent advances in understanding idiopathic pulmonary fibrosis. *F1000Res* 2016;5.

32. Selman M, King TE, Pardo A. Idiopathic pulmonary fibrosis: prevailing and evolving hypotheses about its pathogenesis and implications for therapy. *Ann Intern Med* 2001;134:136-151.

33. Garcia-Alvarez J, Ramirez R, Sampieri CL, Nuttall RK, Edwards DR, Selman M, Pardo A. Membrane type-matrix metalloproteinases in idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2006;23:13-21.

34. Kim KK, Kugler MC, Wolters PJ, Robillard L, Galvez MG, Brumwell AN, Sheppard D, Chapman HA. Alveolar epithelial cell mesenchymal transition develops in vivo during pulmonary fibrosis and is regulated by the extracellular matrix. *Proc Natl Acad Sci U S A* 2006;103:13180-13185.

35. Strieter RM, Mehrad B. New mechanisms of pulmonary fibrosis. *Chest* 2009;136:1364-1370.

36. Chilosi M, Poletti V, Zamò A, Lestani M, Montagna L, Piccoli P, Pedron S, Bertaso M, Scarpa A, Murer B, Cancellieri A, Maestro R, Semenzato G, Doglioni C. Aberrant Wnt/beta-catenin pathway activation in idiopathic pulmonary fibrosis. *Am J Pathol* 2003;162:1495-1502.

37. Plataki M, Koutsopoulos AV, Darivianaki K, Delides G, Siafakas NM, Bouros D. Expression of apoptotic and antiapoptotic markers in epithelial cells in idiopathic pulmonary fibrosis. *Chest* 2005;127:266-274.

38. Kuwano K, Hagimoto N, Kawasaki M, Yatomi T, Nakamura N, Nagata S, Suda T, Kunitake R, Maeyama T, Miyazaki H, Hara N. Essential roles of the Fas-Fas ligand pathway in the development of pulmonary fibrosis. *J Clin Invest* 1999;104:13-19.

39. Chandler DB. Possible mechanisms of bleomycin-induced fibrosis. *Clin Chest Med* 1990;11:21-30.

40. Mathai SK, Newton CA, Schwartz DA, Garcia CK. Pulmonary fibrosis in the era of stratified medicine. *Thorax* 2016;71:1154-1160.

41. Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, Fingerlin TE, Zhang W, Gudmundsson G, Groshong SD, Evans CM, Garantziotis S, Adler KB, Dickey BF, du Bois RM, Yang IV, Herron A, Kervitsky D, Talbert JL, Markin C, Park J, Crews AL, Slifer SH, Auerbach S, Roy MG, Lin J, Hennessy CE, Schwarz MI, Schwartz DA. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503-1512.

42. Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scuirba J, Richards TJ, Juan-Guardela BM, Vij R, Han MK, Martinez FJ, Kossen K, Seiwert SD, Christie JD, Nicolae D, Kaminski N, Garcia JGN. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013;1:309-317.

43. Hodgson U, Pulkkinen V, Dixon M, Peyrard-Janvid M, Rehn M, Lahermo P, Ollikainen V, Salmenkivi K, Kinnula V, Kere J, Tukiainen P, Laitinen T. ELMOD2 is a candidate gene for familial idiopathic pulmonary fibrosis. *Am J Hum Genet* 2006;79:149-154.

44. Lawson WE, Loyd JE. The genetic approach in pulmonary fibrosis: can it provide clues to this complex disease? *Proc Am Thorac Soc* 2006;3:345-349.

45. Garcia-Sancho C, Buendia-Roldan I, Fernandez-Plata MR, Navarro C, Perez-Padilla R, Vargas MH, Loyd JE, Selman M. Familial pulmonary fibrosis is the strongest risk factor for idiopathic pulmonary fibrosis. *Respir Med* 2011;105:1902-1907.

46. Nogee LM, Dunbar AE, 3rd, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001;344:573-579.

47. Wang Y, Kuan PJ, Xing C, Cronkhite JT, Torres F, Rosenblatt RL, DiMaio JM, Kinch LN, Grishin NV, Garcia CK. Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer. *Am J Hum Genet* 2009;84:52-59.

48. Nathan N, Giraud V, Picard C, Nunes H, Dastot-Le Moal F, Copin B, Galeron L, De Ligniville A, Kuziner N, Reynaud-Gaubert M, Valeyre D, Couderc LJ, Chinet T, Borie R, Crestani B, Simansour M, Nau V, Tissier S, Duquesnoy P, Mansour-Hendili L, Legendre M, Kannengiesser C, Coulomb-L'Hermine A, Gouya L, Amselem S, Clement A. Germline SFTPA1 mutation in familial idiopathic interstitial pneumonia and lung cancer. *Hum Mol Genet* 2016;25:1457-1467.

49. Nayak A, Dodagatta-Marri E, Tsolaki AG, Kishore U. An Insight into the Diverse Roles of Surfactant Proteins, SP-A and SP-D in Innate and Adaptive Immunity. *Front Immunol* 2012;3:131.

50. Pastva AM, Wright JR, Williams KL. Immunomodulatory roles of surfactant proteins A and D: implications in lung disease. *Proc Am Thorac Soc* 2007;4:252-257.

51. Mulugeta S, Nureki S, Beers MF. Lost after translation: insights from pulmonary surfactant for understanding the role of alveolar epithelial dysfunction and cellular quality control in fibrotic lung disease. *Am J Physiol Lung Cell Mol Physiol* 2015;309:L507-525.

52. Maitra M, Wang Y, Gerard RD, Mendelson CR, Garcia CK. Surfactant protein A2 mutations associated with pulmonary fibrosis lead to protein instability and endoplasmic reticulum stress. *J Biol Chem* 2010;285:22103-22113.

53. Yanagihara T, Sato S, Upagupta C, Kolb M. What have we learned from basic science studies on idiopathic pulmonary fibrosis? *Eur Respir Rev* 2019;28.

54. Peljto AL, Zhang Y, Fingerlin TE, Ma SF, Garcia JG, Richards TJ, Silveira LJ, Lindell KO, Steele MP, Loyd JE, Gibson KF, Seibold MA, Brown KK, Talbert JL, Markin C, Kossen K, Seiwert SD, Murphy E, Noth I, Schwarz MI, Kaminski N, Schwartz DA. Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *Jama* 2013;309:2232-2239.

55. Nakano Y, Yang IV, Walts AD, Watson AM, Helling BA, Fletcher AA, Lara AR, Schwarz MI, Evans CM, Schwartz DA. MUC5B Promoter Variant rs35705950 Affects MUC5B Expression in the Distal Airways in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2016;193:464-466.

56. Roy MG, Livraghi-Butrico A, Fletcher AA, McElwee MM, Evans SE, Boerner RM, Alexander SN, Bellinghausen LK, Song AS, Petrova YM, Tuvim MJ, Adachi R, Romo I, Bordt AS, Bowden MG, Sisson JH, Woodruff PG, Thornton DJ, Rousseau K, De la Garza MM, Moghaddam SJ, Karmouty-Quintana H, Blackburn MR, Drouin SM, Davis CW, Terrell KA, Grubb BR, O'Neal WK, Flores SC, Cota-Gomez A, Lozupone CA, Donnelly JM, Watson AM, Hennessy CE, Keith RC, Yang IV, Barthel L, Henson

PM, Janssen WJ, Schwartz DA, Boucher RC, Dickey BF, Evans CM. Muc5b is required for airway defence. *Nature* 2014;505:412-416.

57. Hancock LA, Hennessy CE, Solomon GM, Dobrinskikh E, Estrella A, Hara N, Hill DB, Kissner WJ, Markovetz MR, Grove Villalon DE, Voss ME, Tearney GJ, Carroll KS, Shi Y, Schwarz MI, Thelin WR, Rowe SM, Yang IV, Evans CM, Schwartz DA. Muc5b overexpression causes mucociliary dysfunction and enhances lung fibrosis in mice. *Nat Commun* 2018;9:5363.

58. Calado RT, Young NS. Telomere diseases. *N Engl J Med* 2009;361:2353-2365.

59. Armanios MY, Chen JJ, Cogan JD, Alder JK, Ingersoll RG, Markin C, Lawson WE, Xie M, Vulto I, Phillips JA, 3rd, Lansdorp PM, Greider CW, Loyd JE. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007;356:1317-1326.

60. Cronkhite JT, Xing C, Raghu G, Chin KM, Torres F, Rosenblatt RL, Garcia CK. Telomere shortening in familial and sporadic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008;178:729-737.

61. Silhan LL, Shah PD, Chambers DC, Snyder LD, Riise GC, Wagner CL, Hellstrom-Lindberg E, Orens JB, Mewton JF, Danoff SK, Arcasoy MO, Armanios M. Lung transplantation in telomerase mutation carriers with pulmonary fibrosis. *Eur Respir J* 2014;44:178-187.

62. Townsley DM, Dumitriu B, Young NS. Danazol Treatment for Telomere Diseases. *N Engl J Med* 2016;375:1095-1096.

63. Tsubouchi K, Araya J, Kuwano K. PINK1-PARK2-mediated mitophagy in COPD and IPF pathogeneses. *Inflamm Regen* 2018;38:18.

64. Ryu C, Sun H, Gulati M, Herazo-Maya JD, Chen Y, Osafo-Addo A, Brandsdorfer C, Winkler J, Blaul C, Faunce J, Pan H, Woolard T, Tzouvelekis A, Antin-Ozerkis DE, Puchalski JT, Slade M, Gonzalez AL, Bogenhagen DF, Kirillov V, Feghali-Bostwick C, Gibson K, Lindell K, Herzog RI, Dela Cruz CS, Mehal W, Kaminski N, Herzog EL, Trujillo G. Extracellular Mitochondrial DNA Is Generated by Fibroblasts and Predicts Death in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2017;196:1571-1581.

65. Bueno M, Lai YC, Romero Y, Brands J, St Croix CM, Kamga C, Corey C, Herazo-Maya JD, Sembrat J, Lee JS, Duncan SR, Rojas M, Shiva S, Chu CT, Mora AL. PINK1 deficiency impairs mitochondrial homeostasis and promotes lung fibrosis. *J Clin Invest* 2015;125:521-538.

66. Yu G, Tzouvelekis A, Wang R, Herazo-Maya JD, Ibarra GH, Srivastava A, de Castro JPW, DeIuliis G, Ahangari F, Woolard T, Aurelien N, Arrojo EDR, Gan Y, Graham M, Liu X, Homer RJ, Scanlan TS, Mannam P, Lee PJ, Herzog EL, Bianco AC, Kaminski N. Thyroid hormone inhibits lung fibrosis in mice by improving epithelial mitochondrial function. *Nat Med* 2018;24:39-49.

67. Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, Chen W, Bis JC, Fitzpatrick AL, Smith E, Johnson AD, Gardner JP, Srinivasan SR, Schork N, Rotter JI, Herbig U, Psaty BM, Sastrasinh M, Murray

SS, Vasan RS, Province MA, Glazer NL, Lu X, Cao X, Kronmal R, Mangino M, Soranzo N, Spector TD, Berenson GS, Aviv A. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. *Proc Natl Acad Sci U S A* 2010;107:9293-9298.

68. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 2007;8:253-262.

69. Yang IV, Pedersen BS, Rabinovich E, Hennessy CE, Davidson EJ, Murphy E, Guardela BJ, Tedrow JR, Zhang Y, Singh MK, Correll M, Schwarz MI, Geraci M, Sciurba FC, Quackenbush J, Spira A, Kaminski N, Schwartz DA. Relationship of DNA methylation and gene expression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014;190:1263-1272.

70. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, Crossman DK, Bray M, Zhang K, Thannickal VJ, Hagood JS. Altered DNA methylation profile in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186:525-535.

71. Flynt AS, Lai EC. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat Rev Genet* 2008;9:831-842.

72. Pandit KV, Milosevic J, Kaminski N. MicroRNAs in idiopathic pulmonary fibrosis. *Transl Res* 2011;157:191-199.

73. Oak SR, Murray L, Herath A, Sleeman M, Anderson I, Joshi AD, Coelho AL, Flaherty KR, Toews GB, Knight D, Martinez FJ, Hogaboam CM. A micro RNA processing defect in rapidly progressing idiopathic pulmonary fibrosis. *PLoS One* 2011;6:e21253.

74. Ding Q, Luckhardt T, Hecker L, Zhou Y, Liu G, Antony VB, deAndrade J, Thannickal VJ. New insights into the pathogenesis and treatment of idiopathic pulmonary fibrosis. *Drugs* 2011;71:981-1001.

75. Taskar VS, Coultas DB. Is idiopathic pulmonary fibrosis an environmental disease? *Proc Am Thorac Soc* 2006;3:293-298.

76. Garcia-Sancho Figueroa MC, Carrillo G, Perez-Padilla R, Fernandez-Plata MR, Buendia-Roldan I, Vargas MH, Selman M. Risk factors for idiopathic pulmonary fibrosis in a Mexican population. A case-control study. *Respir Med* 2010;104:305-309.

77. Antoniou KM, Hansell DM, Rubens MB, Marten K, Desai SR, Siafakas NM, Nicholson AG, du Bois RM, Wells AU. Idiopathic pulmonary fibrosis: outcome in relation to smoking status. *Am J Respir Crit Care Med* 2008;177:190-194.

78. Steele MP, Speer MC, Loyd JE, Brown KK, Herron A, Slifer SH, Burch LH, Wahidi MM, Phillips JA, 3rd, Sporn TA, McAdams HP, Schwarz MI, Schwartz DA. Clinical and pathologic features of familial interstitial pneumonia. *Am J Respir Crit Care Med* 2005;172:1146-1152.

79. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R. [The Montreal definition and classification of gastroesophageal reflux disease: a global, evidence-based consensus paper]. *Z Gastroenterol* 2007;45:1125-1140.

80. D'Ovidio F, Mura M, Tsang M, Waddell TK, Hutcheon MA, Singer LG, Hadjiliadis D, Chaparro C, Gutierrez C, Pierre A, Darling G, Liu M, Keshavjee S. Bile acid aspiration and the development of bronchiolitis obliterans after lung transplantation. *J Thorac Cardiovasc Surg* 2005;129:1144-1152.

81. McGrath EE, Millar AB. Hot off the breath: triple therapy for idiopathic pulmonary fibrosis--hear the PANTHER roar. *Thorax* 2012;67:97-98.

82. Raghu G, Anstrom KJ, King TE, Jr., Lasky JA, Martinez FJ. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012;366:1968-1977.

83. Shulgina L, Cahn AP, Chilvers ER, Parfrey H, Clark AB, Wilson EC, Twentyman OP, Davison AG, Curtin JJ, Crawford MB, Wilson AM. Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: a randomised controlled trial. *Thorax* 2013;68:155-162.

84. Zhang Y, Noth I, Garcia JG, Kaminski N. A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. *N Engl J Med* 2011;364:1576-1577.

85. Yang IV, Luna LG, Cotter J, Talbert J, Leach SM, Kidd R, Turner J, Kummer N, Kervitsky D, Brown KK, Boon K, Schwarz MI, Schwartz DA, Steele MP. The peripheral blood transcriptome identifies the presence and extent of disease in idiopathic pulmonary fibrosis. *PLoS One* 2012;7:e37708.

86. Conway SP, Simmonds EJ, Littlewood JM. Acute severe deterioration in cystic fibrosis associated with influenza A virus infection. *Thorax* 1992;47:112-114.

87. Wootton SC, Kim DS, Kondoh Y, Chen E, Lee JS, Song JW, Huh JW, Taniguchi H, Chiu C, Boushey H, Lancaster LH, Wolters PJ, DeRisi J, Ganem D, Collard HR. Viral infection in acute exacerbation of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011;183:1698-1702.

88. Moore BB, Moore TA. Viruses in Idiopathic Pulmonary Fibrosis. Etiology and Exacerbation. *Ann Am Thorac Soc* 2015;12 Suppl 2:S186-192.

89. Vergnon JM, Vincent M, de The G, Mornex JF, Weynants P, Brune J. Cryptogenic fibrosing alveolitis and Epstein-Barr virus: an association? *Lancet* 1984;2:768-771.

90. Manika K, Alexiou-Daniel S, Papakosta D, Papa A, Kontakiotis T, Patakas D, Antoniadis A. Epstein-Barr virus DNA in bronchoalveolar lavage fluid from patients with idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2007;24:134-140.

91. Tang W, Pavlish OA, Spiegelman VS, Parkhitko AA, Fuchs SY. Interaction of Epstein-Barr virus latent membrane protein 1 with SCFHOS/beta-TrCP E3 ubiquitin ligase regulates extent of NF-kappaB activation. *J Biol Chem* 2003;278:48942-48949.

92. Stoolman JS, Vannella KM, Coomes SM, Wilke CA, Sisson TH, Toews GB, Moore BB. Latent infection by gammaherpesvirus stimulates profibrotic mediator release from multiple cell types. *Am J Physiol Lung Cell Mol Physiol* 2011;300:L274-285.

93. Lawson WE, Crossno PF, Polosukhin VV, Roldan J, Cheng DS, Lane KB, Blackwell TR, Xu C, Markin C, Ware LB, Miller GG, Loyd JE, Blackwell TS. Endoplasmic reticulum stress in alveolar epithelial cells is prominent in IPF: association with altered surfactant protein processing and herpesvirus infection. *Am J Physiol Lung Cell Mol Physiol* 2008;294:L1119-1126.

94. Kropski JA, Pritchett JM, Zoz DF, Crossno PF, Markin C, Garnett ET, Degryse AL, Mitchell DB, Polosukhin VV, Rickman OB, Choi L, Cheng DS, McConaha ME, Jones BR, Gleaves LA, McMahon FB, Worrell JA, Solus JF, Ware LB, Lee JW, Massion PP, Zaynagetdinov R, White ES, Kurtis JD, Johnson JE, Groshong SD, Lancaster LH, Young LR, Steele MP, Phillips Iii JA, Cogan JD, Loyd JE, Lawson WE, Blackwell TS. Extensive phenotyping of individuals at risk for familial interstitial pneumonia reveals clues to the pathogenesis of interstitial lung disease. *Am J Respir Crit Care Med* 2015;191:417-426.

95. Kim J, Chun J, Lee C, Han K, Choi S, Lee J, Soh H, Choi J, Park S, Kang E, Lee H, Kim J. INcreased risk of idiopathic pulmonary fibrosis in inflammatory bowel disease: A nationwide study. *J Gastroenterol Hepatol* 2020; 35(2): 249-255.

96. Bando M, Ohno S, Oshikawa K, Takahashi M, Okamoto H, Sugiyama Y. Infection of TT virus in patients with idiopathic pulmonary fibrosis. *Respir Med* 2001;95:935-942.

97. Ushiki A, Yamazaki Y, Hama M, Yasuo M, Hanaoka M, Kubo K. Viral infections in patients with an acute exacerbation of idiopathic interstitial pneumonia. *Respir Investig* 2014;52:65-70.

98. Santos GC, Parra ER, Stegun FW, Cirqueira CS, Capelozzi VL. Immunohistochemical detection of virus through its nuclear cytopathic effect in idiopathic interstitial pneumonia other than acute exacerbation. *Braz J Med Biol Res* 2013;46:985-992.

99. Weng D, Chen XQ, Qiu H, Zhang Y, Li QH, Zhao MM, Wu Q, Chen T, Hu Y, Wang LS, Wei YR, Du YK, Chen SS, Zhou Y, Zhang F, Shen L, Su YL, Kolb M, Li HP. The Role of Infection in Acute Exacerbation of Idiopathic Pulmonary Fibrosis. *Mediators Inflamm* 2019;2019:5160694.

100. Calabrese F, Kipar A, Lunardi F, Balestro E, Perissinotto E, Rossi E, Nannini N, Marulli G, Stewart JP, Rea F. Herpes virus infection is associated with vascular remodeling and pulmonary hypertension in idiopathic pulmonary fibrosis. *PLoS One* 2013;8:e55715.

101. McMillan TR, Moore BB, Weinberg JB, Vannella KM, Fields WB, Christensen PJ, van Dyk LF, Toews GB. Exacerbation of established pulmonary fibrosis in a murine model by gammaherpesvirus. *Am J Respir Crit Care Med* 2008;177:771-780.

102. Umeda Y, Morikawa M, Anzai M, Sumida Y, Kadowaki M, Ameshima S, Ishizaki T. Acute exacerbation of idiopathic pulmonary fibrosis after pandemic influenza A (H1N1) vaccination. *Intern Med* 2010;49:2333-2336.

103. Hunninghake GM. A new hope for idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2142-2143.

104. Woodcock HV, Molyneaux PL, Maher TM. Reducing lung function decline in patients with idiopathic pulmonary fibrosis: potential of nintedanib. *Drug Des Devel Ther* 2013;7:503-510.

105. Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, Cottin V, Flaherty KR, Hansell DM, Inoue Y, Kim DS, Kolb M, Nicholson AG, Noble PW, Selman M, Taniguchi H, Brun M, Le Maulf F, Girard M, Stowasser S, Schlenker-Herceg R, Disse B, Collard HR. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071-2082.

106. Richeldi L, Costabel U, Selman M, Kim DS, Hansell DM, Nicholson AG, Brown KK, Flaherty KR, Noble PW, Raghu G, Brun M, Gupta A, Juhel N, Kluglich M, du Bois RM. Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. *N Engl J Med* 2011;365:1079-1087.

107. King TE, Jr., Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster L, Lederer DJ, Nathan SD, Pereira CA, Sahn SA, Sussman R, Swigris JJ, Noble PW. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2083-2092.

108. Collard HR, Richeldi L, Kim DS, Taniguchi H, Tschoepe I, Luisetti M, Roman J, Tino G, Schlenker-Herceg R, Hallmann C, du Bois RM. Acute exacerbations in the INPULSIS trials of nintedanib in idiopathic pulmonary fibrosis. *Eur Respir J* 2017;49.

109. Clarridge JE, 3rd. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev* 2004;17:840-862, table of contents.

110. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The Microbiome and the Respiratory Tract. *Annu Rev Physiol* 2016;78:481-504.

111. O'Dwyer DN, Dickson RP, Moore BB. The Lung Microbiome, Immunity, and the Pathogenesis of Chronic Lung Disease. *J Immunol* 2016;196:4839-4847.

112. Dickson RP, Erb-Downward JR, Prescott HC, Martinez FJ, Curtis JL, Lama VN, Huffnagle GB. Analysis of culture-dependent versus culture-independent techniques for identification of bacteria in clinically obtained bronchoalveolar lavage fluid. *J Clin Microbiol* 2014;52:3605-3613.

113. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L, Pachter L, Moffatt MF, Cookson WO. Disordered microbial communities in asthmatic airways. *PLoS One* 2010;5:e8578.

114. Han MK, Zhou Y, Murray S, Tayob N, Noth I, Lama VN, Moore BB, White ES, Flaherty KR, Huffnagle GB, Martinez FJ. Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014;2:548-556.

115. Molyneaux PL, Cox MJ, Willis-Owen SA, Mallia P, Russell KE, Russell AM, Murphy E, Johnston SL, Schwartz DA, Wells AU, Cookson WO, Maher TM, Moffatt MF. The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014;190:906-913.

116. Huang YJ, LiPuma JJ. The Microbiome in Cystic Fibrosis. *Clin Chest Med* 2016;37:59-67.

117. Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, Cooper J, Sin DD, Mohn WW, Hogg JC. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;185:1073-1080.

118. Gollwitzer ES, Marsland BJ. Impact of Early-Life Exposures on Immune Maturation and Susceptibility to Disease. *Trends Immunol* 2015;36:684-696.

119. Pattaroni C, Watzenboeck ML, Schneidegger S, Kieser S, Wong NC, Bernasconi E, Pernot J, Mercier L, Knapp S, Nicod LP, Marsland CP, Roth-Kleiner M, Marsland BJ. Early-Life Formation of the Microbial and Immunological Environment of the Human Airways. *Cell Host Microbe* 2018;24:857-865.e854.

120. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. *Sci Transl Med* 2014;6:237ra265.

121. Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci Rep* 2016;6:23129.

122. Bosch A, Levin E, van Houten MA, Hasrat R, Kalkman G, Biesbroek G, de Steenhuijsen Piters WAA, de Groot PCM, Pernet P, Keijser BJF, Sanders EAM, Bogaert D. Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* 2016;9:336-345.

123. Chu DM, Ma J, Prince AL, Antony KM, Seferovic MD, Aagaard KM. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med* 2017;23:314-326.

124. Bosch A, de Steenhuijsen Piters WAA, van Houten MA, Chu M, Biesbroek G, Kool J, Pernet P, de Groot PCM, Eijkemans MJC, Keijser BJF, Sanders EAM, Bogaert D. Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences. A Prospective Cohort Study. *Am J Respir Crit Care Med* 2017;196:1582-1590.

125. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* 2010;107:11971-11975.

126. Suárez-Arrabal MC, Mella C, Lopez SM, Brown NV, Hall MW, Hammond S, Shiels W, Groner J, Marcon M, Ramilo O, Mejias A. Nasopharyngeal bacterial burden and antibiotics: Influence on inflammatory markers and disease severity in infants with respiratory syncytial virus bronchiolitis. *J Infect* 2015;71:458-469.

127. Gollwitzer ES, Saglani S, Trompette A, Yadava K, Sherburn R, McCoy KD, Nicod LP, Lloyd CM, Marsland BJ. Lung microbiota promotes tolerance to allergens in neonates via PD-L1. *Nat Med* 2014;20:642-647.

128. Herbst T, Sichelstiel A, Schär C, Yadava K, Bürki K, Cahenzli J, McCoy K, Marsland BJ, Harris NL. Dysregulation of allergic airway inflammation in the absence of microbial colonization. *Am J Respir Crit Care Med* 2011;184:198-205.

129. Dickson RP, Huffnagle GB. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. *PLoS Pathog* 2015;11:e1004923.

130. Venkataraman A, Bassis CM, Beck JM, Young VB, Curtis JL, Huffnagle GB, Schmidt TM. Application of a neutral community model to assess structuring of the human lung microbiome. *mBio* 2015;6.

131. Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. *Expert Rev Respir Med* 2013;7:245-257.

132. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, Curtis JL. Bacterial Topography of the Healthy Human Lower Respiratory Tract. *mBio* 2017;8.

133. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio* 2015;6:e00037.

134. Huffnagle GB, Dickson RP, Lukacs NW. The respiratory tract microbiome and lung inflammation: a two-way street. *Mucosal Immunol* 2017;10:299-306.

135. Mathieu E, Escribano-Vazquez U, Descamps D, Cherbuy C, Langella P, Riffault S, Remot A, Thomas M. Paradigms of Lung Microbiota Functions in Health and Disease, Particularly, in Asthma. *Front Physiol* 2018;9:1168.

136. Costa AN, Costa FMD, Campos SV, Salles RK, Athanazio RA. The pulmonary microbiome: challenges of a new paradigm. *J Bras Pneumol* 2018;44:424-432.

137. Invernizzi R, Lloyd CM, Molyneaux PL. Respiratory microbiome and epithelial interactions shape immunity in the lungs. *Immunology* 2020.

138. Durack J, Lynch SV, Nariya S, Bhakta NR, Beigelman A, Castro M, Dyer AM, Israel E, Kraft M, Martin RJ, Mauger DT, Rosenberg SR, Sharp-King T, White SR, Woodruff PG, Avila PC, Denlinger LC, Holguin F, Lazarus SC, Lugogo N, Moore WC, Peters SP, Que L, Smith LJ, Sorkness CA, Wechsler ME, Wenzel SE, Boushey HA, Huang YJ. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. *J Allergy Clin Immunol* 2017;140:63-75.

139. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, Holt BJ, Hales BJ, Walker ML, Hollams E, Bochkov YA, Grindle K, Johnston SL, Gern JE, Sly PD, Holt PG, Holt KE, Inouye M. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015;17:704-715.

140. Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. *PLoS One* 2012;7:e47305.

141. Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, Wolfgang MC. The adult cystic fibrosis airway microbiota is stable over time and infection type, and highly resilient to antibiotic treatment of exacerbations. *PLoS One* 2012;7:e45001.

142. Feigelman R, Kahlert CR, Baty F, Rassouli F, Kleiner RL, Kohler P, Brutsche MH, von Mering C. Sputum DNA sequencing in cystic fibrosis: non-invasive access to the lung microbiome and to pathogen details. *Microbiome* 2017;5:20.

143. Frayman KB, Armstrong DS, Carzino R, Ferkol TW, Grimwood K, Storch GA, Teo SM, Wylie KM, Ranganathan SC. The lower airway microbiota in early cystic fibrosis lung disease: a longitudinal analysis. *Thorax* 2017;72:1104-1112.

144. Molyneaux PL, Cox MJ, Wells AU, Kim HC, Ji W, Cookson WO, Moffatt MF, Kim DS, Maher TM. Changes in the respiratory microbiome during acute exacerbations of idiopathic pulmonary fibrosis. *Respir Res* 2017;18:29.

145. Invernizzi R, Molyneaux PL. The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2019;28.

146. Invernizzi R, Barnett J, Rawal B, Nair A, Ghai P, Kingston S, Chua F, Wu Z, Wells AU, Renzoni ER, Nicholson AG, Rice A, Lloyd CM, Byrne AJ, Maher TM, Devaraj A, Molyneaux PL. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020.

147. Richter AG, Stockley RA, Harper L, Thickett DR. Pulmonary infection in Wegener granulomatosis and idiopathic pulmonary fibrosis. *Thorax* 2009;64:692-697.

148. Knippenberg S, Ueberberg B, Maus R, Bohling J, Ding N, Tort Tarres M, Hoymann HG, Jonigk D, Izykowski N, Paton JC, Ogunniyi AD, Lindig S, Bauer M, Welte T, Seeger W, Guenther A, Sisson TH, Gauldie J, Kolb M, Maus UA. Streptococcus pneumoniae triggers progression of pulmonary fibrosis through pneumolysin. *Thorax* 2015;70:636-646.

149. O'Dwyer DN, Ashley SL, Gurczynski SJ, Xia M, Wilke C, Falkowski NR, Norman KC, Arnold KB, Huffnagle GB, Salisbury ML, Han MK, Flaherty KR, White ES, Martinez FJ, Erb-Downward JR, Murray S, Moore BB, Dickson RP. Lung Microbiota Contribute to Pulmonary Inflammation and Disease Progression in Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2019;199:1127-1138.

150. Metsala J, Lundqvist A, Virta LJ, Kaila M, Gissler M, Virtanen SM. Mother's and offspring's use of antibiotics and infant allergy to cow's milk. *Epidemiology* 2013;24:303-309.

151. Korpela K, Salonen A, Virta LJ, Kekkonen RA, Forslund K, Bork P, de Vos WM. Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nat Commun* 2016;7:10410.

152. Varney VA, Parnell HM, Salisbury DT, Ratnatheepan S, Tayar RB. A double blind randomised placebo controlled pilot study of oral co-trimoxazole in advanced fibrotic lung disease. *Pulm Pharmacol Ther* 2008;21:178-187.

153. Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, Flint HJ, Louis P. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *Isme j* 2014;8:1323-1335.

154. Qin Y, Wade PA. Crosstalk between the microbiome and epigenome: messages from bugs. *J Biochem* 2018;163:105-112.

155. Mortensen PB, Clausen MR. Short-chain fatty acids in the human colon: relation to gastrointestinal health and disease. *Scand J Gastroenterol Suppl* 1996;216:132-148.

156. Shukla SD, Budden KF, Neal R, Hansbro PM. Microbiome effects on immunity, health and disease in the lung. *Clin Transl Immunology* 2017;6:e133.

157. Correa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MA. Regulation of immune cell function by short-chain fatty acids. *Clin Transl Immunology* 2016;5:e73.

158. Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, Tuohy K. Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* 2018;57:1-24.

159. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 2016;165:1332-1345.

160. Gill PA, van Zelm MC, Muir JG, Gibson PR. Review article: short chain fatty acids as potential therapeutic agents in human gastrointestinal and inflammatory disorders. *Aliment Pharmacol Ther* 2018;48:15-34.

161. Li M, van Esch B, Wagenaar GTM, Garssen J, Folkerts G, Henricks PAJ. Pro- and anti-inflammatory effects of short chain fatty acids on immune and endothelial cells. *Eur J Pharmacol* 2018;831:52-59.

162. Ang Z, Ding JL. GPR41 and GPR43 in Obesity and Inflammation - Protective or Causative? *Front Immunol* 2016;7:28.

163. Marsland BJ, Trompette A, Gollwitzer ES. The Gut-Lung Axis in Respiratory Disease. *Ann Am Thorac Soc* 2015;12 Suppl 2:S150-156.

164. Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L. The role of short-chain fatty acids in health and disease. *Adv Immunol* 2014;121:91-119.

165. Anand S, Mande SS. Diet, Microbiota and Gut-Lung Connection. *Front Microbiol* 2018;9:2147.

166. Bingula R, Filaire M, Radosevic-Robin N, Bey M, Berthon JY, Bernalier-Donadille A, Vasson MP, Filaire E. Desired Turbulence? Gut-Lung Axis, Immunity, and Lung Cancer. *J Oncol* 2017;2017:5035371.

167. Trompette A, Gollwitzer ES, Yadava K, Sichelstiel AK, Sprenger N, Ngom-Bru C, Blanchard C, Junt T, Nicod LP, Harris NL, Marsland BJ. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med* 2014;20:159-166.

168. Vael C, Nelen V, Verhulst SL, Goossens H, Desager KN. Early intestinal Bacteroides fragilis colonisation and development of asthma. *BMC Pulm Med* 2008;8:19.

169. Groves HT, Higham SL, Moffatt MF, Cox MJ, Tregoning JS. Respiratory Viral Infection Alters the Gut Microbiota by Inducing Inappetence. *mBio* 2020;11.

170. Southam DS, Dolovich M, O'Byrne PM, Inman MD. Distribution of intranasal instillations in mice: effects of volume, time, body position, and anesthesia. *Am J Physiol Lung Cell Mol Physiol* 2002;282:L833-839.

171. Tomoda K, Kubo K, Dairiki K, Yamaji T, Yamamoto Y, Nishii Y, Nakamura A, Yoshikawa M, Hamada K, Kimura H. Whey peptide-based enteral diet attenuated elastase-induced emphysema with increase in short chain fatty acids in mice. *BMC Pulm Med* 2015;15:64.

172. Thorburn AN, McKenzie CI, Shen S, Stanley D, Macia L, Mason LJ, Roberts LK, Wong CH, Shim R, Robert R, Chevalier N, Tan JK, Mariño E, Moore RJ, Wong L, McConville MJ, Tull DL, Wood LG, Murphy VE, Mattes J, Gibson PG, Mackay CR. Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. *Nat Commun* 2015;6:7320.

173. Maslowski KM, Mackay CR. Diet, gut microbiota and immune responses. *Nat Immunol* 2011;12:5-9.

174. Sze MA, Tsuruta M, Yang SW, Oh Y, Man SF, Hogg JC, Sin DD. Changes in the bacterial microbiota in gut, blood, and lungs following acute LPS instillation into mice lungs. *PLoS One* 2014;9:e111228.

175. Mirković B, Murray MA, Lavelle GM, Molloy K, Azim AA, Gunaratnam C, Healy F, Slattery D, McNally P, Hatch J, Wolfgang M, Tunney MM, Muhlebach MS, Devery R, Greene CM, McElvaney NG. The Role of Short-Chain Fatty Acids, Produced by Anaerobic Bacteria, in the Cystic Fibrosis Airway. *Am J Respir Crit Care Med* 2015;192:1314-1324.

176. Rutting S, Xenaki D, Malouf M, Horvat JC, Wood LG, Hansbro PM, Oliver BG. Short-chain fatty acids increase TNFα-induced inflammation in primary human lung mesenchymal cells through the activation of p38 MAPK. *Am J Physiol Lung Cell Mol Physiol* 2019;316:L157-l174.

177. Hiemstra PS, McCray PB, Jr., Bals R. The innate immune function of airway epithelial cells in inflammatory lung disease. *Eur Respir J* 2015;45:1150-1162.

178. E.R.W. Morphometry of the Human Lung: Springer Berlin Heidelberg; 1963.

179. Leiva-Juárez MM, Kolls JK, Evans SE. Lung epithelial cells: therapeutically inducible effectors of antimicrobial defense. *Mucosal Immunol* 2018;11:21-34.

180. Sacco O, Silvestri M, Sabatini F, Sale R, Defilippi AC, Rossi GA. Epithelial cells and fibroblasts: structural repair and remodelling in the airways. *Paediatr Respir Rev* 2004;5 Suppl A:S35-40.

181. Whitsett JA, Weaver TE. Alveolar development and disease. *Am J Respir Cell Mol Biol* 2015;53:1-7.

182. Zepp JA, Morrisey EE. Cellular crosstalk in the development and regeneration of the respiratory system. *Nat Rev Mol Cell Biol* 2019;20:551-566.

