UNDERSTANDING SHARED PATHOGENESIS BETWEEN CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD) AND LUNG CANCER BY MEANS OF CELL SPECIFIC GENOMICS

CLARA EMILY GREEN

UNIVERSITY OF BIRMINGHAM

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

The Institute of Inflammation and Ageing

College of Medical and Dental Sciences

University of Birmingham

February 2018
Abstract

Introduction

COPD (Chronic Obstructive Pulmonary Disease) and lung cancer are related conditions associated with inflammation. Relatively little focus has been given to the endothelium, through which inflammatory cells transmigrate to reach the lung. We sought to determine if coding and non-coding alterations in pulmonary endothelium exist in COPD and lung cancer.

Methods

Patients with and without COPD undergoing thoracic surgery were recruited. Pulmonary Endothelial Cells were isolated from lung and tumour and extracted RNA (ribonucleic acid) used for miRNA (micro-RNA) and mRNA (messenger RNA) microarrays. Ingenuity pathway analysis (IPA) was also carried out.

Results

2071 genes and 43 miRNAs were significantly upregulated in COPD. 4 targets were validated by quantitative polymerase chain reaction, of which miR-181b-3p was chosen for functional validation. Another target, miR-429, was also increased in lung tumour. Several cancer-related pathways such as transforming growth factor- β were altered in the IPA. There was significantly reduced tube formation and endothelial sprouting in Human umbilical vein endothelial cells transfected with miR-181b-3p, consistent with an effect on angiogenesis.
Conclusions

Upregulation of miR-181b-3p reduces tube formation and sprouting by endothelial cells. This might be significant in the development of emphysema as lung vasculature is important in the structural maintenance of alveoli.
Dedication

I dedicate this thesis to my husband Robert Green for all of your encouragement and support.
Acknowledgements

Firstly I would like to thank my supervisors Dr. Alice Turner and Prof. Roy Bicknell for the invaluable advice and support you have given me throughout the past four years. It has been a pleasure to work with both of you. I am also grateful for the help and assistance of members of both the Respiratory and Bicknell/Heath groups without whom I would have been unable to complete the necessary laboratory work.

Thank you to all of the staff responsible for the managing of, and collection of tissue for, the Midlands Lung Tissue Consortium, including Prof. David Thickett, Prof. Babu Naidu, Dr. Gerald Langman, Thoracic surgeons and Thoracic research nurses at Birmingham Heartlands Hospital.

I am also grateful to Dr. Zsuzsanna Nagy for her advice regarding microarray analysis and for the use of Ingenuity Pathway Analysis.

Finally, I would like to thank my family for their endless help and support throughout the course of this study. Thank you to my husband Rob for listening and providing advice throughout my degree. Thank you to my sister Elena who provided me with support over the phone. Thank you to my parents John and Olivia for providing me with hours of childcare so I could complete this work. Thank you to my sons Eric and Arthur for providing me with much needed love and joy during the course of study.
Table of Contents

Chapter One

Introduction ................................................................. 1

1.1 Chronic obstructive Pulmonary Disease (COPD)...................... 2

1.1.1 Clinical Features.......................................................... 3

1.1.2 Management............................................................... 9

1.1.3 Pathogenesis............................................................. 12

1.2 Lung Cancer................................................................. 16

1.2.1 Clinical Features.......................................................... 16

1.2.2 Management............................................................... 19

1.2.3 Pathogenesis............................................................. 21

1.3 Evidence for the shared pathogenesis of COPD and lung cancer.... 26

1.4 The role of the endothelium in the lung.................................. 28

1.4.1 The role of the endothelium in COPD................................. 34

1.4.2 The role of the endothelium in lung cancer......................... 39
Chapter Two

Methods................................................................. 63

2.1 Patient data collection........................................... 64

2.1.1 Patient selection............................................... 64

2.1.2 Clinical data.................................................... 64

2.2 Statistical analysis of clinical data........................... 67

2.3 Collection of lung tissue........................................ 67

2.3.1 Extraction of macrophages.................................. 67
2.3.2 Extraction of endothelium.................................................70

2.4 Preparation of samples for microarray analysis........................73

2.4.1 RNA Extraction...............................................................73

2.4.2 Quality control of RNA prior to microarray analysis..............75

2.4.3 Labelling RNA for mRNA microarray analysis....................75

2.4.4 Hybridisation of mRNA for microarray..............................78

2.5 Microarray Analysis................................................................81

2.5.1 Quality analysis of microarray results.................................81

2.5.2 Assessment of differential gene expression using Limma.......86

2.5.3 Assessment of false discovery rate (FDR).........................89

2.5.4 Assessment of differential gene expression using Significance Analysis of Microarrays (SAM)..............................90

2.5.5 Ingenuity Pathway Analysis (IPA)....................................92

2.6 Reverse transcription..........................................................97

2.6.1 Reverse transcription of RNA to cDNA for RNA samples of concentration >300 ng/µl.....................................................97
2.6.2 Reverse transcription of RNA to cDNA for RNA samples of concentration <300 ng/µl ................................................................. 97

2.6.3 Reverse transcription of miRNA ................................................. 98

2.7 Real-time Polymerase chain reaction (qPCR) .................................. 98

2.7.1 qPCR of mRNA ........................................................................ 98

2.7.2 qPCR of miRNA ....................................................................... 99

2.7.3 Analysis of qPCR results ............................................................ 101

2.8 Culture of Mouse Lung Endothelial Cells (MLECs) ...................... 103

2.8.1 Collagenase digestion of lung tissue .......................................... 103

2.8.2 Negative selection of macrophages using magnetic beads ........ 104

2.8.3 Positive selection of endothelial cells by magnetic beads ........... 106

2.9 Culture of Human Pulmonary Endothelial Cells (HPECs) ............ 108

2.9.1 Collagenase digestion of lung tissue .......................................... 108

2.9.2 Positive selection of endothelial cells by magnetic beads .......... 110

2.10 Cell Storage .............................................................................. 111

2.11 Flow cytometry .......................................................................... 111
2.11.1 Flow cytometry to look for CD31 expression in cells isolated from mouse lung................................................................. 111

2.11.2 Flow cytometry to look for CD31 expression in cells isolated from the lung........................................................................ 113

2.11.3 Flow cytometry for cell cycle analysis........................................... 115

2.12 Cell culture for cellular functional work........................................ 120

2.13 miRNA Inhibition and Overexpression........................................ 120

2.13.1: Transfection of a miRNA inhibitor or mimic into HUVEC in a 6 well plate.................................................................. 120

2.14 Cell growth assay........................................................................ 125

2.15 Matrigel tube formation assay........................................................ 125

2.16 Scratch wound assay...................................................................... 126

2.17 Spheroid assay................................................................................ 129

2.17.1 Preparation of the spheroid assay............................................... 129

2.17.2 Analysis of spheroids using ImageJ.......................................... 131

Chapter Three

MicroRNA and Messenger RNA expression in lung endothelium........133
3.1 Hypothesis

3.2 Aims of this chapter

3.3 miRNA expression in the lung endothelium

3.3.1 Subjects

3.3.2 Combined analysis of 2014 and 2015 microarrays

3.3.3 Computer prediction of targets of identified miRNAs

3.3.4 qPCR validation of miRNA targets

3.3.5 Validation of potential miRNA targets using qPCR in lung cancer

3.4 mRNA expression in the lung endothelium

3.4.1 Subjects

3.4.2 Combined analysis of 2014 and 2016 microarrays

3.4.3 Tumour mRNA analysis

3.4.4 Validation of potential mRNA targets using qPCR

3.5 Discussion

3.5.1 miRNA expression in pulmonary endothelium

3.5.2 mRNA expression in pulmonary endothelium
Chapter Four

Pathway analysis of microRNA and messenger RNA expression data

4.1 Hypothesis

4.2 Aims of this chapter

4.3 Introduction

4.4 Results

4.4.1 Diseases and biological functions associated with COPD

4.4.2: Downstream disease analysis of miRNA and mRNA microarrays

4.4.3 miRNA and mRNA networks in COPD

4.4.4 Canonical pathways in COPD

4.5 Discussion

4.5.1 Diseases associated with COPD
4.5.2 Pathways associated with COPD...............................................225

4.5.3 Limitations................................................................................232

4.6 Conclusions....................................................................................234

Chapter Five

Isolation and culture of Human Pulmonary Endothelial Cells (HPECs)

........................................................................................................236

5.1 Hypothesis.....................................................................................237

5.2 Aims of this chapter.......................................................................237

5.3 Extraction of MLEC (mouse lung endothelial cells)......................237

5.4 Extraction of Human Pulmonary Endothelial Cells (HPECs)........239

5.4.1 Development of the extraction method.....................................242

5.4.2 Flow cytometry of positively selected lung cells......................253

5.5 Discussion.....................................................................................259

5.6 Conclusions....................................................................................263

Chapter Six

Identifying the role of microRNA targets in angiogenesis..................265
6.1 Hypothesis.......................................................................................................................... 266

6.2 Aims of this chapter............................................................................................................... 266

6.3 The role of miR-181b-3p in angiogenesis................................................................. 266

6.3.1 Optimisation of miR-181b-3p mimic into HUVECs......................................................... 267

6.3.2 Optimisation of miR-181b-3p inhibitor into HUVECs..................................................... 270

6.3.3 The effect of miR-181b-3p overexpression on cell growth.............................................. 274

6.3.4 The effect of miR-181b-3p overexpression on the cell cycle........................................... 274

6.3.5 The effect of miR-181b-3p overexpression on HUVEC tube formation................................. 277

6.3.6 The effect of miR-181b-3p overexpression on HUVEC wound healing.............................. 281

6.3.7 The effect of miR-181b-3p overexpression on HUVEC spheroid sprouting...................... 284

6.3.8 Identification of miR-181b-3p targets............................................................................... 288

6.4 Discussion................................................................................................................................ 290

6.4.1 The evidence for miR-181b-3p’s involvement in angiogenesis........................................... 290

6.4.2 Limitations.......................................................................................................................... 293
Chapter Seven