183. Niessen CM. Tight junctions/adherens junctions: basic structure and function. *J Invest Dermatol* 2007;127:2525-2532.

184. Shen L, Weber CR, Raleigh DR, Yu D, Turner JR. Tight junction pore and leak pathways: a dynamic duo. *Annu Rev Physiol* 2011;73:283-309.

185. Loxham M, Davies DE, Blume C. Epithelial function and dysfunction in asthma. *Clin Exp Allergy* 2014;44:1299-1313.

186. Steed E, Balda MS, Matter K. Dynamics and functions of tight junctions. *Trends Cell Biol* 2010;20:142-149.

187. De Rose V, Molloy K, Gohy S, Pilette C, Greene CM. Airway Epithelium Dysfunction in Cystic Fibrosis and COPD. *Mediators Inflamm* 2018;2018:1309746.

188. Aghapour M, Raee P, Moghaddam SJ, Hiemstra PS, Heijink IH. Airway Epithelial Barrier Dysfunction in Chronic Obstructive Pulmonary Disease: Role of Cigarette Smoke Exposure. *Am J Respir Cell Mol Biol* 2018;58:157-169.

189. Wine JJ, Joo NS. Submucosal glands and airway defense. *Proc Am Thorac Soc* 2004;1:47-53.

190. Mason RJ, Williams MC. Type II alveolar cell. Defender of the alveolus. *Am Rev Respir Dis* 1977;115:81-91.

191. Rokicki W, Rokicki M, Wojtacha J, Dżeljijli A. The role and importance of club cells (Clara cells) in the pathogenesis of some respiratory diseases. *Kardiochir Torakochirurgia Pol* 2016;13:26-30.

192. Fehrenbach H. Alveolar epithelial type II cell: defender of the alveolus revisited. *Respir Res* 2001;2:33-46.

193. Adler KB, Li Y. Airway epithelium and mucus: intracellular signaling pathways for gene expression and secretion. *Am J Respir Cell Mol Biol* 2001;25:397-400.

194. Proud D, Leigh R. Epithelial cells and airway diseases. *Immunol Rev* 2011;242:186-204.

195. Okuda K, Chen G, Subramani DB, Wolf M, Gilmore RC, Kato T, Radicioni G, Kesimer M, Chua M, Dang H, Livraghi-Butrico A, Ehre C, Doerschuk CM, Randell SH, Matsui H, Nagase T, O'Neal WK, Boucher RC. Localization of Secretory Mucins MUC5AC and MUC5B in Normal/Healthy Human Airways. *Am J Respir Crit Care Med* 2019;199:715-727.

196. Fahy JV, Dickey BF. Airway mucus function and dysfunction. *N Engl J Med* 2010;363:2233-2247.

197. Evans CM, Raclawska DS, Ttofali F, Liptzin DR, Fletcher AA, Harper DN, McGing MA, McElwee MM, Williams OW, Sanchez E, Roy MG, Kindrachuk KN, Wynn TA, Eltzschig HK, Blackburn MR, Tuvim MJ, Janssen WJ, Schwartz DA, Dickey BF. The polymeric mucin Muc5ac is required for allergic airway hyperreactivity. *Nat Commun* 2015;6:6281.

198. Regnis JA, Robinson M, Bailey DL, Cook P, Hooper P, Chan HK, Gonda I, Bautovich G, Bye PT. Mucociliary clearance in patients with cystic fibrosis and in normal subjects. *Am J Respir Crit Care Med* 1994;150:66-71.

199. Smaldone GC, Foster WM, O'Riordan TG, Messina MS, Perry RJ, Langenback EG. Regional impairment of mucociliary clearance in chronic obstructive pulmonary disease. *Chest* 1993;103:1390-1396.

200. Flynn JM, Niccum D, Dunitz JM, Hunter RC. Evidence and Role for Bacterial Mucin Degradation in Cystic Fibrosis Airway Disease. *PLoS Pathog* 2016;12:e1005846.

201. Hartsock A, Nelson WJ. Adherens and tight junctions: structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta* 2008;1778:660-669.

202. Lei J, Sun L, Huang S, Zhu C, Li P, He J, Mackey V, Coy DH, He Q. The antimicrobial peptides and their potential clinical applications. *Am J Transl Res* 2019;11:3919-3931.

203. Schutte BC, McCray PB, Jr. [beta]-defensins in lung host defense. *Annu Rev Physiol* 2002;64:709-748.

204. Mallia P, Webber J, Gill SK, Trujillo-Torralbo MB, Calderazzo MA, Finney L, Bakhsoliani E, Farne H, Singanayagam A, Footitt J, Hewitt R, Kebadze T, Aniscenko J, Padmanaban V, Molyneaux PL, Adcock IM, Barnes PJ, Ito K, Elkin SL, Kon OM, Cookson WO, Moffat MF, Johnston SL, Tregoning JS. Role of airway glucose in bacterial infections in patients with chronic obstructive pulmonary disease. *J Allergy Clin Immunol* 2018;142:815-823.e816.

205. Mariathasan S, Monack DM. Inflammasome adaptors and sensors: intracellular regulators of infection and inflammation. *Nat Rev Immunol* 2007;7:31-40.

206. Medzhitov R. Recognition of microorganisms and activation of the immune response. *Nature* 2007;449:819-826.

207. Ye Z, Ting JP. NLR, the nucleotide-binding domain leucine-rich repeat containing gene family. *Curr Opin Immunol* 2008;20:3-9.

208. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell* 2010;140:805-820.

209. Xiang M, Fan J. Pattern recognition receptor-dependent mechanisms of acute lung injury. *Mol Med* 2010;16:69-82.

210. Takeda K, Akira S. Toll-like receptors. *Curr Protoc Immunol* 2015;109:14.12.11-14.12.10.

211. Yamamoto M, Sato S, Hemmi H, Hoshino K, Kaisho T, Sanjo H, Takeuchi O, Sugiyama M, Okabe M, Takeda K, Akira S. Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science* 2003;301:640-643.

212. Kumar H, Kawai T, Akira S. Toll-like receptors and innate immunity. *Biochem Biophys Res Commun* 2009;388:621-625.

213. Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, Matsumoto M, Hoshino K, Wagner H, Takeda K, Akira S. A Toll-like receptor recognizes bacterial DNA. *Nature* 2000;408:740-745.

214. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* 2001;413:732-738.

215. Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM, Aderem A. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 2001;410:1099-1103.

216. de Bouteiller O, Merck E, Hasan UA, Hubac S, Benguigui B, Trinchieri G, Bates EE, Caux C. Recognition of double-stranded RNA by human toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J Biol Chem* 2005;280:38133-38145.

217. Latz E, Schoenemeyer A, Visintin A, Fitzgerald KA, Monks BG, Knetter CF, Lien E, Nilsen NJ, Espevik T, Golenbock DT. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* 2004;5:190-198.

218. Lee J, Chuang TH, Redecke V, She L, Pitha PM, Carson DA, Raz E, Cottam HB. Molecular basis for the immunostimulatory activity of guanine nucleoside analogs: activation of Toll-like receptor 7. *Proc Natl Acad Sci U S A* 2003;100:6646-6651.

219. Matsumoto M, Funami K, Tanabe M, Oshiumi H, Shingai M, Seto Y, Yamamoto A, Seya T. Subcellular localization of Toll-like receptor 3 in human dendritic cells. *J Immunol* 2003;171:3154-3162.

220. Nishiya T, DeFranco AL. Ligand-regulated chimeric receptor approach reveals distinctive subcellular localization and signaling properties of the Toll-like receptors. *J Biol Chem* 2004;279:19008-19017.

221. Oosting M, Cheng SC, Bolscher JM, Vestering-Stenger R, Plantinga TS, Verschueren IC, Arts P, Garritsen A, van Eenennaam H, Sturm P, Kullberg BJ, Hoischen A, Adema GJ, van der Meer JW, Netea MG, Joosten LA. Human TLR10 is an anti-inflammatory pattern-recognition receptor. *Proc Natl Acad Sci U S A* 2014;111:E4478-4484.

222. Sha Q, Truong-Tran AQ, Plitt JR, Beck LA, Schleimer RP. Activation of airway epithelial cells by toll-like receptor agonists. *Am J Respir Cell Mol Biol* 2004;31:358-364.

223. Hammad H, Chieppa M, Perros F, Willart MA, Germain RN, Lambrecht BN. House dust mite allergen induces asthma via Toll-like receptor 4 triggering of airway structural cells. *Nat Med* 2009;15:410-416.

224. Hewson CA, Jardine A, Edwards MR, Laza-Stanca V, Johnston SL. Toll-like receptor 3 is induced by and mediates antiviral activity against rhinovirus infection of human bronchial epithelial cells. *J Virol* 2005;79:12273-12279.

225. Ritchie AI, Farne HA, Singanayagam A, Jackson DJ, Mallia P, Johnston SL. Pathogenesis of Viral Infection in Exacerbations of Airway Disease. *Ann Am Thorac Soc* 2015;12 Suppl 2:S115-132.

226. Ghaemmaghami AM, Gough L, Sewell HF, Shakib F. The proteolytic activity of the major dust mite allergen Der p 1 conditions dendritic cells to produce less interleukin-12: allergen-induced Th2 bias determined at the dendritic cell level. *Clin Exp Allergy* 2002;32:1468-1475.

227. Adam E, Hansen KK, Astudillo Fernandez O, Coulon L, Bex F, Duhant X, Jaumotte E, Hollenberg MD, Jacquet A. The house dust mite allergen Der p 1, unlike Der p 3, stimulates the expression of interleukin-8 in human airway epithelial cells via a proteinase-activated receptor-2-independent mechanism. *J Biol Chem* 2006;281:6910-6923.

228. Kauffman HF, Tamm M, Timmerman JA, Borger P. House dust mite major allergens Der p 1 and Der p 5 activate human airway-derived epithelial cells by protease-dependent and protease-independent mechanisms. *Clin Mol Allergy* 2006;4:5.

229. Chiu LL, Perng DW, Yu CH, Su SN, Chow LP. Mold allergen, pen C 13, induces IL-8 expression in human airway epithelial cells by activating protease-activated receptor 1 and 2. *J Immunol* 2007;178:5237-5244.

230. Wang Q, Nagarkar DR, Bowman ER, Schneider D, Gosangi B, Lei J, Zhao Y, McHenry CL, Burgens RV, Miller DJ, Sajjan U, Hershenson MB. Role of double-stranded RNA pattern recognition receptors in rhinovirus-induced airway epithelial cell responses. *J Immunol* 2009;183:6989-6997.

231. Lipinski S, Rosenstiel P. Debug Your Bugs - How NLRs Shape Intestinal Host-Microbe Interactions. *Front Immunol* 2013;4:479.

232. Nathan AT, Peterson EA, Chakir J, Wills-Karp M. Innate immune responses of airway epithelium to house dust mite are mediated through beta-glucan-dependent pathways. *J Allergy Clin Immunol* 2009;123:612-618.

233. Lee HM, Yuk JM, Shin DM, Jo EK. Dectin-1 is inducible and plays an essential role for mycobacteria-induced innate immune responses in airway epithelial cells. *J Clin Immunol* 2009;29:795-805.

234. Sun WK, Lu X, Li X, Sun QY, Su X, Song Y, Sun HM, Shi Y. Dectin-1 is inducible and plays a crucial role in Aspergillus-induced innate immune responses in human bronchial epithelial cells. *Eur J Clin Microbiol Infect Dis* 2012;31:2755-2764.

235. Page K, Lierl KM, Hughes VS, Zhou P, Ledford JR, Wills-Karp M. TLR2-mediated activation of neutrophils in response to German cockroach frass. *J Immunol* 2008;180:6317-6324.

236. Boldogh I, Bacsi A, Choudhury BK, Dharajiya N, Alam R, Hazra TK, Mitra S, Goldblum RM, Sur S. ROS generated by pollen NADPH oxidase provide a signal that augments antigen-induced allergic airway inflammation. *J Clin Invest* 2005;115:2169-2179.

237. Brown GD. Dectin-1: a signalling non-TLR pattern-recognition receptor. *Nat Rev Immunol* 2006;6:33-43.

238. Salo PM, Arbes SJ, Jr., Sever M, Jaramillo R, Cohn RD, London SJ, Zeldin DC. Exposure to Alternaria alternata in US homes is associated with asthma symptoms. *J Allergy Clin Immunol* 2006;118:892-898.

239. Hsu AC, See HV, Hansbro PM, Wark PA. Innate immunity to influenza in chronic airways diseases. *Respirology* 2012;17:1166-1175.

240. Hansbro PM, Beagley KW, Horvat JC, Gibson PG. Role of atypical bacterial infection of the lung in predisposition/protection of asthma. *Pharmacol Ther* 2004;101:193-210.

241. Hansbro NG, Horvat JC, Wark PA, Hansbro PM. Understanding the mechanisms of viral induced asthma: new therapeutic directions. *Pharmacol Ther* 2008;117:313-353.

242. Bosch AA, Biesbroek G, Trzcinski K, Sanders EA, Bogaert D. Viral and bacterial interactions in the upper respiratory tract. *PLoS Pathog* 2013;9:e1003057.

243. Hallstrand TS, Hackett TL, Altemeier WA, Matute-Bello G, Hansbro PM, Knight DA. Airway epithelial regulation of pulmonary immune homeostasis and inflammation. *Clin Immunol* 2014;151:1-15.

244. Ganesan S, Comstock AT, Sajjan US. Barrier function of airway tract epithelium. *Tissue Barriers* 2013;1:e24997.

245. Sajjan U, Wang Q, Zhao Y, Gruenert DC, Hershenson MB. Rhinovirus disrupts the barrier function of polarized airway epithelial cells. *Am J Respir Crit Care Med* 2008;178:1271-1281.

246. Comstock AT, Ganesan S, Chattoraj A, Faris AN, Margolis BL, Hershenson MB, Sajjan US. Rhinovirus-induced barrier dysfunction in polarized airway epithelial cells is mediated by NADPH oxidase 1. *J Virol* 2011;85:6795-6808.

247. Rezaee F, DeSando SA, Ivanov AI, Chapman TJ, Knowlden SA, Beck LA, Georas SN. Sustained protein kinase D activation mediates respiratory syncytial virus-induced airway barrier disruption. *J Virol* 2013;87:11088-11095.

248. Xiao C, Puddicombe SM, Field S, Haywood J, Broughton-Head V, Puxeddu I, Haitchi HM, Vernon-Wilson E, Sammut D, Bedke N, Cremin C, Sones J, Djukanović R, Howarth PH, Collins JE,

Holgate ST, Monk P, Davies DE. Defective epithelial barrier function in asthma. *J Allergy Clin Immunol* 2011;128:549-556.e541-512.

249. Hackett TL, de Bruin HG, Shaheen F, van den Berge M, van Oosterhout AJ, Postma DS, Heijink IH. Caveolin-1 controls airway epithelial barrier function. Implications for asthma. *Am J Respir Cell Mol Biol* 2013;49:662-671.

250. Shaykhiev R, Otaki F, Bonsu P, Dang DT, Teater M, Strulovici-Barel Y, Salit J, Harvey BG, Crystal RG. Cigarette smoking reprograms apical junctional complex molecular architecture in the human airway epithelium in vivo. *Cell Mol Life Sci* 2011;68:877-892.

251. Wonderlich ER, Swan ZD, Bissel SJ, Hartman AL, Carney JP, O'Malley KJ, Obadan AO, Santos J, Walker R, Sturgeon TJ, Frye LJ, Jr., Maiello P, Scanga CA, Bowling JD, Bouwer AL, Duangkhae PA, Wiley CA, Flynn JL, Wang J, Cole KS, Perez DR, Reed DS, Barratt-Boyes SM. Widespread Virus Replication in Alveoli Drives Acute Respiratory Distress Syndrome in Aerosolized H5N1 Influenza Infection of Macaques. *J Immunol* 2017;198:1616-1626.

252. Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SA, Homola D, Trujillo-Torralbo MB, Elkin S, Kon OM, Cookson WO, Moffatt MF, Johnston SL. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013;188:1224-1231.

253. Coyne CB, Shen L, Turner JR, Bergelson JM. Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe* 2007;2:181-192.

254. Grainge C, Dennison P, Lau L, Davies D, Howarth P. Asthmatic and normal respiratory epithelial cells respond differently to mechanical apical stress. *Am J Respir Crit Care Med* 2014;190:477-480.

255. Saatian B, Rezaee F, Desando S, Emo J, Chapman T, Knowlden S, Georas SN. Interleukin-4 and interleukin-13 cause barrier dysfunction in human airway epithelial cells. *Tissue Barriers* 2013;1:e24333.

256. Gause WC, Wynn TA, Allen JE. Type 2 immunity and wound healing: evolutionary refinement of adaptive immunity by helminths. *Nat Rev Immunol* 2013;13:607-614.

257. Heijink IH, Kies PM, Kauffman HF, Postma DS, van Oosterhout AJ, Vellenga E. Down-regulation of E-cadherin in human bronchial epithelial cells leads to epidermal growth factor receptor-dependent Th2 cell-promoting activity. *J Immunol* 2007;178:7678-7685.

258. de Boer WI, Sharma HS, Baelemans SM, Hoogsteden HC, Lambrecht BN, Braunstahl GJ. Altered expression of epithelial junctional proteins in atopic asthma: possible role in inflammation. *Can J Physiol Pharmacol* 2008;86:105-112.

259. Hackett TL, Singhera GK, Shaheen F, Hayden P, Jackson GR, Hegele RG, Van Eeden S, Bai TR, Dorscheid DR, Knight DA. Intrinsic phenotypic differences of asthmatic epithelium and its inflammatory responses to respiratory syncytial virus and air pollution. *Am J Respir Cell Mol Biol* 2011;45:1090-1100.

260. Koppelman GH, Meyers DA, Howard TD, Zheng SL, Hawkins GA, Ampleford EJ, Xu J, Koning H, Bruinenberg M, Nolte IM, van Diemen CC, Boezen HM, Timens W, Whittaker PA, Stine OC, Barton SJ, Holloway JW, Holgate ST, Graves PE, Martinez FD, van Oosterhout AJ, Bleecker ER, Postma DS. Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness. *Am J Respir Crit Care Med* 2009;180:929-935.

261. Ierodiakonou D, Postma DS, Koppelman GH, Boezen HM, Gerritsen J, Ten Hacken N, Timens W, Vonk JM. E-cadherin gene polymorphisms in asthma patients using inhaled corticosteroids. *Eur Respir J* 2011;38:1044-1052.

262. Sörby M, Ostman A. Protein-tyrosine phosphatase-mediated decrease of epidermal growth factor and platelet-derived growth factor receptor tyrosine phosphorylation in high cell density cultures. *J Biol Chem* 1996;271:10963-10966.

263. Symons JR, LeVea CM, Mooney RA. Expression of the leucocyte common antigen-related (LAR) tyrosine phosphatase is regulated by cell density through functional E-cadherin complexes. *Biochem J* 2002;365:513-519.

264. Qian X, Karpova T, Sheppard AM, McNally J, Lowy DR. E-cadherin-mediated adhesion inhibits ligand-dependent activation of diverse receptor tyrosine kinases. *Embo j* 2004;23:1739-1748.

265. Wendt MK, Smith JA, Schiemann WP. Transforming growth factor-β-induced epithelial-mesenchymal transition facilitates epidermal growth factor-dependent breast cancer progression. *Oncogene* 2010;29:6485-6498.

266. Casalino-Matsuda SM, Monzón ME, Forteza RM. Epidermal growth factor receptor activation by epidermal growth factor mediates oxidant-induced goblet cell metaplasia in human airway epithelium. *Am J Respir Cell Mol Biol* 2006;34:581-591.

267. Mazieres J, He B, You L, Xu Z, Jablons DM. Wnt signaling in lung cancer. *Cancer Lett* 2005;222:1-10.

268. Wang K, Zhang Y, Li X, Chen L, Wang H, Wu J, Zheng J, Wu D. Characterization of the Kremen-binding site on Dkk1 and elucidation of the role of Kremen in Dkk-mediated Wnt antagonism. *J Biol Chem* 2008;283:23371-23375.

269. Pfaff EM, Becker S, Günther A, Königshoff M. Dickkopf proteins influence lung epithelial cell proliferation in idiopathic pulmonary fibrosis. *Eur Respir J* 2011;37:79-87.

270. Sucre JMS, Deutsch GH, Jetter CS, Ambalavanan N, Benjamin JT, Gleaves LA, Millis BA, Young LR, Blackwell TS, Kropski JA, Guttentag SH. A Shared Pattern of β-Catenin Activation in Bronchopulmonary Dysplasia and Idiopathic Pulmonary Fibrosis. *Am J Pathol* 2018;188:853-862.

271. Shi J, Li F, Luo M, Wei J, Liu X. Distinct Roles of Wnt/β-Catenin Signaling in the Pathogenesis of Chronic Obstructive Pulmonary Disease and Idiopathic Pulmonary Fibrosis. *Mediators Inflamm* 2017;2017:3520581.

272. Baarsma HA, Skronska-Wasek W, Mutze K, Ciolek F, Wagner DE, John-Schuster G, Heinzelmann K, Günther A, Bracke KR, Dagouassat M, Boczkowski J, Brusselle GG, Smits R, Eickelberg O, Yildirim A, Königshoff M. Noncanonical WNT-5A signaling impairs endogenous lung repair in COPD. *J Exp Med* 2017;214:143-163.

273. Oh JE, Kim RH, Shin KH, Park NH, Kang MK. DeltaNp63α protein triggers epithelial-mesenchymal transition and confers stem cell properties in normal human keratinocytes. *J Biol Chem* 2011;286:38757-38767.

274. Jonsdottir HR, Arason AJ, Palsson R, Franzdottir SR, Gudbjartsson T, Isaksson HJ, Gudmundsson G, Gudjonsson T, Magnusson MK. Basal cells of the human airways acquire mesenchymal traits in idiopathic pulmonary fibrosis and in culture. *Lab Invest* 2015;95:1418-1428.

275. Manning EP, Losier A, Emeagwali N, Ryu C, Honiden S. New Applications of Old Drugs as Novel Therapies in Idiopathic Pulmonary Fibrosis. Metformin, Hydroxychloroquine, and Thyroid Hormone. *Am J Respir Crit Care Med* 2019;199:1561-1563.

276. Raghu G, Rochwerg B, Zhang Y, Garcia CA, Azuma A, Behr J, Brozek JL, Collard HR, Cunningham W, Homma S, Johkoh T, Martinez FJ, Myers J, Protzko SL, Richeldi L, Rind D, Selman M, Theodore A, Wells AU, Hoogsteden H, Schünemann HJ. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med* 2015;192:e3-19.

277. Invernizzi R, Barnett J, Rawal B, Nair A, Ghai P, Kingston S, Chua F, Wu Z, Wells AU, Renzoni ER, Nicholson AG, Rice A, Lloyd CM, Byrne AJ, Maher TM, Devaraj A, Molyneaux PL. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020;55.

278. Desjardins P, Conklin D. NanoDrop microvolume quantitation of nucleic acids. *J Vis Exp* 2010.

279. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme j* 2012;6:1621-1624.

280. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335-336.

281. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998;8:186-194.

282. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13:581-583.

283. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 2016;4:e2584.

284. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006;72:5069-5072.

285. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 2018;6:90.

286. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.

287. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.

288. Team RC. R: A language and environment for statistical computing. R Foundation for Stastical Computing. 2017.

289. Cox MJ, Cookson WO, Moffatt MF. Sequencing the human microbiome in health and disease. *Hum Mol Genet* 2013;22:R88-94.

290. Shannon CE. The mathematical theory of communication. 1963. *MD Comput* 1997;14:306-317.

291. Chao A. Nonparametric Estimation of the Number of Classes in a Population.

292. RH W. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol Mono* 1960;30:279-338.

293. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.

294. Team RC. A language and environment for statistical computing Vienna, Austria: R Foundation for Statistical Computing; 2013.

295. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci U S A* 1985;82:6955-6959.

296. H Gn. NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry. 3. Siegen: Wiley-VCH; 2013.

297. Nicholson JK, Connelly J, Lindon JC, Holmes E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov* 2002;1:153-161.

298. Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. *J Proteome Res* 2007;6:469-479.

299. Nicholson JK, Lindon JC. Systems biology: Metabonomics. *Nature* 2008;455:1054-1056.

300. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* 2006;78:4281-4290.

301. Tsang. High resolution magic angle spinning NMR spectroscopy and pattern recognition studies on neuronal damage: University of London; 2003.

302. Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal Chem* 2005;77:1282-1289.

303. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci* 2016;374:20150202.

304. Chen LS, Paul D, Prentice RL, Wang P. A regularized Hotelling's T(2) test for pathway analysis in proteomic studies. *J Am Stat Assoc* 2011;106:1345-1360.

305. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, KA LC. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* 2019;35:3055-3062.

306. Rohart F, Gautier B, Singh A, KA LC. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13:e1005752.

307. Keerthisingam CB, Jenkins RG, Harrison NK, Hernandez-Rodriguez NA, Booth H, Laurent GJ, Hart SL, Foster ML, McAnulty RJ. Cyclooxygenase-2 deficiency results in a loss of the anti-proliferative response to transforming growth factor-beta in human fibrotic lung fibroblasts and promotes bleomycin-induced pulmonary fibrosis in mice. *Am J Pathol* 2001;158:1411-1422.

308. O'Dwyer DN, Habiel D, Hogaboam C. Host-Microbial Interactions: Idiopathic Pulmonary Fibrosis in Technicolor. *Am J Respir Crit Care Med* 2017;195:1554-1556.

309. Molyneaux PL, Willis-Owen SAG, Cox MJ, James P, Cowman S, Loebinger M, Blanchard A, Edwards LM, Stock C, Daccord C, Renzoni EA, Wells AU, Moffatt MF, Cookson WOC, Maher TM. Host-Microbial Interactions in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2017;195:1640-1650.

310. Huang Y, Ma SF, Espindola MS, Vij R, Oldham JM, Huffnagle GB, Erb-Downward JR, Flaherty KR, Moore BB, White ES, Zhou T, Li J, Lussier YA, Han MK, Kaminski N, Garcia JGN, Hogaboam CM, Martinez FJ, Noth I. Microbes Are Associated with Host Innate Immune Response in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2017;196:208-219.

311. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246:697-722.

312. Jacob J, Bartholmai BJ, Rajagopalan S, van Moorsel CHM, van Es HW, van Beek FT, Struik MHL, Kokosi M, Egashira R, Brun AL, Nair A, Walsh SLF, Cross G, Barnett J, de Lauretis A, Judge EP, Desai S, Karwoski R, Ourselin S, Renzoni E, Maher TM, Altmann A, Wells AU. Predicting Outcomes in Idiopathic Pulmonary Fibrosis Using Automated Computed Tomographic Analysis. *Am J Respir Crit Care Med* 2018;198:767-776.

313. Jacob J, Odink A, Brun AL, Macaluso C, de Lauretis A, Kokosi M, Devaraj A, Desai S, Renzoni E, Wells AU. Functional associations of pleuroparenchymal fibroelastosis and emphysema with hypersensitivity pneumonitis. *Respir Med* 2018;138:95-101.

314. Laurence M, Hatzis C, Brash DE. Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS One* 2014;9:e97876.

315. Grahn N, Olofsson M, Ellnebo-Svedlund K, Monstein HJ, Jonasson J. Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiol Lett* 2003;219:87-91.

316. Barton HA, Taylor NM, Lubbers BR, Pemberton AC. DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. *J Microbiol Methods* 2006;66:21-31.

317. Tanner MA, Goebel BM, Dojka MA, Pace NR. Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants. *Appl Environ Microbiol* 1998;64:3110-3113.

318. Moffatt MF, Cookson WO. The lung microbiome in health and disease. *Clin Med (Lond)* 2017;17:525-529.

319. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 2011;184:957-963.

320. Takahashi Y, Saito A, Chiba H, Kuronuma K, Ikeda K, Kobayashi T, Ariki S, Takahashi M, Sasaki Y, Takahashi H. Impaired diversity of the lung microbiome predicts progression of idiopathic pulmonary fibrosis. *Respir Res* 2018;19:34.

321. Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2013;22:376-381.

322. Cox MJ, Allgaier M, Taylor B, Baek MS, Huang YJ, Daly RA, Karaoz U, Andersen GL, Brown R, Fujimura KE, Wu B, Tran D, Koff J, Kleinhenz ME, Nielson D, Brodie EL, Lynch SV. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. *PLoS One* 2010;5:e11044.

323. Garcia-Nuñez M, Millares L, Pomares X, Ferrari R, Pérez-Brocal V, Gallego M, Espasa M, Moya A, Monsó E. Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *J Clin Microbiol* 2014;52:4217-4223.

324. Huang C, Shi G. Smoking and microbiome in oral, airway, gut and some systemic diseases. *J Transl Med* 2019;17:225.

325. Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, Sinha R, Hwang J, Bushman FD, Collman RG. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *PLoS One* 2010;5:e15216.

326. Simpson JL, Daly J, Baines KJ, Yang IA, Upham JW, Reynolds PN, Hodge S, James AL, Hugenholtz P, Willner D, Gibson PG. Airway dysbiosis: Haemophilus influenzae and Tropheryma in poorly controlled asthma. *Eur Respir J* 2016;47:792-800.

327. Kinder BW, Brown KK, Schwarz MI, Ix JH, Kervitsky A, King TE, Jr. Baseline BAL neutrophilia predicts early mortality in idiopathic pulmonary fibrosis. *Chest* 2008;133:226-232.

328. Ghebre Y, Raghu G. Proton pump inhibitors in IPF: beyond mere suppression of gastric acidity. *Qjm* 2016;109:577-579.

329. Robbie H, Daccord C, Chua F, Devaraj A. Evaluating disease severity in idiopathic pulmonary fibrosis. *Eur Respir Rev* 2017;26.

330. Yagihashi K, Huckleberry J, Colby TV, Tazelaar HD, Zach J, Sundaram B, Pipavath S, Schwarz MI, Lynch DA. Radiologic-pathologic discordance in biopsy-proven usual interstitial pneumonia. *Eur Respir J* 2016;47:1189-1197.

331. Reddy TL, Tominaga M, Hansell DM, von der Thusen J, Rassl D, Parfrey H, Guy S, Twentyman O, Rice A, Maher TM, Renzoni EA, Wells AU, Nicholson AG. Pleuroparenchymal fibroelastosis: a spectrum of histopathological and imaging phenotypes. *Eur Respir J* 2012;40:377-385.

332. Dickson RP, Huffnagle GB, Flaherty KR, White ES, Martinez FJ, Erb-Downward JR, Moore BB, O'Dwyer DN. Radiographic Honeycombing and Altered Lung Microbiota in Patients with Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2019;200:1544-1547.

333. Macaluso C, Maritano Furcada J, Alzaher O, Chaube R, Chua F, Wells AU, Maher TM, George PM, Renzoni EA, Molyneaux PL. The potential impact of azithromycin in idiopathic pulmonary fibrosis. *Eur Respir J* 2019;53.

334. Sibila O, Laserna E, Shoemark A, Keir HR, Finch S, Rodrigo-Troyano A, Perea L, Lonergan M, Goeminne PC, Chalmers JD. Airway Bacterial Load and Inhaled Antibiotic Response in Bronchiectasis. *Am J Respir Crit Care Med* 2019;200:33-41.

335. Jo JH, Kennedy EA, Kong HH. Research Techniques Made Simple: Bacterial 16S Ribosomal RNA Gene Sequencing in Cutaneous Research. *J Invest Dermatol* 2016;136:e23-e27.

336. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methe B, DeSantis TZ, Petrosino JF, Knight R, Birren BW. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;21:494-504.

337. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.

338. Scheiermann J, Klinman DM. Three distinct pneumotypes characterize the microbiome of the lung in BALB/cJ mice. *PLoS One* 2017;12:e0180561.

339. Denning DW, Pashley C, Hartl D, Wardlaw A, Godet C, Del Giacco S, Delhaes L, Sergejeva S. Fungal allergy in asthma-state of the art and research needs. *Clin Transl Allergy* 2014;4:14.

340. Knutsen AP, Bush RK, Demain JG, Denning DW, Dixit A, Fairs A, Greenberger PA, Kariuki B, Kita H, Kurup VP, Moss RB, Niven RM, Pashley CH, Slavin RG, Vijay HM, Wardlaw AJ. Fungi and allergic lower respiratory tract diseases. *J Allergy Clin Immunol* 2012;129:280-291; quiz 292-283.

341. Busse WW, Lemanske RF, Jr., Gern JE. Role of viral respiratory infections in asthma and asthma exacerbations. *Lancet* 2010;376:826-834.

342. Delhaes L, Monchy S, Frealle E, Hubans C, Salleron J, Leroy S, Prevotat A, Wallet F, Wallaert B, Dei-Cas E, Sime-Ngando T, Chabe M, Viscogliosi E. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community--implications for therapeutic management. *PLoS One* 2012;7:e36313.

343. Charlson ES, Diamond JM, Bittinger K, Fitzgerald AS, Yadav A, Haas AR, Bushman FD, Collman RG. Lung-enriched organisms and aberrant bacterial and fungal respiratory microbiota after lung transplant. *Am J Respir Crit Care Med* 2012;186:536-545.

344. Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB. Modulation of post-antibiotic bacterial community reassembly and host response by Candida albicans. *Sci Rep* 2013;3:2191.

345. Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. *Nat Rev Immunol* 2014;14:827-835.

346. Koh A, De Vadder F, Kovatcheva-Datchary P, Backhed F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* 2016;165:1332-1345.

347. Madan JC, Koestler DC, Stanton BA, Davidson L, Moulton LA, Housman ML, Moore JH, Guill MF, Morrison HG, Sogin ML, Hampton TH, Karagas MR, Palumbo PE, Foster JA, Hibberd PL, O'Toole GA. Serial analysis of the gut and respiratory microbiome in cystic fibrosis in infancy: interaction between intestinal and respiratory tracts and impact of nutritional exposures. *mBio* 2012;3.

348. Noverr MC, Falkowski NR, McDonald RA, McKenzie AN, Huffnagle GB. Development of allergic airway disease in mice following antibiotic therapy and fungal microbiota increase: role of host genetics, antigen, and interleukin-13. *Infect Immun* 2005;73:30-38.

349. Trompette A, Gollwitzer ES, Pattaroni C, Lopez-Mejia IC, Riva E, Pernot J, Ubags N, Fajas L, Nicod LP, Marsland BJ. Dietary Fiber Confers Protection against Flu by Shaping Ly6c(-) Patrolling Monocyte Hematopoiesis and CD8(+) T Cell Metabolism. *Immunity* 2018;48:992-1005.e1008.

350. Peng L, Li ZR, Green RS, Holzman IR, Lin J. Butyrate enhances the intestinal barrier by facilitating tight junction assembly via activation of AMP-activated protein kinase in Caco-2 cell monolayers. *J Nutr* 2009;139:1619-1625.

351. Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 1987;28:1221-1227.

352. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud DJ, Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* 2013;54:2325-2340.

353. Ghorbani P, Santhakumar P, Hu Q, Djiadeu P, Wolever TM, Palaniyar N, Grasemann H. Short-chain fatty acids affect cystic fibrosis airway inflammation and bacterial growth. *Eur Respir J* 2015;46:1033-1045.

354. Imoto Y, Kato A, Takabayashi T, Sakashita M, Norton JE, Suh LA, Carter RG, Weibman AR, Hulse KE, Stevens W, Harris KE, Peters AT, Grammer LC, Tan BK, Welch K, Conley DB, Kern RC, Fujieda S, Schleimer RP. Short-chain fatty acids induce tissue plasminogen activator in airway epithelial cells via GPR41&43. *Clin Exp Allergy* 2018;48:544-554.

355. Mirkovic B, Murray MA, Lavelle GM, Molloy K, Azim AA, Gunaratnam C, Healy F, Slattery D, McNally P, Hatch J, Wolfgang M, Tunney MM, Muhlebach MS, Devery R, Greene CM, McElvaney NG. The Role of Short-Chain Fatty Acids, Produced by Anaerobic Bacteria, in the Cystic Fibrosis Airway. *Am J Respir Crit Care Med* 2015;192:1314-1324.

356. Zhao Y, Chen F, Wu W, Sun M, Bilotta AJ, Yao S, Xiao Y, Huang X, Eaves-Pyles TD, Golovko G, Fofanov Y, D'Souza W, Zhao Q, Liu Z, Cong Y. GPR43 mediates microbiota metabolite SCFA regulation of antimicrobial peptide expression in intestinal epithelial cells via activation of mTOR and STAT3. *Mucosal Immunol* 2018;11:752-762.

357. Vinolo MA, Rodrigues HG, Nachbar RT, Curi R. Regulation of inflammation by short chain fatty acids. *Nutrients* 2011;3:858-876.

358. Nakaji S, Sugawara K, Saito D, Yoshioka Y, MacAuley D, Bradley T, Kernohan G, Baxter D. Trends in dietary fiber intake in Japan over the last century. *Eur J Nutr* 2002;41:222-227.

359. Team RC. R: A language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2013.

360. Zhao YD, Yin L, Archer S, Lu C, Zhao G, Yao Y, Wu L, Hsin M, Waddell TK, Keshavjee S, Granton J, de Perrot M. Metabolic heterogeneity of idiopathic pulmonary fibrosis: a metabolomic study. *BMJ Open Respir Res* 2017;4:e000183.

361. Nigdelioglu R, Hamanaka RB, Meliton AY, O'Leary E, Witt LJ, Cho T, Sun K, Bonham C, Wu D, Woods PS, Husain AN, Wolfgeher D, Dulin NO, Chandel NS, Mutlu GM. Transforming Growth Factor (TGF)-β Promotes de Novo Serine Synthesis for Collagen Production. *J Biol Chem* 2016;291:27239-27251.

362. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, Le Cao KA. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* 2019;35:3055-3062.

363. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13:e1005752.

364. Zhao H, Dennery PA, Yao H. Metabolic reprogramming in the pathogenesis of chronic lung diseases, including BPD, COPD, and pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2018;314:L544-l554.

365. Nambiar S, Bong How S, Gummer J, Trengove R, Moodley Y. Metabolomics in chronic lung diseases. *Respirology* 2020;25:139-148.

366. Kumar N, Shahjaman M, Mollah MNH, Islam SMS, Hoque MA. Serum and Plasma Metabolomic Biomarkers for Lung Cancer. *Bioinformation* 2017;13:202-208.

367. Güemes M, Rahman SA, Hussain K. What is a normal blood glucose? *Arch Dis Child* 2016;101:569-574.

368. Zhao X, Psarianos P, Ghoraie LS, Yip K, Goldstein D, Gilbert R, Witterick I, Pang H, Hussain A, Lee JH, Williams J, Bratman SV, Ailles L, Haibe-Kains B, Liu FF. Metabolic regulation of dermal fibroblasts contributes to skin extracellular matrix homeostasis and fibrosis. *Nat Metab* 2019;1:147-157.

369. Yin X, Choudhury M, Kang JH, Schaefbauer KJ, Jung MY, Andrianifahanana M, Hernandez DM, Leof EB. Hexokinase 2 couples glycolysis with the profibrotic actions of TGF-β. *Sci Signal* 2019;12.

370. DeBerardinis RJ, Thompson CB. Cellular metabolism and disease: what do metabolic outliers teach us? *Cell* 2012;148:1132-1144.

371. Adeva-Andany M, López-Ojén M, Funcasta-Calderón R, Ameneiros-Rodríguez E, Donapetry-García C, Vila-Altesor M, Rodríguez-Seijas J. Comprehensive review on lactate metabolism in human health. *Mitochondrion* 2014;17:76-100.

372. Kottmann RM, Kulkarni AA, Smolnycki KA, Lyda E, Dahanayake T, Salibi R, Honnons S, Jones C, Isern NG, Hu JZ, Nathan SD, Grant G, Phipps RP, Sime PJ. Lactic acid is elevated in idiopathic pulmonary fibrosis and induces myofibroblast differentiation via pH-dependent activation of transforming growth factor-β. *Am J Respir Crit Care Med* 2012;186:740-751.

373. Effros RM, Hoagland KW, Bosbous M, Castillo D, Foss B, Dunning M, Gare M, Lin W, Sun F. Dilution of respiratory solutes in exhaled condensates. *Am J Respir Crit Care Med* 2002;165:663-669.

374. O'Neil JJ, Tierney DF. Rat lung metabolism: glucose utilization by isolated perfused lungs and tissue slices. *Am J Physiol* 1974;226:867-873.

375. De Backer D, Creteur J, Zhang H, Norrenberg M, Vincent JL. Lactate production by the lungs in acute lung injury. *Am J Respir Crit Care Med* 1997;156:1099-1104.

376. Hamanaka RB, Nigdelioglu R, Meliton AY, Tian Y, Witt LJ, O'Leary E, Sun KA, Woods PS, Wu D, Ansbro B, Ard S, Rohde JM, Dulin NO, Guzy RD, Mutlu GM. Inhibition of Phosphoglycerate

Dehydrogenase Attenuates Bleomycin-induced Pulmonary Fibrosis. *Am J Respir Cell Mol Biol* 2018;58:585-593.

377. Liu G, Summer R. Cellular Metabolism in Lung Health and Disease. *Annu Rev Physiol* 2019;81:403-428.

378. de Brouwer B, Drent M, van den Ouweland JMW, Wijnen PA, van Moorsel CHM, Bekers O, Grutters JC, White ES, Janssen R. Increased circulating desmosine and age-dependent elastinolysis in idiopathic pulmonary fibrosis. *Respir Res* 2018;19:45.

379. Gaugg MT, Engler A, Bregy L, Nussbaumer-Ochsner Y, Eiffert L, Bruderer T, Zenobi R, Sinues P, Kohler M. Molecular breath analysis supports altered amino acid metabolism in idiopathic pulmonary fibrosis. *Respirology* 2019;24:437-444.

380. Akgul Kalkan E, Sahiner M, Ulker Cakir D, Alpaslan D, Yilmaz S. Quantitative Clinical Diagnostic Analysis of Acetone in Human Blood by HPLC: A Metabolomic Search for Acetone as Indicator. *J Anal Methods Chem* 2016;2016:5176320.

381. Miller ON, Bazzano G. Propanediol metabolism and its relation to lactic acid metabolism. *Ann N Y Acad Sci* 1965;119:957-973.

382. Ruddick JA. Toxicology, metabolism, and biochemistry of 1,2-propanediol. *Toxicol Appl Pharmacol* 1972;21:102-111.

383. Glew RH. You can get there from here: acetone, anionic ketones and even-carbon fatty acids can provide substrates for gluconeogenesis. *Niger J Physiol Sci* 2010;25:2-4.

384. Montuschi P, Paris D, Melck D, Lucidi V, Ciabattoni G, Raia V, Calabrese C, Bush A, Barnes PJ, Motta A. NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax* 2012;67:222-228.

385. Wolak JE, Esther CR, Jr., O'Connell TM. Metabolomic analysis of bronchoalveolar lavage fluid from cystic fibrosis patients. *Biomarkers* 2009;14:55-60.

386. Anderson JC, Lamm WJ, Hlastala MP. Measuring airway exchange of endogenous acetone using a single-exhalation breathing maneuver. *J Appl Physiol (1985)* 2006;100:880-889.

387. Widmark EM. Studies in the Acetone Concentration in Blood, Urine, and Alveolar Air. III: The Elimination of Acetone through the Lungs. *Biochem J* 1920;14:379-394.

388. Reichard GA, Jr., Haff AC, Skutches CL, Paul P, Holroyde CP, Owen OE. Plasma acetone metabolism in the fasting human. *J Clin Invest* 1979;63:619-626.

389. Schuchmann K, Müller V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat Rev Microbiol* 2014;12:809-821.

390. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol* 2014;12:661-672.

391. Vital M, Howe AC, Tiedje JM. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio* 2014;5:e00889.

392. Hegazi FZ, Abo-Elnaga IG. Degradation of organic acids by dairy lactic acid bacteria. *Zentralbl Bakteriol Naturwiss* 1980;135:212-222.

393. Roussos A, Koursarakos P, Patsopoulos D, Gerogianni I, Philippou N. Increased prevalence of irritable bowel syndrome in patients with bronchial asthma. *Respir Med* 2003;97:75-79.

394. Baral V, Connett G. Acute intestinal obstruction as a presentation of cystic fibrosis in infancy. *J Cyst Fibros* 2008;7:277-279.

395. Keely S, Hansbro PM. Lung-gut cross talk: a potential mechanism for intestinal dysfunction in patients with COPD. *Chest* 2014;145:199-200.

396. Chelakkot C, Ghim J, Ryu SH. Mechanisms regulating intestinal barrier integrity and its pathological implications. *Exp Mol Med* 2018;50:103.

397. Zaiss MM, Rapin A, Lebon L, Dubey LK, Mosconi I, Sarter K, Piersigilli A, Menin L, Walker AW, Rougemont J, Paerewijck O, Geldhof P, McCoy KD, Macpherson AJ, Croese J, Giacomin PR, Loukas A, Junt T, Marsland BJ, Harris NL. The Intestinal Microbiota Contributes to the Ability of Helminths to Modulate Allergic Inflammation. *Immunity* 2015;43:998-1010.

398. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, Harmsen HJM, Faber KN, Hermoso MA. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol* 2019;10:277.

399. Rutting S, Xenaki D, Malouf M, Horvat JC, Wood LG, Hansbro PM, Oliver BG. Short-chain fatty acids increase TNFalpha-induced inflammation in primary human lung mesenchymal cells through the activation of p38 MAPK. *Am J Physiol Lung Cell Mol Physiol* 2019;316:L157-l174.

400. Martinez FJ, Collard HR, Pardo A, Raghu G, Richeldi L, Selman M, Swigris JJ, Taniguchi H, Wells AU. Idiopathic pulmonary fibrosis. *Nat Rev Dis Primers* 2017;3:17074.

401. Tang YW, Johnson JE, Browning PJ, Cruz-Gervis RA, Davis A, Graham BS, Brigham KL, Oates JA, Jr., Loyd JE, Stecenko AA. Herpesvirus DNA is consistently detected in lungs of patients with idiopathic pulmonary fibrosis. *J Clin Microbiol* 2003;41:2633-2640.

402. Han MK. High prevalence of abnormal acid gastro-oesophageal reflux in idiopathic pulmonary fibrosis. *Eur Respir J* 2006;28:884-885; author reply 885.

403. Le Poul E, Loison C, Struyf S, Springael JY, Lannoy V, Decobecq ME, Brezillon S, Dupriez V, Vassart G, Van Damme J, Parmentier M, Detheux M. Functional characterization of human receptors

for short chain fatty acids and their role in polymorphonuclear cell activation. *J Biol Chem* 2003;278:25481-25489.

404. Brown AJ, Goldsworthy SM, Barnes AA, Eilert MM, Tcheang L, Daniels D, Muir AI, Wigglesworth MJ, Kinghorn I, Fraser NJ, Pike NB, Strum JC, Steplewski KM, Murdock PR, Holder JC, Marshall FH, Szekeres PG, Wilson S, Ignar DM, Foord SM, Wise A, Dowell SJ. The Orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *J Biol Chem* 2003;278:11312-11319.

405. Fusunyan RD, Quinn JJ, Fujimoto M, MacDermott RP, Sanderson IR. Butyrate switches the pattern of chemokine secretion by intestinal epithelial cells through histone acetylation. *Mol Med* 1999;5:631-640.

406. Srinivasan B, Kolli AR, Esch MB, Abaci HE, Shuler ML, Hickman JJ. TEER measurement techniques for in vitro barrier model systems. *J Lab Autom* 2015;20:107-126.

407. Vancheri C. Idiopathic pulmonary fibrosis: an altered fibroblast proliferation linked to cancer biology. *Proc Am Thorac Soc* 2012;9:153-157.

408. Jeffery PK. Remodeling and inflammation of bronchi in asthma and chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2004;1:176-183.

409. Kato A, Schleimer RP. Beyond inflammation: airway epithelial cells are at the interface of innate and adaptive immunity. *Curr Opin Immunol* 2007;19:711-720.

410. Tisoncik JR, Korth MJ, Simmons CP, Farrar J, Martin TR, Katze MG. Into the eye of the cytokine storm. *Microbiol Mol Biol Rev* 2012;76:16-32.

411. Böcker U, Nebe T, Herweck F, Holt L, Panja A, Jobin C, Rossol S, R BS, Singer MV. Butyrate modulates intestinal epithelial cell-mediated neutrophil migration. *Clin Exp Immunol* 2003;131:53-60.

412. Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Yu D, Schilter HC, Rolph MS, Mackay F, Artis D, Xavier RJ, Teixeira MM, Mackay CR. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 2009;461:1282-1286.

413. Macia L, Thorburn AN, Binge LC, Marino E, Rogers KE, Maslowski KM, Vieira AT, Kranich J, Mackay CR. Microbial influences on epithelial integrity and immune function as a basis for inflammatory diseases. *Immunol Rev* 2012;245:164-176.

414. Jan G, Belzacq AS, Haouzi D, Rouault A, Métivier D, Kroemer G, Brenner C. Propionibacteria induce apoptosis of colorectal carcinoma cells via short-chain fatty acids acting on mitochondria. *Cell Death Differ* 2002;9:179-188.

415. Wang H, He L, Liu B, Feng Y, Zhou H, Zhang Z, Wu Y, Wang J, Gan Y, Yuan T, Wu M, Xie X, Feng Z. Establishment and comparison of air-liquid interface culture systems for primary and immortalized swine tracheal epithelial cells. *BMC Cell Biol* 2018;19:10.

416. Grainger CI, Greenwell LL, Lockley DJ, Martin GP, Forbes B. Culture of Calu-3 cells at the air interface provides a representative model of the airway epithelial barrier. *Pharm Res* 2006;23:1482-1490.

417. Krunkosky TM, Jordan JL, Chambers E, Krause DC. Mycoplasma pneumoniae host-pathogen studies in an air-liquid culture of differentiated human airway epithelial cells. *Microb Pathog* 2007;42:98-103.

418. Bateman AC, Karasin AI, Olsen CW. Differentiated swine airway epithelial cell cultures for the investigation of influenza A virus infection and replication. *Influenza Other Respir Viruses* 2013;7:139-150.

419. Delgado-Ortega M, Olivier M, Sizaret PY, Simon G, Meurens F. Newborn pig trachea cell line cultured in air-liquid interface conditions allows a partial in vitro representation of the porcine upper airway tissue. *BMC Cell Biol* 2014;15:14.

420. Xue D, Ma Y, Li M, Li Y, Luo H, Liu X, Wang Y. Mycoplasma ovipneumoniae induces inflammatory response in sheep airway epithelial cells via a MyD88-dependent TLR signaling pathway. *Vet Immunol Immunopathol* 2015;163:57-66.

421. Dvorak A, Tilley AE, Shaykhiev R, Wang R, Crystal RG. Do airway epithelium air-liquid cultures represent the in vivo airway epithelium transcriptome? *Am J Respir Cell Mol Biol* 2011;44:465-473.

422. Pezzulo AA, Starner TD, Scheetz TE, Traver GL, Tilley AE, Harvey BG, Crystal RG, McCray PB, Jr., Zabner J. The air-liquid interface and use of primary cell cultures are important to recapitulate the transcriptional profile of in vivo airway epithelia. *Am J Physiol Lung Cell Mol Physiol* 2011;300:L25-31.

423. Gray TE, Guzman K, Davis CW, Abdullah LH, Nettesheim P. Mucociliary differentiation of serially passaged normal human tracheobronchial epithelial cells. *Am J Respir Cell Mol Biol* 1996;14:104-112.

424. Kim K, Kwon O, Ryu TY, Jung CR, Kim J, Min JK, Kim DS, Son MY, Cho HS. Propionate of a microbiota metabolite induces cell apoptosis and cell cycle arrest in lung cancer. *Mol Med Rep* 2019;20:1569-1574.

425. Min KA, Rosania GR, Shin MC. Human Airway Primary Epithelial Cells Show Distinct Architectures on Membrane Supports Under Different Culture Conditions. *Cell Biochem Biophys* 2016;74:191-203.

426. Forbes B, Ehrhardt C. Human respiratory epithelial cell culture for drug delivery applications. *Eur J Pharm Biopharm* 2005;60:193-205.

427. Lin H, Li H, Cho HJ, Bian S, Roh HJ, Lee MK, Kim JS, Chung SJ, Shim CK, Kim DD. Air-liquid interface (ALI) culture of human bronchial epithelial cell monolayers as an in vitro model for airway drug transport studies. *J Pharm Sci* 2007;96:341-350.

428. Selman M, Pardo A. Role of epithelial cells in idiopathic pulmonary fibrosis: from innocent targets to serial killers. *Proc Am Thorac Soc* 2006;3:364-372.

429. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci* 2010;123:4195-4200.

430. Laurent GJ, Harrison NK, McAnulty RJ. The regulation of collagen production in normal lung and during interstitial lung disease. *Postgrad Med J* 1988;64 Suppl 4:26-34.

431. Breton JD, Heydet D, Starrs LM, Veldre T, Ghildyal R. Molecular changes during TGFβ-mediated lung fibroblast-myofibroblast differentiation: implication for glucocorticoid resistance. *Physiol Rep* 2018;6:e13669.

432. Klingberg F, Hinz B, White ES. The myofibroblast matrix: implications for tissue repair and fibrosis. *J Pathol* 2013;229:298-309.

433. Kulkarni T, O'Reilly P, Antony VB, Gaggar A, Thannickal VJ. Matrix Remodeling in Pulmonary Fibrosis and Emphysema. *Am J Respir Cell Mol Biol* 2016;54:751-760.

434. Gras D, Petit A, Charriot J, Knabe L, Alagha K, Gamez AS, Garulli C, Bourdin A, Chanez P, Molinari N, Vachier I. Epithelial ciliated beating cells essential for ex vivo ALI culture growth. *BMC Pulm Med* 2017;17:80.

435. Stewart CE, Torr EE, Mohd Jamili NH, Bosquillon C, Sayers I. Evaluation of differentiated human bronchial epithelial cell culture systems for asthma research. *J Allergy (Cairo)* 2012;2012:943982.

436. Cait A, Hughes MR, Antignano F, Cait J, Dimitriu PA, Maas KR, Reynolds LA, Hacker L, Mohr J, Finlay BB, Zaph C, McNagny KM, Mohn WW. Microbiome-driven allergic lung inflammation is ameliorated by short-chain fatty acids. *Mucosal Immunol* 2018;11:785-795.

437. Raghu G. Idiopathic pulmonary fibrosis: lessons from clinical trials over the past 25 years. *Eur Respir J* 2017;50.

438. Collard HR, Moore BB, Flaherty KR, Brown KK, Kaner RJ, King TE, Jr., Lasky JA, Loyd JE, Noth I, Olman MA, Raghu G, Roman J, Ryu JH, Zisman DA, Hunninghake GW, Colby TV, Egan JJ, Hansell DM, Johkoh T, Kaminski N, Kim DS, Kondoh Y, Lynch DA, Müller-Quernheim J, Myers JL, Nicholson

AG, Selman M, Toews GB, Wells AU, Martinez FJ. Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2007;176:636-643.

439. Vergnon JM, Vincent M, de Thé G, Mornex JF, Weynants P, Brune J. Cryptogenic fibrosing alveolitis and Epstein-Barr virus: an association? *Lancet* 1984;2:768-771.

440. Stoolman JS, Vannella KM, Coomes SM, Wilke CA, Sisson TH, Toews GB, Moore BB. Latent infection by γherpesvirus stimulates profibrotic mediator release from multiple cell types. *Am J Physiol Lung Cell Mol Physiol* 2011;300:L274-285.

441. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE, Jr., Kondoh Y, Myers J, Müller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schünemann HJ. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183:788-824.

442. Vasakova M, Morell F, Walsh S, Leslie K, Raghu G. Hypersensitivity Pneumonitis: Perspectives in Diagnosis and Management. *Am J Respir Crit Care Med* 2017;196:680-689.

443. Cappelli S, Bellando Randone S, Martinović D, Tamas MM, Pasalić K, Allanore Y, Mosca M, Talarico R, Opris D, Kiss CG, Tausche AK, Cardarelli S, Riccieri V, Koneva O, Cuomo G, Becker MO, Sulli A, Guiducci S, Radić M, Bombardieri S, Aringer M, Cozzi F, Valesini G, Ananyeva L, Valentini G, Riemekasten G, Cutolo M, Ionescu R, Czirják L, Damjanov N, Rednic S, Matucci Cerinic M. "To be or not to be," ten years after: evidence for mixed connective tissue disease as a distinct entity. *Semin Arthritis Rheum* 2012;41:589-598.

444. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. Standardisation of spirometry. *Eur Respir J* 2005;26:319-338.

445. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P, Gustafsson P, Hankinson J, Jensen R, McKay R, Miller MR, Navajas D, Pedersen OF, Pellegrino R, Wanger J. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;26:720-735.

446. de Oliveira S, Reyes-Aldasoro CC, Candel S, Renshaw SA, Mulero V, Calado A. Cxcl8 (IL-8) mediates neutrophil recruitment and behavior in the zebrafish inflammatory response. *J Immunol* 2013;190:4349-4359.

447. Tong X, Su F, Xu X, Xu H, Yang T, Xu Q, Dai H, Huang K, Zou L, Zhang W, Pei S, Xiao F, Li Y, Wang C. Alterations to the Lung Microbiome in Idiopathic Pulmonary Fibrosis Patients. *Front Cell Infect Microbiol* 2019;9:149.

448. Brown JS. Community-acquired pneumonia. *Clin Med (Lond)* 2012;12:538-543.

449. Halnes I, Baines KJ, Berthon BS, MacDonald-Wicks LK, Gibson PG, Wood LG. Soluble Fibre Meal Challenge Reduces Airway Inflammation and Expression of GPR43 and GPR41 in Asthma. *Nutrients* 2017;9.

450. Chung KF. Cytokines in chronic obstructive pulmonary disease. *Eur Respir J Suppl* 2001;34:50s-59s.

451. Sakamoto K, Taniguchi H, Kondoh Y, Wakai K, Kimura T, Kataoka K, Hashimoto N, Nishiyama O, Hasegawa Y. Acute exacerbation of IPF following diagnostic bronchoalveolar lavage procedures. *Respir Med* 2012;106:436-442.

452. Vannella KM, Moore BB. Viruses as co-factors for the initiation or exacerbation of lung fibrosis. *Fibrogenesis Tissue Repair* 2008;1:2.

453. King TE, Jr., Albera C, Bradford WZ, Costabel U, du Bois RM, Leff JA, Nathan SD, Sahn SA, Valeyre D, Noble PW. All-cause mortality rate in patients with idiopathic pulmonary fibrosis. Implications for the design and execution of clinical trials. *Am J Respir Crit Care Med* 2014;189:825-831.

454. Dickson RP, Singer BH, Newstead MW, Falkowski NR, Erb-Downward JR, Standiford TJ, Huffnagle GB. Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat Microbiol* 2016;1:16113.

455. Singanayagam A, Ritchie AI, Johnston SL. Role of microbiome in the pathophysiology and disease course of asthma. *Curr Opin Pulm Med* 2017;23:41-47.

456. Corrêa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MA. Regulation of immune cell function by short-chain fatty acids. *Clin Transl Immunology* 2016;5:e73.

457. Papaioannou O, Karampitsakos T, Barbayianni I, Chrysikos S, Xylourgidis N, Tzilas V, Bouros D, Aidinis V, Tzouvelekis A. Metabolic Disorders in Chronic Lung Diseases. *Front Med (Lausanne)* 2017;4:246.

9 Appendix

Appendix 1. Bioinformatics pipeline script

```
# 16S Bioinformatics Analysis
# Multiple Sequencing runs (180216_MSQ75;180719_2_MSQ81;180719_MSQ82)
# QIIME2 (version 2018.8)
# Run on HPC cluster
# Rachele Invernizzi - Imperial College London

# Log in to cluster

ssh ri216@login.cx1.hpc.ic.ac.uk

# PREPARING DATA

# Activate

module load anaconda3/personal
source activate py35
source activate qiime2-2018.8

# Set and change directory to main folder

cd $WORK/16S-Analysis

# Rename the sequences (3 fastq files: R1 forward reads; R2 barcodes; R3 reverse reads)
# GZIP the files (converts in .gz format)

cd $WORK/16S-Analysis/180216_MSQ75/2_fastq

mv lane1_NoIndex_L001_R1_001.fastq forward.fastq
mv lane1_NoIndex_L001_R2_001.fastq barcodes.fastq
mv lane1_NoIndex_L001_R3_001.fastq reverse.fastq

gzip barcodes.fastq
gzip forward.fastq
gzip reverse.fastq

cd $WORK/16S-Analysis/180719_2_MSQ81/2_fastq

mv lane1_NoIndex_L001_R1_001.fastq forward.fastq
mv lane1_NoIndex_L001_R2_001.fastq barcodes.fastq
mv lane1_NoIndex_L001_R3_001.fastq reverse.fastq

gzip barcodes.fastq
gzip forward.fastq
gzip reverse.fastq

cd $WORK/16S-Analysis/180719_MSQ82/2_fastq

mv lane1_NoIndex_L001_R1_001.fastq forward.fastq
mv lane1_NoIndex_L001_R2_001.fastq barcodes.fastq
mv lane1_NoIndex_L001_R3_001.fastq reverse.fastq

gzip barcodes.fastq
gzip forward.fastq
gzip reverse.fastq


# IMPORTING DATA

cd $WORK/16S-Analysis/180216_MSQ75

qiime tools import \
--type EMPPairedEndSequences \
--input-path 2_fastq \
--output-path 2_fastq/emp-paired-end-sequences.qza
```

```
cd $WORK/16S-Analysis/180719_2_MSQ81

qiime tools import \
--type EMPPairedEndSequences \
--input-path 2_fastq \
--output-path 2_fastq/emp-paired-end-sequences.qza

cd $WORK/16S-Analysis/180719_MSQ82

qiime tools import \
--type EMPPairedEndSequences \
--input-path 2_fastq \
--output-path 2_fastq/emp-paired-end-sequences.qza

# DEMULTIPLEXING

# Demupltiplex the read files based on sample names
# BarcodeSequence: indicates which barcode sequence is associated with each sample
# Standard quality filtering parameters used

cd $WORK/16S-Analysis/180216_MSQ75/2_fastq

qiime demux emp-paired \
  --i-seqs emp-paired-end-sequences.qza \
  --m-barcodes-file MSQ.75.Map.tsv \
  --m-barcodes-column BarcodeSequence \
  --p-rev-comp-mapping-barcodes \
  --o-per-sample-sequences demux.qza

cd $WORK/16S-Analysis/180719_2_MSQ81/2_fastq

qiime demux emp-paired \
  --i-seqs emp-paired-end-sequences.qza \
  --m-barcodes-file MSQ.81.Map.tsv \
  --m-barcodes-column BarcodeSequence \
  --p-rev-comp-mapping-barcodes \
  --o-per-sample-sequences demux.qza

cd $WORK16S-Analysis/180719_MSQ82/2_fastq

qiime demux emp-paired \
  --i-seqs emp-paired-end-sequences.qza \
  --m-barcodes-file MSQ.82.Map.tsv \
  --m-barcodes-column BarcodeSequence \
  --p-rev-comp-mapping-barcodes \
  --o-per-sample-sequences demux.qza

# Summarise data and visualise

qiime demux summarize --i-data demux.qza --o-visualization demux.qzv

qiime tools view demux.qzv

# DADA2 and DENOISE

# The dada2 denoise-paired requires four parameters: --p-trim-left-f, --p-trim-left-r, --
p-trunc-len-f, and --p-trunc-len-r. The --p-trim-left m trims off the first m bases of
each sequence, and --p-trunc-len n truncates each sequence at position n. The f and r in
each parameter stand for forward and reverse read, correspondingly.

cd $WORK/16S-Analysis/180216_MSQ75/2_fastq

qiime dada2 denoise-paired \
  --i-demultiplexed-seqs demux.qza \
  --p-trim-left-f 13 \
  --p-trim-left-r 13 \
  --p-trunc-len-f 140 \
  --p-trunc-len-r 140 \
  --p-n-threads 32 \
  --o-representative-sequences rep-seqs-dada2.qza \
  --o-table table-dada2.qza \
```

```
  --o-denoising-stats stats.qza \
  --output-dir Merged_Runs
cd $WORK/16S-Analysis/180719_2_MSQ81/2_fastq

qiime dada2 denoise-paired \
  --i-demultiplexed-seqs demux.qza \
  --p-trim-left-f 13 \
  --p-trim-left-r 13 \
  --p-trunc-len-f 140 \
  --p-trunc-len-r 140 \
  --p-n-threads 32 \
  --o-representative-sequences rep-seqs-dada2.qza \
  --o-table table-dada2.qza \
  --o-denoising-stats stats.qza \
  --output-dir Merged_Runs

cd $WORK/16S-Analysis/180719_MSQ82/2_fastq

qiime dada2 denoise-paired \
  --i-demultiplexed-seqs demux.qza \
  --p-trim-left-f 13 \
  --p-trim-left-r 13 \
  --p-trunc-len-f 140 \
  --p-trunc-len-r 140 \
  --p-n-threads 32 \
  --o-representative-sequences rep-seqs-dada2.qza \
  --o-table table-dada2.qza \
  --o-denoising-stats stats.qza \
  --output-dir Merged_Runs

# Rename files

cd $WORK/16S-Analysis/Merged_Runs

mv rep-seqs-dada2.qza rep-seqs-1.qza
mv table-dada2.qza table-1.qza

mv rep-seqs-dada2.qza rep-seqs-2.qza
mv table-dada2.qza table-2.qza

mv rep-seqs-dada2.qza rep-seqs-3.qza
mv table-dada2.qza table-3.qza

# Output:
FeatureTable [Frequency]: containing counts (frequencies) of each unique sequence in each
sample in the dataset
FeatureData [Sequence]: maps feature identifiers in the FeatureTable to the sequences they
represent

# EXCLUDING SEQUENCES BY ALIGNMENT

# Aligns a set of query sequences contained in a FeatureData[Sequence] file against a set
of reference sequences
# DOWNLOAD reference sequences (greengenes gg_13_8_otus; rep_set folder (unaligned);
99_otus.fasta)
# Import 99_otus.fasta (output: 99_otus.qza)

cd $WORK/16S-Analysis/Merged_Runs

qiime tools import \
--type 'FeatureData[Sequence]' \
--input-path 99_otus.fasta \
--output-path 99_otus.qza

# Employs vsearch to decide whether a sequence "hits" the reference sequences, and outputs
separate files of sequences that hit/do not hit the reference sequences

cd $WORK/16S-Analysis/Merged_Runs

qiime quality-control exclude-seqs \
--i-query-sequences rep-seqs-1.qza \
--i-reference-sequences 99_otus.qza \
--p-method vsearch \
```

```
--p-perc-identity 0.99 \
--p-perc-query-aligned 0.99 \
--p-threads 4 \
--o-sequence-hits sequence-hits-1.qza \
--o-sequence-misses sequence-misses-1.qza

cd $WORK/CHPManuscriptRI/180719_2_MSQ81/2_fastq

qiime quality-control exclude-seqs \
--i-query-sequences rep-seqs-2.qza \
--i-reference-sequences 99_otus.qza \
--p-method vsearch \
--p-perc-identity 0.99 \
--p-perc-query-aligned 0.99 \
--p-threads 4 \
--o-sequence-hits sequence-hits-2.qza \
--o-sequence-misses sequence-misses-2.qza

cd $WORK/CHPManuscriptRI/180719_MSQ82/2_fastq

qiime quality-control exclude-seqs \
--i-query-sequences rep-seqs-3.qza \
--i-reference-sequences 99_otus.qza \
--p-method vsearch \
--p-perc-identity 0.99 \
--p-perc-query-aligned 0.99 \
--p-threads 4 \
--o-sequence-hits sequence-hits-3.qza \
--o-sequence-misses sequence-misses-3.qza


# FILTERING SEQUENCES POST-ALIGNMENT

# Filter feature table to remove hits or misses
# Filter out contaminants or sequences that align to host DNA

cd $WORK/16S-Analysis/180216_MSQ75/2_fastq

qiime feature-table filter-features \
--i-table table-1.qza \
--m-metadata-file sequence-misses-1.qza \
--o-filtered-table no-hits-filtered-table-1.qza \
--p-exclude-ids

cd $WORK/16S-Analysis/180719_2_MSQ81/2_fastq
qiime feature-table filter-features \
--i-table table-2.qza \
--m-metadata-file sequence-misses-2.qza \
--o-filtered-table no-hits-filtered-table-2.qza \
--p-exclude-ids

cd $WORK/16S-Analysis/180719_MSQ82/2_fastq

qiime feature-table filter-features \
--i-table table-3.qza \
--m-metadata-file sequence-misses-3.qza \
--o-filtered-table no-hits-filtered-table-3.qza \
--p-exclude-ids

# SUMMARISING FEATURE TABLE AND FEATURE DATA

cd $WORK/16S-Analysis/Merged_Runs

qiime feature-table summarize \
--i-table filtered-table-1.qza \
--o-visualization filtered-table-1.qzv \
--m-sample-metadata-file mapping.tsv

qiime feature-table tabulate-seqs \
--i-data rep-seqs-1.qza \
--o-visualization rep-seqs-1.qzv
```

```
cd $WORK/16S-Analysis/Merged_Runs

qiime feature-table summarize \
--i-table filtered-table-2.qza \
--o-visualization filtered-table-2.qzv \
--m-sample-metadata-file mapping.tsv

qiime feature-table tabulate-seqs \
--i-data rep-seqs-2.qza \
--o-visualization rep-seqs-2.qzv

cd $WORK/16S-Analysis/Merged_Runs

qiime feature-table summarize \
--i-table filtered-table-3.qza \
--o-visualization filtered-table-3.qzv \
--m-sample-metadata-file mapping.tsv

qiime feature-table tabulate-seqs \
--i-data rep-seqs-3.qza \
--o-visualization rep-seqs-3.qzv

# Output visualisations

qiime tools view table-1.qzv
qiime tools view table-2.qzv
qiime tools view table-3.qzv
qiime tools view rep-seqs-1.qzv
qiime tools view rep-seqs-2.qzv
qiime tools view rep-seqs-3.qzv

# MERGING TABLES AND SEQUENCES

cd $WORK/16S-Analysis/MergedRuns

qiime feature-table merge \
--i-tables filtered-table-1.qza \
--i-tables filtered-table-2.qza \
--i-tables filtered-table-3.qza \
--o-merged-table merged_table.qza

qiime feature-table merge-seqs \
--i-data sequence-hits-1.qza \
--i-data sequence-hits-2.qza \
--i-data sequence-hits-3.qza \
--o-merged-data merged_sequences.qza

# SUMMARISING MERGED FILES

cd $WORK/16S-Analysis/MergedRuns

qiime feature-table summarize \
--i-table merged_table.qza \
--o-visualization merged_table.qzv \
--m-sample-metadata-file mapping.tsv

qiime feature-table tabulate-seqs \
--i-data merged_sequences.qza \
--o-visualization merged_sequences.qzv

# Output visualisations

qiime tools view merged_table.qzv
qiime tools view merged_sequences.qzv

# TAXONOMIC CLASSIFICATION

qiime feature-classifier classify-sklearn \
--i-reads merged_sequences.qza \
--i-classifier gg-13-8-99-515-806-nb-classifier.qza \
--o-classification taxonomy-unfiltered-merged-final.qza
```

```
qiime metadata tabulate \
--m-input-file taxonomy-unfiltered-merged-final.qza \
--o-visualization taxonomy-unfiltered-merged-final.qzv

qiime tools view taxonomy-unfiltered-merged-final.qzv

qiime taxa barplot \
--i-table merged_table.qza \
--i-taxonomy taxonomy-unfiltered-merged-final.qza \
--m-metadata-file mapping.tsv \
--o-visualization taxa-bar-plots-unfiltered-merged-final.qzv

# Output visualisation:

qiime tools view taxa-bar-plots-unfiltered-merged-final.qzv


# GENERATE A TREE FOR PHYLOGENETIC DIVERSITY

qiime alignment mafft \
--i-sequences unfiltered-rep-seqs.qza \
--o-alignment unfiltered-aligned-seqs.qza

qiime alignment mask \
--i-alignment unfiltered-aligned-seqs.qza \
--o-masked-alignment unfiltered-aligned-masked-seqs.qza

qiime phylogeny fasttree \
--i-alignment unfiltered-aligned-masked-seqs.qza \
--o-tree unfiltered-unrooted-tree.qza

qiime phylogeny midpoint-root \
--i-tree unfiltered-unrooted-tree.qza \
--o-rooted-tree unfiltered-rooted-tree.qza

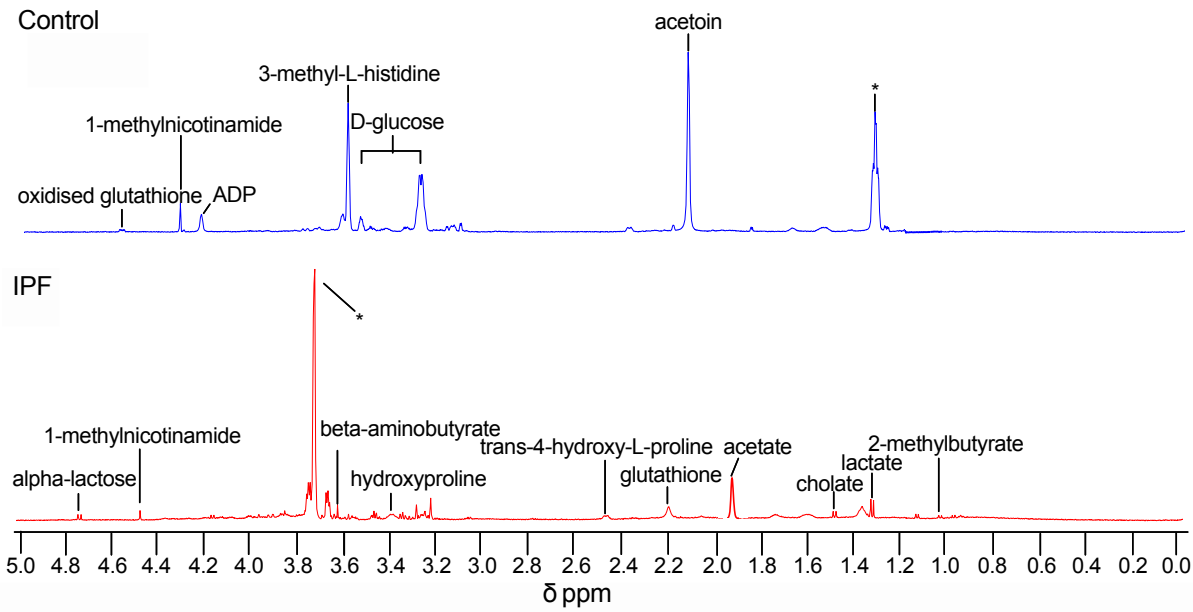
# EXPORT

qiime tools export \
--input-path unfiltered-rooted-tree.qza \
--output-path unfiltered-exported-tree

qiime tools export \
--input-path unfiltered-table.qza \
--output-path unfiltered-final-table


qiime tools export \
--input-path taxonomy-unfiltered-merged-final.qza \
--output-path unfiltered-exported-taxonomy

# IMPORT FILES ("unfiltered-table.qza","unfiltered-rooted-tree.qza","taxonomy-unfiltered-
merged-final.qza",  "mapping.tsv",  "CH3_mapping_diagnosis.csv")  INTO  R  STUDIO  FOR
DOWNSTREAM PART OF ANALYSIS
```



Appendix 2 Representative 600 MHz ^{1}H nuclear magnetic resonance (NMR) spectra of bronchoalveolar lavage (BAL) samples obtained from healthy controls and subjects with idiopathic pulmonary fibrosis (IPF).