Final discussion and future work

7.1 Summary of project findings

7.1.1 Key project findings

7.1.2 Summary of the analysis of microRNA and messenger RNA expression in lung endothelium in COPD and lung cancer

7.1.3 Summary of findings from Ingenuity Pathway Analysis

7.1.4 Summary of findings from Human Pulmonary Endothelial Cell extractions

7.1.5 Summary of findings from functional studies investigating the role of miR-181b-3p in angiogenesis

7.2 Possible mechanisms to explain miR-181b-3p's role in angiogenesis

7.3 The potential of miR-181b-3p as a treatment target

7.4 Conclusions

7.4.1 Reflections on the study hypothesis

7.5 Future work
7.5.1 Endothelial functional assays..................................................305

7.5.2 MiRNA target identification......................................................307

7.5.3 Immunohistochemistry..............................................................309

7.5.4 Neutrophil Transendothelial Migration studies.........................310

7.5.5 \textit{In vitro} drug response......................................................311

7.5.6 \textit{In vivo} drug response..........................................................311

7.5.7 Macrophages.............................................................................312

References..........................................................................................313

Appendices.........................................................................................365
List of figures

Figure 1.1: GOLD classification of COPD.................................................. 7

Figure 1.2: Appearances of emphysema on CT................................. 8

Figure 1.3: Appearances of bronchiectasis on CT.......................... 8

Figure 1.4: COPD value pyramid....................................................... 11

Figure 1.5: Pathogenesis of COPD..................................................... 15

Figure 1.6: A summary of the relationship between COPD and lung cancer..... 29

Figure 1.7: The endothelium............................................................ 32

Figure 1.8: Transendothelial migration.......................................... 32

Figure 1.9: miRNA processing......................................................... 42

Figure 1.10: MiRNAs in the setting of hypoxia, NSCLC and COPD and their effect on angiogenesis through the HIF-1α/VEGF pathway................................................. 49

Figure 2.1: Extraction of endothelial cells for future RNA work............. 69

Figure 2.2: RNA extraction using the Qiagen miRNeasy kit................ 72

Figure 2.3: Quality control of extracted RNA.................................... 76

Figure 2.4: Diagram of a loaded Agilent gasket................................. 80

Figure 2.5: Magnified view of a sample loaded into a gasket................. 80
Figure 2.6: A “targets.txt” file used in arrayqualitymetrics

Figure 2.7: Output from arrayqualitymetrics

Figure 2.8: An example SAM plot

Figure 2.9: The process by which Ingenuity identifies significant associations with diseases and biological functions

Figure 2.10: Example of a canonical pathway significantly altered in a microarray analysis

Figure 2.11: Culture of Mouse Lung Endothelial Cells (MLECs)

Figure 2.12: Culture of Human Pulmonary Endothelial Cells (HPECs)

Figure 2.13: A cytogram plot from the program ‘FloJo’

Figure 2.14: A histogram plot from flow cytometry

Figure 2.15: Cytogram plots from the program ‘FloJo’ used for cell cycle analysis

Figure 2.16: A histogram plot of cell counts for cell cycle analysis

Figure 2.17: Overview of the spheroid assay

Figure 2.18: Analysis of an endothelial spheroid using the ‘Spheroid Analysis’ plugin in ImageJ
Figure 3.1: SAM plot for the analysis comparing 2014 and 2015 miRNA microarrays

Figure 3.2: SAM plots and corresponding heatmaps for the miRNA microarray analysis using SAM

Figure 3.3: qPCR validation of endothelial enrichment

Figure 3.4: qPCR validation of potential miRNA targets

Figure 3.5: qPCR validation of miR-429 in lung cancer

Figure 3.6: SAM plot for the analysis comparing 2014 and 2016 mRNA microarrays

Figure 3.7: SAM plots and corresponding heatmaps for the mRNA microarray analysis using SAM

Figure 3.8: SAM plots and corresponding heatmaps for the mRNA lung tumour microarray analysis using SAM

Figure 3.9: qPCR validation of potential mRNA targets

Figure 4.1: The top 10 diseases and biological functions altered in the 2014 miRNA microarray data

Figure 4.2: The top 10 diseases and biological functions altered in the 2015 miRNA microarray data
Figure 4.3: The top 10 diseases and biological functions altered in the 2014 mRNA microarray data .................................................................179

Figure 4.4: The top 10 diseases and biological functions altered in the 2016 mRNA microarray data .................................................................179

Figure 4.5: Molecular and cellular functions associated with differentially expressed miRNAs in COPD in the 2014 miRNA microarray data.................................182

Figure 4.6: Molecular and cellular functions associated with differentially expressed miRNAs in COPD in the 2015 miRNA microarray data.................................183

Figure 4.7: Molecular and cellular functions associated with differentially expressed mRNAs in COPD in the 2014 mRNA microarray data.................................184

Figure 4.8: Molecular and cellular functions associated with differentially expressed mRNAs in COPD in the analysis: 2016 mRNA microarray data.................................185

Figure 4.9: Diseases associated with differentially expressed miRNAs in COPD in the 2014 miRNA microarray data. .................................................................187

Figure 4.10: Diseases associated with differentially expressed miRNAs in COPD in the 2015 miRNA microarray data. .................................................................188

Figure 4.11: Diseases associated with differentially expressed mRNAs in COPD in the 2014 mRNA microarray data .................................................................189

Figure 4.12: Diseases associated with differentially expressed mRNAs in COPD in the 2016 mRNA microarray data. .................................................................190
Figure 4.13: Downstream prediction of miRNA changes on cell proliferation of tumour cell lines in the 2014 miRNA microarray analysis

Figure 4.14: Downstream prediction of miRNA changes on neoplasia of tumour cell lines in the 2014 miRNA microarray analysis

Figure 4.15: Downstream prediction of miRNA changes on cell proliferation of carcinoma cell lines in the 2014 miRNA microarray analysis

Figure 4.16: Downstream prediction of miRNA changes on invasion of cell lines in the 2014 miRNA microarray analysis

Figure 4.17: Downstream prediction of miRNA changes on cytokinesis of ventricular myocytes in the 2014 miRNA microarray analysis

Figure 4.18: Downstream prediction of miRNA changes on maturation of chondrocyte cell lines in the 2014 miRNA microarray analysis

Figure 4.19: Downstream prediction of mRNA changes on hypersensitivity reactions in the 2016 mRNA microarray analysis

Figure 4.20: Downstream prediction of mRNA changes on allergy in the 2014 mRNA microarray analysis

Figure 4.21: Downstream prediction of mRNA changes on metabolism of eicosanoid in the 2014 mRNA microarray analysis

Figure 4.22: Downstream prediction of mRNA changes on synthesis of fatty acid in the 2014 mRNA microarray analysis
Figure 4.23: Downstream prediction of mRNA changes on synthesis of lipid in the 2014 mRNA microarray analysis........................................................................208

Figure 4.24: A network identified in the 2015 miRNA microarrays.................................210

Figure 4.25: A network identified in the 2015 miRNA microarrays.................................211

Figure 4.26: A network identified in the 2014 miRNA microarrays.................................212

Figure 4.27: A network identified in the 2014 miRNA microarrays.................................213

Figure 4.28: A network identified in the 2016 mRNA microarrays.................................215

Figure 4.29: A network identified in the 2014 mRNA microarrays.................................216

Figure 4.30: Eicosanoid Signalling pathway according to results from the 2014 mRNA microarray analysis.........................................................................................218

Figure 4.31: The TGF-β pathway illustrated in a cell..........................................................227

Figure 4.32: LTA4H biology in the setting of COPD.............................................................231

Figure 5.1: Appearance of mouse lung culture after negative selection......................238

Figure 5.2: Day one after positive selection for CD31 in mouse lung cell culture...240

Figure 5.3: A histogram of flow cytometry performed on isolated MLECs...............241

Figure 5.4: Appearance of mixed lung cell culture on day two, extraction one (different areas of the same 10 cm plate).................................................................243
Figure 5.5: Appearance of mixed lung cell culture on day five, extraction four……247

Figure 5.6: Appearance of mixed lung cell culture on day five, extraction five……248

Figure 5.7: Appearance of lung cells on day 10 after positive selection using Ulex coated magnetic beads……………………………………………………………………255

Figure 5.8: A cytogram plot from the program ‘FloJo’ for HUVECs……………256

Figure 5.9: A histogram plot from flow cytometry performed on HUVECs……256

Figure 5.10: A cytogram plot from the program ‘FloJo’ for lung cells……………257

Figure 5.11: A histogram plot from flow cytometry performed on lung cells……257

Figure 6.1: Optimisation of miR-181b-3p mimic concentration………………268

Figure 6.2: Relative expression of miR-181b following mimic transfection at 10 nM concentration …………………………………………………………………………………269

Figure 6.3: Optimisation of miR-181b-3p inhibitor concentration………………271

Figure 6.4: Inhibition of miR-181b-3p using an inhibitor at 10 nM………………272

Figure 6.5: Relative expression of miR-181b following inhibitor transfection……273

Figure 6.6: Cell growth assay……………………………………………………………275

Figure 6.7: Cell cycle analysis according to flow cytometry……………………275

Figure 6.8: Histogram plots of cell counts for cell cycle analysis………………276
Figure 6.9: Number of nodes formed by HUVECs imbedded in matrigel........... 278

Figure 6.10: Number of junctions formed by HUVECs imbedded in matrigel........278

Figure 6.11: Number of meshes formed by HUVECs imbedded in matrigel.........279

Figure 6.12: Mesh area formed by HUVECs imbedded in matrigel...................279

Figure 6.13: An example of images from a matrigel experiment....................280

Figure 6.14: Percentage wound area remaining during scratch wound assay......282

Figure 6.15: An example of images from a scratch wound experiment.............283

Figure 6.16: Number of endothelial sprouts from spheroids.......................285

Figure 6.17: Total length of endothelial sprouts from spheroids...................285

Figure 6.18: Average length of endothelial sprouts from spheroids...............286

Figure 6.19: An example of an endothelial spheroid assay..........................287