Each identified peak is labelled accordingly. Unidentified metabolites are marked (*).

Appendix 3. Spearman's rank correlations (*rho*) between the metabolites detected in the bronchoalveolar lavage (BAL) of idiopathic pulmonary fibrosis (IPF) subjects and BAL cell differentials.

Comparison	Spearman's *rho*	*P*-value
Macrophages vs oxidised glutathione	0.028	NS
Macrophages vs propionate	0.041	NS
Macrophages vs ethanol	0.005	NS
Macrophages vs butyrate	-0.006	NS
Macrophages vs lactate	-0.026	NS
Macrophages vs acetone	0.048	NS
Macrophages vs glucose	-0.052	NS
Macrophages vs alanine	-0.058	NS
Macrophages vs choline	0.031	NS
Macrophages vs glycine	-0.091	NS
Macrophages vs scyllo-inositol	0.087	0.22
Macrophages vs 3-ketobutyrate	0.020	NS
Macrophages vs acetate	-0.070	NS
Macrophages vs sarcosine	-0.0008	NS
Macrophages vs dimethyl sulfone	-0.081	NS
Macrophages vs formate	-0.080	NS

Lymphocytes vs oxidised glutathione	-0.130	NS
Lymphocytes vs propionate	-0.130	NS
Lymphocytes vs ethanol	0.029	NS
Lymphocytes vs butyrate	-0.015	NS
Lymphocytes vs lactate	-0.013	NS
Lymphocytes vs acetone	-0.011	NS
Lymphocytes vs glucose	-0.022	NS
Lymphocytes vs alanine	0.016	NS
Lymphocytes vs choline	0.007	NS
Lymphocytes vs glycine	-0.007	NS
Lymphocytes vs scyllo-inositol	0.037	NS
Lymphocytes vs 3-ketobutyrate	-0.092	NS
Lymphocytes vs acetate	0.100	NS
Lymphocytes vs sarcosine	-0.111	NS
Lymphocytes vs dimethyl sulfone	0.085	NS
Lymphocytes vs formate	0.091	NS
Neutrophils vs oxidised glutathione	0.034	NS
Neutrophils vs propionate	0.025	NS
Neutrophils vs ethanol	-0.049	NS
Neutrophils vs butyrate	0.032	NS
Neutrophils vs lactate	0.075	NS
Neutrophils vs acetone	-0.053	NS
Neutrophils vs glucose	0.120	NS
Neutrophils vs alanine	0.093	NS
Neutrophils vs choline	-0.021	NS
Neutrophils vs glycine	0.140	NS
Neutrophils vs scyllo-inositol	-0.048	NS
Neutrophils vs 3-ketobutyrate	0.033	NS
Neutrophils vs acetate	0.093	NS
Neutrophils vs sarcosine	0.097	NS
Neutrophils vs dimethyl sulfone	0.099	NS
Neutrophils vs formate	0.069	NS

Appendix 4. Spearman's rank correlations (*rho*) between the metabolites detected in the bronchoalveolar lavage (BAL) of idiopathic pulmonary fibrosis (IPF) subjects and healthy controls by ^{1}H nuclear magnetic resonance (^{1}H NMR) spectroscopy and gas-chromatography mass spectrometry (GC-MS).

Diagnosis	Comparison	Spearman's *rho*	*P*-value
Controls	Acetate vs Acetate	0.81	$P < 0.01$
	Propionate vs Propionate	0.83	$P < 0.01$
	Valerate vs Valerate	0.78	$P < 0.05$
	Lactate vs Lactate	0.79	$P < 0.01$
	Butyrate vs Butyrate	0.85	$P < 0.001$
IPF	Acetate vs Acetate	0.83	$P < 0.01$
	Propionate vs Propionate	0.89	$P < 0.001$
	Valerate vs Valerate	0.77	$P < 0.05$
	Lactate vs Lactate	0.81	$P < 0.01$
	Butyrate vs Butyrate	0.88	$P < 0.01$

LUNG SCIENCE CONFERENCE
IDIOPATHIC PULMONARY FIBROSIS

The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis

Rachele Invernizzi[1] and Philip L. Molyneaux[1,2]

Affiliations: [1]Fibrosis Research Group, National Heart and Lung Institute, Imperial College, London, UK. [2]Dept of Respiratory Medicine, Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK.

Correspondence: Philip L. Molyneaux, Fibrosis Research Group, National Heart and Lung Institute, Imperial College, London, SW7 2AZ, UK. E-mail: p.molyneaux@imperial.ac.uk.

@ERSpublications
Microbial dysbosis plays a role in the pathogenesis of IPF and further changes occur during acute exacerbations despite exclusion of overt infection. http://bit.ly/2MQkfBN

Cite this article as: Invernizzi R, Molyneaux PL. The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2019; 28: 190045 [https://doi.org/10.1183/16000617.0045-2019].

ABSTRACT Idiopathic pulmonary fibrosis (IPF) arises in genetically susceptible individuals as a result of an aberrant wound-healing response following repetitive alveolar injury. The clinical course of the disease remains both variable and unpredictable with periods of more rapid decline, termed acute exacerbation of IPF (AE-IPF), often punctuating the disease trajectory. Exacerbations carry a significant morbidity and mortality, and their exact pathogenesis remains unclear. Given the emerging evidence that disruption and alteration in the lung microbiome plays a role in the pathogenesis and progression of IPF, this review discusses the current knowledge of the contribution of infection and the respiratory microbiome to AE-IPF.

Introduction

Although the pathogenesis of idiopathic pulmonary fibrosis (IPF) remains poorly understood the condition is thought to arise in genetically susceptible ageing individuals as a consequence of repeated microinjuries to the alveolar epithelium by environmental triggers [1]. Over time, this process results in an aberrant wound-healing response, which leads to fibrosis and causes the progressive structural destruction of the lung [2, 3]. This suggests both environmental and host factors play a role in IPF, likely with bidirectional interaction between the two. The clinical course of the disease is both variable and unpredictable with significant heterogeneity among individual patients [4]. Whatever trajectory the disease adopts it can be punctuated by unpredictable periods of acute deterioration; episodes termed acute exacerbations of IPF (AE-IPF). These are defined as acute, clinically significant respiratory deterioration with no identifiable cause, characterised by new, widespread alveolar abnormalities [5]. The exact incidence of acute exacerbations is unclear but seems to range between 4–20% per year with these episodes conferring a substantial morbidity and mortality (often >50%) [6, 7]. Known triggers for AE-IPF include surgical lung biopsy, lung resection and drugs [8–11]. Nonetheless, our current understanding of the pathophysiology of AE-IPF is limited and it remains unclear whether AE-IPF represent an accelerated

Submitted article, peer reviewed.

Received: April 24 2019 | Accepted after revision: June 16 2019

Copyright ©ERS 2019. This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.

phase of the underlying fibrotic condition or a response to an occult trigger. Given the growing evidence for the role of infection in the pathogenesis and progression of IPF [12, 13], in this review we seek to explore the potential contribution of infection and the microbiome in acute exacerbations.

The lung microbiome: a complex and dynamic community of microbes

The epithelial surfaces of the human respiratory tract have historically been described as sterile. More recently, novel culture-independent techniques have debunked this notion by demonstrating that the respiratory tract harbours a complex and dynamic community of microbes which play a role both in health and disease [14]. High-throughput sequencing of the 16S rRNA gene is the most commonly used approach to study bacterial communities. Groups of bacteria sharing similar gene sequences are clustered together and identified by comparison to 16S rRNA reference databases. The commensal, symbiotic and pathogenic microorganisms which inhabit the lower airways are collectively defined as the "lung microbiome". In health, the primary routes of microbial immigration to the lungs include microaspiration, inhalation of air-borne bacteria and direct dispersion along mucosal surfaces [15]. Of these, microaspiration is probably the dominant mechanism, as the composition of the oral microbiome most closely resembles the lung microbiome. It is also expected to be the most relevant to IPF given the strong association with reflux [16]. Elimination is achieved through mucociliary clearance, coughing, and innate and adaptive immune defences which selectively recognise, kill and clear the microbes [17]. The quantity and composition of the microbiome is influenced by the balance between the rate of immigration and clearance. The discovery of the lung microbiome has prompted the exploration of its role in a variety of respiratory conditions, including IPF. These molecular techniques have not only enabled researchers to characterise niche-specific microbial communities in the respiratory tract of patients and healthy individuals, but have also demonstrated associations which point towards the interaction between the microbiome and the host as key in the aetiology and progression of lung disease [15, 18–20].

The role of infection and the microbiome in AE-IPF

Historically, research investigating the role of infections in the pathogenesis of IPF and as triggers of AE-IPF has focused primarily on viruses. When considering the role of infection in AE-IPF, it is important to remember these are events that, prior to the latest revised diagnostic criteria [5], specifically required the exclusion of any infective trigger [21]. Therefore, the following findings are all in cohorts of patients where overt clinical infection has been excluded.

It has been hypothesised that chronic viral infection may act as a persistent antigenic stimulus or co-factor in susceptible hosts, triggering fibrosis, and there is growing evidence both from human and animal studies to support this [22, 23]. The human herpes virus (HHV) family has received the most attention as either an aetiologic or exacerbating agent of IPF [24, 25], but a number of other viral infections have been consistently associated with IPF and AE-IPF (table 1).

There is growing evidence, albeit associative, suggesting that subclinical or occult viral infections may play a pathogenic role in the development of AE-IPF [32]. Molecular analysis of viral particles and bronchoalveolar lavage (BAL) revealed infection in 19 out of 43 cases of AE-IPF, including 12 cases of torque teno virus (TTV) in the exacerbation cohort with no viruses detected in the stable IPF group [26]. Supporting this observation, IPF subjects who were found to be seropositive for TTV-DNA demonstrated a worse survival in an independent cohort [30]. More common respiratory viral pathogens have also been detected using molecular techniques during exacerbations, including respiratory syncytial virus and cytomegalovirus [27]. Having excluded infection clinically, a cohort of 37 IPF patients experiencing an exacerbation underwent surgical lung biopsy and over one-third exhibited evidence of viral infection based on immunohistochemistry or microarray [28]. More recently, viral sequences were profiled in nasopharyngeal swabs of 30 stable IPF and 30 AE-IPF patients which revealed higher virus-positive rates in the AE-IPF cohort (60%) compared to stable disease (43%) [29]. HHV was the most prominent virus in the AE-IPF group as well as influenza A [29]. As AE-IPF have been reported following influenza A vaccination, its presence suggests that cold-associated infection may potentially trigger these events [31]. Overall these observational studies highlight a potential role for viral infections, although the supporting mechanistic role for airborne viruses is insufficient to prove their causal role in AE-IPF.

While the role of viruses has been extensively explored, the long-held, but incorrect, theory of sterility of the lung outside of clinical infection means historically little work has evaluated the role of bacteria in AE-IPF. Using quantitative methods, the first study to investigate the presence of bacterial growth in the lower airways of 22 stable IPF patients confirmed the presence of known pathogens including *Haemophilus*, *Pseudomonas* and *Streptococcus* in 36% of the cases [33]. More recently, molecular culture-independent techniques have been used to characterise the microbial composition in the lower

TABLE 1 Summary of studies linking viruses with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis (IPF)

Virus	Main conclusions	[Ref.]
HHV	EBV detected in serum of 12 out of 13 subjects with IPF but not in patients with other forms of ILD	[22]
	MHV-68 triggers an exaggerated fibrotic response in mice	[23]
	Increased incidence of EBV in BAL and lung biopsies of IPF subjects compared to controls	[24, 25]
	EBV detected in BAL of two out of 43 AE-IPF subjects	[26]
	CMV detected in BAL of two out of 43 AE-IPF subjects	[27]
	A total of 38% of AE-IPF subjects exhibited evidence of CMV infection	[28]
	HHV detected in nasopharyngeal swabs of 15 out of 30 AE-IPF subjects and four out of 30 individuals with stable IPF	[29]
TTV	TTV detected in BAL of 12 out of 43 AE-IPF subjects	[26]
	No evidence of TTV in BAL of stable IPF subjects	[30]
	Increased mortality in IPF subjects with presence of TTV-DNA in serum compared to IPF subjects with no TTV-DNA	
Influenza A	Influenza A detected in nasopharyngeal swabs of 12 out of 30 AE-IPF subjects but not in individuals with stable IPF	[29]
	A case of AE-IPF was reported following pandemic influenza A vaccination	[31]

HHV: human herpes virus; TTV: torque teno virus; EBV: Epstein–Barr virus; ILD: interstitial lung disease; MHV-68: murine γherpes virus-68; BAL: bronchoalveolar lavage; AE-IPF: acute exacerbation of IPF; CMV: cytomegalovirus.

airways of IPF patients. Recent studies suggest that microbial dysbiosis may be linked to disease outcome (table 2).

The Correlating Outcomes with Biochemical Markers to Estimate Time-Progression (COMET) study was the first to truly evaluate the lung microbiome in IPF patients [34]. The authors demonstrated an increased abundance of either *Streptococcus* or *Staphylococcus* was associated with a significant reduction in progression-free survival time [34]. The enrichment of these two bacteria was observed in less than half of the cohort; therefore, it remains unlikely that these organisms alone can explain the disease pathogenesis. Consistent with the COMET study, Molyneaux *et al.*[36] demonstrated that the lower airways of patients with IPF were not sterile and indeed were more likely to harbour potentially pathogenic *Haemophilus*, *Neisseria* and *Streptococcus* species than healthy controls. The authors also

TABLE 2 Summary of studies linking the respiratory microbiome with the pathogenesis, progression and acute exacerbation of idiopathic pulmonary fibrosis (IPF)

Diagnosis	Main conclusions	[Ref.]
IPF	Positive BAL cultures in eight out of 22 stable IPF subjects: *Haemophilus influenzae* (n=2), *Haemophilus parainfluenzae* (n=2), *Moraxella catarrhalis* (n=1), *Pseudomonas aeruginosa* (n=1), *Proteus mirabilis* (n=1), *Streptococcus pneumonia* (n=1)	[33]
	Increased abundance of *Streptococcus* OTU1345 and *Staphylococcus* OTU1348 is associated with a significant reduction in progression-free survival in IPF	[34]
	Streptococcus pneumoniae triggers progression of pulmonary fibrosis through pneumolysin in two different mouse models	[35]
	Increased bacterial burden in IPF subjects compared with COPD and healthy controls	[36]
	Higher bacterial burden at the time of diagnosis predicts disease progression in IPF	[37]
	Germ-free mice protected from mortality following bleomycin exposure	
AE-IPF	Four-fold increase in bacterial burden in AE-IPF subjects compared to stable IPF	[38]
	Increased abundance of *Campylobacter* and *Stenotrophomonas* and decreased abundance of *Veillonella* in AE-IPF compared to stable IPF	
	Positive sputum cultures in nine out of 48 AE-IPF subjects: *Klebsiella pneumoniae* (n=2), *Mycobacterium tuberculosis* (n=4), *Pseudomonas aeruginosa* (n=1), *Loffi Acinetobacter* (n=1), other (n=1)	[29]

AE-IPF: acute exacerbation of IPF; BAL: bronchoalveolar lavage.

identified the presence of a higher overall bacterial burden in IPF, which was associated with a reduced progression-free survival time in subjects with IPF. Having characterised the microbiome in stable IPF, the same authors then set out to examine for potential changes during acute exacerbations. Using a cohort of matched stable and exacerbation patients who underwent bronchoscopy, the authors were able to demonstrate a further increase in bacterial burden during an exacerbation, despite the presence of negative clinical cultures [38]. Subjects experiencing an AE-IPF exhibited a marked change in the respiratory microbiome with an increase in *Campylobacter* and *Stenotrophomonas* species compared to stable disease. Both of these organisms are potential respiratory pathogens but the presence of *Campylobacter* also leads the authors to suggest their findings support the role of reflux and aspiration with AE-IPF. One weakness of the study was that there were only two paired samples from the same individuals when stable and experiencing an exacerbation. The stark changes in these two pairs of samples is intriguing, especially as in one individual there was a marked outgrowth of streptococcal species from the baseline microbiota which then dominated the exacerbation state. This result is concordant with findings from two distinct mouse models, where infection with *Streptococcus pneumoniae* induced exacerbation of established lung fibrosis [35].

More recently, Weng *et al.* [29] have employed molecular techniques to examine sputum cultures of 170 patients experiencing an exacerbation. They detected 38 different bacterial strains with gram-negative bacteria accounting for 89% of pathogens: *Klebsiella pneumonia* (26%), *Mycobacterium tuberculosis* (21%), and *Acinetobacter baumannii* (10%) dominated the sputum samples [29]. Together these studies demonstrate that alterations in the respiratory microbiome occur during acute exacerbations but they do not provide any functional or mechanistic insights. In a recent study, O'Dwyer *et al.* [37] tried to overcome this issue in an attempt to move from correlation to causation. A germ-free mouse model of fibrosis was used to explore the role of bacteria on lung inflammation and fibrinogenesis. Following bleomycin exposure, germ-free mice were protected from mortality, even though they exhibited similar severity of pulmonary fibrosis when compared to conventional mice [37]. The authors postulate that this may reflect the clinical observation that subjects with IPF not only die from progressive fibrosis but also from distinct inflammatory causes of acute-on-chronic respiratory failure including respiratory infections and AE-IPF. While there are several limitations of animal models in IPF, collectively, this study shows how preclinical germ-free models are a useful tool to investigate potential mechanisms of host-microbiota interactions and provides the first causal evidence that the microbiome participates in the pathogenesis and mortality of fibrotic lung disease.

Conclusions and future directions

The role of the respiratory microbiome in the pathogenesis and progression of IPF is just beginning to be elucidated. Clinically, acute exacerbations are important events in the natural course of IPF which carry a significant morbidity and mortality. While there is a paucity of information regarding the aetiology, pathophysiology and clinical management of acute exacerbations, an increasing number of findings indicate that infection, both viral and bacterial might be involved. Recent advances in sequencing technologies have allowed the use of molecular microbial technologies to characterise the respiratory microbiota in these patients demonstrating a clear change in the respiratory microbiome, as well as an increased bacterial burden in BAL compared to stable disease. However, the microbiome has not been fully characterised and further research should also investigate organisms other than bacteria and viruses, including fungi. It is unclear whether a higher bacterial load and an altered microbiome during an acute exacerbation reflect an active infection, increased aspiration or occur as a result of widespread diffuse alveolar damage. As sequencing DNA from BAL provides a snapshot in time of the microbial diversity, future research should incorporate longitudinal sampling before, during and after an AE-IPF. Furthermore, there is a pressing need to move from correlation to causation [39]. In order to do so, functional studies examining host-microbiome interactions as well as better models of AE-IPF are needed, given existing models fail to fully recapitulate physiologic findings in IPF. Current consensus is that all patients experiencing an exacerbation should receive broad spectrum empirical antibiotics even if no overt infection can be identified [40]. While this is likely driven by the use of immunosuppressive agents, the molecular microbial data may support this empiric therapy. Indeed, there is already some evidence that azithromycin may be beneficial in AE-IPF [41]. Ultimately, understanding the role of the microbiome in AE-IPF will allow its manipulation and may provide an opportunity for targeted therapeutic intervention.

Conflict of interest: R. Invernizzi has nothing to disclose. P.L. Molyneaux has, *via* his institution, received speaker's fees from Roche and Boehringer Ingelheim.

References

1 Maher TM, Wells AU, Laurent GJ. Idiopathic pulmonary fibrosis: multiple causes and multiple mechanisms? *Eur Respir J* 2007; 30: 835–839.
2 Daccord C, Maher TM. Recent advances in understanding idiopathic pulmonary fibrosis. *F1000Res* 2016; 5: 1046.

3 Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med* 2018; 378: 1811–1823.
4 Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 431–440.
5 Collard HR, Ryerson CJ, Corte TJ, *et al.* Acute exacerbation of idiopathic pulmonary fibrosis. An International Working Group Report. *Am J Respir Crit Care Med* 2016; 194: 265–275.
6 Moua T, Westerly BD, Dulohery MM, *et al.* Patients with fibrotic interstitial lung disease hospitalized for acute respiratory worsening: a large cohort analysis. *Chest* 2016; 149: 1205–1214.
7 Ryerson CJ, Collard HR. Acute exacerbations complicating interstitial lung disease. *Curr Opin Pulm Med* 2014; 20: 436–441.
8 Kreider ME, Hansen-Flaschen J, Ahmad NN, *et al.* Complications of video-assisted thoracoscopic lung biopsy in patients with interstitial lung disease. *Ann Thorac Surg* 2007; 83: 1140–1144.
9 Bando M, Ohno S, Hosono T, *et al.* Risk of acute exacerbation after video-assisted thoracoscopic lung biopsy for interstitial lung disease. *J Bronchology Interv Pulmonol* 2009; 16: 229–235.
10 Sakamoto S, Homma S, Mun M, *et al.* Acute exacerbation of idiopathic interstitial pneumonia following lung surgery in 3 of 68 consecutive patients: a retrospective study. *Intern Med* 2011; 50: 77–85.
11 Kondoh Y, Cottin V, Brown KK. Recent lessons learned in the management of acute exacerbation of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2017; 26: 170050.
12 Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2013; 22: 376–381.
13 Hewitt RJ, Molyneaux PL. The respiratory microbiome in idiopathic pulmonary fibrosis. *Ann Transl Med* 2017; 5: 250.
14 Dickson RP, Erb-Downward JR, Martinez FJ, *et al.* The microbiome and the respiratory tract. *Annu Rev Physiol* 2016; 78: 481–504.
15 Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. *Expert Rev Respir Med* 2013; 7: 245–257.
16 Mackintosh JA, Desai SR, Adamali H, *et al.* In patients with idiopathic pulmonary fibrosis the presence of hiatus hernia is associated with disease progression and mortality. *Eur Respir J* 2019; 53: 1802412.
17 Dickson RP, Erb-Downward JR, Freeman CM, *et al.* Bacterial topography of the healthy human lower respiratory tract. *MBio* 2017; 8: e02287.
18 Hilty M, Burke C, Pedro H, *et al.* Disordered microbial communities in asthmatic airways. *PLoS One* 2010; 5: e8578.
19 Molyneaux PL, Mallia P, Cox MJ, *et al.* Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; 188: 1224–1231.
20 Molyneaux PL, Willis-Owen SAG, Cox MJ, *et al.* Host-microbial interactions in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017; 195: 1640–1650.
21 Collard HR, Moore BB, Flaherty KR, *et al.* Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2007; 176: 636–643.
22 Vergnon JM, Vincent M, de The G, *et al.* Cryptogenic fibrosing alveolitis and Epstein–Barr virus: an association? *Lancet* 1984; 2: 768–771.
23 Stoolman JS, Vannella KM, Coomes SM, *et al.* Latent infection by γherpesvirus stimulates profibrotic mediator release from multiple cell types. *Am J Physiol Lung Cell Mol Physiol* 2011; 300: L274–L285.
24 Manika K, Alexiou-Daniel S, Papakosta D, *et al.* Epstein–Barr virus DNA in bronchoalveolar lavage fluid from patients with idiopathic pulmonary fibrosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2007; 24: 134–140.
25 Tang YW, Johnson JE, Browning PJ, *et al.* Herpesvirus DNA is consistently detected in lungs of patients with idiopathic pulmonary fibrosis. *J Clin Microbiol* 2003; 41: 2633–2640.
26 Wootton SC, Kim DS, Kondoh Y, *et al.* Viral infection in acute exacerbation of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 1698–1702.
27 Ushiki A, Yamazaki Y, Hama M, *et al.* Viral infections in patients with an acute exacerbation of idiopathic interstitial pneumonia. *Respir Investig* 2014; 52: 65–70.
28 Santos GC, Parra ER, Stegun FW, *et al.* Immunohistochemical detection of virus through its nuclear cytopathic effect in idiopathic interstitial pneumonia other than acute exacerbation. *Braz J Med Biol Res* 2013; 46: 985–992.
29 Weng D, Chen XQ, Qiu H, *et al.* The role of infection in acute exacerbation of idiopathic pulmonary fibrosis. *Mediators Inflamm* 2019; 2019: 5160694.
30 Bando M, Ohno S, Oshikawa K, *et al.* Infection of TT virus in patients with idiopathic pulmonary fibrosis. *Respir Med* 2001; 95: 935–942.
31 Umeda Y, Morikawa M, Anzai M, *et al.* Acute exacerbation of idiopathic pulmonary fibrosis after pandemic influenza A (H1N1) vaccination. *Intern Med* 2010; 49: 2333–2336.
32 Vannella KM, Moore BB. Viruses as co-factors for the initiation or exacerbation of lung fibrosis. *Fibrogenesis Tissue Repair* 2008; 1: 2.
33 Richter AG, Stockley RA, Harper L, *et al.* Pulmonary infection in Wegener granulomatosis and idiopathic pulmonary fibrosis. *Thorax* 2009; 64: 692–697.
34 Han MK, Zhou Y, Murray S, *et al.* Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014; 2: 548–556.
35 Knippenberg S, Ueberberg B, Maus R, *et al. Streptococcus pneumoniae* triggers progression of pulmonary fibrosis through pneumolysin. *Thorax* 2015; 70: 636–646.
36 Molyneaux PL, Cox MJ, Willis-Owen SA, *et al.* The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014; 190: 906–913.
37 O'Dwyer DN, Ashley SL, Gurczynski SJ, *et al.* Lung microbiota contribute to pulmonary inflammation and disease progression in pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 199: 1127–1138.
38 Molyneaux PL, Cox MJ, Wells AU, *et al.* Changes in the respiratory microbiome during acute exacerbations of idiopathic pulmonary fibrosis. *Respir Res* 2017; 18: 29.
39 Segal LN, Molyneaux PL. The challenging road of moving from association to causation for microbiome research in IPF. *Am J Respir Crit Care Med* 2019; 199: 1054–1056.

40 Maher TM, Whyte MK, Hoyles RK, *et al.* Development of a consensus statement for the definition, diagnosis, and treatment of acute exacerbations of idiopathic pulmonary fibrosis using the Delphi technique. *Adv Ther* 2015; 32: 929–943.
41 Macaluso C, Furcada JM, Alzaher O, *et al.* The potential impact of azithromycin in idiopathic pulmonary fibrosis. *Eur Respir J* 2018; 53: 1800628.

Immunology
The Journal of cells, molecules, systems and technologies

British Society for immunology

doi:10.1111/imm.13195

IMMUNOLOGY REVIEW ARTICLE

REVIEW SERIES: BARRIER IMMUNITY
Series Editors: James A. Harker and Laura J. Pallett

Respiratory microbiome and epithelial interactions shape immunity in the lungs

OTHER ARTICLES PUBLISHED IN THIS REVIEW SERIES
Immunological fortification at our barrier organs: Protecting us as we age. Immunology 2020, 160: 103-105.
Immune responses in the human female reproductive tract. Immunology 2020, 160: 106-155.
Skin barrier immunity and ageing. Immunology 2020, 160: 116-125.
Origin and ontogeny of lung macrophages: from mice to humans. Immunology 2020, 160: 126-138.
Regulation of barrier immunity and homeostasis by integrin-mediated transforming growth factor β activation. Immunology 2020, 160: 139-148.
The developing immune network in human prenatal skin. Immunology 2020, 160: 149-156.
The liver as an immunological barrier redefined by single-cell analysis. Immunology 2020, 160: 157-170.

Rachele Invernizzi,[1] Clare M. Lloyd[1] and Philip L. Molyneaux[1,2]

[1]*Inflammation, Repair and Development Section, National Heart and Lung Institute, Imperial College, London, and* [2]*Department of Respiratory Medicine, Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK*

doi:10.1111/imm.13195
Received 14 January 2020; revised 17 February 2020; accepted 17 February 2020.
Correspondence: Philip L. Molyneaux, Inflammation, Repair and Development Section, National Heart and Lung Institute, Imperial College, London SW7 2AZ, UK.
Email: p.molyneaux@imperial.ac.uk
Senior author: Philip L. Molyneaux

Summary

The airway epithelium represents a physical barrier to the external environment acting as the first line of defence against potentially harmful environmental stimuli including microbes and allergens. However, lung epithelial cells are increasingly recognized as active effectors of microbial defence, contributing to both innate and adaptive immune function in the lower respiratory tract. These cells express an ample repertoire of pattern recognition receptors with specificity for conserved microbial and host motifs. Modern molecular techniques have uncovered the complexity of the lower respiratory tract microbiome. The interaction between the microbiota and the airway epithelium is key to understanding how stable immune homeostasis is maintained. Loss of epithelial integrity following exposure to infection can result in the onset of inflammation in susceptible individuals and may culminate in lung disease. Here we discuss the current knowledge regarding the molecular and cellular mechanisms by which the pulmonary epithelium interacts with the lung microbiome in shaping immunity in the lung. Specifically, we focus on the interactions between the lung microbiome and the cells of the conducting airways in modulating immune cell regulation, and how defects in barrier structure and function may culminate in lung disease. Understanding these interactions is fundamental in the search for more effective therapies for respiratory diseases.

Keywords: immunity; lungs; mucosal immunology; respiratory epithelium; respiratory microbiome.

The epithelial barrier

The human respiratory tract is the main portal of entry for the continuous immigration and elimination of a vast array of airborne microorganisms and particles, including viruses, bacteria and fungi. The lungs are continually exposed to a multitude of microorganisms, some of which are able to persist and colonize the respiratory

Abbreviations: AEC1, alveolar epithelial type 1 cells; AEC2, alveolar epithelial type 2 cells; BAL, bronchoalveolar lavage; CF, cystic fibrosis; CLRs, C-type lectin receptors; COPD, chronic obstructive pulmonary disease; DKK, Dikkopf; EGFR, epidermal growth factor receptor; GORD, gastro-oesophageal reflux disease; HHV, human herpes virus; ICAM-1, intracellular adhesion molecule 1; IPF, idiopathic pulmonary fibrosis; KRM, Kreme; Man-6-P, mannose-6-phosphate; NLRs, NOD-like receptors; NOD, nucleotide-binding and oligomerization domain; PAF, platelet-activating factor; PAMPs, pathogen-associated molecular patterns; PARs, protease-activated receptors; PRR, pattern recognition receptors; RLRs, retinoic acid-inducible gene-I-like receptors; RSV, respiratory syncytial virus; SCFA, short-chain fatty acid; TARC, thymus- and activation-regulated chemokine; TEER, transepithelial electrical resistance; Th2, T helper 2; TIR, toll/IL-1 receptor; TLRs, toll-like receptors; TRIF, TIR-domain-containing adapter-inducing interferon-β; TSLP, thymic stromal lymphopoietin

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

tract. These microbes that reside in the lower airways constitute the lung microbiome, which is distinct in composition from the microbiota observed in the oral and nasal cavities. The surface of the lungs presents a continuous layer of epithelial cells that constitute a physical and biological barrier for inhaled substances and pathogens. The specialized respiratory epithelium is required for maintenance of immune homeostasis, and epithelial dysfunction is involved in the development of many inflammatory disorders of the airways and lungs.[1] The respiratory tract is a complex organ system, and its primary function is the exchange of oxygen and carbon dioxide. It is divided into the upper respiratory tract including the nasal cavity, pharynx, larynx, and the lower respiratory tract that comprises the conducting airways (trachea and bronchi), the small airways (bronchioles) and the respiratory zone (the alveoli). The adult human airways have a surface area of approximately 70 m^2, which harbours a vast range of bacterial communities, with the highest bacterial load observed in the upper respiratory tract.[2] The respiratory tract spans different anatomical sites, and the types of epithelial cells that comprise these sites are varied both in composition and in structure, reflecting the distinct functions of the airway epithelial cells lining the conducting airways and the alveolar regions of the lungs (Fig. 1).

In the lower respiratory tract, the respiratory epithelium provides a highly effective barrier acting as the first line of defence against potentially harmful environmental stimuli including microbes and allergens.[1] The pseudostratified epithelium of the conducting airways is primarily comprised of multi-ciliated cells, mucus-secreting goblet cells, neuroendocrine cells and basal cells, which secrete surfactant.[1] It constitutes the first site of interaction with inhaled environmental compounds, and is designed to facilitate effective mucociliary clearance of particles and microbes.[3] Multi-ciliated cells have cilia on the apical surface that beat co-ordinately to shuttle inhaled particulates and mucus out of the airways whilst goblet and secretory cells trap inhaled particulates and microorganisms. Basal cells are progenitor cells that act as resident stem cells for the trachea and proximal airways. These cells are capable of self-renewal as well as repopulating the pseudostratified epithelium during homeostasis and after injury.[4] The tracheal region and large airways harbour stromal cells such as interstitial fibroblasts that aid in the regulation of the regenerative response of the airway epithelium after injury.[5] In contrast, the alveolar surfaces in the peripheral lung are lined by flat alveolar epithelial type 1 cells (AEC1) that form a continuous cell layer and are specialized in gas exchange, while cuboidal alveolar epithelial type 2 cells (AEC2) act as progenitor cells and secrete pulmonary surfactant, which reduces surface tension in order to prevent alveolar collapse during respiration.[6,7] A common feature shared by these cell types is the presence of intracellular tight junctions that are localized at the apical surface, and are pivotal for epithelial adhesion and barrier function.[8] These tight junctions ensure the cells adhere together to form a regulated impermeable barrier. This intracellular adhesion complex consists of interconnections of proteins and receptors including, zonular occluden (ZO)-1, -2 and -3, occludins, claudins, and transmembrane junctional adhesion molecules. Tight junctions control paracellular permeability, and immediately below them are the adherens junctions, composed of β-catenin and E-cadherin, which mechanically connect the adjacent cells and initiate proliferation and differentiation.[8,9] The presence and function of these complexes can be influenced by cellular differentiation, exposure to a range of stimuli, such as allergens and pollutants, and may be altered in disease conditions, contributing to immunopathology.[10–13]

The lung microbiota in health and disease

The long-held, but incorrect, theory of sterility of the lung outside of clinical infection means historically little work has evaluated the role of the lung microbiota in chronic lung disease. Advances in high-throughput molecular sequencing technologies have debunked this notion, demonstrating that the epithelial surfaces of the human respiratory tract are colonized by a complex and dynamic microbiota, termed the 'lung microbiome', which plays a role both in health and disease.[14] The most commonly used approach to the study of bacterial communities exploits high-throughput sequencing of amplicons of the 16S rRNA gene, a highly conserved locus in the bacterial genome that contains specific variable regions that can be used for phylogenetic classification.[15] Sequencing studies have demonstrated stark differences in the lung microbiome between health and disease, and it is becoming increasingly apparent that the lung microbiome may not only contribute to explain the pathogenesis of these diseases, but also serve as a prognostic marker and therapeutic target.[16–20]

In health, the lung microbiota has a low density but harbours a prominent diversity of interacting microbiota. The airway microbiota in healthy lungs is dominated by the *Bacteroidetes*, *Firmicutes* and *Proteobacteria*. Prominent genera in the lower airways include *Prevotella*, *Veillonella* and *Streptococcus*.[21] The composition of the lung microbiome is determined by three factors: microbial immigration; microbial elimination; and the relative reproduction rates of its members. The main routes of microbial immigration to the lungs are inhalation of air-borne bacteria, direct dispersion along mucosal surfaces and microaspiration.[14,22] The similarities in community composition between the oral and lung microbiota suggest that microaspiration is likely to be the dominant route of immigration.[14,23] Microbial elimination is an active process that is achieved through mucociliary clearance, cough

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

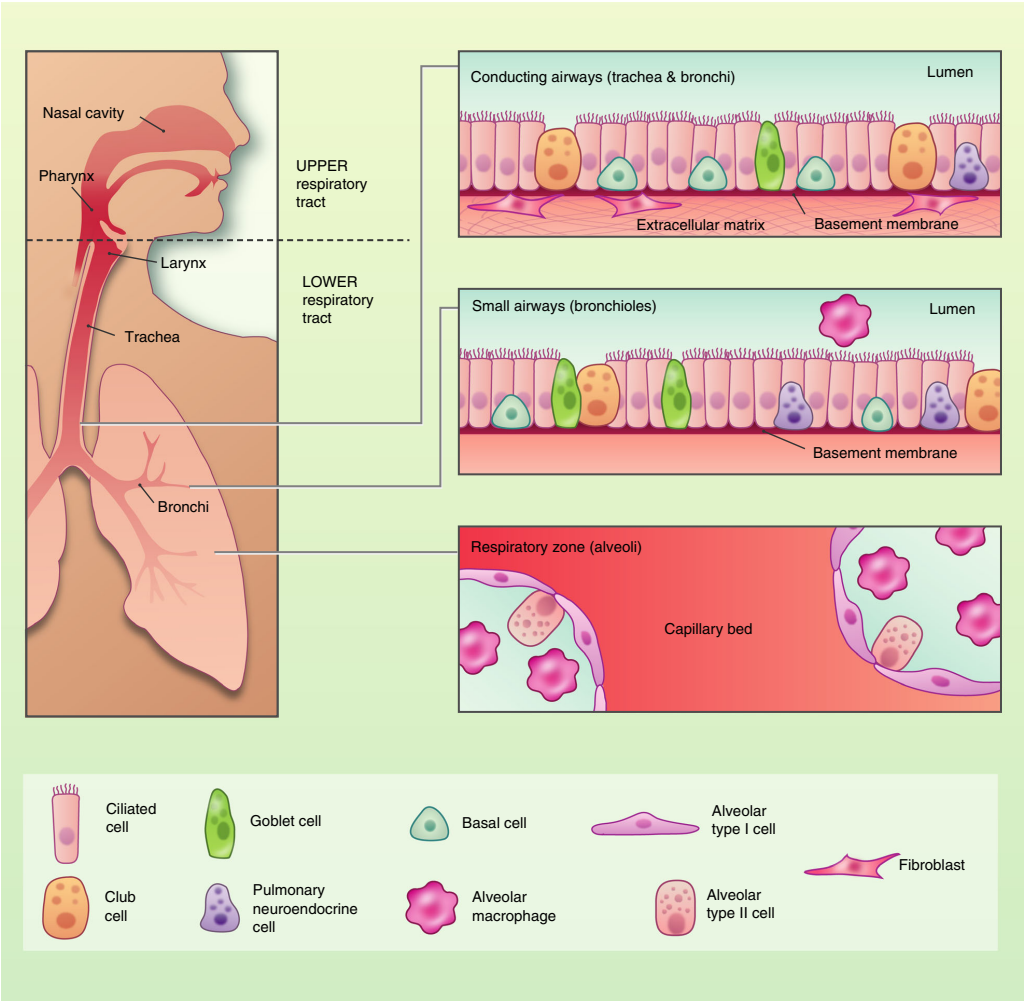


Figure 1. Composition of the respiratory epithelium in health. The human respiratory tract is divided into the upper respiratory tract including the nasal cavity, pharynx, larynx, and the lower respiratory tract that comprises the conducting airways (trachea and bronchi), the small airways (bronchioles) and the respiratory zone (the alveoli). The trachea and large airways are composed of ciliated cells, mucus-secreting goblet cells, neuroendocrine cells and undifferentiated basal cells. Stromal cells including fibroblasts provide extracellular matrix that modulate airway epithelial cell turnover. The alveoli are comprised of alveolar macrophages and alveolar type I and type II cells.

and host immune defences.[24] The environmental conditions necessary for microbial growth within the respiratory tract (e.g. pH, temperature, nutrient availability, oxygen tension and activation of host inflammatory cells) are heterogenous, and considerable regional variation can be observed in a single healthy lung. During lung disease, the balance between immigration and elimination is perturbed, leading to alterations in the lung microbiota, with bacteria exhibiting competitive advantages becoming predominant[14] (Table 1). This overgrowth of bacterial species leads to a decrease in overall richness, and is associated with the progression of chronic lung diseases such as asthma, chronic obstructive pulmonary disease (COPD), cystic fibrosis (CF) and idiopathic pulmonary fibrosis (IPF).[16,17,19,20,25–28] Furthermore, these studies have linked microbial dysbiosis with an increased morbidity and mortality in a multitude of chronic respiratory diseases.

In addition to the bacterial component of the microbiome, there is evidence that other non-bacterial

Table 1. Summary of studies illustrating changes in the respiratory microbiome that occur during chronic lung disease

Disease	Taxonomic changes at the genus level[1]	Sample type	Cohort size	[Ref.]
Asthma	Increased abundance of *Haemophilus*	Bronchoscopy	11	16
	Increased abundance of *Haemophilus*, *Neisseria*, *Fusobacterium* and *Porphyromonas*	Bronchoscopy	42	29
	Increased abundance of *Streptococcus*	Nasopharyngeal	234	30
COPD	Increased abundance of *Lactobacillus*	Lung tissue	24	17
	Increased abundance of *Fusobacteria*, *Leptotrichia* and *Fusobacterium*	BAL	32	31
CF	Increased abundance of *Burkholderia*, *Streptococcus* and *Staphylococcus*	Sputum	23	25
	Increased abundance of *Pseudomonas*, *Staphylococcus*, *Stenotrophomonas* and *Achromobacter*	Sputum	17	32
	Increased abundance of *Staphylococcus*, *Streptococcus* and *Pseudomonas*	BAL	95	33
IPF	Increased abundance of *Haemophilus*, *Streptococcus*, *Neisseria* and *Veillonell*a	BAL	65	19
	Increased abundance of *Staphylococcus* and *Streptococcus*	BAL	55	20
	Increased abundance of *Streptococcus*, *Prevotella*, *Veillonella*, *Haemophilus* and *Pseudomonas*	BAL	35	26

BAL, bronchoalveolar lavage; CF, cystic fibrosis; COPD, chronic obstructive pulmonary disease; IPF, idiopathic pulmonary fibrosis.
[1]Changes in the microbiota in cases compared with controls.

organisms, including viruses and fungi, may play an important role in health and disease. Emerging evidence suggests that the lung virome and mycobiome have a significant impact on clinical outcome of chronic respiratory diseases such as asthma, COPD, CF and IPF.[27,34,35] It is also increasingly recognized that respiratory viruses are a major cause of acute exacerbations of chronic pulmonary diseases, and that the host response to viral infection is dysregulated in these diseases. For example, studies support a mechanistic role for viruses in the initiation and progression of IPF. The human herpes virus (HHV) family, which includes cytomegalovirus, Epstein–Barr virus, HHV-7 and HHV-8, has received considerable attention as either an aetiological or exacerbating agent of IPF.[36] In asthma, infection by respiratory viruses is a major trigger of wheezing in infants as well as exacerbations of asthma in older children. The infectious agents associated with these wheezing events include respiratory syncytial virus (RSV), human rhinovirus, parainfluenza, coronavirus and other viruses.[37] Despite the fact that very little is known about the lung mycobiome compared with the bacterial microbiome, in recent years specific fungi have been identified in respiratory diseases.[38,39] It has also been shown that commensal fungi can affect the immune system and also modulate bacterial communities, thus contributing to the restoration of the bacterial microbiota following antibiotic treatment.[40,41] Although studies into the composition and function of lung viral and fungal communities are gaining momentum, research in this field is still in its infancy, and future mechanistic work is needed to yield novel insights into their role in the pathogenesis of pulmonary diseases.

A homeostatic balance between the lung and its microbiota

Airway epithelial cells are in continuous contact with the environment, and this interaction is a crucial factor in maintaining stable homeostasis.[42] The three essential components that contribute to the barrier function of airway epithelium are: mucociliary escalators that trap and remove inhaled microbes and noxious stimuli from the airways; intracellular tight and adherens junctions that regulate epithelial paracellular permeability; and secreted antimicrobial peptides that kill inhaled pathogens.[1] The physiological features of the respiratory tract are pivotal for maintaining a balance between the lung and its microbiota. The pulmonary epithelium is composed of distinct cell types and is not continuous from the upper respiratory tract to the alveoli.[43] In the large airways, the submucosal gland harbours the mucus and serous cells that are responsible for the production of most airway mucus.[44] Towards the bronchioles, goblet and club cells produce mucus, whereas in the alveoli AEC1 and AEC2 secrete pulmonary surfactant.[45–47] Mucus is a gel composed mainly of water and glycosylated proteins, such as mucins.[48] Mucins are the large glycoprotein constituents of airway mucus.[49] MUC5AC from goblet cells and MUC5B from submucosal glands are the predominant gel-forming mucins in the human airways.[50,51] Mucus is

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

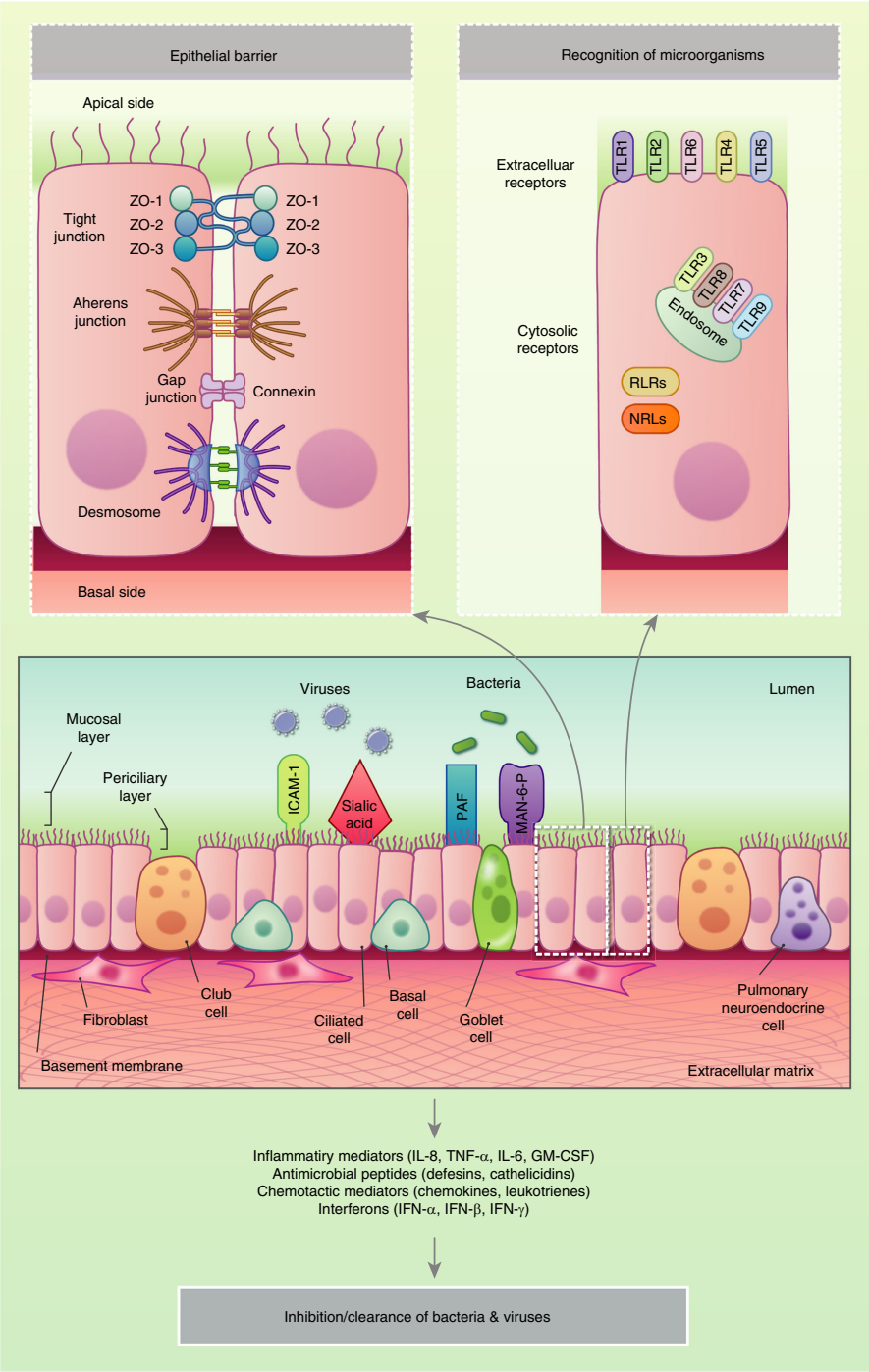
a major ecological niche for the microbiota; it promotes healthy interactions between microorganisms and epithelial cells through the formation of a thin mobile layer that is supported by the periciliary layer covering the cilia.[49,52] In health, the mucus layer provides an effective defence against epithelial injury; however, excessive mucus production is associated with increased susceptibility to viral infections, respiratory allergies, and contributes to obstruction in several respiratory diseases such as pneumonia, asthma, COPD and CF.[53,54] This obstruction hinders the ability of the cilia to transport mucus out of the lungs, and favours the selection of certain pathogenic bacteria that thrive in mucus and establish themselves in this obstructed niche.[55] Furthermore, microaspiration is commonly observed in chronic lung disease, and gastro-oesophageal reflux disease (GORD) has emerged as a frequent co-morbidity in chronic respiratory disorders. Although GORD is typically confined to the lower oesophagus, in some individuals it can be associated with pulmonary microaspiration of gastric contents.[56] The occurrence of aspiration can lead to increased levels of bile acid in the lungs, which can favour the establishment of microbial biofilms, composed of microbial cells that form structured communities embedded in an extracellular matrix.[57] Bacterial and fungal biofilms have emerged as a mechanism of virulence for opportunistic pathogens including *Pseudomonas aeruginosa* and staphylococcal species.[58,59] A hallmark of biofilms is their remarkable antimicrobial tolerance, which allows them to persist in the host and evade the host immune system.[59] Therefore, GORD-derived bile and subsequent biofilm growth on host tissue may contribute to chronic respiratory infection. In addition, direct seeding of bacteria from the intestinal microbiota into the airways may play a role in influencing the interaction between the epithelium and the microbiome.[60]

The intracellular tight and adherens junctions provide fundamental adhesive contacts between neighbouring epithelial cells.[61] Tight junctions are the main regulators of paracellular permeability, whereas the adherens junctions mechanically connect adjacent cells, and initiate the formation and maturation of cell−cell contacts.[8,9] During disease, barrier function is impaired leading to increased epithelial permeability and pathogen entry.[10−13] Upon infection, epithelial cells are able to respond by the release of humoral factors such as antimicrobial peptides that exert broad-spectrum antimicrobial activity, and a number of these are upregulated in response to the microbial dysbiosis seen in disease.[62] Defensins are the most abundant antimicrobial peptides, and comprise α-defensin in neutrophils and β-defensin in epithelial cells. Human β-defensin 1, 2, 3 and 4 are largely expressed on airway epithelial cells and function in airway mucosal defence by targeting a wide range of respiratory pathogens. Their expression in airway epithelia may be constitutive or inducible by bacterial products or proinflammatory cytokines.[63] Nutrient availability in each specific niche will also play a role in the maintenance and selection of the lung microbiota. Host compounds such as mucins, cytokines, defensins and lactoferrins are the nutrient source available for bacteria in the lungs.[1,64] Microbial growth rates can also be influenced by changes in oxygen tension. Within the lungs there is significant regional variation in oxygen tension that can alter population dynamics by affecting bacterial proliferation.[65] In chronic pulmonary diseases, an increased mucus production promotes bacterial growth and leads to the formation of zones with low oxygen concentration and high temperature favouring the selection and maintenance of particular bacteria.[14] Collectively, these differences in the cellular function and the physiological characteristics of the lungs (e.g. temperature, mucus, pulmonary surfactant and oxygen tension) have a major impact on the establishment and persistence of the bacterial communities.

The airway epithelium as a sensor of microbial presence

The airway epithelium not only represents a structural barrier but is also integral for innate host defence. Airway epithelial cells express several pattern recognition receptors (PRRs), and secrete antimicrobial molecules and mucins that aid in the defence against invading pathogens.[3] PRRs are germ-line encoded receptors that recognize conserved pathogen-associated molecular patterns (PAMPs) or cell damage-associated molecular patterns.[66−68] These transmembrane receptors can be activated independently of the adaptive immune response and are classified into four families: the toll-like receptors (TLRs), the cytoplasmic proteins retinoic acid-inducible gene-I-like receptors (RLRs), the NOD-like receptors (NLRs) and the C-type lectin receptors (CLRs).[69]

To date, 10 (TLR1−TLR10) have been identified in the human genome, and are characterized by N-terminal leucine-rich repeats, a transmembrane region and a cytoplasmic Toll/IL-1 receptor (TIR) homology domain that mediates signalling via MyD88 and TIR-domain-containing adapter-inducing interferon-β (TRIF).[70] Toll-like receptors play an important role in pathogen recognition and innate immunity. The main function of TLRs in mammals is to recognize distinct PAMPs, including components of microbial cell walls (TLR1, TLR2, TLR4 and TLR6), flagellin (TLR5), single-stranded or double-stranded viral RNA (TLR3, TLR7 and TLR8) or CpG-containing DNA (TLR9).[71−74] The TLR3, 7, 8 and 9 are specifically localized in endosomes and detect pathogen nucleic acids following endocytosis.[75−79] Lastly, TLR10 is expressed in humans, but the specific ligand for it remains to be elucidated.[80] Signalling through TLRs is key in early inflammatory responses to bacteria through



© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

Figure 2. Response of the airway epithelium to infection. Bacteria and viruses bind to cellular receptors, including intercellular adhesion molecule 1 (ICAM-1), sialic acid residues, platelet-activating factor (PAF) and mannose-6-phosphate (Man-6-P). The pathogens are internalized in airway epithelial cells, which induces the production of innate immune defensins and the stimulation of intracellular and extracellular immune receptors such as the toll-like receptors (TLR), retinoic acid-inducible gene-I-like receptor (RLRs) and NOD-like receptors (NLRs). In healthy individuals, pathogen clearance is achieved following the generation of pro-inflammatory innate and adaptive immune responses.

the activation and expression of host cytokines, chemokines and antimicrobial peptides. TLRs are localized on immune and inflammatory cells; however, given that airway epithelial cells come into contact with several potential pathogens, the expression of TLRs is also relevant to immunity in the airways. Indeed, human bronchial epithelial cells express most TLRs; TLR2, TLR3, TLR5 and TLR6 being the most highly expressed.[81]

The PRR system can be also activated in response to allergens and viruses.[82–84] PRRs within the epithelium are able to recognize allergen-associated danger signals via protease-dependent and protease-independent pathways. Protease-dependent pathways specifically require direct cleavage of cell-surface molecules or the activation of G-protein-coupled protease-activated receptors (PARs).[85–88] RLRs are cytoplasmic receptors that recognize the presence of genomic RNA or double-stranded RNA viruses or double-stranded RNA intermediates from single-stranded RNA viruses. Melanoma differentiation-associated gene-5 is expressed in human bronchial epithelial cells, and is involved in the detection of rhinovirus and other viruses.[89] NLRs are intracellular PRRs characterized by a central nucleotide-binding and oligomerization domain (NOD) and C-terminal leucine-rich repeats.[90] More work is needed to define the role of NOD function in airway epithelial cells. Although the contribution of epithelial cells to antibacterial and antiviral immunity is well established, there is a paucity of information regarding the role of airway epithelial cells in immunity to fungi. CLRs belong to the dectin family and recognize fungal cell wall components. For example, Dectin-1 is expressed in airway epithelial cells, and is responsible for their activation by mycobacterium, *Aspergillus* and allergens such as house dust mite.[91–93] Dectin receptors can be activated by β-glucans via TLR4 signalling, TLR2-dependant signalling, or induction of oxidative stress by intrinsic NADPH oxidase activity.[94–96] Because fungal spores are potent allergens and are associated with allergic and chronic diseases of the respiratory tract such as asthma, the importance of the lung mycobiome should not be overlooked.[97]

Respiratory infections and the initiation of epithelial immune response

Infections play a crucial role in the induction and exacerbation of a multitude of chronic respiratory diseases such as asthma and COPD.[98] Such infections represent a major cause of morbidity and mortality globally, and may involve bacteria (e.g. *Streptococcus pneumoniae*, *Haemophilus influenzae, Chlamydia pneumoniae*) and viruses (e.g. RSV, influenza).[99,100] The initiation of infection typically occurs through the exposure of the respiratory tract to bacterial cells or viral particles (Fig. 2).

During infection, viruses can bind to specific cellular receptors such as intracellular adhesion molecule 1 (ICAM-1) and sialic acid residues, whereas bacteria bind to platelet-activating factor (PAF) and mannose-6 phosphate (Man-6-P) receptors.[101] The interaction between these infectious agents and the epithelial cells induces the stimulation of the PRRs, and the secretion of antimicrobial effector molecules, peptides, enzymes, reactive oxygen species together with a range of chemokines and growth factors. Collectively, these molecules facilitate the recruitment of and communication between immune cells, and contribute to the initiation of the innate immune responses that are pivotal for early control of an infection. In healthy individuals, this response promptly results in pathogen clearance. In contrast, in patients with chronic respiratory diseases these defences may be impaired leading to increased susceptibility to infection that drives inflammatory responses that in turn exacerbate the underlying disease. During chronic disease there is evidence that this barrier function is impaired, with disruption of tight junctions and increased epithelial permeability. This loss of barrier function leads to the entry of pathogens and other particles. Compromised barrier function is intimately linked to triggering of abnormal repair and remodelling.[102] Respiratory viruses, air pollutants and allergens all have the capacity to inflict such damage on the epithelium, either directly due to inherent proteolytic activity or indirectly by binding to PAR-2 expressed on the apical surface of bronchial epithelial cells. For example, *in vitro* studies have shown that infection of the cells with rhinovirus results in the loss of ZO-1 from the junctional complex resulting in decreased transepithelial electrical resistance (TEER) indicative of barrier 'leakiness'. Both coxsackie and rhinovirus affect occludins[103,104] whereas RSV causes disassembly of junctional complexes.[105] In asthma, both decreases in ZO-1 and TEER that correlate with asthma severity[106] or E-cadherin and β-catenin have been shown *in vitro*.[107] Furthermore, reduced expression of occludins and claudins has been shown in COPD.[108] *In vivo*, leakage of proteins such as albumin into the alveolar space has been reported following influenza infection.[109] Viruses, pollutants and allergens are all associated with junction dysfunction and

exacerbated clinical symptoms across the spectrum of these diseases.[103–105,110,111]

Defects in barrier structure and function have been observed in primary epithelial cell cultures from asthmatics even after being propagated for several weeks at air–liquid interface, suggesting that reduced barrier function is a stable property of these cells.[112] The type 2 cytokines IL-4 and 13 induce barrier dysfunction by inhibiting expression of ZO-1, occludins, E-cadherin and β-catenin.[113] It has been proposed that junctional disruption could be a T helper 2 (Th2) promoting signal in the airway to restore mucosal integrity.[114] This idea is supported by the fact that siRNA knockdown of E-cadherin results in epidermal growth factor receptor (EGFR)-dependent production of Th2-promoting cytokines thymus- and activation-regulated chemokine (TARC) and thymic stromal lymphopoietin (TSLP) by airway epithelial cells that explains the association between barrier dysfunction and Th2-driven pulmonary disease.[115] If the wound healing response is not curtailed following restoration of mucosal integrity, detrimental airway remodelling can ensue. E-cadherin, in addition to having a structural role at adherens junctions, is also the ligand for CD103, which is expressed on innate and adaptive immune cells and a subset of dendritic cells. Furthermore, it also binds to KLRG1, which is expressed by natural killer cells and regulatory T-cells and group 2 innate lymphoid cells. E-cadherin and occludin are known to be downregulated in asthmatics,[106,116,117] and the genes for E-cadherin and protocadherin-1 have been associated with airway hyperresponsiveness.[118,119] E-cadherin anchors EGFR within adherens junctions,[120] preventing activation of the receptor.[121,122] EGFR can redistribute to the luminal surface of the cell as a result of reduced expression of E-cadherin culminating in activation.[123] Excessive activation of EGFR can result in cell proliferation and goblet cell hyperplasia/metaplasia.[124] Increased expression of EGFR has been observed in asthma and correlated with impaired mucosal barrier function.[106]

When dissociated from E-cadherin, β-catenin can translocate to the nucleus, activating Wnt signalling, which results in cell proliferation.[125] Dikkopf (DKK) proteins regulate Wnt signalling via interaction with Kreme (KRM) receptors,[126] and the Wnt/β-catenin pathway has been implicated in IPF.[127,128] DKK1 expression is increased in lung and bronchoalveolar lavage (BAL) fluid of IPF patients. In addition to being located in hyperplastic alveolar cells, the strongest expression is observed in basal bronchial epithelial cells. KRM1 expression is also elevated in the lung tissue of IPF patients, particularly in areas of bronchiolization.[127] The phosphorylation signatures of β-catenin in IPF are the same as in normal lung development, again suggesting that repair of lung injury involves reactivation of developmental programmes.[129] Wnt/β-catenin signalling is also implicated in the pathogenesis of COPD, but in direct contrast to what is observed in IPF, signalling in the mesenchymal compartment is inactivated rather than activated.[130] Interestingly a shift from canonical Wnt signalling (β-catenin-dependent) to non-canonical pathways is suggested to drive pathology in IPF.[131]

Basal cells, which are candidate stem cells in the conducting airways, are found in highest concentration in the trachea, and numbers diminish further down the bronchial tree to the bronchioles. Upon damage these cells upregulate p63, a basal cell-restricted transcription factor, which has been implicated in epithelial mesenchymal transition.[132] It has been shown that basal epithelial cells overlying fibroblastic foci in IPF acquire increased reactivity, determined by p63 expression, and gain a partial mesenchymal phenotype,[133] further evidence of the involvement of bronchial-derived cells in the pathogenesis of IPF.

Collectively, it is likely that the dialogue between specific bacterial communities and host is a dynamic phenomenon. The installation and location of the bacterial communities within the lungs is affected by the cellular components within the lungs as well as the abiotic environment (temperature, pH, mucus and pulmonary surfactant). Specifically, loss of barrier function can influence the composition of the respiratory microbiota by allowing the entry of pathogens and other particles. However, it is evident that more information is needed to elucidate the nature of the dialogue occurring between the respiratory microbiota and epithelial cells in the lungs to further understand the function and mechanisms through which the microbiota shapes immunity in the lung.

The gut–lung axis

An increasing number of studies have provided evidence of the cross-talk that occurs between the intestinal microbiota and the lungs, termed the 'gut–lung axis'. Similar to the gut, the airway microbiota in healthy lungs is dominated by the short-chain fatty acid (SCFA)-producing bacteria: *Bacteroidetes*, *Firmicutes* and *Proteobacteria*. These are metabolically active bacteria that produce SCFAs via the bacterial fermentation of non-digestible carbohydrates by the gut microbiota; acetate, propionate and butyrate being the most abundant in the human intestinal tract.[134] It is now well documented that the intestinal microbiota influences the microbial composition within the lung either by direct seeding of bacteria into the airways or by the distribution of SCFAs. These SCFAs not only act as a fuel for intestinal epithelial cells but also exert distinct physiological effects, acting as a link between the microbiota and the immune system, and playing a key role in maintaining mucosal immunity. Although there are limited studies investigating the impact of these SCFAs in the lungs, in the gut they have been shown to be integral for tight junction formation and to promote anti-inflammatory mechanisms to sustain

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

intestinal homeostasis.[135] Indeed, SCFAs interact with immune cells, and modulate their recruitment, differentiation, activation and survival at different tissues.[136] The proposed pathway of this gut–lung interaction involves the entry of the microbiota and its products from the airways into the intestinal mucosa. These are phagocytosed and transferred to the mesenteric lymph nodes by antigen-presenting cells that stimulate the priming of B- and T-cells. Once activated, these cells can then migrate back to the original site or to distal locations such as the lung epithelium where they can exert directly on their target. Another proposed pathway involves the direct migration of bacterial products from the intestinal mucosa to the lungs via the bloodstream where they act to stimulate the immune system. Similarly, the same route is proposed beginning in the lung and culminating in the intestinal mucosa, although the influence of the lung microbiota and its metabolites on the gut is yet to be elucidated.[137]

The composition of a stable gut microbiota has been shown to have an influence on lung immunity. Mice devoid of their intestinal microbiota during early life are at an increased risk of developing allergic airway disease.[138,139] An increased risk of asthma has been associated with a predominance of *Bacteroidetes fragilis* and total anaerobes in early life,[140] decreases in the total abundances of *Escherichia coli*, *Faecalibacterium*, *Lachnospira*, *Rothia* and *Veillonella* bacterial genera.[141] In adults, taxa-specific differences in the gut microbiota were observed, such as an increased relative abundance of *Bifidobacterium adolescentis*, which negatively correlated with the time since asthma diagnosis.[142] Furthermore, mice fed a peptide-based enteral diet that increased the production of SCFAs attenuated elastase-induced inflammation and emphysema in the lung.[143] The axis is by no means one-directional, with respiratory infection driving perturbations in the gut microbiome, but further work is needed to validate this.[144,145] A study reported that the intranasal delivery of a non-absorbable tracer into the nasal cavity of mice appeared in the gastrointestinal tract shortly after.[144] In addition, work by Sze et al.[146] reported that intratracheal administration of lipopolysaccharide not only disrupted the airway microbiota but also led to the translocation of airway bacteria belonging to the *Clostridium* and *Lachnospiraceae* genera into the bloodstream and affected the gut microbiota within 24 hr, resulting in an evident increase in total bacterial load. Collectively, these studies provide evidence of the bidirectional communication happening between these two sites. However, our understanding of this cross-talk is still in its infancy, and there is a pressing need to uncover whether these changes observed in the microbiota are the cause or the effect of disease. Furthermore, there is growing evidence highlighting the use of oral antibiotics and their effect on lung immunity. Antibiotics could affect the lung microbiome either directly or indirectly through their effects on the gut microbiome and subsequently on lung immunity. To support this hypothesis, murine studies have shown that depletion of certain species within the gut microbiota due to antibiotic intake influences lung diseases and allergic inflammation.[147] For example, microbial dysbiosis due to the use of macrolides in early life correlated with an increased risk of asthma in Finnish pre-school children.[148] It is clear that we need to appreciate the interactions between the gut and airway microbiota in order to understand the impact of antibiotics on lung health.

Conclusions and future directions

The epithelial surfaces of the lung, including the proximal and distal airways, are now considered to play a pivotal role in maintenance of immune homeostasis. The epithelium acts as the first line of defence against inhaled viral, bacterial and fungal pathogens. Studies have shown that the epithelial surfaces of the human respiratory tract are colonized by a complex and dynamic microbiota that plays a role both in health and disease. It is becoming increasingly recognized that the cellular and physiological characteristics of the lungs impact the establishment and persistence of these bacterial communities. Dysregulation of the epithelial immune response and barrier function along with microbial dysbiosis can contribute to chronic inflammatory lung diseases such as asthma, COPD, CF and IPF. However, the precise nature of the relationship between the respiratory microbiome and the epithelium in the lungs remains an active area of investigation. A better understanding of this interaction in maintaining airway homeostasis will provide new insights into the pathogenesis of several respiratory conditions as well as novel diagnostic and therapeutic opportunities. Emerging evidence, particularly involving SCFAs, has highlighted a link between the gut and the lung and the microbial communities within. These results highlight the need for future studies that will characterize the microbiota and metabolites comprising these two mucosal sites and determine their interaction with the host in order to develop more targeted treatments for lung diseases.

Disclosures

The authors have no competing interests or conflicts of interest to declare.

References

1 Hiemstra PS, McCray PB Jr, Bals R. The innate immune function of airway epithelial cells in inflammatory lung disease. *Eur Respir J* 2015; **45**:1150–62.
2 Weibel ER. Morphometry of the Human Lung. Berlin Heidelberg: Springer, 1963.
3 Leiva-Juarez MM, Kolls JK, Evans SE. Lung epithelial cells: therapeutically inducible effectors of antimicrobial defense. *Mucosal Immunol* 2018; **11**:21–34.
4 Rock JR, Randell SH, Hogan BL. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. *Dis Model Mech* 2010; **3**:545–56.