Figure 6.20: Relative expression of ELDT1 and IL8 in cells transfected with miR-181b-3p mimic and negative siRNA......................................................289
List of tables

Table 1.1: Severity of COPD by spirometry…………………………………………… 7
Table 1.2: Histological classifications of lung cancer……………………………….. 18
Table 1.3: Treatments used in the management of NSCLC…………………………… 19
Table 1.4: Palliative therapies used in the management of lung cancer……………… 20
Table 1.5: New therapies in lung cancer………………………………………………… 22
Table 1.6: Examples of genes commonly mutated in lung cancer………………….. 23
Table 1.7: Methods of measuring endothelial dysfunction in COPD……………….. 38
Table 2.1: Summary of information collected for each patient……………………… 66
Table 2.2: Output from limma……………………………………………………………… 88
Table 2.3: Example of a SAM input file………………………………………………… 91
Table 2.4: Probes (Roche Universal Probe library) and primers used for mRNA qPCR…………………………………………………………………………………. 100
Table 2.5: The number of HUVEC plated and volumes of inhibitor, mimic and lipofectamine used for different plates in various mimic or inhibitor experiments…. 101
Table 3.1: Patient demographics comparing 2014 to 2015 miRNA microarrays……… 136
Table 3.2: Number of patients with each tumour stage in the 2014 and 2015 miRNA microarrays. ………………………………………………………………………………….. 136
Table 3.3: Patient demographics comparing COPD to non-COPD patients in the combined miRNA microarray analysis

Table 3.4: Number of patients with each tumour stage in the combined miRNA microarray analysis

Table 3.5: Limma analysis results from the combined miRNA microarray analysis

Table 3.6: Significantly upregulated miRNAs in COPD

Table 3.7: miRNAs selected for validation work

Table 3.8: Patient demographics comparing 2014 to 2016 mRNA microarrays

Table 3.9: Number of patients in each group with each tumour stage in the 2014 and 2016 mRNA microarrays

Table 3.10: Patient demographics comparing COPD to non-COPD patients in the combined mRNA microarray analysis

Table 3.11: Number of patients in each group with each tumour stage in the combined mRNA microarray analysis

Table 3.12: Limma analysis results from the combined mRNA microarray analysis

Table 3.13: Patient demographics comparing 2014 non-COPD patients to tumour tissue patients

Table 3.14: Number of patients in each group (non-COPD 2014 and tumour) with each tumour stage
Table 3.15: Patient demographics comparing 2016 non-COPD patients to tumour tissue patients……………………………………………………………………………………………………158

Table 3.16: Number of patients in each group (non-COPD 2016 and tumour) with each tumour stage…………………………………………………………………………………………158

Table 3.17: A comparison of the results of the COPD mRNA microarray analysis and the two tumour mRNA microarray analyses for the 6 possible targets identified in the COPD analysis (section 3.4.2)…………………………………………………………………………………………161

Table 4.1: Molecular and cellular processes altered in COPD in each microarray analysis…………………………………………………………………………………………181

Table 4.2: Diseases associated with COPD in each microarray analysis………………186

Table 4.3: Diseases identified as significant in IPA and their relation to known comorbidities of patients included in the study and the current literature………………191

Table 4.4: Downstream analysis predictions for diseases/cellular functions significant in the 2016 mRNA analysis…………………………………………………………………………………………202

Table 4.5: Downstream analysis predictions for diseases/cellular functions significant in the 2014 mRNA analysis…………………………………………………………………………………………204

Table 5.1: Flow cytometry results for HUVECs…………………………………………………………258

Table 5.2: Flow cytometry results for lung cells…………………………………………………………258
CHAPTER ONE
INTRODUCTION

Please note that part of this chapter has already been published in the form of two review articles.(1, 2)
Chronic obstructive pulmonary disease (COPD) and lung cancer are two important smoking related conditions. COPD has 10.1% prevalence (with Forced expiratory volume in one second (FEV1) < 80%) in adults over 40 years worldwide (3) and is also an important cause of morbidity and mortality, resulting in over 29 000 deaths in the UK in 2012. (4) Lung cancer is the leading cause of cancer morbidity in North America, and is amongst the top 5 cancers ranked by disability adjusted life years lost (DALYs) in all regions of the world, except Sub-Saharan Africa. (5) It is the most common cause of cancer death accounting for 35 895 deaths in the UK in 2014. (6) Survival rates have changed little despite improvements in screening (7), surgery (8) and chemotherapy (9, 10). Focussing screening to high-risk groups could identify patients early, and thus raise cure rates. Patients with COPD are one such group; lung cancer incidence in COPD patients is high, even after adjustment for smoke exposure (11) suggesting possible shared pathogenesis. By understanding the pathogenesis of COPD and lung cancer in detail it is possible that new treatments may be developed and the risk of lung cancer in COPD may be reduced.

1.1 Chronic Obstructive Pulmonary Disease (COPD)

COPD is a condition characterized by airflow obstruction which is not normally fully reversible with the use of bronchodilators, and is generally thought to be progressive over time. (12) Damage in both the airways and lung parenchyma contribute to the airway obstruction. (12, 13) An abnormal chronic inflammatory response to noxious particles or gases also occurs. (12) Spirometry is used to diagnose airway obstruction which is defined as a ratio of Forced Expiratory Volume in one second (FEV1) to Forced Vital Capacity (FVC; the total volume expired) of less than 0.7. (14) The
airflow obstruction usually occurs as a result of smoking but genetic factors (e.g., alpha 1 antitrypsin deficiency), the burning of biomass fuels in developing countries and occupational factors also play a role. (12, 15) In fact only 20%-30% of smokers develop COPD suggesting an important role for other factors in the development of the disease. (16)

1.1.1 Clinical Features

COPD patients are typically aged over 35 at presentation with a history of a known risk factor (usually smoking). (17) Common presenting symptoms include exertional breathlessness, chronic cough, sputum production and wheeze. (17) Occasionally systemic symptoms such as weight loss may also feature. (17) British and American/European guidelines classify COPD by means of severity ranging from mild (FEV1 ≥ 80% predicted) to very severe (FEV1 < 30% predicted) (see Table 1.1). (17, 18) However, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) have suggested that a patient’s symptoms and exacerbation history should also be taken into consideration in determining severity. (12) GOLD classifies patients into 4 classes using airway obstruction, exacerbation history and results from symptom scoring systems for breathlessness (Modified British Medical Research Council Questionnaire; mMRC) and health status impairment (COPD Assessment Test; CAT) (see Figure 1.1). (12) In theory by using patient symptoms and history as part of the severity assessment one should be able to identify patients at higher risk of disease progression who may not initially look high risk based on FEV1 alone. This should allow the assessing clinician to start relevant treatment to reduce risk earlier. However, the GOLD classification does have its limitations. For example, the cut-offs chosen for mMRC and CAT scoring were not evidence based and the presumed
The cause of airway obstruction seen in COPD varies between patients, but is primarily the result of a mixture of airways disease (chronic bronchiolitis) and destruction of the parenchyma (emphysema). (12) The contribution of airways and parenchymal disease varies between COPD patients and this is reflected in a range of clinical presentations including chronic bronchitis, emphysema and bronchiectasis. (12, 18) A combination of these conditions may be present in a single patient. (18)

Chronic bronchitis is defined as a productive cough occurring on most days for a period of at least three months in two consecutive years. (23) Emphysema is a pathological diagnosis of permanent enlargement of the airspaces distal to the terminal bronchioles resulting from destruction of the lung parenchyma. (24) Bronchiectasis is also defined pathologically as the persistent and progressive development of dilated and thick-walled bronchi. (25)
Spirometry (including reversibility testing using bronchodilators) should be performed in all patients to confirm the diagnosis of COPD. (17, 18) Although most patients will demonstrate fixed airway obstruction a degree of reversibility may be seen. (18) However, a large increase in FEV1 in response to bronchodilation should prompt consideration of a diagnosis of asthma. (18) Lung function tests may also demonstrate reduced maximal expiratory flows in the mid-volume range suggesting small airway obstruction. (26) These changes can sometimes be seen in early COPD before a change in FEV1/FVC. (26) However, these values vary highly between individuals and consequently are not used routinely in the diagnosis of COPD. (26, 27) Patients with advanced COPD may also show an increase in static lung volumes (Total Lung Capacity, TLC and Residual Volume, RV). (18) Increased static lung volumes in combination with a reduced transfer factor (relative to lung volume, KCO) is suggestive of emphysema in these patients. (28) However, emphysema is more accurately diagnosed using computed tomography (CT). (29) In classical COPD emphysema commonly has a centrilobular appearance and is primarily distributed in the upper zones. (30) It is characterized by enlargement of the centrilobular airspace and primarily effects the proximal respiratory bronchioles. (31) Another distribution of emphysema, panlobular emphysema, consists of a uniformly dilated airspace from the respiratory bronchioles to the alveoli and is typical of emphysema occurring in alpha-1-antitrypsin deficiency (A1ATD; section 1.1.3). (31) Emphysema in this group of patients is also typically in the lower zones in contrast to usual COPD patients. (32) Finally, paraseptal emphysema is characterized by enlarged airspaces in the periphery of the lobules and is commonly associated with fibrosis or other types of emphysema. (31)
CT can be used to measure the density of the lungs and CT densitometry has been demonstrated as a useful measure of the severity of emphysema in vivo.\footnote{33} Various measures of density can be used but the 15\textsuperscript{th} percentile point (Perc15) (the cut-off value in Hounsfield Units below which 15\% of lung voxels are distributed) is a true measure of density and decreases as the severity of emphysema worsens.\footnote{33} For this reason Perc15 is the most commonly used measure of CT density and has been demonstrated to be the most robust at measuring emphysema progression.\footnote{33} This has led to the use of Perc15 as an outcome measure in COPD clinical trials such as alpha-1-antitrypsin augmentation therapy.\footnote{34} Appearances of emphysema on CT are shown in figure 1.2.

High Resolution CT (HRCT) is the investigation of choice to diagnose bronchiectasis.\footnote{25} Classically bronchial wall dilatation (lumen greater than accompanying pulmonary artery or lack of tapering) is seen.\footnote{25, 35} This can be classified according to the shape of the bronchi into cylindrical, varicose and cystic forms.\footnote{36} Bronchial wall thickening and mucous plugging of smaller airways (‘tree-in-bud’ pattern) may also be visualised.\footnote{35, 37} Patients typically present with a daily productive cough.\footnote{38} The colour of the sputum produced is related to purulence (release of neutrophil myeloperoxidase)\footnote{39} which relates to changes seen on HRCT.\footnote{40} Cystic and varicose variants of bronchiectasis are associated with greater sputum purulence.\footnote{40} Typical images of bronchiectasis on HRCT are seen in figure 1.3.
Table 1.1: Severity of COPD by spirometry

<table>
<thead>
<tr>
<th>Severity</th>
<th>FEV1 % predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>&gt; 80</td>
</tr>
<tr>
<td>Moderate</td>
<td>50-80</td>
</tr>
<tr>
<td>Severe</td>
<td>30-50</td>
</tr>
<tr>
<td>Very Severe</td>
<td>&lt;30</td>
</tr>
</tbody>
</table>

This table outlines how the severity of COPD is graded according to results recorded by spirometry. This classification is used in both British and American/European guidelines. FEV1 = Forced expiratory volume in one second.

Figure 1.1: GOLD classification of COPD: Severity of COPD according to the international GOLD classification. Patients are placed into one of four groups (A-D) based on their severity of airflow obstruction on spirometry (table 1.1), symptoms (according to one of two symptom-severity scores, mMRC or CAT) and history of exacerbations over the past 12 months. mMRC = Modified British Medical Research Council Questionnaire. CAT = COPD Assessment Test.
Figure 1.2: Appearances of emphysema on CT. This is a scan of the thorax at the level of the heart. It represents a cross-sectional slice of tissue viewed as if you are looking towards the patient’s head. There is enlargement of airspaces and tissue destruction throughout both lungs (highlighted).

Figure 1.3: Appearances of bronchiectasis on CT. This is a scan of the thorax at the level of the heart. It represents a cross-sectional slice of tissue viewed as if you are looking towards the patient’s head. Thickened and enlarged bronchioles can be seen. An example of such a bronchiole is highlighted.
1.1.2 Management

Management of COPD can be divided into non-pharmacological and pharmacological aspects. The most important treatment intervention is smoking cessation that both slows the decline of FEV1 and improves patient symptoms and overall survival. (17, 41-43) Therefore, current NICE guidance recommends all COPD patients should be given encouragement to help stop smoking regardless of age. (17) Patients should be offered attendance at a support programme for smoking cessation which may also include smoking cessation therapy: nicotine replacement therapy (NRT), bupropion or varenicline. (17)

Pulmonary rehabilitation, an individualised programme of care including education and exercise, has been shown to improve quality of life and symptoms. (44) It is recommended in patients who are functionally disabled by their COPD. (17)

Pharmacological treatment in COPD primarily consists of inhaled therapy targeting airflow obstruction and inflammation. (17) Short acting bronchodilators (beta2 agonists and muscarinic antagonists) are recommended in the first instance to relieve symptoms. (17) NICE recommend treatment is then increased in a stepwise fashion in patients who continue to have symptoms. If FEV1 ≥50% long-acting bronchodilators are recommended (long-acting beta2 agonists (LABA) or long acting muscarinic antagonists (LAMA)). (17) If FEV1 < 50% combined therapy with inhaled steroids (ICS) (LABA+ICS) or LAMA alone is recommended. (17) Finally, LABA, LAMA and ICS can be used at once in patients with ongoing symptoms or exacerbations. (17)
Recently, several combination inhalers of LABA/LAMA have been licensed for the management of COPD. These have been shown to improve lung function and symptoms in comparison with monotherapies. (45) A meta-analysis of 6 studies comparing LABA/LAMA to LABA/ICS have also shown that these two combinations appear to be equivalent at preventing exacerbations. (46) However, evidence that LABA/LAMAs are more effective at preventing exacerbations compared to LAMAs alone is lacking. (45, 47) This class of medications is not recommended by NICE, but is recommended by GOLD for patients with breathlessness without a history of exacerbations. (12)

Other treatments used in the management of stable COPD include oral theophylline (another bronchodilator with potential anti-inflammatory properties (48)) which can be trialled after inhaled therapy has been optimized and oral mucolytics (in patients with a productive cough). (17) Long term oxygen therapy has also been shown to improve prognosis in patients with an arterial partial pressure of oxygen of less than 7.3 or 8 with co-morbidities (such as pulmonary hypertension). (17, 49)

Exacerbations of COPD (defined as a worsening of a patient’s symptoms) may be infective or non-infective. (17) Similarly the main focus of treatment is bronchodilation and anti-inflammatory treatment with steroids. Antibiotics may also be required if a patient is producing purulent sputum. (12, 17) A summary of interventions in COPD and their cost is presented in figure 1.4.
Figure 1.4: COPD value pyramid. The cost in quality-adjusted life years (QALYs) for common COPD interventions.(50, 51) Interventions at the top of the pyramid are more expensive and have a higher QALY than interventions in the bottom part of the pyramid. Simple interventions such as smoking cessation represent better value (in terms of QALYs) in comparison to medications.
1.1.3 Pathogenesis

Numerous mechanisms have been proposed in the pathogenesis of COPD. One of the key mechanisms is the balance between anti-proteases and proteases. Different proteases break down connective tissue in the lung such as elastin; this enzyme release is prevented from doing tissue damage in healthy patients by protective antiproteases. A good clinical example of the importance of protease balance is alpha1-antitrypsin deficiency (A1ATD); a genetically determined anti-protease deficiency which predisposes to COPD. A1AT (alpha-1-antitrypsin) is an anti-protease that inhibits the actions of serine proteases such as neutrophil elastase (NE). Cysteine proteases (cathepsins) which are found in lysosomes may also be involved in the pathogenesis of COPD. For example, cathepsin inhibitors reduce the development of emphysema in transgenic mice. The third class of proteases implicated in the development of COPD are matrix metalloproteases (MMPs). Several MMPs are found in increased amounts in patients with COPD and transgenic MMP-12−/− mice are also resistant to emphysema induced by chronic smoke exposure.

As well as increased protease activity, reduced anti-protease activity may also play a role in the development of usual COPD. For example, oxidative stress in COPD patients (discussed below) may inactivate both A1AT and secretory leukocyte protease inhibitor (SLPI); the other major serine protease inhibitor in the airways. MMPs are inhibited by tissue inhibitors of MMPs (TIMPs) of which there are 4 types. Although there is evidence that 2 of the TIMPs are up-regulated in smokers there is also an increase in the frequency of loss-of-function mutations in TIMP-2 in COPD patients suggesting that TIMP activity may be reduced.
There is a wealth of evidence demonstrating systemic inflammation in COPD, as shown by increased levels of chemokines, cytokines and acute phase reactants. Smoking can cause inflammation, but the degree seen in COPD is higher than in smokers alone and persists despite smoking cessation. Inflammation in COPD also appears to be greater in severe disease, although its variability over time and with exacerbations has thwarted attempts to find biomarkers in the blood that relate consistently to clinical features. Interleukin-6 (IL-6), C reactive protein (CRP), IL-8 and surfactant protein D (SP-D) levels are typically high in COPD and are important in the recruitment of inflammatory cells, although fibrinogen seems the most reliable biomarker to date. Fibrinogen levels are associated with increased exacerbation rates and poorer outcomes, however it is a non-specific marker and as such has inherent weaknesses. More specific related markers, such as Aa-Val360 may prove more useful in the future. There is an accumulation of inflammatory cells in COPD lungs, including macrophages, neutrophils, B cells and CD4+ and CD8+ T cells. Macrophages release multiple inflammatory mediators including reactive oxygen species, cytokines, chemokines, extracellular matrix proteins and MMPs. In COPD their function may be impaired, for instance they show impaired phagocytosis of bacteria, which may result in an increased inflammatory response to bacteria in the lower airways. Neutrophils also produce reactive oxygen species, elastase and cytokines, which play a role in emphysema and COPD development. Lymphocytes, including both B and T cells, are also found in high numbers in COPD lung and may be involved in immune activation, leading to perpetuation of inflammation and on-going parenchymal destruction. Such a reaction is typical of autoimmune disease, and
characteristics of autoimmunity have been reported in COPD (76) although whether they are cause or effect is a matter of debate.(77)

Oxidative stress has also been proposed as a causative factor in COPD. The normal metabolism of oxygen results in the development of reactive oxygen species (ROS) - these are usually removed from the cell by enzymes or anti-oxidants such as superoxide dismutase.(53, 78) If the balance between the formation and removal of ROS is disturbed oxidative stress can occur.(78) Oxidative stress is well recognized in COPD and is particularly elevated during exacerbations.(79) Cigarette smoke is a key driver of oxidative stress.(80) It contains noxious chemicals which are metabolized to benign and/or toxic metabolites, the latter of which can damage tissue and predispose to disease.(80) Other oxidants such as ozone may also play a role in oxidative stress in these patients.(80) Inflammatory cells which are activated in COPD such as neutrophils also provide an endogenous source of oxidative stress by producing ROS.(81) Oxidative stress may play a role in the development of COPD via different mechanisms. For example, oxidants activate intracellular pathways by upregulating the expression of multiple inflammatory genes resulting in damage to the lung. (53) Oxidative stress also damages cells in the lung directly resulting in cell lysis and increased epithelial permeability. Increased oxidants cause mucous hypersecretion and impaired mucociliary clearance further contributing to the disease process.(82)

The pathogenesis of COPD is summarized in figure 1.5.
Figure 1.5: Pathogenesis of COPD: Inflammatory cells release reactive oxygen species causing oxidative stress. Inflammatory cells also release proteases disrupting protease/antiprotease balance. Inflammation results in parenchymal destruction and COPD. Pollutants such as cigarette smoke cause inflammation directly and indirectly by producing oxidative stress. Oxidative stress also results in lung damage directly. Oxidative stress reduces antiprotease activity further disrupting protease/antiprotease balance resulting in parenchymal damage.
1.2 Lung Cancer