5 Sacco O, Silvestri M, Sabatini F, Sale R, Defilippi AC, Rossi GA. Epithelial cells and fibroblasts: structural repair and remodelling in the airways. *Paediatr Respir Rev* 2004;**5**(Suppl. A):S35–40.
6 Whitsett JA, Weaver TE. Alveolar development and disease. *Am J Respir Cell Mol Biol* 2015; **53**:1–7.
7 Zepp JA, Morrisey EE. Cellular crosstalk in the development and regeneration of the respiratory system. *Nat Rev Mol Cell Biol* 2019; **20**:551–66.
8 Niessen CM. Tight junctions/adherens junctions: basic structure and function. *J Invest Dermatol* 2007; **127**:2525–32.
9 Shen L, Weber CR, Raleigh DR, Yu D, Turner JR. Tight junction pore and leak pathways: a dynamic duo. *Annu Rev Physiol* 2011; **73**:283–309.
10 Loxham M, Davies DE, Blume C. Epithelial function and dysfunction in asthma. *Clin Exp Allergy* 2014; **44**:1299–313.
11 Steed E, Balda MS, Matter K. Dynamics and functions of tight junctions. *Trends Cell Biol* 2010; **20**:142–9.
12 De Rose V, Molloy K, Gohy S, Pilette C, Greene CM. Airway epithelium dysfunction in cystic fibrosis and COPD. *Mediators Inflamm* 2018; **2018**:1–20.
13 Aghapour M, Raee P, Moghaddam SJ, Hiemstra PS, Heijink IH. Airway epithelial barrier dysfunction in chronic obstructive pulmonary disease: role of cigarette smoke exposure. *Am J Respir Cell Mol Biol* 2018; **58**:157–69.
14 Dickson RP, Huffnagle GB. The lung microbiome: new principles for respiratory bacteriology in health and disease. *PLoS Pathog* 2015; **11**:e1004923.
15 Jo JH, Kennedy EA, Kong HH. Research techniques made simple: bacterial 16S ribosomal RNA gene sequencing in cutaneous research. *J Invest Dermatol* 2016; **136**:e23–7.
16 Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C *et al*. Disordered microbial communities in asthmatic airways. *PLoS ONE* 2010; **5**:e8578.
17 Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV *et al*. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012; **185**:1073–80.
18 Huang YJ, LiPuma JJ. The microbiome in cystic fibrosis. *Clin Chest Med* 2016; **37**:59–67.
19 Molyneaux PL, Cox MJ, Willis-Owen SA, Mallia P, Russell KE, Russell AM *et al*. The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014; **190**:906–13.
20 Han MK, Zhou Y, Murray S, Tayob N, Noth I, Lama VN *et al*. Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014; **2**:548–56.
21 Salisbury ML, Han MK, Dickson RP, Molyneaux PL. Microbiome in interstitial lung disease: from pathogenesis to treatment target. *Curr Opin Pulm Med* 2017; **23**:404–10.
22 Gleeson K, Eggli DF, Maxwell SL. Quantitative aspiration during sleep in normal subjects. *Chest* 1997; **111**:1266–72.
23 Segal LN, Clemente JC, Tsay JC, Koralov SB, Keller BC, Wu BG *et al*. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. *Nat Microbiol* 2016; **1**:16031.
24 Mathieu E, Escribano-Vazquez U, Descamps D, Cherbuy C, Langella P, Riffault S *et al*. Paradigms of lung microbiota functions in health and disease, particularly, in asthma. *Front Physiol* 2018; **9**:1168.
25 Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM *et al*. The adult cystic fibrosis airway microbiota is stable over time and infection type, and highly resilient to antibiotic treatment of exacerbations. *PLoS ONE* 2012; **7**:e45001.
26 Molyneaux PL, Cox MJ, Wells AU, Kim HC, Ji W, Cookson WO *et al*. Changes in the respiratory microbiome during acute exacerbations of idiopathic pulmonary fibrosis. *Respir Res* 2017; **18**:29.
27 Invernizzi R, Molyneaux PL. The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2019; **28**:190 045.
28 Invernizzi R, Barnett J, Rawal B, Nair A, Ghai P, Kingston S *et al*. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020; **55**:1–9.
29 Durack J, Lynch SV, Nariya S, Bhakta NR, Beigelman A, Castro M *et al*. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. *J Allergy Clin Immunol* 2017; **140**:63–75.
30 Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N *et al*. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015; **17**:704–15.
31 Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. *PLoS ONE* 2012; **7**:e47305.
32 Feigelman R, Kahlert CR, Baty F, Rassouli F, Kleiner RL, Kohler P *et al*. Sputum DNA sequencing in cystic fibrosis: non-invasive access to the lung microbiome and to pathogen details. *Microbiome* 2017; **5**:20.
33 Frayman KB, Armstrong DS, Carzino R, Ferkol TW, Grimwood K, Storch GA *et al*. The lower airway microbiota in early cystic fibrosis lung disease: a longitudinal analysis. *Thorax* 2017; **72**:1104–12.
34 Denning DW, Pashley C, Hartl D, Wardlaw A, Godet C, Del Giacco S *et al*. Fungal allergy in asthma-state of the art and research needs. *Clin Transl Allergy* 2014; **4**:14.
35 Knutsen AP, Bush RK, Demain JG, Denning DW, Dixit A, Fairs A *et al*. Fungi and allergic lower respiratory tract diseases. *J Allergy Clin Immunol* 2012; **129**:280–91; quiz 292–283.
36 Moore BB, Moore TA. Viruses in idiopathic pulmonary fibrosis. Etiology and exacerbation. *Ann Am Thorac Soc* 2015; **12**(Suppl. 2):S186–192.
37 Busse WW, Lemanske RF Jr, Gern JE. Role of viral respiratory infections in asthma and asthma exacerbations. *Lancet* 2010; **376**:826–34.
38 Delhaes L, Monchy S, Frealle E, Hubans C, Salleron J, Leroy S *et al*. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community–implications for therapeutic management. *PLoS ONE* 2012; **7**:e36313.
39 Charlson ES, Diamond JM, Bittinger K, Fitzgerald AS, Yadav A, Haas AR *et al*. Lung-enriched organisms and aberrant bacterial and fungal respiratory microbiota after lung transplant. *Am J Respir Crit Care Med* 2012; **186**:536–45.
40 Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB. Modulation of post-antibiotic bacterial community reassembly and host response by *Candida albicans*. *Sci Rep* 2013; **3**:2191.
41 Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. *Nat Rev Immunol* 2014; **14**:827–35.
42 Song X, Gao H, Lin Y, Yao Y, Zhu S, Wang J *et al*. Alterations in the microbiota drive interleukin-17C production from intestinal epithelial cells to promote tumorigenesis. *Immunity* 2014; **40**:140–52.
43 Evans SE, Xu Y, Tuvim MJ, Dickey BF. Inducible innate resistance of lung epithelium to infection. *Annu Rev Physiol* 2010; **72**:413–35.
44 Wine JJ, Joo NS. Submucosal glands and airway defense. *Proc Am Thorac Soc* 2004; **1**:47–53.
45 Rokicki W, Rokicki M, Wojtacha J, Dzeljijli A. The role and importance of club cells (Clara cells) in the pathogenesis of some respiratory diseases. *Kardiochir Torakochirurgia Pol* 2016; **13**:26–30.
46 Mason RJ, Williams MC. Type II alveolar cell. Defender of the alveolus. *Am Rev Respir Dis* 1977; **115**:81–91.
47 Fehrenbach H. Alveolar epithelial type II cell: defender of the alveolus revisited. *Respir Res* 2001; **2**:33–46.
48 Adler KB, Li Y. Airway epithelium and mucus: intracellular signaling pathways for gene expression and secretion. *Am J Respir Cell Mol Biol* 2001; **25**:397–400.
49 Fahy JV, Dickey BF. Airway mucus function and dysfunction. *N Engl J Med* 2010; **363**:2233–47.
50 Proud D, Leigh R. Epithelial cells and airway diseases. *Immunol Rev* 2011; **242**:186–204.
51 Okuda K, Chen G, Subramani DB, Wolf M, Gilmore RC, Kato T *et al*. Localization of secretory mucins MUC5AC and MUC5B in normal/healthy human airways. *Am J Respir Crit Care Med* 2019; **199**:715–27.
52 Evans CM, Raclawska DS, Ttofali F, Liptzin DR, Fletcher AA, Harper DN *et al*. The polymeric mucin Muc5ac is required for allergic airway hyperreactivity. *Nat Commun* 2015; **6**:6281.
53 Regnis JA, Robinson M, Bailey DL, Cook P, Hooper P, Chan HK *et al*. Mucociliary clearance in patients with cystic fibrosis and in normal subjects. *Am J Respir Crit Care Med* 1994; **150**:66–71.
54 Smaldone GC, Foster WM, O'Riordan TG, Messina MS, Perry RJ, Langenback EG. Regional impairment of mucociliary clearance in chronic obstructive pulmonary disease. *Chest* 1993; **103**:1390–6.
55 Flynn JM, Niccum D, Dunitz JM, Hunter RC. Evidence and role for bacterial mucin degradation in cystic fibrosis airway disease. *PLoS Pathog* 2016; **12**:e1005846.
56 Lee AL, Goldstein RS. Gastroesophageal reflux disease in COPD: links and risks. *Int J Chron Obstruct Pulmon Dis* 2015; **10**:1935–49.
57 Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* 2016; **14**:563–75.
58 Maurice NM, Bedi B, Sadikot RT. *Pseudomonas aeruginosa* biofilms: host response and clinical implications in lung infections. *Am J Respir Cell Mol Biol* 2018; **58**:428–39.
59 Thurlow LR, Hanke ML, Fritz T, Angle A, Aldrich A, Williams SH *et al*. *Staphylococcus aureus* biofilms prevent macrophage phagocytosis and attenuate inflammation in vivo. *J Immunol* 2011; **186**:6585–96.
60 Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. *Annu Rev Physiol* 2016; **78**:481–504.
61 Hartsock A, Nelson WJ. Adherens and tight junctions: structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta* 2008; **1778**:660–9.
62 Molyneaux PL, Willis-Owen SAG, Cox MJ, James P, Cowman S, Loebinger M *et al*. Host-microbial interactions in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017; **195**:1640–50.

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

63 Schutte BC, McCray PB Jr. [beta]-defensins in lung host defense. *Annu Rev Physiol* 2002; **64**:709–48.
64 Mallia P, Webber J, Gill SK, Trujillo-Torralbo MB, Calderazzo MA, Finney L *et al.* Role of airway glucose in bacterial infections in patients with chronic obstructive pulmonary disease. *J Allergy Clin Immunol* 2018; **142**:815–23.e6.
65 O'Dwyer DN, Dickson RP, Moore BB. The lung microbiome, immunity, and the pathogenesis of chronic lung disease. *J Immunol* 2016; **196**:4839–47.
66 Mariathasan S, Monack DM. Inflammasome adaptors and sensors: intracellular regulators of infection and inflammation. *Nat Rev Immunol* 2007; **7**:31–40.
67 Ye Z, Ting JP. NLR, the nucleotide-binding domain leucine-rich repeat containing gene family. *Curr Opin Immunol* 2008; **20**:3–9.
68 Medzhitov R. Recognition of microorganisms and activation of the immune response. *Nature* 2007; **449**:819–26.
69 Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell* 2010; **140**:805–20.
70 Takeda K, Kaisho T, Akira S. Toll-like receptors. *Annu Rev Immunol* 2003; **21**:335–76.
71 Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* 2000; **408**:740–45.
72 Hemmi H, Kaisho T, Takeuchi O, Sato S, Sanjo H, Hoshino K *et al.* Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. *Nat Immunol* 2002; **3**:196–200.
73 Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* 2001; **413**:732–8.
74 Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 2001; **410**:1099–103.
75 de Bouteiller O, Merck E, Hasan UA, Hubac S, Benguigui B, Trinchieri G *et al.* Recognition of double-stranded RNA by human toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J Biol Chem* 2005; **280**:38 133–45.
76 Latz E, Schoenemeyer A, Visintin A, Fitzgerald KA, Monks BG, Knetter CF *et al.* TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* 2004; **5**:190–8.
77 Lee J, Chuang TH, Redecke V, She L, Pitha PM, Carson DA *et al.* Molecular basis for the immunostimulatory activity of guanine nucleoside analogs: activation of Toll-like receptor 7. *Proc Natl Acad Sci USA* 2003; **100**:6646–51.
78 Matsumoto M, Funami K, Tanabe M, Oshiumi H, Shingai M, Seto Y *et al.* Subcellular localization of Toll-like receptor 3 in human dendritic cells. *J Immunol* 2003; **171**:3154–62.
79 Nishiya T, DeFranco AL. Ligand-regulated chimeric receptor approach reveals distinctive subcellular localization and signaling properties of the Toll-like receptors. *J Biol Chem* 2004; **279**:19 008–17.
80 Oosting M, Cheng SC, Bolscher JM, Vestering-Stenger R, Plantinga TS, Verschueren IC *et al.* Human TLR10 is an anti-inflammatory pattern-recognition receptor. *Proc Natl Acad Sci USA* 2014; **111**:E4478–84.
81 Sha Q, Truong-Tran AQ, Plitt JR, Beck LA, Schleimer RP. Activation of airway epithelial cells by toll-like receptor agonists. *Am J Respir Cell Mol Biol* 2004; **31**:358–64.
82 Hammad H, Chieppa M, Perros F, Willart MA, Germain RN, Lambrecht BN. House dust mite allergen induces asthma via Toll-like receptor 4 triggering of airway structural cells. *Nat Med* 2009; **15**:410–16.
83 Hewson CA, Jardine A, Edwards MR, Laza-Stanca V, Johnston SL. Toll-like receptor 3 is induced by and mediates antiviral activity against rhinovirus infection of human bronchial epithelial cells. *J Virol* 2005; **79**:12 273–79.
84 Ritchie AI, Farne HA, Singanayagam A, Jackson DJ, Mallia P, Johnston SL. Pathogenesis of viral infection in exacerbations of airway disease. *Ann Am Thorac Soc* 2015; **12** (Suppl. 2):S115–32.
85 Ghaemmaghami AM, Gough L, Sewell HF, Shakib F. The proteolytic activity of the major dust mite allergen Der p 1 conditions dendritic cells to produce less interleukin-12: allergen-induced Th2 bias determined at the dendritic cell level. *Clin Exp Allergy* 2002; **32**:1468–75.
86 Adam E, Hansen KK, Astudillo Fernandez O, Coulon L, Bex F, Duhant X *et al.* The house dust mite allergen Der p 1, unlike Der p 3, stimulates the expression of interleukin-8 in human airway epithelial cells via a proteinase-activated receptor-2-independent mechanism. *J Biol Chem* 2006; **281**:6910–23.
87 Kauffman HF, Tamm M, Timmerman JA, Borger P. House dust mite major allergens Der p 1 and Der p 5 activate human airway-derived epithelial cells by protease-dependent and protease-independent mechanisms. *Clin Mol Allergy* 2006; **4**:5.
88 Chiu LL, Perng DW, Yu CH, Su SN, Chow LP. Mold allergen, pen C 13, induces IL-8 expression in human airway epithelial cells by activating protease-activated receptor 1 and 2. *J Immunol* 2007; **178**:5237–44.
89 Wang Q, Nagarkar DR, Bowman ER, Schneider D, Gosangi B, Lei J *et al.* Role of double-stranded RNA pattern recognition receptors in rhinovirus-induced airway epithelial cell responses. *J Immunol* 2009; **183**:6989–97.
90 Lipinski S, Rosenstiel P. Debug your bugs - how NLRs shape intestinal host-microbe interactions. *Front Immunol* 2013; **4**:479.
91 Lee HM, Yuk JM, Shin DM, Jo EK. Dectin-1 is inducible and plays an essential role for mycobacteria-induced innate immune responses in airway epithelial cells. *J Clin Immunol* 2009; **29**:795–805.
92 Sun WK, Lu X, Li X, Sun QY, Su X, Song Y *et al.* Dectin-1 is inducible and plays a crucial role in Aspergillus-induced innate immune responses in human bronchial epithelial cells. *Eur J Clin Microbiol Infect Dis* 2012; **31**:2755–64.
93 Nathan AT, Peterson EA, Chakir J, Wills-Karp M. Innate immune responses of airway epithelium to house dust mite are mediated through beta-glucan-dependent pathways. *J Allergy Clin Immunol* 2009; **123**:612–8.
94 Trompette A, Divanovic S, Visintin A, Blanchard C, Hegde RS, Madan R *et al.* Allergenicity resulting from functional mimicry of a Toll-like receptor complex protein. *Nature* 2009; **457**:585–8.
95 Page K, Lierl KM, Hughes VS, Zhou P, Ledford JR, Wills-Karp M. TLR2-mediated activation of neutrophils in response to German cockroach frass. *J Immunol* 2008; **180**:6317–24.
96 Boldogh I, Bacsi A, Choudhury BK, Dharajiya N, Alam R, Hazra TK *et al.* ROS generated by pollen NADPH oxidase provide a signal that augments antigen-induced allergic airway inflammation. *J Clin Invest* 2005; **115**:2169–79.
97 Salo PM, Arbes SJ Jr, Sever M, Jaramillo R, Cohn RD, London SJ *et al.* Exposure to *Alternaria alternata* in US homes is associated with asthma symptoms. *J Allergy Clin Immunol* 2006; **118**:892–8.
98 Hsu AC, See HV, Hansbro PM, Wark PA. Innate immunity to influenza in chronic airways diseases. *Respirology* 2012; **17**:1166–75.
99 Hansbro NG, Horvat JC, Wark PA, Hansbro PM. Understanding the mechanisms of viral induced asthma: new therapeutic directions. *Pharmacol Ther* 2008; **117**:313–53.
100 Hansbro PM, Beagley KW, Horvat JC, Gibson PG. Role of atypical bacterial infection of the lung in predisposition/protection of asthma. *Pharmacol Ther* 2004; **101**:193–210.
101 Bosch AA, Biesbroek G, Trzcinski K, Sanders EA, Bogaert D. Viral and bacterial interactions in the upper respiratory tract. *PLoS Pathog* 2013; **9**:e1003057.
102 Ganesan S, Comstock AT, Sajjan US. Barrier function of airway tract epithelium. *Tissue Barriers* 2013; **1**:e24997.
103 Sajjan U, Wang Q, Zhao Y, Gruenert DC, Hershenson MB. Rhinovirus disrupts the barrier function of polarized airway epithelial cells. *Am J Respir Crit Care Med* 2008; **178**:1271–81.
104 Comstock AT, Ganesan S, Chattoraj A, Faris AN, Margolis BL, Hershenson MB *et al.* Rhinovirus-induced barrier dysfunction in polarized airway epithelial cells is mediated by NADPH oxidase 1. *J Virol* 2011; **85**:6795–08.
105 Rezaee F, DeSando SA, Ivanov AI, Chapman TJ, Knowlden SA, Beck LA *et al.* Sustained protein kinase D activation mediates respiratory syncytial virus-induced airway barrier disruption. *J Virol* 2013; **87**:11 088–95.
106 Xiao C, Puddicombe SM, Field S, Haywood J, Broughton-Head V, Puxeddu I *et al.* Defective epithelial barrier function in asthma. *J Allergy Clin Immunol* 2011; **128**: 549–56.e12.
107 Hackett TL, de Bruin HG, Shaheen F, van den Berge M, van Oosterhout AJ, Postma DS *et al.* Caveolin-1 controls airway epithelial barrier function. Implications for asthma. *Am J Respir Cell Mol Biol* 2013; **49**:662–71.
108 Shaykhiev R, Otaki F, Bonsu P, Dang DT, Teater M, Strulovici-Barel Y *et al.* Cigarette smoking reprograms apical junctional complex molecular architecture in the human airway epithelium in vivo. *Cell Mol Life Sci* 2011; **68**:877–92.
109 Wonderlich ER, Swan ZD, Bissel SJ, Hartman AL, Carney JP, O'Malley KJ *et al.* Widespread virus replication in alveoli drives acute respiratory distress syndrome in aerosolized H5N1 influenza infection of macaques. *J Immunol* 2017; **198**:1616–26.
110 Coyne CB, Shen L, Turner JR, Bergelson JM. Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe* 2007; **2**:181–92.
111 Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SA, Homola D *et al.* Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; **188**:1224–31.
112 Grainge C, Dennison P, Lau L, Davies D, Howarth P. Asthmatic and normal respiratory epithelial cells respond differently to mechanical apical stress. *Am J Respir Crit Care Med* 2014; **190**:477–80.
113 Saatian B, Rezaee F, Desando S, Emo J, Chapman T, Knowlden S *et al.* Interleukin-4 and interleukin-13 cause barrier dysfunction in human airway epithelial cells. *Tissue Barriers* 2013; **1**:e24333.
114 Gause WC, Wynn TA, Allen JE. Type 2 immunity and wound healing: evolutionary refinement of adaptive immunity by helminths. *Nat Rev Immunol* 2013; **13**:607–14.
115 Heijink IH, Kies PM, Kauffman HF, Postma DS, van Oosterhout AJ, Vellenga E. Down-regulation of E-cadherin in human bronchial epithelial cells leads to epidermal growth factor receptor-dependent Th2 cell-promoting activity. *J Immunol* 2007; **178**:7678–85.

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

116 de Boer WI, Sharma HS, Baelemans SM, Hoogsteden HC, Lambrecht BN, Braunstahl GJ. Altered expression of epithelial junctional proteins in atopic asthma: possible role in inflammation. *Can J Physiol Pharmacol* 2008; **86**:105–12.

117 Hackett TL, Singhera GK, Shaheen F, Hayden P, Jackson GR, Hegele RG *et al.* Intrinsic phenotypic differences of asthmatic epithelium and its inflammatory responses to respiratory syncytial virus and air pollution. *Am J Respir Cell Mol Biol* 2011; **45**:1090–100.

118 Ierodiakonou D, Postma DS, Koppelman GH, Boezen HM, Gerritsen J, Ten Hacken N *et al.* E-cadherin gene polymorphisms in asthma patients using inhaled corticosteroids. *Eur Respir J* 2011; **38**:1044–52.

119 Koppelman GH, Meyers DA, Howard TD, Zheng SL, Hawkins GA, Ampleford EJ *et al.* Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness. *Am J Respir Crit Care Med* 2009; **180**:929–35.

120 Qian X, Karpova T, Sheppard AM, McNally J, Lowy DR. E-cadherin-mediated adhesion inhibits ligand-dependent activation of diverse receptor tyrosine kinases. *Embo J* 2004; **23**:1739–48.

121 Sorby M, Ostman A. Protein-tyrosine phosphatase-mediated decrease of epidermal growth factor and platelet-derived growth factor receptor tyrosine phosphorylation in high cell density cultures. *J Biol Chem* 1996; **271**:10 963–66.

122 Symons JR, LeVea CM, Mooney RA. Expression of the leucocyte common antigen-related (LAR) tyrosine phosphatase is regulated by cell density through functional E-cadherin complexes. *Biochem J* 2002; **365**:513–9.

123 Wendt MK, Smith JA, Schiemann WP. Transforming growth factor-beta-induced epithelial-mesenchymal transition facilitates epidermal growth factor-dependent breast cancer progression. *Oncogene* 2010; **29**:6485–98.

124 Casalino-Matsuda SM, Monzon ME, Forteza RM. Epidermal growth factor receptor activation by epidermal growth factor mediates oxidant-induced goblet cell metaplasia in human airway epithelium. *Am J Respir Cell Mol Biol* 2006; **34**:581–91.

125 Mazieres J, He B, You L, Xu Z, Jablons DM. Wnt signaling in lung cancer. *Cancer Lett* 2005; **222**:1–10.

126 Wang K, Zhang Y, Li X, Chen L, Wang H, Wu J *et al.* Characterization of the Kremen-binding site on Dkk1 and elucidation of the role of Kremen in Dkk-mediated Wnt antagonism. *J Biol Chem* 2008; **283**:23 371–75.

127 Pfaff EM, Becker S, Gunther A, Konigshoff M. Dickkopf proteins influence lung epithelial cell proliferation in idiopathic pulmonary fibrosis. *Eur Respir J* 2011; **37**:79–87.

128 Chilosi M, Poletti V, Zamo A, Lestani M, Montagna L, Piccoli P *et al.* Aberrant Wnt/beta-catenin pathway activation in idiopathic pulmonary fibrosis. *Am J Pathol* 2003; **162**:1495–502.

129 Sucre JMS, Deutsch GH, Jetter CS, Ambalavanan N, Benjamin JT, Gleaves LA *et al.* A shared pattern of beta-catenin activation in bronchopulmonary dysplasia and idiopathic pulmonary fibrosis. *Am J Pathol* 2018; **188**:853–62.

130 Shi J, Li F, Luo M, Wei J, Liu X. Distinct roles of wnt/beta-catenin signaling in the pathogenesis of chronic obstructive pulmonary disease and idiopathic pulmonary fibrosis. *Mediators Inflamm* 2017; **2017**:3 520 581.

131 Baarsma HA, Skronska-Wasek W, Mutze K, Ciolek F, Wagner DE, John-Schuster G *et al.* Noncanonical WNT-5A signaling impairs endogenous lung repair in COPD. *J Exp Med* 2017; **214**:143–63.

132 Oh JE, Kim RH, Shin KH, Park NH, Kang MK. DeltaNp63alpha protein triggers epithelial-mesenchymal transition and confers stem cell properties in normal human keratinocytes. *J Biol Chem* 2011; **286**:38 757–67.

133 Jonsdottir HR, Arason AJ, Palsson R, Franzdottir SR, Gudbjartsson T, Isaksson HJ *et al.* Basal cells of the human airways acquire mesenchymal traits in idiopathic pulmonary fibrosis and in culture. *Lab Invest* 2015; **95**:1418–28.

134 Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 1987; **28**:1221–7.

135 Peng L, Li ZR, Green RS, Holzman IR, Lin J. Butyrate enhances the intestinal barrier by facilitating tight junction assembly via activation of AMP-activated protein kinase in Caco-2 cell monolayers. *J Nutr* 2009; **139**:1619–25.

136 Dang AT, Marsland BJ. Microbes, metabolites, and the gut-lung axis. *Mucosal Immunol* 2019; **12**:843–50.

137 Samuelson DR, Welsh DA, Shellito JE. Regulation of lung immunity and host defense by the intestinal microbiota. *Front Microbiol* 2015; **6**:1085.

138 Noverr MC, Falkowski NR, McDonald RA, McKenzie AN, Huffnagle GB. Development of allergic airway disease in mice following antibiotic therapy and fungal microbiota increase: role of host genetics, antigen, and interleukin-13. *Infect Immun* 2005; **73**:30–38.

139 Fagundes CT, Amaral FA, Vieira AT, Soares AC, Pinho V, Nicoli JR *et al.* Transient TLR activation restores inflammatory response and ability to control pulmonary bacterial infection in germfree mice. *J Immunol* 2012; **188**:1411–20.

140 Vael C, Nelen V, Verhulst SL, Goossens H, Desager KN. Early intestinal Bacteroides fragilis colonisation and development of asthma. *BMC Pulm Med* 2008; **8**:19.

141 Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S *et al.* infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med* 2015; **7**:307ra152.

142 Hevia A, Milani C, Lopez P, Donado CD, Cuervo A, Gonzalez S *et al.* Allergic patients with long-term asthma display low levels of bifidobacterium adolescentis. *PLoS ONE* 2016; **11**:e0147809.

143 Tomoda K, Kubo K, Dairiki K, Yamaji T, Yamamoto Y, Nishii Y *et al.* Whey peptide-based enteral diet attenuated elastase-induced emphysema with increase in short chain fatty acids in mice. *BMC Pulm Med* 2015; **15**:64.

144 Southam DS, Dolovich M, O'Byrne PM, Inman MD. Distribution of intranasal instillations in mice: effects of volume, time, body position, and anesthesia. *Am J Physiol Lung Cell Mol Physiol* 2002; **282**:L833–39.

145 Groves HT, Higham SL, Moffatt MF, Cox MJ, Tregoning JS. Respiratory viral infection alters the gut microbiota by inducing inappetence. *mBio* 2020; **11**:2150–74511.

146 Sze MA, Tsuruta M, Yang SW, Oh Y, Man SF, Hogg JC *et al.* Changes in the bacterial microbiota in gut, blood, and lungs following acute LPS instillation into mice lungs. *PLoS ONE* 2014; **9**:e111228.

147 Metsala J, Lundqvist A, Virta LJ, Kaila M, Gissler M, Virtanen SM. Mother's and offspring's use of antibiotics and infant allergy to cow's milk. *Epidemiology* 2013; **24**:303–9.

148 Korpela K, Salonen A, Virta LJ, Kekkonen RA, Forslund K, Bork P *et al.* Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nat Commun* 2016; **7**:10 410.

© 2020 The Authors. *Immunology* published by John Wiley & Sons Ltd.,

ORIGINAL ARTICLE
INTERSTITIAL LUNG DISEASE

Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent

Rachele Invernizzi[1,4], Joseph Barnett[1,2,4], Bhavin Rawal[2], Arjun Nair[3], Poonam Ghai[1], Shaun Kingston[1], Felix Chua[2], Zhe Wu[1,2], Athol U. Wells[2], Elizabeth R. Renzoni[2], Andrew G. Nicholson[1,2], Alexandra Rice[2], Clare M. Lloyd[1], Adam J. Byrne[1], Toby M. Maher[1,2], Anand Devaraj[1,2] and Philip L. Molyneaux[1,2]

Affiliations: [1]National Heart and Lung Institute, Imperial College London, London, UK. [2]Royal Brompton Hospital, London, UK. [3]Dept of Radiology, University College Hospital, London, UK. [4]Contributed equally as first authors.

Correspondence: Philip L. Molyneaux, Fibrosis Research Group, National Heart and Lung Institute, Imperial College London, UK. E-mail: p.molyneaux@imperial.ac.uk

@ERSpublications
Elevated bacterial burden predicts mortality in IPF and is independent of both radiological features and extent of disease http://bit.ly/2RaDbdv

Cite this article as: Invernizzi R, Barnett J, Rawal B, *et al*. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020; 55: 1901519 [https://doi.org/10.1183/13993003.01519-2019].

ABSTRACT Increasing bacterial burden in the lower airways of patients with idiopathic pulmonary fibrosis confers an increased risk of disease progression and mortality. However, it remains unclear whether this increased bacterial burden directly influences progression of fibrosis or simply reflects the magnitude of the underlying disease extent or severity.

We prospectively recruited 193 patients who underwent bronchoscopy and received a multidisciplinary diagnosis of idiopathic pulmonary fibrosis. Quantification of the total bacterial burden in bronchoalveolar lavage fluid was performed by 16S rRNA gene qPCR. Imaging was independently evaluated by two readers assigning quantitative scores for extent, severity and topography of radiographic changes and relationship of these features with bacterial burden was assessed.

Increased bacterial burden significantly associated with disease progression (HR 2.1; 95% CI 1.287–3.474; p=0.0028). Multivariate stepwise regression demonstrated no relationship between bacterial burden and radiological features or extent of disease. When specifically considering patients with definite or probable usual interstitial pneumonia there was no difference in bacterial burden between these two groups. Despite a postulated association between pleuroparenchymal fibroelastosis and clinical infection, there was no relationship between either the presence or extent of pleuroparenchymal fibroelastosis and bacterial burden.

We demonstrate that bacterial burden in the lower airways is not simply secondary to the extent of the underlying architectural destruction of the lung parenchyma seen in idiopathic pulmonary fibrosis. The independent nature of this association supports a relationship with the underlying pathogenic mechanisms and highlights the urgent need for functional studies.

This article has an editorial commentary: https://doi.org/13993003.00318-2020

This article has supplementary material available from erj.ersjournals.com

Received: 1 Aug 2019 | Accepted after revision: 29 Dec 2019

Copyright ©ERS 2020. This version is distributed under the terms of the Creative Commons Attribution Licence 4.0.

Introduction

The incidence of idiopathic pulmonary fibrosis (IPF) continues to rise as does the burden of related mortality [1]. Whilst a precise understanding of the pathogenesis remains elusive, there is a growing appreciation of the importance played by the interaction between the lung and environment in the development of IPF [2–4]. Over the past decade, our understanding of the dynamic and complex bacterial communities in the lower airways (the microbiome) and their role both in health and disease has increased dramatically [5]. Although, historically, infection was not considered a major driver in IPF it is now clear that infective episodes carry the same devastating mortality as acute exacerbations of the disease [3, 6]. Furthermore, even at diagnosis and in the absence of infection there is a dramatic increase in bacterial load with altered composition of the respiratory microbiome in the lower airways of patients with IPF [7, 8].

We and others have highlighted differences in microbiome diversity and levels of individual bacteria when comparing both healthy individuals and those with IPF and when comparing individuals with IPF who are stable or progressive. A number of associations between components of the microbial communities and cytokines, peripheral blood-transcriptomic profile and disease behaviour have been identified [9, 10]. More recently, animal models have begun to tease out mechanisms by which dysbiosis in the lower airways can drive fibrosis [11]. The best validated microbial signal identified in IPF is that of the lower airway bacterial burden. At the time of diagnosis, individuals with IPF have a higher bacterial burden than healthy individuals and subjects with chronic obstructive pulmonary disease. Importantly, across IPF subjects, the level of bacterial burden relates to survival and increases further during culture-negative acute exacerbations [12].

It remains unclear whether bacterial burden directly influences progression of fibrosis or if bacterial numbers instead reflect other underlying disease processes. While previous studies have accounted for disease severity using physiological parameters (including forced vital capacity (FVC) and measures of gas transfer (diffusing capacity of the lung for CO (D_{LCO})), there has been no assessment of the relationship between bacterial burden and radiological markers of fibrotic distortion of the lung or disease extent or severity. We hypothesised that if the bacterial burden is elevated secondary to the underlying architectural destruction of the lung parenchyma then there might be a clear association between radiographic findings such as honeycombing or traction bronchiectasis and bacterial burden in the lower airways of individuals with IPF.

Methods

Patient recruitment

Patients undergoing diagnostic bronchoscopy with bronchoalveolar lavage (BAL) for suspected IPF were prospectively recruited between November 2015 and January 2017. Only subjects receiving a multidisciplinary team diagnosis (MDT) of IPF according to current American Thoracic Society (ATS)/ European Respiratory Society (ERS)/Japanese Respiratory Society (JRS)/ Latin American Thoracic Association (ALAT) guideline definitions were included [13]. Subjects were excluded if they had a history of self-reported upper or lower respiratory tract infection, antibiotic use in the prior 3 months, acute exacerbation of IPF, or other respiratory disorders. Written informed consent was obtained from all subjects, and the study was approved by the local Research Ethics Committee (10/HO720/12 and 15/SC/0101).

Bronchoscopy

Fibreoptic bronchoscopy with BAL was performed *via* the oropharyngeal route as previously described [7]. Briefly, 60-mL aliquots of warmed saline, to a total volume of 240 mL, were separately instilled into a segment of the right middle lobe. Post-collection, an aliquot of unfiltered and unprocessed BAL was immediately placed on ice, snap frozen and stored at −80°C. Negative control samples were collected by aspirating buffered saline through the bronchoscope suction channel before use. Cell differentials of macrophages, lymphocytes, neutrophils and eosinophils were performed as previously described [14].

Radiology

The CT variable definitions were based on the Fleischner Glossary of terms [15] and current ATS/ERS/ JRS/ALAT guideline definitions [13]. Computed tomography (CT) data were scored using the following scales: to the nearest 5% per lobe (fibrosis, honeycombing), on a binary basis (usual interstitial pneumonia (UIP)), on a lobar extent with the lingua defined as the sixth lobe (pleuroparenchymal fibroelastosis (PPFE)) [16] and on an 18-point scale (traction bronchiectasis) [17].

Bacterial DNA extraction

BAL samples (2 mL) were centrifuged at 21 000×*g* for 30 min to pellet cell debris and bacteria. Pellets were resuspended in 100 µL of supernatant and added to lysing matrix E tubes (MP Biomedicals, Solon, OH, USA) containing 500 µL cetyl trimethylammonium bromide (CTAB) buffer (10% w/v CTAB in 0.5 M phosphate buffered NaCl) and 500 µL phenol:chloroform:isoamyl alcohol (25:24:1), and shaken in a

FastPrep Instrument (MP Biomedicals) at 5.5 ms^{-1} for 60 s. Following bead-beating, samples were extracted with an equal volume of chloroform:isoamyl alcohol (24:1), DNA was precipitated with 2 volumes of precipitation solution (30% w/v PEG6000 in NaCl) and, following ethanol washing, was resuspended in 100 µL Tris-EDTA. The quality and quantity of the isolated DNA was measured using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Hemel Hempstead, UK) and the DNA was stored at −80°C until further use.

16S rRNA gene quantitative PCR

Triplicate 10 µL quantitative PCRs (qPCR) were set up, containing 1 µL of bacterial DNA and 9 µL of Femto bacterial qPCR premix (Cambridge bioscience, Cambridge, UK). Each run contained a 10-fold dilution series of the *Vibrio natriegens DSM 759* gene cloned into a plasmid of known size and a non-template control. For data acquisition, the following cycling parameters were used: 1 cycle of 95°C for 10 min; 40 cycles of 95°C for 30 s, 50°C for 30 s, 72°C for 1 min; and 1 cycle of 72°C for 7 min [18].

Statistical analysis

Continuous variables are presented as mean±SD and categorical variables as proportions. The time-to-event curves were calculated using the Kaplan–Meier method and compared with the use of the log-rank test. Differences between subject groups were evaluated with the use of the Mann–Whitney test for continuous variables and Fisher's exact test for categorical variables. Spearman's *rho* was used to calculate correlations between continuous variables. Assessment for collinearity of predictor variables was made using correlation coefficients. Bacterial burden data were found to be skewed and were log transformed. Transformed bacterial burden data were examined using univariate and stepwise multivariate linear regression models against CT features and cell differentials. All analyses were performed with the use of R [19]. A two-sided p-value <0.05 was considered to indicate statistical significance.

Results

Subjects and sampling

Between 2014 and 2017, 193 subjects undergoing diagnostic bronchoscopy, resulting in a final MDT of IPF, were prospectively recruited. 19 subjects underwent cryo- or surgical lung biopsy and this information was available to the MDT when assigning a diagnosis. Reflecting the demographics of IPF, the cohort was predominantly male, with an average age of 70 years and a significant smoking history (table 1). Of the 193 patients included in this study, 107 had concomitant gastro-oesophageal reflux (GORD) (defined as either self-reported symptoms or asymptomatic use of proton pump inhibitors (PPI)). 97 subjects in total were on anti-reflux therapy. 36 subjects were receiving inhaled corticosteroids at the time of bronchoscopy and 21 were on oral corticosteroids (<10 $mg{\cdot}day^{-1}$) (supplementary table S1).

Bacterial burden predicts mortality in IPF

Bacterial DNA was successfully extracted from all subjects and quantified using qPCR. The mean bacterial burden was 2.87×10^5 ($\pm1.63\times10^6$) 16S rRNA gene copies per mL of BAL. This is elevated compared to

TABLE 1 Clinical and radiological characteristics of patients with idiopathic pulmonary fibrosis

Demographics	**N=193**
Sex male	140 (73)
Age years	70±8
Smoking status	
Never	70 (36)
Ex	113 (59)
Current	10 (5)
Pulmonary function tests	
FVC % predicted	79±18
D_{LCO} % predicted	48±14
Radiology	
UIP probable/definite	57.7
Honeycombing	30
PPFE	21

Data are presented as n (%), mean±SD or %. Radiological variables are expressed as presence or absence. FVC: forced vital capacity; D_{LCO}: diffusing capacity of the lung for CO; UIP: usual interstitial pneumonia; PPFE: pleuroparenchymal fibroelastosis.

historical controls and similar to previously published burdens [11]. Technical controls (saline flushed through the working channel of the bronchoscope), as anticipated, demonstrated counts close to or below the lower limit of qPCR quantification (100 copies·mL^{-1}) (supplementary figure S1). To confirm the association of bacterial burden with disease progression, a composite end point of all-cause mortality within the first year or a decline of 10% or more in FVC at 12 months was employed. A total of 102 patients fulfilled the criteria for progressive disease. Using an unadjusted Cox proportional hazards model, bacterial burden significantly associated with disease progression (HR 1.20; 95% CI 1.06–1.35; p=0.0034). This association remained significant after adjustment for age, sex, baseline % predicted FVC, baseline % predicted D_{LCO} and smoking status in a multivariable Cox proportional hazards model (HR 1.16; 95% CI 1.02–1.30; p=0.0022). A survival analysis confirmed that subjects in the tertile with the highest bacterial burden had a significantly increased hazard compared with those in the lowest tertile (HR 2.1; 95% CI 1.287–3.474; p=0.0028). Moreover, individuals in the middle tertile were also at a substantially increased risk of mortality compared with subjects in the tertile with the lowest bacterial burden (HR 2.0; 95% CI 1.197–3.307; p=0.008) (figure 1). There were no significant differences in age, sex, smoking history or disease severity between the patients with IPF within these tertiles of bacterial burden.

Bacterial burden and BAL differential cell counts

Next, the relationship between BAL differential cell count and bacterial burden was examined. As anticipated, the BAL cell differential for this cohort of individuals with IPF showed a predominance of alveolar macrophages 71.3±10.86% and neutrophils 9.86±7.78% (supplementary figure S2). A linear model was used to assess for any association between bacterial burden and BAL inflammatory differential cell count (supplementary table S2). Interestingly, there was no association (Spearman's rho: −0.093, p=0.24) between bacterial burden and total numbers of leukocytes per mL of BAL fluid (supplementary figure S3) or the proportion of any specific airway-inflammatory cell type (figure 2).

Bacterial burden and physiological disease severity

We then sought to investigate whether there was a relationship between baseline disease severity, measured by FVC (% predicted) and D_{LCO} (% predicted), and bacterial burden. However, no association (FVC, % predicted: Spearman's rho: −0.013, p=0.86; D_{LCO}, % predicted: Spearman's rho: −0.002 p=0.98) was found between the two (supplementary figure S4). While we did not set out to address the issues of GORD or concomitant medication use directly, we saw no effect on the bacterial burden based on the use of low-dose corticosteroids, inhaled steroids or PPI.

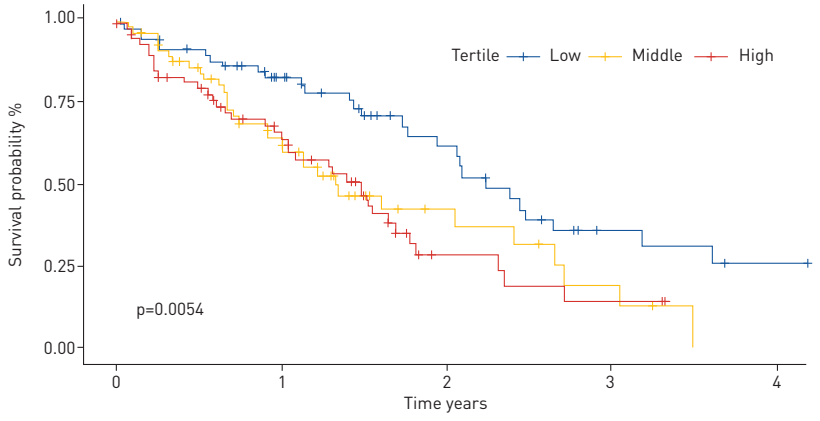


FIGURE 1 Increased bacterial burden in idiopathic pulmonary fibrosis at the time of diagnosis increases mortality. Kaplan–Meier curve generated by Cox proportional-hazards model stratified by bacterial burden (low tertile: 4.15×10^{3}–3.17×10^{4} 16S rRNA gene copies·mL^{-1} of bronchoalveolar lavage (BAL) fluid; middle tertile: 3.33×10^{4}–1.90×10^{5} 16S rRNA gene copies·mL^{-1} of BAL fluid; high tertile: 1.90×10^{5}–5.44×10^{6} 16S rRNA gene copies·mL^{-1} BAL fluid). Log rank p test value reported.

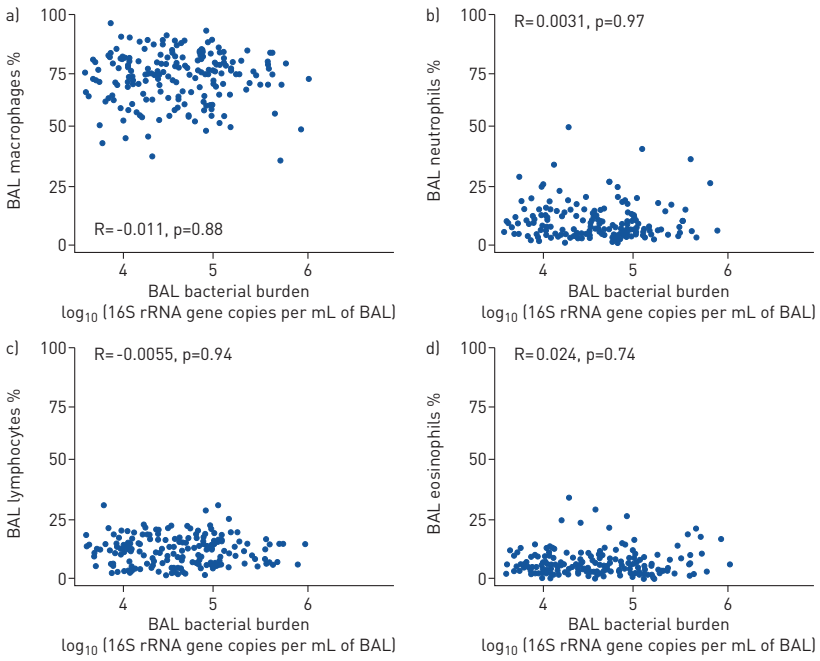


FIGURE 2 No correlation between bronchoalveolar lavage (BAL) differential cell counts and bacterial burden of subjects with idiopathic pulmonary fibrosis (IPF). Illustrating correlation between bacterial burden and percentage of (a) macrophages, (b) neutrophils, (c) lymphocytes and (d) eosinophils in BAL of subjects with IPF (n=193). Bacterial burden calculated by qPCR and expressed as log10 16S rRNA gene copies·mL^{-1} of BAL.

Bacterial burden and radiological disease extent

Of the overall cohort, 140 subjects had CT imaging performed within 4 months of bronchoscopy. The median (range) time from bronchoscopy to imaging was −31 (−117–120) days. CTs were scored by two independent radiologists [A. Nair, B. Rawal], blinded to all clinical data. In cases of disagreement >20% between radiologists (incorporating >2 lobe difference in lobar scores, traction bronchiectasis score of >3 or disagreement of a binary variable), data were arbitrated by a third radiologist [A. Devaraj]. For scores evaluated to the nearest 5%, the whole lung score was defined as the mean lobar score. The mean score between radiologists, following arbitration, was used as the final score for analysis. In general, there was a reasonable agreement between radiologists (supplementary table S3) [20, 21]. Over half of the cohort had probable or definite UIP (57.7%) according to the current ATS/ERS/JRS/ALAT guidelines [13]. 21% (30 out of 141) of subjects had radiological evidence of PPFE and 30% (43 out of 141) had honeycombing

TABLE 2 Radiological features and their univariate prediction of bronchoalveolar lavage bacterial burden

	Value	β estimate	SE	Adjusted p-value
Traction bronchiectasis extent score	6.7±3.0	0.691	1.750	0.492
Fibrosis extent %	22.0±10.7	1.050	0.921	0.297
Honeycombing extent %	2.6±6.8	−0.028	1.830	0.978
Definite or probable UIP pattern	88.0 (62.4)	−0.597	0.278	0.552
PPFE lobar extent	0.4±0.9	0.414	1.140	0.680

Data are presented as mean±SD or n (%), unless otherwise stated. Traction bronchiectasis scored on an 18-point scale. PPFE scored by number of lobes involved. Univariate regression performed against log transformed 16S rRNA gene copy number. SE: standard error; UIP: usual interstitial pneumonia; PPFE: pleuroparenchymal fibroelastosis.

(table 1). The relationship between radiological markers of disease severity and bacterial burden was explored using multivariate stepwise regression. CT fibrosis extent and CT traction bronchiectasis score were strongly collinear (Spearman's rho: 0.57), and so were included in separate models. There was no association between bacterial burden and radiological extent of disease or extent of either traction bronchiectasis or honeycombing (table 2). When specifically considering patients with definite or probable UIP, there was no difference in bacterial burden between these two groups (figure 3a). Furthermore, no difference in bacterial burden was observed when comparing IPF patients with either indeterminate or probable/definite UIP (supplementary figure S5). Despite a postulated association between PPFE and clinical infection [22], there was no relationship between either the presence or extent of PPFE and bacterial burden (figure 3b). Furthermore, no difference in bacterial burden was found when considering the presence or absence of honeycombing (figure 3c).

We then incorporated radiographic features of disease severity into the survival model to assess if bacterial burden predicts survival independent of radiographic features. Bacterial burden remained an independent predictor of survival even when incorporating radiographic features into the survival models and following adjustment for age, sex, baseline % predicted FVC, baseline % predicted D_{LCO} and smoking status. Again, given that CT fibrosis extent and CT traction bronchiectasis score were strongly collinear (Spearman's rho: 0.57), they were included in separate models. In both models, bacterial burden remained predictive of disease progression (HR 1.16; 95% CI 1.01–1.35; p=0.04).

Discussion

This study validates previous findings that, for individuals with IPF, an increased bacterial burden at the time of diagnosis confers a worse survival. However, we demonstrate that this increased bacterial burden is independent of radiological markers of disease morphology and severity. This work, the largest prospective study of bacteria in the lower airways of individuals with IPF performed to date, confirms the importance of airway bacterial burden in IPF. Furthermore, the lack of association seen with measures of disease severity and fibrotic destruction of the lung suggests that changes in bacterial burden are disease relevant and unlikely simply secondary to the architectural distortion.

One of the many paradoxes in fibrotic lung disease is that, despite the existence of often extensive traction bronchiectasis, the productive symptoms seen in bronchiectasis rarely develop [23]. While the bacterial

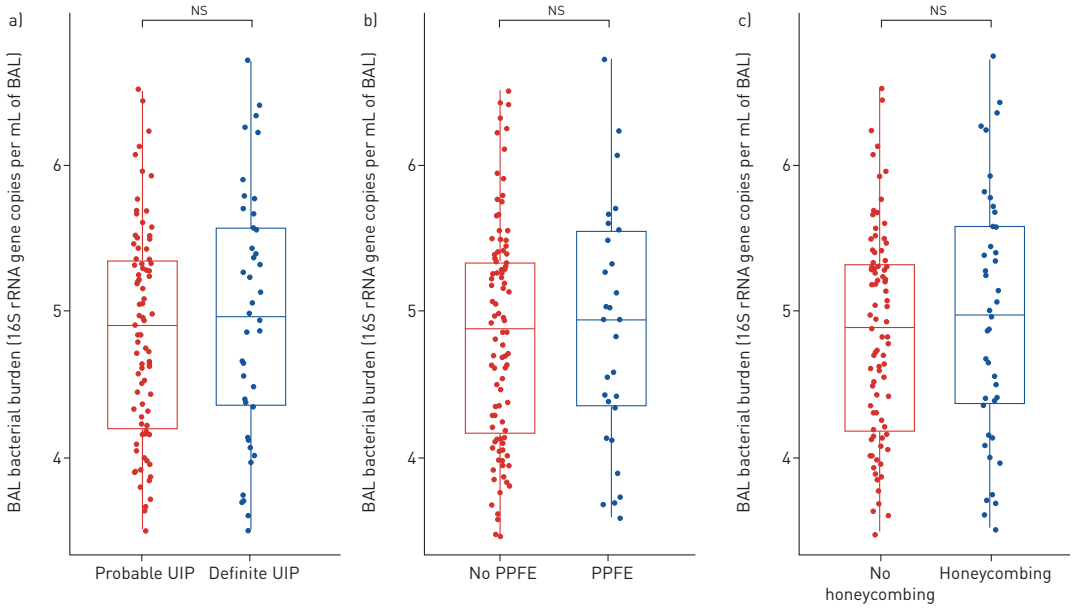


FIGURE 3 Relationship between radiological features and bronchoalveolar lavage (BAL) bacterial burden. No difference in bacterial burden log10 (16S rRNA gene copies·mL^{-1} of BAL) based on (a) a definite or probable pattern of usual interstitial pneumonia (UIP), (b) the presence or absence of radiological pleuroparenchymal fibroelastosis (PPFE) or (c) the presence or absence of honeycombing. Statistical significance tested with Mann–Whitney test.

burden in the lower airways in IPF is magnitudes lower than that seen in cystic fibrosis or bronchiectasis, it is elevated compared with health. The quantity and composition of the respiratory microbiome is influenced by the balance between the rate of immigration and clearance. The primary routes of immigration are microaspiration and inhalation of airborne bacteria while elimination is achieved through a combination of the cough response, mucociliary clearance and the innate and adaptive immune response. In disease, the balance between these two opposing factors is lost, resulting in dysbiosis [24]. Given the strong association with GORD and IPF, microaspiration is likely to be a large driver in this cohort of patients [25]. Indeed, gastric bacteria have been identified in the respiratory microbiome supporting this [12]. However, it is not only immigration which may be abnormal in IPF, given the widespread abnormalities in the mucin genetic architecture [26–28]. The mucin glycoproteins are a major structural component of the mucus barrier, maintaining the hydration of the airway epithelium and crucially entrapping particles for removal by mucociliary clearance. In mice, *MUC5B* appears essential for normal macrophage function and effective mucociliary clearance of bacteria [29]. This leaves the hypothesis that a loss of balance between bacterial immigration and clearance results in the increased bacterial burden in the lower airways resulting in prolonged and repetitive exposure, triggering an exaggerated interstitial injury and eventually leading to the development of fibrosis.

Despite the increased microbial burden, there was no host response at a cellular level seen when examining BAL differential cell counts. There are conflicting data regarding the prognostic implications of BAL neutrophilia in IPF. The previous largest study to date demonstrated an increased risk of mortality at 12 months with increased BAL neutrophil counts, but this association was lost after a year, leading the investigators to hypothesise they may be identifying a subgroup of individuals with more active disease [30]. Here, we demonstrate the presence of increased neutrophils in the airways of patients with IPF but no association with their counts and overall bacterial burden. We also find no association between BAL cellular profiles and the bacterial burden in the airways, highlighting the difference between the stable microbiome and periods of infection, in which a cellular response is seen. This conforms with the theory that repetitive microinjuries to the epithelium trigger the disease, and you would not expect such a large cellular response. Therefore, more nuanced methods of detecting response, such as characterisation of the epithelial and airway transcriptome, may be required and should be included in any future studies.

While traction bronchiectasis and honeycombing are clearly defined and distinct entities, radiologically, they are likely to represent two aspects of the continuous spectrum of pathological remodelling that occurs in IPF. Both features have been associated with disease behaviour and clinical outcome, as has the extent of overall fibrosis [31]. There was, however, no correlation between the extent of honeycombing or traction bronchiectasis and bacterial burden, which remains a predictor of progression even accounting for these. A number of the subjects had radiology indeterminant for UIP and the multidisciplinary meeting had access to clinical data, serology and BAL cell differential counts to support the diagnosis of IPF. The relative proportions of CT UIP morphologies (definite, probable, indeterminate or inconsistent) in our cohort of MDT IPF diagnoses are similar to previously published proportions [32]. There were no links with extent of fibrosis or traction bronchiectasis, nor differences between the radiological diagnostic cohorts or probable *versus* definite UIP. We also considered if concomitant disease cofounders could have been driving this burden and given the clinical association between PPFE and infection, this was specifically interrogated. Almost one-third of the cohort had radiological evidence of PPFE and there was no difference in bacterial burden between those with and without disease or when simply considering disease extent. Although this is an interesting finding, the number of subjects with PPFE were small and this lack of association will have to be validated in larger cohorts, in which it would also be interesting to examine the microbiome in subjects with isolated PPFE.

Recently the Correlating Outcomes With Biochemical Markers to Estimate Time-progression in IPF (COMET) authors sought to examine the relationship between the microbiome and the presence or absence of honeycombing in a subgroup of 68 subjects with IPF [33]. Simply dichotomising the cohort based on the presence or absence of honeycombing they found no association with the lower airways bacterial burden. The prospective nature and quantitative radiological analysis undertaken here enabled us to build and expand upon their work. Our study design specifically allowed us to answer questions about the graded relationships between radiological change and bacterial burden. This has enabled us to undertake a far more detailed and nuanced analysis and strengthens our conclusions that the increased bacterial burden is not simply related to the extent of radiological disease.

Our work [34], and that of O'Dwyer *et al.* [33], also raises a number of issues related to the potential for antimicrobial therapies in IPF [34, 35], some of which are already being trialled (NCT02759120 and 17464641). While there is clear microbial dysbiosis in the lower airways in patients with IPF, the only consistent finding has been that of an increased bacterial burden. None of the therapies being trialled at present has ever demonstrated efficacy at reducing bacterial burden. Cautionary tales from the

bronchiectasis field must be heeded, as heterogeneous responses to antibiotics have led to multiple antibiotic trials failing to meet their primary endpoints. Indeed, recently Sibila *et al.* [36] have demonstrated, in a *post hoc* analysis of multiple trials of inhaled antibiotics, that patients with higher bacterial burden consistently demonstrated a clinical response and improved quality of life. It is therefore conceivable that a more targeted approach may also be required for patients with IPF rather than simply treating indiscriminately. However, this study showed no evidence that bacterial burden can be predicted from radiology or differential cell counts.

This work has a number of limitations. First, BAL was undertaken in the right middle lobe, while radiographic extent of disease was assessed globally. This standardised approach was undertaken in line with current diagnostic and research standards; however, we appreciate an in-depth geographical survey of multiple lobes may have provided more information and should be considered in any future studies. This would also provide the perfect opportunity to assess the regional differences in both bacterial burden and microbial composition in fibrotic lung disease. This study was designed to investigate the relationship between bacterial burden and radiological disease extent and we did not undertake any characterisation of the bacterial communities. While, to date, increased bacterial burden is the most consistent microbiome-related finding in IPF this may simply reflect the ease and reproducibility of this measurement, compared with more complex 16S amplicon analysis. Larger cohorts and more standardised sequencing/analysis pipelines [37] (allowing meta-analysis) will be required to reveal more consistent relationships between specific community composition and clinical outcomes. Given the known association between the single nucleotide polymorphism rs35705950 of the *MUC5B* mucin gene [38], these prospective studies should also factor in the effect of genotype on bacterial burden. Another limitation is that radiology interpretation of CT forms a key part of the MDT, potentially confounding any study on CT findings. We attempted to minimise this impact by ensuring that at least one scoring radiologist was independent from the MDT involved in making the final diagnosis of IPF.

While our hypothesis was that that architectural destruction acts as a reservoir to increase bacterial burden, an alternate exists where the increased bacterial burden may indeed cause the architectural destruction. Our data are not supportive of either of these two scenarios but longitudinal and mechanistic studies will be required to support this. There is a known association between GORD reflux and IPF. Many patients are on PPI which have been linked to an increased risk of pneumonia [21, 39]. While we did not set out to address this issue prospectively, we demonstrated no association between GORD or PPI use and bacterial burden. Though more formal prospective studies with objective measurements of acid and non-acid reflux will be required to answer this question directly.

We demonstrate that an elevated bacterial burden predicts mortality in IPF but is independent of radiological features, differential cell counts and extent of disease. We would argue that the independent nature of this association supports a relationship with the underlying pathogenetic mechanisms. However, functional studies are urgently needed, especially as interventional studies are already enrolling and the potential for bacterial load to guide antibiotic therapy for patients with IPF should be prospectively tested.

Conflict of interest: R. Invernizzi has nothing to disclose. J. Barnett has nothing to disclose. B. Rawal has nothing to disclose. A. Nair has nothing to disclose. P. Ghai has nothing to disclose. S. Kingston has nothing to disclose. F. Chua has nothing to disclose. Z. Wu has nothing to disclose. A.U. Wells reports speakers and consultancy fees from Intermune Roche, Boehringer Ingelheim and Bayer. E.R. Renzoni has received speaker fees from BI, Roche and Mundipharma. A.G. Nicholson reports personal fees for consultancy from Medical Quantitative Image Analysis and Sanofi, personal fees for consultancy and lectures from Boehringer Ingelheim, personal fees for lectures from Roche, outside the submitted work. A. Rice is a shareholder in Pfizer. C.M. Lloyd has nothing to disclose. A.J. Byrne has nothing to disclose. T.M. Maher has, *via* his institution, received industry-academic funding from GlaxoSmithKline R&D and UCB, and has received consultancy or speakers fees from Apellis, AstraZeneca, Bayer, Blade Therapeutics, Boehringer Ingelheim, Bristol-Myers Squibb, Galapagos, GlaxoSmithKline R&D, Indalo, Novartis, Pliant, ProMetic, Respivnat, Roche, Samumed and UCB. A. Devaraj reports personal fees from GSK, Roche and Boehringer Ingelheim, outside the submitted work. P.L. Molyneaux has, *via* his institution, received industry-academic funding from Roche, Boehringer Ingelheim and Galapagos, and has received speaker fees from Roche.

Support statement: P.L. Molyneaux is supported by an Action for Pulmonary Fibrosis Mike Bray fellowship. T.M. Maher is supported by a National Institute for Health Research Clinician Scientist Fellowship (NIHR ref: CS-2013-13-017) and is a British Lung Foundation Chair in Respiratory Research (C17-3). C.M. Lloyd is supported by a Wellcome Senior Fellowship in Basic Biomedical Science (107059/Z/15/Z). A.J. Byrne is supported by a Wellcome Trust award (205949/Z/17/Z) and an Asthma UK senior fellowship (AUK-SNF-2017-381). Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med* 2018; 378: 1811–1823.
2 Molyneaux PL, Maher TM. Respiratory microbiome in IPF: cause, effect, or biomarker? *Lancet Respir Med* 2014; 2600: 10–11.
3 Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2013; 22: 376–381.

4 O'Dwyer DN, Moore BB, Molyneaux PL. Interstitial lung disease. *In*: Cox MJ, Ege MJ, von Mutius E, eds. The Lung Microbiome (ERS Monograph). Sheffield, European Respiratory Society, 2019; pp. 173–187.
5 Faner R, Sibila O, Agustí A, *et al.* The microbiome in respiratory medicine: Current challenges and future perspectives. *Eur Respir J* 2017; 49: 1602086.
6 Kim DS, Park JH, Park BK, *et al.* Acute exacerbation of idiopathic pulmonary fibrosis: Frequency and clinical features. *Eur Respir J* 2006; 27: 143–150.
7 Molyneaux PL, Cox MJ, Willis-Owen SAG, *et al.* The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014; 190: 906–913.
8 Han MKLK, Zhou Y, Murray S, *et al.* Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014; 2: 548–556.
9 Molyneaux PL, Willis-Owen SAG, Cox MJ, *et al.* Host–microbial interactions in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017; 195: 1640–1650.
10 Chan P, Bax L, Chen C, *et al.* Model-based meta-analysis on the efficacy of pharmacological treatments for idiopathic pulmonary fibrosis. *CPT Pharmacometrics Syst Pharmacol* 2017; 6: 695–704.
11 O'Dwyer DN, Ashley SL, Gurczynski SJ, *et al.* Lung microbiota contribute to pulmonary inflammation and disease progression in pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 199: 1127–1138.
12 Molyneaux PL, Cox MJ, Wells AU, *et al.* Changes in the respiratory microbiome during acute exacerbations of idiopathic pulmonary fibrosis. *Respir Res* 2017; 18: 29.
13 Raghu G, Remy-Jardin M, Myers JLL, *et al.* Diagnosis of idiopathic pulmonary fibrosis. an official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018; 198: e44–e68.
14 Goh NSL, Veeraraghavan S, Desai SR, *et al.* Bronchoalveolar lavage cellular profiles in patients with systemic sclerosis-associated interstitial lung disease are not predictive of disease progression. *Arthritis Rheum* 2007; 56: 2005–2012.
15 Hansell DM, Bankier AA, MacMahon H, *et al.* Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008; 246: 697–722.
16 Jacob J, Odink A, Brun AL, *et al.* Functional associations of pleuroparenchymal fibroelastosis and emphysema with hypersensitivity pneumonitis. *Respir Med* 2018; 138: 95–101.
17 Jacob J, Bartholmai BJ, Rajagopalan S, *et al.* Predicting outcomes in idiopathic pulmonary fibrosis using automated CT analysis. *Am J Respir Crit Care Med* 2018; 198: 767–776.
18 Molyneaux PL, Mallia P, Cox MJ, *et al.* Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013; 188: 1224–1231.
19 Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2018. www.r-project.org/
20 Walsh SLF, Calandriello L, Sverzellati N, *et al.* Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax* 2016; 71: 45–51.
21 Bhat M, Pasini E, Copeland J, *et al.* Impact of immunosuppression on the metagenomic composition of the intestinal microbiome: a systems biology approach to post-transplant diabetes. *Sci Rep* 2017; 7: 1–12.
22 Reddy TL, Tominaga M, Hansell DM, *et al.* Pleuroparenchymal fibroelastosis: A spectrum of histopathological and imaging phenotypes. *Eur Respir J* 2012; 40: 377–385.
23 van Manen MJG, Birring SS, Vancheri C, *et al.* Cough in idiopathic pulmonary fibrosis. *Eur Respir Rev* 2016; 25: 278–286.
24 Dickson RP, Erb-Downward JR, Freeman CM, *et al.* Bacterial topography of the healthy human lower respiratory tract. *MBio* 2017; 8: 1–12.
25 Mackintosh JA, Desai SR, Adamali H, *et al.* In patients with idiopathic pulmonary fibrosis the presence of hiatus hernia is associated with disease progression and mortality. *Eur Respir J* 2019; 53: 1802412.
26 Zhang Y, Noth I, Garcia JGN, *et al.* A variant in the promoter of *MUC5B* and idiopathic pulmonary fibrosis nt5e mutations and arterial calcifications. *N Engl J Med* 2011; 364: 1576–1577.
27 Allen RJ, Porte J, Braybrooke R, *et al.* Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med* 2017; 5: 869–880.
28 Seibold MA, Wise AL, Speer MC, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
29 Roy MG, Livraghi-Butrico A, Fletcher AA, *et al. Muc5b* is required for airway defence. *Nature* 2014; 505: 412–416.
30 Kinder BW, Brown KK, Schwarz MI, *et al.* Baseline BAL neutrophilia predicts early mortality in idiopathic pulmonary fibrosis. *Chest* 2008; 133: 226–232.
31 Robbie H, Daccord C, Chua F, *et al.* Evaluating disease severity in idiopathic pulmonary fibrosis. *Eur Respir Rev* 2017; 26: 170051.
32 Yagihashi K, Huckleberry J, Colby TV, *et al.* Radiologic–pathologic discordance in biopsy-proven usual interstitial pneumonia. *Eur Respir J* 2016; 47: 1189–1197.
33 Dickson RP, Huffnagle GB, Flaherty KR, *et al.* Radiographic honeycombing and altered lung microbiota in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 200: 1544.
34 Macaluso C, Maritano Furcada J, Alzaher O, *et al.* The potential impact of azithromycin in idiopathic pulmonary fibrosis. *Eur Respir J* 2019; 53: 1800628.
35 Shulgina L, Cahn AP, Chilvers ER, *et al.* Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: a randomised controlled trial. *Thorax* 2013; 68: 155–162.
36 Sibila O, Laserna E, Shoemark A, *et al.* Airway bacterial load and inhaled antibiotic response in bronchiectasis. *Am J Respir Crit Care Med* 2019; 200: 33–41.
37 Carney SM, Clemente JC, Cox MJ, *et al.* Methods in lung microbiome research. *Am J Respir Cell Mol Biol* 2020; 62: 283–299.
38 Stock CJ, Sato H, Fonseca C, *et al.* Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. *Thorax* 2013; 68: 436–441.
39 Zirk-Sadowski J, Masoli JA, Delgado J, *et al.* Proton-pump inhibitors and long-term risk of community-acquired pneumonia in older adults. *J Am Geriatr Soc* 2018; 66: 1332–1338.

ORIGINAL ARTICLE

The Respiratory Microbiome in Chronic Hypersensitivity Pneumonitis Is Distinct from That of Idiopathic Pulmonary Fibrosis

Rachele Invernizzi[1], Benjamin G. Wu[2], Joseph Barnett[1,3], Poonam Ghai[1], Shaun Kingston[1,3], Richard J. Hewitt[1,3], Johanna Feary[1,3], Yonghua Li[2], Felix Chua[3], Zhe Wu[1,3], Athol U. Wells[3], Peter M. George[3], Elisabetta A. Renzoni[3], Andrew G. Nicholson[1,3], Alexandra Rice[3], Anand Devaraj[1,3], Leopoldo N. Segal[2], Adam J. Byrne[1], Toby M. Maher[1]*, Clare M. Lloyd[1], and Philip L. Molyneaux[1,3]

[1]National Heart and Lung Institute, Imperial College London, London, United Kingdom; [2]Division of Pulmonary and Critical Care Medicine, New York University, New York, New York; and [3]Royal Brompton Hospital, London, United Kingdom

ORCID IDs: 0000-0002-6957-9626 (R.I.); 0000-0003-4294-0052 (R.J.H.).

Abstract

Rationale: Chronic hypersensitivity pneumonitis (CHP) is a condition that arises after repeated exposure and sensitization to inhaled antigens. The lung microbiome is increasingly implicated in respiratory disease, but, to date, no study has investigated the composition of microbial communities in the lower airways in CHP.

Objectives: To characterize and compare the airway microbiome in subjects with CHP, subjects with idiopathic pulmonary fibrosis (IPF), and control subjects.

Methods: We prospectively recruited individuals with a CHP diagnosis ($n = 110$), individuals with an IPF diagnosis ($n = 45$), and control subjects ($n = 28$). Subjects underwent BAL and bacterial DNA was isolated, quantified by quantitative PCR and the 16S ribosomal RNA gene was sequenced to characterize the bacterial communities in the lower airways.

Measurements and Main Results: Distinct differences in the microbial profiles were evident in the lower airways of subjects with CHP and IPF. At the phylum level, the prevailing microbiota of both subjects with IPF and subjects with CHP included *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria*. However, in IPF, *Firmicutes* dominated, whereas the percentage of reads assigned to *Proteobacteria* in the same group was significantly lower than the percentage found in subjects with CHP. At the genus level, the *Staphylococcus* burden was increased in CHP, and *Actinomyces* and *Veillonella* burdens were increased in IPF. The lower airway bacterial burden in subjects with CHP was higher than that in control subjects but lower than that of those with IPF. In contrast to IPF, there was no association between bacterial burden and survival in CHP.

Conclusions: The microbial profile of the lower airways in subjects with CHP is distinct from that of IPF, and, notably, the bacterial burden in individuals with CHP fails to predict survival.

Keywords: lung microbiota; 16S; fibrosis

(Received in original form February 28, 2020; accepted in final form July 21, 2020)

This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (http://creativecommons.org/licenses/by-nc-nd/4.0/). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

*T.M.M. is Associate Editor of *AJRCCM*. His participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works.

Supported by an Action for Pulmonary Fibrosis Mike Bray Fellowship (P.L.M.), a National Institute for Health Research Clinician Scientist Fellowship (CS-2013-13-017 to T.M.M.), a British Lung Foundation Chair in Respiratory Research grant (C17-3 to T.M.M.), a Wellcome Senior Fellowship in Basic Biomedical Science (107059/Z/15/Z to C.M.L.), a Wellcome Trust award (205949/Z/17/Z to A.J.B.), and an Asthma UK senior fellowship (AUK-SNF-2017-381 to A.J.B.).

Author Contributions: Conception and design: R.I. and P.L.M. Acquisition of data and recruitment: J.B., R.J.H., J.F., F.C., Z.W., A.U.W., E.A.R., A.G.N., A.R., A.D., and P.L.M. Analysis and interpretation: R.I., B.G.W., P.G., S.K., Y.L., L.N.S., A.J.B., T.M.M., C.M.L., and P.L.M. Drafting the manuscript: R.I. and P.L.M. Editing manuscript: R.I., B.G.W., F.C., A.G.N., A.R., L.N.S., A.J.B., T.M.M., C.M.L., and P.L.M. Final approval: all authors.

Correspondence and requests for reprints should be addressed to Philip L. Molyneaux, M.B. B.S., Ph.D., National Heart and Lung Institute, Imperial College London, London, SW7 2AZ, UK. E-mail: p.molyneaux@imperial.ac.uk.

This article has a related editorial.

This article has an online supplement, which is accessible from the issue's table of contents at www.atsjournals.org.

Am J Respir Crit Care Med Vol ■■, Iss ■■, pp 1–9, ■■ ■■, 2020
Copyright © 2020 by the American Thoracic Society
Originally Published in Press as DOI: 10.1164/rccm.202002-0460OC on July 21, 2020
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: Numerous studies of the respiratory microbiome in idiopathic pulmonary fibrosis (IPF) have reported associations between changes in the bacterial burden, microbial communities, disease progression, and mortality. Chronic hypersensitivity pneumonitis (CHP) is a pathogenically and prognostically distinct disorder compared with IPF but is still characterized phenotypically by irreversible fibrotic destruction of the lung parenchyma. Despite the phenotypic overlap, no study has examined whether the respiratory microbiota differs between individuals with these conditions.

What This Study Adds to the Field: Despite the phenotypic overlap between CHP and IPF, there are clear differences in the bacterial burden and microbial composition in the lower airways of patients with CHP compared with healthy subjects and patients with IPF. In CHP there are increased proportions of *Proteobacteria* and lower proportions of *Firmicutes* compared with subjects with IPF. A higher bacterial burden has been consistently demonstrated in the lower airways in IPF and is associated with an increased risk of mortality. However, the bacterial burden in subjects with CHP was far lower than in those with IPF and did not associate with survival. Our findings suggest that the observed alterations in the lung microbiome are disease specific and do not simply reflect the presence of fibrosis within the lung.

Hypersensitivity pneumonitis is an immune-mediated interstitial lung disease that develops in genetically susceptible individuals after repeated inhalation of organic antigens, including fungal, bacterial, animal, and insect proteins (1). It can be classified into acute and chronic forms, depending on the nature of the inciting antigen, the intensity and duration of exposure, host factors, and radiological findings (2). Chronic hypersensitivity pneumonitis (CHP) carries the most significant morbidity and mortality. It arises from protracted exposure to a triggering antigen, which initially elicits inflammation and ultimately evolves into irreversible and often progressive interstitial fibrosis (3, 4). As yet, there are no established international guidelines for the diagnosis of CHP, and it remains unclear why only some individuals exposed to recognized triggers develop the disease (5). This diagnostic uncertainty is further confounded in advanced disease because the clinical features of CHP and idiopathic pulmonary fibrosis (IPF) are often indistinguishable (6).

The existence of a microbiome in the healthy lung is generally accepted, and its potential role in respiratory pathology is increasingly appreciated (7–9).The advent of next-generation sequencing technologies has allowed direct and culture-independent sequence-based interrogation of microbial communities in both homeostatic and perturbed states (10, 11). Changes in the respiratory microbiome have been associated with the clinical, physiological, and therapeutic aspects of chronic lung diseases such as asthma, bronchiectasis, chronic obstructive pulmonary disease (COPD), and IPF (12–18). Studies of the respiratory microbiome in IPF have reported associations between changes in microbial communities and disease progression and mortality (16, 17, 19, 20). A higher bacterial burden has been reported during acute exacerbations of IPF (21, 22) and is consistently associated with an increased risk of mortality (16–18). Differences in bacterial community composition and diversity have also been correlated with disease outcomes, peripheral blood host-defense gene expression (19), and indices of alveolar inflammation (18). These observational studies support the hypothesis that the respiratory microbiome may be implicated in the pathogenesis of fibrotic lung disease (23, 24).

CHP and IPF are pathogenically and prognostically distinct disorders that are both characterized by irreversible fibrotic destruction of the lung parenchyma; we set out to test the hypothesis that observed alterations in the lung microbiome are disease specific and do not simply reflect the presence of fibrosis within the lung. In this study, we prospectively recruited individuals with CHP and IPF diagnoses as well as healthy control subjects to characterize the bacterial communities in the lower airways by 16S ribosomal (r)RNA gene sequencing. Despite the phenotypic overlap between CHP and IPF, we demonstrate that there are clear differences in patients with CHP in the microbial composition and bacterial burden of the lower airways and in survival, compared with healthy subjects and patients with IPF.