Lung cancer is common: Over 39 000 patients are diagnosed with lung cancer each year in the UK and it is now the leading cause of female cancer death.(83) 90% of lung cancers are caused by smoking, although other environmental factors such as asbestos exposure also play a role.(84) Lung cancers are classified by histology into small cell and non-small cell cancers (NSCLC); the latter are the more common and include squamous and adenocarcinomas.(85) Generally small cell tumours have the poorest prognosis, with overall survival being 9 months when managed by chemotherapy.(86) Even in NSCLC prognosis is generally poor unless surgery can be offered, and a systematic review reported mean survival of just over 7 months if left untreated.(87)

1.2.1 Clinical Features

The incidence of lung cancer is related to age; patients are at increased risk of lung cancer if they are over 40 years old.(88) In 2012 43% of all patients diagnosed in the UK were over 75 years old.(6) NICE recommend that people should be referred urgently for suspected lung cancer via a 2 week wait referral to hospital if they are over 40 and present with unexplained haemoptysis.(88) An urgent chest x-ray to look for lung cancer should be performed in people over 40 with 2 or more symptoms suggestive of lung cancer (or one symptom if they previously smoked).(88) Symptoms can be non-specific such as cough, weight loss and breathlessness such that clinical suspicion is key to making the diagnosis.(88) A chest x-ray is also required in people over 40 with possible signs of lung cancer on clinical examination.
(such as finger clubbing), recurrent chest infections, lymphadenopathy or thrombocytosis.(88)

Once a patient has been referred to a specialist for suspected lung cancer the initial investigation of choice is a CT scan of the chest, lower neck and upper abdomen.(83, 89) If the CT confirms the presence of a likely lung cancer the next investigations should be tailored to the individual patient depending on the location of the tumour. Investigations commonly used include biopsy (via CT or ultrasound guidance) and Positron emission tomography (PET)-CT scanning.(83, 89)

Once imaging and biopsy results are available the lung cancer is further defined in two ways: histological type and stage. Both of these play a role in the management of the lung cancer. The International Association for the Study of Lung Cancer (IASLC), American Thoracic Society (ATS) and the European Respiratory Society (ERS) released guidelines on the classification of lung cancer in 2011.(90) A summary of this can be found in table 1.2.

Approximately 85% of all lung cancers diagnosed are non-small cell lung carcinoma (NSCLC), the majority of which are adenocarcinoma.(91, 92) Squamous cell carcinoma is most strongly associated with smoking; over 90% of cases are smoking related.(91) Adenocarcinoma is more common in never-smokers than smokers and is the most common type of lung cancer in young patients.(92, 93) Small cell lung carcinoma (SCLC) is distinct from other subgroups as it has a rapid doubling time and tends to metastasize early. SCLC is therefore particularly sensitive to chemotherapy, but the prognosis is poor and most patients die from recurrent disease.(94, 95)
# Table 1.2: Histological classifications of lung cancer (90)

<table>
<thead>
<tr>
<th>Non-small cell carcinoma (NSCLC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Divisions of NSCLC</strong></td>
</tr>
<tr>
<td><em>Adenocarcinoma</em>, further divisions:</td>
</tr>
<tr>
<td>Patterns present: acinar, papillary, micropapillary, lepidic, mucinous, fetal, colloid</td>
</tr>
<tr>
<td>Signet ring features</td>
</tr>
<tr>
<td>Clear cell features</td>
</tr>
<tr>
<td><em>Adenosquamous carcinoma</em></td>
</tr>
<tr>
<td><em>Squamous cell carcinoma</em></td>
</tr>
<tr>
<td><em>Large cell neuroendocrine carcinoma</em> (LNEC)</td>
</tr>
<tr>
<td><em>Poorly differentiated NSCLC with spindle and/or giant cell carcinoma</em> (previously sarcomatoid)</td>
</tr>
<tr>
<td><em>NSCLC, not otherwise specified</em> (previously included large cell carcinoma)</td>
</tr>
</tbody>
</table>

This table outlines the classification of lung cancer according to the International Association for the Study of Lung Cancer (IASLC), American Thoracic Society (ATS) and the European Respiratory Society (ERS). Lung cancer is divided into 2 main groups: Small cell lung carcinoma (SCLC) and Non-small cell lung carcinoma (NSCLC). NSCLC is further divided into several other types based on histological appearances.
Lung cancer can be further divided into stages according to the ‘Tumour, Node, Metastases’ (TNM) staging from the International Union against cancer.(96) This has been used by NICE in the UK since 2010.(83) In this staging system the tumour, lymph nodes and presence/absence of metastases are rated and these ratings are combined to give an overall stage.(96) In general, patients with stage I cancer have the best prognosis and those with stage IV the worst.(96) This has now been validated in different centres.(97)

1.2.2 Management
Lung cancer is managed by groups of different specialists in lung cancer multi-disciplinary teams (MDTs). NICE recommend that all patients with lung cancer are discussed in such MDTs, which are considered quorate when they include respiratory physicians, lung cancer specialist nurses, oncologists, pathologists, radiologists and thoracic surgeons.(83)

Treatment is generally determined by cancer stage, histological subtype and patient fitness.(83, 89) It can be divided into treatment with and without curative intent (table 1.3).(83, 89)

As SCLC behaves differently to NSCLC treatment is considered differently by NICE.(83) Patients with limited disease (T1-4N0-3M0) should be offered platinum based chemotherapy and concurrent radiotherapy if their disease can be encompassed within a radical thoracic radiotherapy volume and their performance status is good.(83, 98) In a very small proportion of patients (T1-2aN0M0) surgery should be considered.(83, 99) However, due to the aggressive nature of the disease
Table 1.3: Treatments used in the management of NSCLC

<table>
<thead>
<tr>
<th>Type of intervention</th>
<th>When used</th>
<th>Curative?</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonectomy</td>
<td>T1a-3N0-1M0</td>
<td>Possibly</td>
<td>(83)</td>
</tr>
<tr>
<td>Lobectomy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited resections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radical Radiotherapy: CHART SABR</td>
<td>T1a-3N0-1M0 with poor patient fitness</td>
<td>Possibly</td>
<td>(100, 101)</td>
</tr>
<tr>
<td>Adjuvant chemotherapy</td>
<td>Tumours &gt;4cm T1-3N2M0</td>
<td>Possibly with radical treatment</td>
<td>(102)</td>
</tr>
<tr>
<td>Palliative chemotherapy</td>
<td>Stage III or IV</td>
<td>No</td>
<td>(83)</td>
</tr>
</tbody>
</table>

This table outlines common forms of treatment for NSCLC. The second column lists when these treatments would be used according to TMN staging criteria. The final column lists whether or not the treatment is curative. CHART: Continuous hyperfractionated accelerated chemotherapy; SABR: Stereotactic ablative radiotherapy.

Table 1.4: Palliative therapies used in the management of lung cancer

<table>
<thead>
<tr>
<th>Reason for treatment</th>
<th>Intervention</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypercalcaemia</td>
<td>Bisphosphonates</td>
<td>(103)</td>
</tr>
<tr>
<td>Pain or cough</td>
<td>Opioids</td>
<td>(83)</td>
</tr>
<tr>
<td>Symptoms arising from primary cancer eg breathlessness, haemoptysis, chest pain</td>
<td>Palliative thoracic radiotherapy</td>
<td>(104)</td>
</tr>
<tr>
<td>Brain metastases</td>
<td>Dexamethasone and cranial radiotherapy</td>
<td>(105)</td>
</tr>
<tr>
<td>Superior vena cava (SVC) obstruction</td>
<td>Stenting +/- chemotherapy and radiotherapy</td>
<td>(106)</td>
</tr>
<tr>
<td>Bony metastases</td>
<td>Radiotherapy</td>
<td>(107)</td>
</tr>
<tr>
<td>Metastatic spinal cord compression</td>
<td>Spinal radiotherapy</td>
<td>(108)</td>
</tr>
<tr>
<td>Breathlessness due to malignant pleural effusion</td>
<td>Pleural aspiration/pleurodesis</td>
<td>(83, 109)</td>
</tr>
<tr>
<td>Endobronchial obstruction by tumour</td>
<td>Endobronchial debulking (surgical/flexible bronchoscopy)</td>
<td>(110)</td>
</tr>
</tbody>
</table>

This table outlines common palliative treatments used in all types of lung cancer and reasons why these treatments might be considered.
it is likely that chemotherapy and/or radiotherapy will still be required in these patients.(89)

In addition to surgery/radiotherapy/chemotherapy patients may require symptomatic or palliative treatment. Palliative treatments are summarized in table 1.4.

New treatments designed to target particular mutations in lung cancer have also been developed. It is possible these treatments will allow clinicians to tailor specific treatments to certain patients rather than treating patients as a homogenous group. Examples of new treatments in development are outlined in table 1.5.

1.2.3 Pathogenesis

Smoking, occupational toxins and air pollution may result in damaging mutations which have potential to induce dysplastic and neoplastic changes in the lung parenchyma due to alterations in cell differentiation, growth and death.(111) Cigarette smoke contains over 60 carcinogens (112) which can be activated by cytochrome P450 (CYP) enzymes; inhalation of these, pollutants and microorganisms can cause damage directly or due to oxidative stress.(111) This results in DNA adducts which in turn may cause mutations if they are not repaired.(113) Commonly these mutations are in oncogenes (such as p53), but they may also affect inflammatory pathways.(114, 115) Mostly these mutations are repaired, but when there is a high rate of damage due to ROS cells are likely to be transformed to a malignant phenotype.(114) Table 1.6 outlines several of the genes known to be mutated in lung cancer. As shown in table 1.5 several of these now have specific drugs targeting them.
Table 1.5: New therapies in lung cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Target mutation</th>
<th>Stage of testing(116)</th>
<th>Licensed</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZD4547</td>
<td>Fibroblast growth factor</td>
<td>Phase II/III</td>
<td>No</td>
<td>(117)</td>
</tr>
<tr>
<td>AZD2014</td>
<td>Mechanistic target of rapamycin (mTOR)</td>
<td>Phase II</td>
<td>No</td>
<td>(118)</td>
</tr>
<tr>
<td>Palbociclib</td>
<td>Cyclin-dependent kinase-4/6 (CDK-4/6)</td>
<td>Phase II/III</td>
<td>No</td>
<td>(119)</td>
</tr>
<tr>
<td>Crizotinib</td>
<td>Anaplastic lymphoma kinase (ALK)</td>
<td>Phase III</td>
<td>Adenocarcinoma</td>
<td>(120)</td>
</tr>
<tr>
<td>Selumetinib</td>
<td>Mitogen-activated protein kinase kinase (MEK)</td>
<td>Phase II</td>
<td>No</td>
<td>(121)</td>
</tr>
<tr>
<td>AZD5363</td>
<td>Protein kinase B (AKT)</td>
<td>Phase II</td>
<td>No</td>
<td>(122)</td>
</tr>
<tr>
<td>AZD9291</td>
<td>Epidermal growth factor receptor (EGFR) T790M</td>
<td>Phase II</td>
<td>No</td>
<td>(123)</td>
</tr>
<tr>
<td>Durvalumab</td>
<td>Programmed death-ligand 1 (PDL-1)</td>
<td>Phase III</td>
<td>No</td>
<td>(124)</td>
</tr>
</tbody>
</table>

This table outlines new treatments for lung cancer based on specific mutations within cancer cells. The third column lists what phase of testing the treatment is at and the fourth column lists whether or not the treatment is currently licensed.
Table 1.6: Examples of genes commonly mutated in lung cancer

<table>
<thead>
<tr>
<th>Gene</th>
<th>Effect</th>
<th>Drug licensing stage</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td>Proliferation, cell survival</td>
<td>Licensed</td>
<td>(125)</td>
</tr>
<tr>
<td>ALK</td>
<td>Proliferation, cell survival</td>
<td>Licensed</td>
<td>(126)</td>
</tr>
<tr>
<td>CDK4</td>
<td>Proliferation, migration</td>
<td>II/III</td>
<td>(127)</td>
</tr>
<tr>
<td>FGF</td>
<td>EMT</td>
<td>II/III</td>
<td>(128)</td>
</tr>
<tr>
<td>AKT</td>
<td>Proliferation, cell survival, migration</td>
<td>II</td>
<td>(129)</td>
</tr>
<tr>
<td>MEK</td>
<td>Proliferation, cell survival, uncontrolled cell differentiation, angiogenesis</td>
<td>II</td>
<td>(130)</td>
</tr>
<tr>
<td>mTOR</td>
<td>Proliferation, migration</td>
<td>II</td>
<td>(131)</td>
</tr>
<tr>
<td>c-Kit</td>
<td>Proliferation</td>
<td></td>
<td>(132)</td>
</tr>
<tr>
<td>FHIT</td>
<td>Proliferation, cell survival</td>
<td></td>
<td>(133)</td>
</tr>
<tr>
<td>FUS1</td>
<td>Proliferation, cell survival</td>
<td></td>
<td>(134)</td>
</tr>
<tr>
<td>MYC</td>
<td>Proliferation, cell survival, uncontrolled cell differentiation, angiogenesis, migration</td>
<td></td>
<td>(135)</td>
</tr>
<tr>
<td>p16INK</td>
<td>Proliferation</td>
<td></td>
<td>(136)</td>
</tr>
<tr>
<td>p53</td>
<td>Cell survival</td>
<td></td>
<td>(137)</td>
</tr>
<tr>
<td>PI3K</td>
<td>Proliferation, cell survival, migration</td>
<td></td>
<td>(138)</td>
</tr>
<tr>
<td>PTEN</td>
<td>Proliferation, cell survival, migration</td>
<td></td>
<td>(139)</td>
</tr>
<tr>
<td>RAS</td>
<td>Proliferation, cell survival, angiogenesis</td>
<td></td>
<td>(140)</td>
</tr>
<tr>
<td>RASSF1A</td>
<td>Proliferation, cell survival, microtubule instability, EMT</td>
<td></td>
<td>(141)</td>
</tr>
<tr>
<td>RB</td>
<td>Proliferation</td>
<td></td>
<td>(142)</td>
</tr>
<tr>
<td>SEMA3B</td>
<td>Cell survival</td>
<td></td>
<td>(143)</td>
</tr>
<tr>
<td>SEMA3F</td>
<td>Cell survival</td>
<td></td>
<td>(144)</td>
</tr>
<tr>
<td>STAT</td>
<td>Proliferation, cell survival, migration, angiogenesis</td>
<td></td>
<td>(145)</td>
</tr>
<tr>
<td>TGFβ</td>
<td>EMT, migration</td>
<td></td>
<td>(146)</td>
</tr>
</tbody>
</table>

This table outlines genes commonly mutated in lung cancer and the effect of the mutation. EGFR = Epidermal growth factor receptor. ALK = Anaplastic lymphoma kinase. CDK4 = Cyclin-dependent kinase-4. FGF = Fibroblast growth factor. AKT = Protein kinase B. MEK = Mitogen-activated protein kinase kinase. mTOR = Mechanistic target of rapamycin. c-Kit = stem cell factor receptor. FHIT = Fragile histidine triad protein. FUS1 = FUS RNA binding protein. MYC = MYC proto-oncogene. bHLH transcription factor. p16INK = cyclin-dependent kinase inhibitor 2A. multiple tumor suppressor 1. P53 = tumour protein 53. PI3K = Phosphatidylinositol-4,5-bisphosphate 3-kinase. PTEN = phosphatase and tensin homolog. RASSF1A = Ras association domain family member 1. SEMA3B = Semaphorin 3B. SEMA3F = Semaphorin 3F. STAT = Signal transducer and activator of transcription. TGFβ = Transforming growth factor-β.
There is evidence that carcinogenesis occurs at sites of chronic inflammation.\(^{(147)}\) For example, hepatocellular carcinoma can occur in patients with chronic hepatitis and colon cancer in the setting of colitis.\(^{(148)}\) There is some evidence that increased inflammation may also be associated with the development of lung cancer. Epidemiologically, a cohort study of 7081 patients showed an increased risk of lung cancer in patients with a CRP of > 3 mg/dL.\(^{(149)}\) Furthermore, a mouse model of chronic inflammation showed increased lung tumourigenesis.\(^{(150)}\) However, clinical studies of anti-inflammatory drugs, such as inhaled corticosteroids, have shown inconsistent results. A cohort study has shown lower rates of lung cancer compared to patients not taking inhaled corticosteroids\(^{(151)}\), whilst a randomized controlled trial (RCT) did not.\(^{(152)}\) Whether targeting pulmonary inflammation to prevent lung cancer will be beneficial therefore remains uncertain. One way inflammation may lead to the development of lung cancer is by activation of the epithelial growth factor (EGFR) cascade. This is activated in response to oxidative stress, neutrophil elastase and other proteases.\(^{(153)}\) Overexpression of EGFR has been associated with a high risk of developing lung cancer and can occur years after smoking cessation.\(^{(154)}\) The arachidonic acid metabolic pathway may also be related to lung cancer development. Inflammatory cells release arachidonic acid metabolites including prostaglandins - this is mediated by cyclooxygenase enzymes (COX) including COX-2.\(^{(147)}\) Prostaglandin E2, the product of COX-2, regulates the inflammatory response, but also has effects on cell proliferation, apoptosis and angiogenesis\(^{(147)}\) and therefore may have a role in cancer development. Raised COX-2 levels relate to survival in NSCLC\(^{(155)}\), inhibition of COX-2 reduces lung cancer in animal models\(^{(155)}\) and patients who regularly take COX-2 inhibitors have
reduced rates of lung cancer.(156) Finally, carriers of a polymorphism of the COX-2 gene have an increased risk of lung cancer.(157)

Epithelial-mesenchymal transition (EMT), a process where epithelial cells gradually acquire a mesenchymal phenotype may have an important role in the development of lung cancer.(158) During this process epithelial cells undergo multiple changes in transcription, motility and cell adhesion properties.(158) Epithelial cells typically have tight cell-cell interactions and lack of mobility whereas mesenchymal cells are mobile with only loose cell-cell interactions.(159) During EMT key transcription factors (such as Snail) (160) are upregulated resulting in expression of different cytoskeletal proteins and cellular junctional components.(158) Typically the expression of E-cadherin reduces and N-cadherin (expressed on mesenchymal cells) increases. This phenomenon can be used to monitor the development of EMT.(161) This change to a mesenchymal cell type is thought to allow the malignant cells to migrate contributing to cancer invasion and metastasis and may also play a role in resistance to chemotherapeutic drugs.(162) Different signalling factors can induce EMT including fibroblast growth factor and Notch although transforming growth factor-β (TGF-β) has been the most commonly studied.(162) TGF-β directly activates known EMT transcription factors thus repressing the expression of E-cadherin.(163) TGF-β may be particularly important in lung cancer as expression of TGF-β is upregulated by smoking.(159) Increased expression of TGF-β has also been shown to be related to increased metastatic potential of NSCLC tissues.(164)
1.3 Evidence for the shared pathogenesis of COPD and lung cancer

The relative risk of lung cancer in COPD is over twice that of the general population. (11) As both diseases are related to smoking it could be assumed that the relationship between COPD and lung cancer might be related to smoking alone. However, in recent years it has been shown that patients with COPD, or CT diagnosed emphysema, have a higher risk for lung cancer even when adjusting for smoking history. (165, 166) Furthermore, increased severity of airway obstruction is associated with a correspondingly elevated risk of lung cancer. (167) Emphysema in non-smokers is also a risk for lung cancer further reinforcing that smoking, alone, cannot explain the relationship between COPD and lung cancer. (111) Taken together this evidence suggests that there is a link between COPD and lung cancer other than smoking.