Methods

Study Design

Treatment-naive subjects were prospectively recruited at the Royal Brompton Hospital (London) between May 2014 and May 2018. A diagnosis of IPF was made after multidisciplinary team discussion, according to the latest international guidelines (25). A consensus diagnosis of CHP was made using multidisciplinary team discussion with integration of clinical, radiological, BAL, and pathological data (5, 26). Pulmonary function testing was performed, clinical measurements were recorded, and subjects underwent fiberoptic bronchoscopy with BAL via the oropharyngeal route as part of standard care according to a standard operating procedure (27). BAL cell profiling was performed by assessing leukocyte differential counts. Computed tomographic data were scored as previously described (28). Approval for the study was obtained from the local research ethics committee (15/SC/0101 and 15-LO-1399), and written informed consent was obtained from all subjects. The IPF cohort was not used in any prior studies. *See* the online supplement for further details.

Bacterial DNA Isolation and 16S rRNA Amplicon Sequencing

Genomic DNA was extracted from BAL cell pellets as previously described, with minor modifications (29). Amplification of the V4 hypervariable region of the 16S rRNA gene was performed using the barcoded universal primer pair 515F/806R and sequenced using an Illumina MiSeq instrument (Illumina) to produce 150 base-paired reads. For full details, *see* the online supplement.

Bioinformatics Analysis

Reads were analyzed using the Quantitative Insights into Microbial Ecology 2 (version 2018.8) bioinformatics pipeline (30). Bioinformatics processing included demultiplexing, denoising, removal of chimeric and short reads, 97% identity binning of reads into amplicon sequence variants (ASV) using Divisive Amplicon Denoising Algorithm 2 (31), and taxonomic annotation using the Greengenes database (32). Absolute ASV sequence counts were normalized to obtain the relative abundances of the microbiota in each sample. After abundance filtering, a rarefied data set was generated and used for downstream analyses. For details on processing and quality control, *see* the online supplement.

16S rRNA Gene Quantitative PCR

Triplicate 10-μl quantitative PCR (qPCR) reactions were set up containing 1 μl of bacterial DNA and 9 μl of Femto bacterial qPCR premix (Cambridge Bioscience). Each run contained a 10-fold dilution series of the *Vibrio natriegens DSM 759* gene cloned into a plasmid of known size and a nontemplate control (28). For a description of cycling parameters, *see* the METHODS section of the online supplement.

Statistical Analysis

All data analysis was performed in R (version 3.4.3; R Foundation for Statistical Computing) (33). Differences in BAL differential cell counts, bacterial burden, and taxonomic composition among subjects with CHP, subjects with IPF, and control subjects were statistically tested by using the Kruskal-Wallis test and the Dunn test for multiple comparisons, which was adjusted using the Bonferroni method. Univariate linear regressions between bacterial burden and BAL differential cell counts were performed and adjusted for age, sex, smoking status, and FVC using a multivariable linear regression model. Changes in microbial biodiversity and richness between subjects with CHP and subjects with IPF were evaluated using the Shannon and Chao1 α-diversity indexes and were tested using the Mann-Whitney test. The β-diversity of subjects with CHP and subjects with IPF was restricted to ASV that were present at >1% of the sample population, calculated using the Euclidean distance and visualized by principal component analysis (PCA), with the significance of differences in community composition being determined by permutational multivariate ANOVA adjusted for age, sex, smoking status, FVC (% predicted), and $D_{L_{CO}}$ (% predicted). The linear discriminant analysis effect size was used to identify differentially abundant bacteria among the groups. To study CHP and IPF disease progression, a Kaplan-Meier curve was assessed using a multivariate Cox proportional-hazard model, adjusted for age, sex, smoking status, FVC (% predicted), and percentage of honeycombing. Survival time was defined as the time from diagnosis to death, loss to follow-up, or the end of the study period. In all analyses, a P value < 0.05 was considered statistically significant.

Data Availability

Sequences are available via the National Center for Biotechnology Information Sequence Read Archive (accession number PRJNA609242). Codes and metadata used for analysis are available at https://github.com/molyneaux-lab.

Results

Subjects, Sampling, and Sequencing

One hundred ten subjects with CHP, 45 subjects with IPF, and 28 control subjects were enrolled in the study and underwent bronchoscopy (Table 1). The demographics demonstrate a number of expected differences based on the known sex, smoking, and clinical profiles of both IPF and CHP (34). Subjects with IPF were predominantly male (82%), with a mean ± SD age of 62 ± 19 years and moderately severe restrictive lung disease (FVC, 75% predicted ± 19% predicted; Hb-adjusted $D_{L_{CO}}$, 44% predicted ± 17% predicted). There were more female subjects in the CHP cohort, and although their mean age (66 ± 9 years) and $D_{L_{CO}}$ (44% predicted ± 16% predicted) were matched, they had a more preserved FVC compared with the IPF cohort (85% predicted ± 23% predicted). The median follow-up in the CHP cohort was 919 days, and the median follow-up in the IPF cohort was 744 days (not significant). Although alveolar macrophages predominated in the BAL cells of all groups, a mixed alveolitis with predominant neutrophilia was observed in subjects with IPF (18% ± 6% neutrophils, 12% ± 5% lymphocytes), whereas subjects with CHP demonstrated an elevated lymphocyte count (32% ± 15% lymphocytes). Radiologically, the subjects with CHP had more ground-glass and low-density lobules and less honeycombing than subjects with IPF. In keeping with the literature, over half (52%) of the patients with CHP had no obvious exposure history. The majority of exposures in the CHP cohort were mold-related (20%), avian (15.45%) or feather down–related (5.45%). In contrast, only eight patients with IPF had an obvious exposure history (17.7%) (Table 1). All 183 samples yielded genomic DNA amplicons and underwent sequencing of the V4 region of the 16S rRNA gene. After demultiplexing, denoising, quality control, and abundance filtering, 8,406,014 sequences were retained, representing 495 unique ASV across 183 samples (35). This final curated data set was used for all subsequent analyses.

Bacterial Burden

The strongest and most robust microbial signature associated with fibrotic lung disease to date is that of the bacterial burden. We and others have already shown that IPF is characterized by an increased bacterial burden in BAL cells, which is associated with disease progression. We therefore initially set out to explore differences in the BAL bacterial burden among control subjects and subjects with CHP and IPF. On average, subjects with IPF had 1.54×10^5 copies of the 16S rRNA gene per milliliter of BAL fluid, which was higher than the average copy number in control subjects (8.57×10^3 copies/ml; $P < 0.00001$). Although subjects with CHP had a higher bacterial load (1.14×10^4) compared with healthy control subjects ($P < 0.0001$), they had a significantly lower burden compared with subjects with IPF ($P < 0.00001$) (Figure 1). No correlation was found between the bacterial load and the BAL differential cell counts (*see* Figure E1 in the online supplement). The negative control samples yielded a bacterial burden close to or below the lower limit of qPCR quantification (100 copies/ml of BAL fluid) (Figure E2).

Survival Analysis

The differences in survival between subjects with CHP and IPF were explored using a multivariate Cox regression analysis, and subjects with CHP were found to survive

Table 1. Baseline Characteristics of Study Subjects

Parameter	Control Subjects (*n* = 28)	CHP (*n* = 110)	IPF (*n* = 45)	*P* Value*
Age, yr	55 ± 15	66 ± 9	62 ± 19	NS
Sex, F, *n* (%)	7 (25)	54 (49)	8 (18)	<0.001
Baseline PFT				
FVC, % predicted	94 ± 18	85 ± 23	75 ± 19	<0.01
FEV_1, % predicted	73 ± 23	86 ± 22	78 ± 17	<0.05
D_{LCO}, % predicted	68 ± 32	44 ± 16	44 ± 17	NS
Smoking status†				
Nonsmoker, *n* (%)	7 (25)	56 (51)	15 (33)	—
Ex-smoker, *n* (%)	10 (36)	54 (49)	26 (58)	NS
Current smoker, *n* (%)	1 (4)	0 (0)	4 (9)	—
BAL cell differentials				
Macrophages, %	70 ± 16	62 ± 15	60 ± 8	NS
Neutrophils, %	3 ± 2	10 ± 9	18 ± 6	<0.0001
Lymphocytes, %	10 ± 6	32 ± 15	12 ± 5	<0.0001
Radiology‡				
Fibrosis present, *n* (%)	NA	98 (89)	44 (100)	<0.05
UIP, *n* (%)	NA	31 (28)	35 (78)	<0.001
Honeycombing, *n* (%)	NA	10 (9)	13 (29)	<0.01
GG extent %, median (IQR)	NA	10 (14)	5 (10)	<0.01
Low-density lobule extent, *n* (%)	NA	5 (12)	5 (8)	<0.05
Antigen exposure§				
Avian, *n* (%)	NA	17 (15)	1 (2)	<0.05
Feather down, *n* (%)	NA	6 (5)	1 (2)	NS
Mold, *n* (%)	NA	22 (20)	3 (7)	NS
Other, *n* (%)	NA	8 (7)	3 (7)	NS
None, *n* (%)	NA	57 (52)	37 (82)	<0.001

Definition of abbreviations: CHP = chronic hypersensitivity pneumonitis; GG = ground glass; IPF = idiopathic pulmonary fibrosis; IQR = interquartile range; NA = not applicable; NS = not significant; PFT = pulmonary function test; UIP = usual interstitial pneumonia.
Data are presented as the mean ± SD unless otherwise noted.
*P values were calculated by using the Mann-Whitney *U* test or the Fisher exact test between subjects with CHP and subjects with IPF.
†Smoking status: statistical testing by nonsmoker versus ex/current smoker; data were available for 18 of 28 control subjects.
‡Fibrosis was defined as the presence of traction bronchiectasis and/or honeycombing. UIP was defined as "definite" or "probable" UIP. Radiology data for low-density lobules and ground glass were scored to the nearest 5% on a lobar basis, with the lingua as a sixth lobe. The final score represents the mean of the lobar scores.
§Antigen exposure: statistical testing by exposure versus no exposure.

longer ($P = 0.0001$) than individuals with IPF, even after adjustment of confounding factors: age, sex, smoking history, baseline % predicted FVC, and percentage of honeycombing (Figure 2 and Table E1). Next, using a univariate analysis, we assessed whether there was an association between bacterial burden and survival. No associations were found between bacterial burden and survival in subjects with CHP (hazard ratio, 1.35; 95% confidence interval, 0.92–1.99; $P = 0.128$), whereas burden was associated with mortality in subjects with IPF (hazard ratio, 1.92; 95% confidence interval, 1.32–2.79; $P = 0.000647$). This association remained significant in analysis of IPF using a Cox proportional hazards model adjusted for age, sex, smoking status, and baseline % predicted FVC (hazard ratio, 3.23; 95% confidence interval, 1.64–6.34; $P < 0.001$). This is consistent with published data, which have shown that patients with IPF have an increased pulmonary bacterial load, the magnitude of which predicts the rate of progression of IPF and risk of mortality (16, 18). Collectively, these results suggest that the association between increased bacterial burden and survival is specific to IPF and is not seen in CHP.

Description of BAL Microbiota Composition in Health and Disease

Across all cohorts, the composition of the microbial communities was dominated by five most abundant phyla: *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, and *Fusobacteria* (Figure 3A). At the phylum level, the microbiota of subjects with CHP was dominated by *Firmicutes* (38%), *Bacteroidetes* (24%), *Proteobacteria* (23%), and *Actinobacteria* (9%). Although the same phyla predominated in the subjects with IPF, there were a number of differences. In IPF, *Firmicutes* accounted for over 43% of the total reads, and the percentage of reads assigned to *Proteobacteria* was significantly lower compared with the percentage assigned in subjects with CHP (15% and 23%, respectively; $P = 0.04$). No significant differences were observed when comparing individuals with CHP with control subjects (Figure E3); in contrast, an increased abundance of *Firmicutes* (43%; $P = 0.04$) and *Actinobacteria* (11%; $P = 0.04$) was noted in subjects with IPF compared with control subjects (Figure E3). Comparison at the genus level identified that the microbial landscape of all groups was dominated by *Prevotella* (control subjects, 24%; subjects with CHP, 23%; subjects with IPF, 25%), *Veillonella* (control subjects, 13%; subjects with CHP, 14%; subjects with IPF, 18%), and *Streptococcus* (control subjects, 19%;

4C/FPO

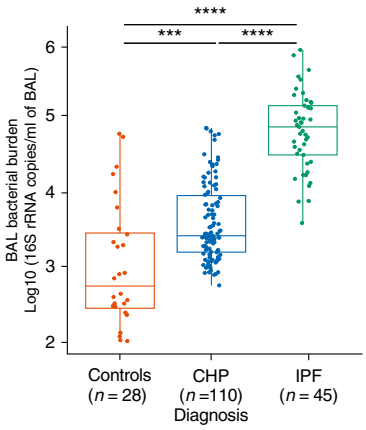


Figure 1. Bacterial burden in BAL fluid of control subjects, subjects with CHP, and subjects with IPF. The Kruskal-Wallis test with the Dunn test for multiple comparisons was used ($P < 2.2 \times 10^{-16}$). Data are presented as the median and interquartile range (***$P < 0.0001$ and ****$P < 0.00001$). The bacterial burden was calculated by using quantitative PCR and was expressed in $\log_{10}$ values (16S ribosomal RNA gene copies/ml of BAL fluid). CHP = chronic hypersensitivity pneumonitis; IPF = idiopathic pulmonary fibrosis.

subjects with CHP, 21%; subjects with IPF, 18%) (Figure 3A). Of these genera, *Veillonella* (18%) was found to be most abundant in the BAL cells of subjects with IPF compared with the BAL cells of control subjects (13%; $P = 0.02$; Figure E4) and subjects with CHP (14%; $P = 0.02$). Within the *Actinobacteria* phylum, there was a lower relative abundance of *Actinomyces* (4%) in subjects with CHP compared with individuals with IPF (6%; $P = 0.001$) but not compared with control subjects (5%; $P = 0.95$; Figure E4). *Staphylococcus* was found to be increased in subjects with an IPF diagnosis (0.5%) compared with control subjects (0.1%; $P = 0.03$; Figure E4) but was most abundant in the CHP cohort (2%; $P = 0.0005$, compared with IPF cohort) (Figure 3A). Given the previous association between disease progression and an increased relative abundance of *Streptococcus* and *Staphylococcus* in IPF, we investigated whether these bacterial genera were associated with an increased risk of mortality in CHP and IPF. Using a univariate analysis, we found that both *Streptococcus* (hazard ratio, 2.07; 95% confidence interval, 0.83–5.16; $P = 0.030$) and *Staphylococcus* (hazard ratio, 1.80; 95% confidence interval, 0.79–4.09; $P = 0.033$) were significantly associated with increased mortality in IPF. However, consistent with the literature (17), the enrichment of these two bacteria was observed in less than half of the cohort; therefore, it remains unlikely that these organisms alone can explain the disease pathogenesis or progression. Repeating the same analysis in subjects with CHP, no associations between an increased abundance of either *Streptococcus* (hazard ratio, 1.04; 95% confidence interval, 0.34–3.19; $P = 0.9$) or *Staphylococcus* (hazard ratio, 0.25; 95% confidence interval, 0.17–1.59; $P = 0.3$) in the BAL cells of these individuals and disease progression were demonstrated (Figure E5). α-Diversity metrics were used to characterize the distribution (evenness) and number (richness) of taxa expected within each sample. The microbial communities of subjects with IPF had a decreased evenness compared with those of the control subjects ($P < 0.03$) but not compared with those of subjects with CHP. There were no differences in evenness when comparing control subjects with individuals with CHP. Similarly, there were no demonstrable changes when comparing microbial richness among the groups (Figure 3B). Changes in the community composition (β diversity) between healthy control subjects and subjects with CHP and IPF were investigated using PCA and were tested by permutational multivariate ANOVA adjusted for age, sex, smoking status, FVC (% predicted) and D_{LCO} (% predicted), which confirmed differences in the microbial profiles of the subjects on the basis of diagnosis ($P = 0.001$) (Figure 3C). The PCA revealed that variation among the groups was mostly driven by changes in the *Proteobacteria* and *Firmicutes* phyla, with a tendency toward an enrichment of *Firmicutes* in the BAL from patients with IPF and of *Proteobacteria* in the BAL cells from patients with CHP. Community composition changes among the groups were further interrogated by looking for differentially abundant bacteria using the linear discriminant analysis effect size. Multiple taxonomic differences were found: the BAL fluid of subjects with IPF was enriched with bacteria belonging to the *Firmicutes* phylum, and the BAL fluid of the group with CHP was enriched with bacteria belonging to the *Actinobacteria* and *Fusobacteria* phyla. In contrast, those enriched in the healthy control group belonged to the *Bacteroidetes* phylum (Figure E6). Analyzing the CHP cohort on the basis of antigen positivity or negativity identified no differences between bacterial burden when comparing subjects with CHP with or without an exposure history to an antigen (Figure E7A). Similarly, no differences were observed in the microbial community composition (Figure E7B).

4C/FPO

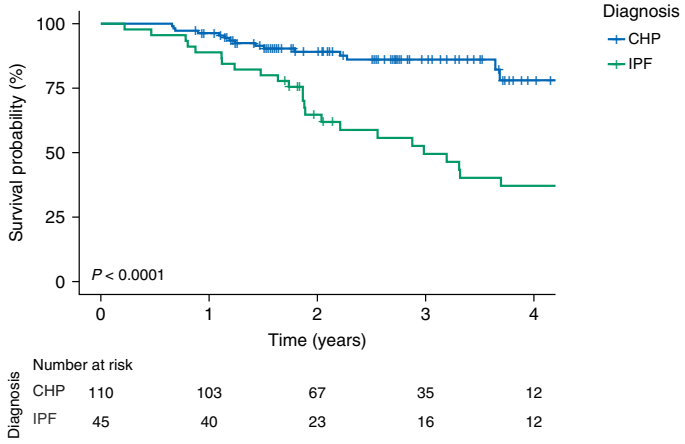


Figure 2. Survival probability of subjects with chronic hypersensitivity pneumonitis (CHP) compared with individuals with idiopathic pulmonary fibrosis (IPF). A Kaplan-Meier curve generated by using a Cox proportional hazards model displaying the survival probability (percentage) of subjects with CHP compared with subjects with IPF is shown. Log-rank *P* test values are reported.

Discussion

There are clear differences in the composition of the lower airway respiratory microbiota in CHP compared with IPF. Although individuals with CHP have a higher bacterial load in BAL fluid compared with healthy control subjects, it remains significantly lower than that observed in subjects with IPF. In contrast to the mounting evidence for the role of the bacterial burden in IPF, no association was found between the bacterial burden and survival in subjects with CHP.

Over the past decade, studies have begun to unravel the causes and consequences of variation within the respiratory microbiota, improving our understanding of how it may be linked to the pathogenesis of lung diseases. Considerable advances in culture-independent technologies have not only enabled researchers to characterize niche-specific microbial ecosystems but have also enabled them to study the role of the microbiome in lung health and disease (10, 11). Although the spotlight in respiratory research has been on asthma, COPD, and IPF, CHP remains an understudied disease (12, 13, 15–18). Research into the lung microbiota in CHP is very limited; to date, no specific study has investigated the composition of microbial communities in the lower airways in CHP.

The baseline bacterial phyla that dominated the BAL fluid of control subjects and subjects with CHP and IPF in the present study were *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, and *Fusobacteria*, which have previously been reported to be present in the airways of healthy subjects and patients with asthma, COPD, and IPF (10, 36). Marked differences were evident in the intergroup distribution of phyla. Specifically, an increased abundance of *Firmicutes* and *Actinobacteria* was noted in subjects with IPF compared with control subjects and subjects with CHP. In contrast, individuals with CHP had an increased proportion of *Proteobacteria* compared with individuals with IPF. This finding is consistent with the results of a recent study of subjects with IPF who had an increase in *Firmicutes* and a decrease in *Proteobacteria*; this reduction in microbial diversity was associated with disease activity in subjects with IPF (36). Furthermore, studies have shown that during respiratory disease, there is an increased prevalence of *Firmicutes* and *Proteobacteria* (12). This observation suggests that an altered state of the lung microbiome may be directly associated with the pathophysiology of fibrotic lung disease.

At the genus level, the microbial profile of all groups was characterized by *Prevotella*, *Veillonella*, and *Streptococcus*, bacteria known to prevail and/or dominate in the airways of healthy subjects and patients with respiratory disease (11). The phylum *Firmicutes* is associated with an increased risk of IPF progression when found at increased relative abundance in the lung (17). Specifically, within the *Firmicutes* phylum an increased abundance of *Streptococcus* and/or *Staphylococcus* species has been shown to be strongly correlated with reduced progression-free survival time in IPF (17). In the present study, *Staphylococcus* was increased in subjects with IPF compared with healthy control subjects and was further increased in subjects with CHP compared with subjects with IPF. In addition, in agreement with previous findings, *Veillonella* was also found to be increased in subjects with IPF compared with control subjects (16) and subjects with CHP. We further delineated differences in microbial evenness and whole community composition between CHP and IPF. Collectively, these findings show that subjects with CHP have a distinct microbiome compared with patients with IPF and healthy control subjects, which suggest that alterations in the microbiome characterized by an increased abundance of bacteria belonging to the *Firmicutes* phylum may play a role in the pathogenesis of these interstitial lung diseases. However, we cannot as yet draw any causal conclusions about this altered respiratory microbiome in CHP ahead of functional and longitudinal studies to comprehensively examine pathogenic host–microbe interactions.

It has previously been shown that IPF is characterized by an increased bacterial burden in BAL fluid compared with control subjects and that the bacterial load at the time of diagnosis predicts rapidly progressive IPF and an increased risk of mortality (16, 17). Given the radiological overlap and similarities in clinical behavior between patients with CHP and patients with IPF, it was surprising to find such striking differences in the lower airway bacterial burden between these two groups of patients. Although the BAL bacterial load in subjects with CHP was higher compared with that in control subjects, it was still lower than that in patients with IPF. Comparable with authors of other studies, we validate the finding that a higher bacterial burden at the time of diagnosis predicts subsequent progression of IPF. This association was not evident in the group with CHP and may explain why immunosuppression has been shown to be harmful in IPF but remains widely used to treat CHP in the absence of similar observations (37). Regardless of bacterial burden, subjects with CHP survived longer than individuals with IPF, a result that mirrors the findings of Vasakova and colleagues (38).

Closely associated with the alterations in the microbial composition is the change in the host immune status of patients with fibrotic lung disease. Specifically, lymphocytes, which are markedly increased in the airways of patients with CHP, play a key role in the inflammatory process underpinning the pathogenesis of the condition (39). An increased proportion of BAL lymphocytes was also present in the subjects with CHP in the present study. However, in agreement with what has been previously shown in IPF (16), no associations were found between BAL differential leukocyte counts and the bacterial burden in these subjects. It is increasingly recognized that a reciprocal communication likely exists between the microbiome and the pulmonary immune system, which may be critical for promoting pathology in a number of respiratory conditions (40). Future elucidation of microbial composition as well as the immune-cell subsets involved will improve our understanding of the cross-talk between microbiome and host.

There are a number of limitations to our work. Procedural and sequencing contamination presents a critical problem when evaluating samples with low microbial biomass, such as samples from the lungs (41). To minimize the risk of contamination, low-biomass protocols for DNA extraction were used, and negative bronchoscopy control subjects and reagent control subjects were sequenced alongside the BAL samples. The inclusion of the negative control samples, which yielded a bacterial burden close to or below the lower limit of qPCR quantification (100 copies/ml of BAL fluid) and had a taxonomic profile

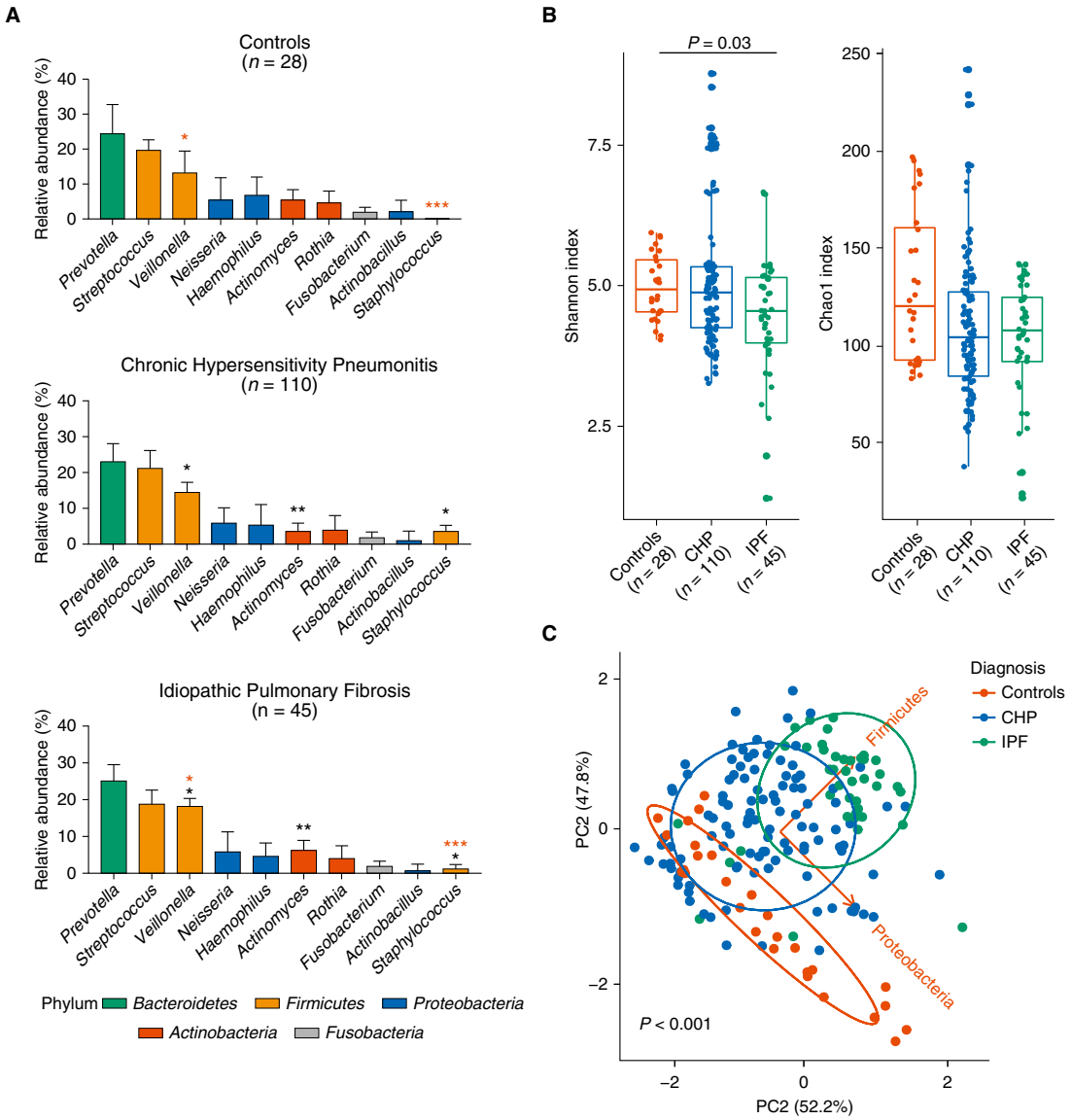


Figure 3. Taxonomic composition of bacteria in BAL of healthy control subjects and subjects with chronic hypersensitivity pneumonitis (CHP) and idiopathic pulmonary fibrosis (IPF). (*A*) Relative abundance of bacteria at the phylum (*Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Fusobacteria*) and genus (*Prevotella*, *Streptococcus*, *Veillonella*, *Neisseria*, *Haemophilus*, *Actinomyces*, *Rothia*, *Fusobacterium*, *Actinobacillus*, and *Staphylococcus*) level. Statistical significance was tested by using the Kruskal-Wallis test with the Dunn multiple-comparison test adjusted with the Bonferroni method. Statistical differences between CHP and IPF are indicated by black asterisks, and those between control subjects and IPF are indicated by red asterisks (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$). Data are presented as the mean ± SD. (*B*) Shannon and Chao1 α-diversity measures comparing healthy control subjects and individuals with CHP and IPF diagnoses. Statistical significance was tested by using the Kruskal-Wallis test with the Dunn multiple-comparison test adjusted with the Bonferroni method. Data are presented as the median and interquartile range. (*C*) PC analysis on Euclidean distance comparing healthy control subjects, subjects with CHP, and subjects with IPF (permutational multivariate ANOVA [PERMANOVA], $P < 0.001$). PERMANOVA was adjusted for age, sex, smoking status, FVC (% predicted), and D_{LCO} (% predicted). PC = principal component.

distinct from that of the BAL specimens, robustly indicates that the sequencing protocol used resulted in no apparent contamination. Furthermore, samples from case patients and control subjects were always processed in a blinded, randomized fashion to uphold identical storage and processing conditions across all samples. Another significant limitation is that 16S rRNA gene sequencing provides a snapshot of the microbial community at a particular moment in time, and it is therefore challenging to draw conclusions regarding whether the altered microbial community is a cause or a consequence of the disease (42). Finally, this exploratory study was based on feasibility of recruitment and sampling, and some of the most severe cases of both CHP and IPF may have not been included, as only subjects able to safely undergo bronchoscopy were enrolled. Therefore, functional and longitudinal studies examining host–microbe interactions are needed to validate these findings.

In conclusion, in this prospective observational cohort study, we provide the first evidence of the composition of the microbial communities in the lower airways in CHP. The microbial profile of subjects with CHP is distinct from that of subjects with IPF. Compared with subjects with IPF, these individuals have a decreased bacterial burden, which does not predict survival. Future work should aim at understanding whether the total bacterial load and relative abundances of the altered microbial genera may be acting synergistically to advance the fibrotic process. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors would like to thank the Clinical Data Warehouse team at Royal Brompton and Harefield National Health Service Foundation Trust.

References

1. Salisbury ML, Myers JL, Belloli EA, Kazerooni EA, Martinez FJ, Flaherty KR. Diagnosis and treatment of fibrotic hypersensitivity pneumonia: where we stand and where we need to go. *Am J Respir Crit Care Med* 2017;196:690–699.
2. Selman M, Pardo A, King TE Jr. Hypersensitivity pneumonitis: insights in diagnosis and pathobiology. *Am J Respir Crit Care Med* 2012;186: 314–324.
3. Campbell JM. Acute symptoms following work with hay. *Br Med J* 1932; 2:1143–1144.
4. Morell F, Villar A, Montero MA, Muñoz X, Colby TV, Pipvath S, *et al*. Chronic hypersensitivity pneumonitis in patients diagnosed with idiopathic pulmonary fibrosis: a prospective case-cohort study. *Lancet Respir Med* 2013;1:685–694.
5. Morisset J, Johannson KA, Jones KD, Wolters PJ, Collard HR, Walsh SLF, *et al*.; HP Delphi Collaborators. Identification of diagnostic criteria for chronic hypersensitivity pneumonitis: an international modified Delphi survey. *Am J Respir Crit Care Med* 2018;197: 1036–1044.
6. Silva CI, Müller NL, Lynch DA, Curran-Everett D, Brown KK, Lee KS, *et al*. Chronic hypersensitivity pneumonitis: differentiation from idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia by using thin-section CT. *Radiology* 2008;246:288–297.
7. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. *Annu Rev Physiol* 2016;78:481–504.
8. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, *et al*. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 2011;184:957–963.
9. Invernizzi R, Lloyd CM, Molyneaux PL. Respiratory microbiome and epithelial interactions shape immunity in the lungs. *Immunology* 2020; 160:171–182.
10. Moffatt MF, Cookson WO. The lung microbiome in health and disease. *Clin Med (Lond)* 2017;17:525–529.
11. Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. *Expert Rev Respir Med* 2013;7:245–257.
12. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, *et al*. Disordered microbial communities in asthmatic airways. *PLoS One* 2010;5:e8578.
13. Stokholm J, Blaser MJ, Thorsen J, Rasmussen MA, Waage J, Vinding RK, *et al*. Maturation of the gut microbiome and risk of asthma in childhood. *Nat Commun* 2018;9:141.
14. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, *et al*. Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PLoS One* 2017;12:e0170622.
15. Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SA, Homola D, *et al*. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013;188:1224–1231.
16. Molyneaux PL, Cox MJ, Willis-Owen SA, Mallia P, Russell KE, Russell AM, *et al*. The role of bacteria in the pathogenesis and progression of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2014;190: 906–913.
17. Han MK, Zhou Y, Murray S, Tayob N, Noth I, Lama VN, *et al*.; COMET Investigators. Lung microbiome and disease progression in idiopathic pulmonary fibrosis: an analysis of the COMET study. *Lancet Respir Med* 2014;2:548–556.
18. O'Dwyer DN, Ashley SL, Gurczynski SJ, Xia M, Wilke C, Falkowski NR, *et al*. Lung microbiota contribute to pulmonary inflammation and disease progression in pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;199:1127–1138.
19. Molyneaux PL, Willis-Owen SAG, Cox MJ, James P, Cowman S, Loebinger M, *et al*. Host-microbial interactions in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017;195: 1640–1650.
20. Huang Y, Ma SF, Espindola MS, Vij R, Oldham JM, Huffnagle GB, *et al*.; COMET-IPF Investigators. Microbes are associated with host innate immune response in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2017;196:208–219.
21. Molyneaux PL, Cox MJ, Wells AU, Kim HC, Ji W, Cookson WO, *et al*. Changes in the respiratory microbiome during acute exacerbations of idiopathic pulmonary fibrosis. *Respir Res* 2017; 18:29.
22. Invernizzi R, Molyneaux PL. The contribution of infection and the respiratory microbiome in acute exacerbations of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2019;28:190045.
23. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet* 2012;13:260–270.
24. Renz H, Autenrieth IB, Brandtzæg P, Cookson WO, Holgate S, von Mutius E, *et al*. Gene-environment interaction in chronic disease: a European Science Foundation Forward Look. *J Allergy Clin Immunol* 2011;128:S27–S49.
25. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, *et al*.; American Thoracic Society; European Respiratory Society; Japanese Respiratory Society; Latin American Thoracic Society. Diagnosis of idiopathic pulmonary fibrosis: an official ATS/ERS/JRS/ ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018; 198:e44–e68.
26. Ryerson CJ, Corte TJ, Lee JS, Richeldi L, Walsh SLF, Myers JL, *et al*. A standardized diagnostic ontology for fibrotic interstitial lung disease: an International Working Group Perspective. *Am J Respir Crit Care Med* 2017;196:1249–1254.

27. Meyer KC, Raghu G, Baughman RP, Brown KK, Costabel U, du Bois RM, *et al*.; American Thoracic Society Committee on BAL in Interstitial Lung Disease. An official American Thoracic Society clinical practice guideline: the clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease. *Am J Respir Crit Care Med* 2012;185:1004–1014.
28. Invernizzi R, Barnett J, Rawal B, Nair A, Ghai P, Kingston S, *et al*. Bacterial burden in the lower airways predicts disease progression in idiopathic pulmonary fibrosis and is independent of radiological disease extent. *Eur Respir J* 2020;55:1901519.
29. DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK. Selective progressive response of soil microbial community to wild oat roots. *ISME J* 2009;3:168–178.
30. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–336.
31. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13:581–583.
32. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, *et al*. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 2012;6:610–618.
33. R Core Team. A language and environment for statistical computing. Version 3.4.3. Vienna, Austria: R Foundation for Statistical Computing; 2013.
34. Spagnolo P, Rossi G, Cavazza A, Bonifazi M, Paladini I, Bonella F, *et al*. Hypersensitivity pneumonitis: a comprehensive review. *J Investig Allergol Clin Immunol* 2015;25:237–250; quiz following 250.
35. Callahan BJ, Sankaran K, Fukuyama JA, McMurdie PJ, Holmes SP. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000 Res* 2016;5:1492.
36. Takahashi Y, Saito A, Chiba H, Kuronuma K, Ikeda K, Kobayashi T, *et al*. Impaired diversity of the lung microbiome predicts progression of idiopathic pulmonary fibrosis. *Respir Res* 2018;19:34.
37. Raghu G, Anstrom KJ, King TE Jr, Lasky JA, Martinez FJ; Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012;366:1968–1977.
38. Vasakova M, Morell F, Walsh S, Leslie K, Raghu G. Hypersensitivity pneumonitis: perspectives in diagnosis and management. *Am J Respir Crit Care Med* 2017;196:680–689.
39. Pereira CA, Gimenez A, Kuranishi L, Storrer K. Chronic hypersensitivity pneumonitis. *J Asthma Allergy* 2016;9:171–181.
40. O'Dwyer DN, Dickson RP, Moore BB. The lung microbiome, immunity, and the pathogenesis of chronic lung disease. *J Immunol* 2016;196: 4839–4847.
41. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, *et al*. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 2014;12:87.
42. Hewitt RJ, Molyneaux PL. The respiratory microbiome in idiopathic pulmonary fibrosis. *Ann Transl Med* 2017;5:250.