Only 20%-30% of smokers develop COPD and 10%-15% develop lung cancer. (16) Familial linkage studies and genome-wide association studies (GWAS) have been performed for COPD, lung function and lung cancer and there is overlap in the chromosomal areas identified, which demonstrates shared genetic risk. (104) For example, Wang et al looked at the gene expression pattern in lung samples from COPD patients and demonstrated that genes involved in extracellular membrane synthesis and apoptosis were up-regulated, whilst genes involved in the anti-inflammatory response were down-regulated. (168) Another study looking at gene expression in squamous cell carcinoma found that in patients with co-existing COPD there was a more frequent loss of 5q or a low expression of genes on 5q than in patients without co-existing COPD. (169)
There is impairment in DNA repair in COPD due to low levels of Ku 86, a protein involved in DNA repair. (170) This suggests that oxidant induced damage in COPD patients is more likely to result in carcinogenesis. It is also possible that processes which control DNA repair may influence a patient’s risk of developing lung cancer and that such processes may be altered in COPD. (171) For example, there is evidence that acetylation of histone H3 on lysine 56 (H3K56) is important in DNA repair. (172) Deacetylation of H3K56 is controlled by histone deactylases (HDACs) 1 and 2 and sirtuin (SIRT) 1. (173) As there are low levels of HDAC2 and SIRT1 in COPD this may reduce the protection against DNA breakage caused by environmental factors further increasing lung cancer risk. (174)

Epigenetics is the regulation of gene expression by heritable mechanisms that do not make direct changes to DNA itself. (111) Examples of epigenetic mechanisms include histone acetylation and methylation. (111) Tumour suppressor genes, including p53, may be rendered inactive in patients with low SIRT1 levels, including COPD patients. (175, 176) Methylation of the promoter of p16 (a tumour suppressor gene) is seen in the sputum of COPD patients and aberrant methylation of p16 can be also seen in the sputum of patients with NSCLC suggesting this too may be a shared mechanism of disease. (177)

It is likely that mechanisms underlying COPD development increase the risk of lung cancer development. For example, inflammation in COPD might predispose to the development of malignancy. (64, 147) Oxidative stress in COPD may also increase cancer risk. Nuclear factor erythroid 2-related factor 2 (Nrf2) is the main transcription factor that regulates phase II detoxifying antioxidant enzymes. (178) Low Nrf2
expression occurs in COPD and this may also predispose patients with COPD to lung cancer due to increased oxidative stress. (179) Oxidative stress also increases p21 expression, a cyclin-dependent kinase inhibitor whose levels are raised in patients with COPD and lung cancer. (180) Elevation of p21 results in hyperproliferation and carcinogenesis. (181) As above, the balance between anti-proteinases and proteinases is an important determinant of emphysema that occurs in COPD. (52) Proteinases are also important in lung cancer development as they release growth factors such as transforming growth factor-β (TGF-β) which can lead to tumorigenesis. (111) TGF-β also drives epithelial cell mesenchymal transition (EMT), which is regulated by galectin 3. (163) There are raised levels of galectin 3 in COPD lung and increased levels are also associated with poor prognosis in NSCLC. (182, 183) A summary of the relationship between COPD and lung cancer is presented in figure 1.6

1.4 The role of the endothelium in the lung

The pulmonary vasculature is critical to gas exchange in the lung, with a total pulmonary vascular surface area of 90 m². (184) The entire vascular system is lined by endothelial cells, which form a continuous monolayer. (184) Endothelial cells are encased by a basement membrane, a thin protein sheet (50 nm thick) that consists of laminins, collagen and proteoglycans. (185) Endothelial cells are also covered on the luminal side by the glycocalyx, a network of proteoglycans and glycoproteins involved in multiple processes such as cell-cell signalling and haemostasis. (186) Finally, embedded in the basement membrane are a non-continuous layer of pericytes, which are key mediators of several microvascular processes such as endothelial cell
Figure 1.6: A summary of the relationship between COPD and lung cancer (1). The figure shows some of the key pathways leading to both chronic obstructive pulmonary disease (COPD) and cancer, and demonstrates the complexity of the interactions between the diseases. Cigarette smoke causes oxidative stress which can both drive inflammation and occur due to inflammation; both processes lead to COPD. Inflammation may in turn lead to activation of matrix metalloproteases (MMPs) and the transforming growth factor-β (TGF-β) pathway, which by way of epithelial mesenchymal transition can promote lung cancer growth. Oxidative stress may also directly activate the epidermal growth factor receptor pathway, which is involved in lung cancer growth. Cigarette smoke also interacts with pre-existing genetic predisposition and causes changes in DNA and miRNA, which lead to processes relevant to cancer growth, such as cell proliferation and apoptosis, as well as to COPD. Finally, COPD may cause hypoxia that may augment angiogenesis, thereby interacting with prostaglandin based pathways to influence cell proliferation further, with the potential to influence cancer risk. EGFR: epithelial growth factor. EMT: epithelial cell mesenchymal transition.
proliferation and angiogenesis. (187, 188) A diagram of the structure of the endothelium is shown in figure 1.7.

The endothelium does not represent a passive lining of the vasculature, but a dynamic tissue involved in vessel homeostasis, response to pathological stimuli and cell-cell interactions.(184)

One particular important function of the pulmonary endothelium is maintenance of pulmonary vascular tone by the production of mediators such as nitric oxide (NO).(189) There is evidence that NO production is important in maintaining the low resistance seen in pulmonary vasculature as systemic delivery of NO synthase inhibitors increases pulmonary vascular resistance.(189) Prostacyclin (PGI$_2$), an eicosanoid, is another potent vasodilator produced by the endothelium.(190) In fact, systemic delivery of PGI$_2$ has been used as a treatment of primary pulmonary hypertension (PPH); a condition characterised by pulmonary vascular constriction.(190) However, endothelial cells can also produce potent vasoconstrictors in addition to vasodilators. For example, endothelin-1 (ET-1), which is produced by endothelial cells in response to several stimuli such as hypoxia, sheer stress, angiotensin and growth factors.(191) Interestingly, ET-1 production appears to be reduced by NO and may therefore partly explain the vasodilatory response to NO.(192) Targeting ET-1 has also been exploited in the treatment of PPH.(191) Another eicosanoid, thromboxane A2 (TXA$_2$), is a potent vasoconstrictor and acts in opposition to PGI$_2$.(193) TXA$_2$ acts via a G protein coupled receptor to increase calcium influx into smooth muscle cells resulting in constriction.(193) Reactive oxygen species (ROS) produced in response to oxidative stress may also act via NO to cause vasoconstriction by reducing NO bioavailability.(194) Therefore, mediators
produced by the endothelium may result in pulmonary vasoconstriction and vasodilation and so perhaps it is the overall balance of these mediators that determines the overall vascular response.

Neutrophil transendothelial migration (TEM) is an important mechanism by which the endothelium plays a role in the lung. As previously discussed (section 1.1.3) neutrophils play an important role in the inflammatory response. (54) In order to reach the lung tissue neutrophils must bind to, and migrate through, the endothelium. (188) Initially neutrophils extend part of themselves (a pseudopod) to invaginate the apical endothelial cell membrane. The neutrophil binds to the endothelial cell through a variety of cell surface proteins before migrating between the endothelial cells. (188) This is illustrated in figure 1.8.

Endothelial mediators may also be important in the pulmonary inflammatory response. For example, increased cellular levels of calcium secondary to acute inflammation activate NO synthase.(184) This results in the production of proteins, which produce a scaffold for inflammatory cells. (195, 196) TXA$_2$ is also important in the inflammatory response and appears to upregulate TEM by enhancing the production of adhesion molecules such as intercellular adhesion molecule (ICAM). (197) In addition, release of ROS may also upregulate TEM through a similar mechanism. (198) Finally, another important endothelial inflammatory mediator is von Willebrand Factor (vWF). This factor is important in the clotting cascade, but platelets bound to vWF express P-selectin, which binds to and recruits leucocytes resulting in increased inflammation and a possible link between inflammation and fibrosis.(199)
Figure 1.7: The endothelium. The endothelium consists of a layer of endothelial cells encased in a basal membrane. The endothelial cells are covered on the luminal side by the glycocalyx, a network of proteoglycans and glycoproteins. Finally, pericytes are also embedded in the basal membrane and are important in some endothelium functions such as angiogenesis.

Figure 1.8: Transendothelial migration: The neutrophil passes along the endothelium before binding to an endothelial cell via adhesion molecules (eg MAC-1, illustrated as red bars). The neutrophil invaginates the endothelial cell membrane before migrating between endothelial cells.
Conversely, endothelial cells may also result in the production of anti-inflammatory mediators. For example, NO may act in an anti-inflammatory manner through the reduction of adhesion molecules, which would reduce inflammatory cell infiltration via TEM.(200) Similarly, PGI$_2$ also has anti-inflammatory actions resulting in the reduced vascular permeability and therefore oedema formation, in response to acute inflammatory stimuli.(201) Thus, in a similar fashion to vasomotor tone control, it is likely that the overall balance of these different mediators controls whether the endothelium does or does not promote inflammation.

Finally, another important role of the pulmonary endothelium is its part in the renin-angiotensin system (RAS). Renin (which is produced by the kidney) converts angiotensinogen from the liver into angiotensin I.(202) This in turn is converted to angiotensin II by angiotensin converting enzyme (ACE), which is predominantly expressed in the pulmonary endothelium.(202) Angiotensin II stimulates the adrenal cortex to produce aldosterone resulting in increased salt and water retention from the kidney thus increasing systemic blood pressure.(202) This is important in the response to acute hypotension.(202) In addition to these systemic effects, angiotensin II acts locally in the lung via its receptor AT$_1$R (Angiotensin receptor 1) resulting in vasoconstriction, increased ROS production and vascular remodelling.(203) The effects of angiotensin binding on its other receptor AT$_2$R (Angiotensin receptor 2) appear to oppose AT$_1$R effects in other organ systems but current research into pulmonary AT$_2$R effects is conflicting and therefore controversial at present.(204)
1.4.1 The role of the endothelium in COPD

In the 1950s Liebow demonstrated that alveolar septa in COPD patients were almost avascular. (205) This led to the hypothesis that vascular atrophy resulted in the destruction of alveoli. (205) Supporting this concept, increased levels of apoptotic endothelial cells have been identified in the lungs of patients with COPD. (206) It is also possible to induce emphysema in rodents by deliberately causing endothelial apoptosis. (207) A reduction in endothelium in patients with emphysema may be caused by reduced levels of Vascular Endothelial Growth Factor (VEGF). (208) VEGF is a highly specific growth factor for endothelial cells that is produced in response to hypoxia. (209) It induces both cell proliferation and migration and prevents endothelial cell apoptosis. (209) VEGF levels might be reduced in such patients as Hypoxia Inducible Factor-1α (HIF-1α), a major transcription factor for VEGF, is also reduced in patients with emphysema. (210) Levels of HIF-1α and VEGF may be related to disease severity: both are correlated with FEV1 percentage predicted in patients with emphysema. (210) Interestingly, similar studies looking at the expression of HIF-1α and VEGF in patients with chronic bronchitis (rather than emphysema) have shown HIF-1α and VEGF are increased in this patient group. (211) This suggests that the endothelium might be involved in different ways depending on the clinical presentation of COPD. There is also evidence that vasculature may be altered in the airways of patients with COPD, in addition to their peripheral lung tissue – several groups have shown an increased vascular area in the airways of patients with COPD. (212, 213) It is possible that this might contribute to airway narrowing. (213)

In addition to altered levels of endothelium in patients with COPD, the endothelium appears to behave in a dysfunctional manner. Endothelial dysfunction is defined as
disturbed endothelial dependent vasodilatation. (214) It results in a breakdown of the microvascular endothelial barrier and loss of the anti-adhesive and anti-thrombotic functions of the endothelium. (214) Endothelial dysfunction is associated with severity of COPD and is related to FEV1. (215, 216) Dysfunction is also related to clinical outcomes: patients with increased endothelial dysfunction have reduced 6 minute walk test results and a worse overall prognosis. (217, 218) Endothelial dysfunction is also increased in patients with exacerbations of COPD. (219) Therefore it has been postulated that increased endothelial dysfunction may induce the development of systemic atherosclerosis and therefore the increased cardiac events seen in these patients. (218)

Endothelial dysfunction was previously measured by arterial catheterization to identify the response of the artery to acetylcholine. (220) Patients with endothelial dysfunction respond with vasoconstriction rather than vasodilatation as expected. (220) However, due to the invasive nature of this technique flow mediated dilation (FMD) of the brachial artery was developed as an alternative measurement of endothelial dysfunction. (221) FMD looks at the response of the brachial artery to reactive hyperaemia using Doppler ultrasound and can be used as a surrogate measure of more central endothelial dysfunction. (221) It is reproducible both within and between days when repeated measures are made in COPD patients (222) and associated with FEV1 and percentage of emphysema on CT scan. (216) These associations were independent of smoking and other major causes of endothelial dysfunction. (216) The relationship between FMD and FEV1 was explained by the percentage of emphysema. (216) This suggests that endothelial dysfunction might be involved in emphysema pathogenesis and COPD. (216) Alternatively, one can look at
the blood level of von Willebrand factor (vWF) as an indication of endothelial dysfunction. (223) This is a glycoprotein synthesized by endothelial cells, with increased levels being related to worsening endothelial dysfunction. (223) Exhaled nitric oxide (NO) is another measure of endothelial dysfunction. NO is reduced in endothelial dysfunction due to a reduction in production and/or inactivation of NO synthase by ROS. (223) Maricic et al demonstrated both increased vWF and reduced exhaled NO levels in patients with COPD (223). Finally, the level of endothelial microparticles (EMPs) in blood can be used as a measurement of endothelial dysfunction. EMPs are membrane vesicles which are shed by activated or apoptotic endothelial cells. (219) EMP levels are increased in patients with COPD who have frequent exacerbations compared to those who do not have frequent exacerbations. (219) EMPs are also positively correlated with the severity of emphysema in patients with COPD. (224) Endothelial dysfunction may be a result of the increased levels of oxidative stress seen in COPD. Patients with COPD and low levels of FMD show improvements in FMD when given anti-oxidants. (225)

The RhoA/Rho-kinase pathway which is upregulated in patients with COPD may also result in endothelial dysfunction. (226) RhoA is a small G-protein and Rho-kinase is its’ downstream effector. (226) This pathway is important in a variety of cell functions including migration and proliferation. (226) Levels of RhoA and Rho-Kinase are associated with the level of endothelial dysfunction in patients with COPD. (226) Angiotensin-converting enzyme (ACE), a regulatory protein with both vascular and collagenolitic effects has different variants. (227) The D variant is associated with both endothelial dysfunction and number of exacerbations in patients with COPD and thus
may also play a role in the development of endothelial dysfunction. (227) A summary of the different ways endothelial dysfunction can be measured is in table 1.7.

Although there have not been specific treatment trials for endothelial dysfunction in COPD, many of the above pathways have potential treatments associated with them such as anti-oxidants mentioned above. Patients on long-term ICS have both lower levels of VEGF in pulmonary tissue (228) and reduced endothelial dysfunction compared to patients not on ICS. (229) Therefore, targeting inflammation may be important to improve endothelial dysfunction in these patients. (229) The ACE D variant is associated with increased production of ACE suggesting that perhaps ACE inhibitor drugs (already available for other conditions such as hypertension) may provide another treatment option for this subgroup of patients. (227) Rho-kinase inhibitors have also improved NO release from endothelial cells in vitro suggesting that blocking this pathway may provide another means to improve endothelial dysfunction. (230)

TEM appears to be upregulated in COPD and macrophage-1 antigen (MAC-1), a protein involved in TEM is upregulated in neutrophils from COPD patients. (231) MAC-1 binds to Intracellular adhesion molecule-1 (ICAM-1) on the surface of endothelial cells. Serum levels of ICAM-1 are inversely related to lung function and are also associated with increased percentages of emphysema on CT scan suggesting that this mechanism may be clinically relevant. (232, 233) Blocking the action of ICAM-1 in rodent models has also reduced pulmonary inflammation, further supporting the possibility that ICAM-1 might be related to inflammation in COPD. (234)
<table>
<thead>
<tr>
<th>Measurement</th>
<th>Alteration in COPD</th>
<th>Associated changes in COPD</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMD of brachial artery</td>
<td>Reduced</td>
<td>Increased severity of emphysema. Reduced FEV1.</td>
<td>(216)</td>
</tr>
<tr>
<td>Peripheral blood vWF levels</td>
<td>Increased</td>
<td></td>
<td>(223)</td>
</tr>
<tr>
<td>Exhaled NO levels</td>
<td>Reduced</td>
<td></td>
<td>(223)</td>
</tr>
<tr>
<td>Peripheral blood EMP levels</td>
<td>Increased</td>
<td>Increased severity of emphysema. Increased exacerbation rate.</td>
<td>(219, 224)</td>
</tr>
<tr>
<td>RhoA / Rho-kinase pulmonary artery levels</td>
<td>Increased</td>
<td></td>
<td>(226)</td>
</tr>
<tr>
<td>ACE levels</td>
<td>D variant associated with worse prognosis.</td>
<td>Increased exacerbation rate.</td>
<td>(227)</td>
</tr>
</tbody>
</table>

This table summarizes different methods of how to assess endothelial dysfunction in patients and how the measurement is altered in COPD. The third column lists any other associated changes seen in COPD with endothelial dysfunction such as exacerbation rate. FMD: Flow-mediated dilatation; vWF: vWillebrand Factor; NO: Nitric oxide; EMP: Endothelial microparticles; ACE: Angiotensin converting enzyme.
1.4.2 The role of the endothelium in lung cancer

Pathologists identified that tumours were vascular in the 1800s suggesting that vessels might be necessary for tumour growth.(235) By the 1970s it was suggested that inhibiting angiogenesis could treat cancer which led to the search for several angiogenic factors in the 1980s.(236, 237) Angiogenesis is defined as the growth of new blood vessels and is required for larger tumours to obtain nutrients and oxygen.(238, 239) This is essential for tumour growth, progression and metastasis.(240) Hypoxia is a key driver and stimulates the production of proangiogenic factors such as HIF-1α.(241) Angiogenesis involves the proliferation, migration and invasion of endothelial cells.(242)

VEGF is involved in one of the major pathways in angiogenesis and has been extensively studied.(242) Activation of VEGF results in endothelial survival, migration and differentiation.(240) VEGF also mobilises endothelial progenitor cells (EPCs) from the bone marrow.(240) There are 6 versions of VEGF and 3 VEGF receptors.(240) Importantly, in NSCLC, VEGF has been shown to be a negative prognostic marker.(243) VEGF is predominantly upregulated in response to HIF-1α expression which is increased in hypoxic cells.(244) However, other growth factor receptors such as EGFR and insulin-like growth factor have also been shown to upregulate VEGF expression. (245, 246) As above, COX-2 plays a role in angiogenesis – this effect is partially mediated through upregulation of VEGF.(247) Finally, mutations in oncogenes and tumour suppressor genes can also increase VEGF expression.(248, 249)
Eventually, in the 2000s drugs were developed to target VEGF in cancer. Bevacizumab, a monoclonal antibody to VEGF, was given to patients with non-operable, non-squamous NSCLC in a phase III study in 2006. Patients with squamous cell lung cancer were excluded due to increased rates of haemorrhage seen in this subgroup during previous phase II trials. Patients were randomised to traditional chemotherapy with and without bevacizumab. Patients in the bevacizumab group had significantly improved progression-free and overall survival. However, benefits of treatment with bevacizumab are limited and eventually patients develop resistance. This may be due to the role of other growth factors in the angiogenesis pathway including fibroblast growth factor (FGF) and platelet derived growth factor (PDGF).

There are 22 ligands in the FGF family, 5 of which have a role in angiogenesis. Endothelial cells express FGF receptors (FGFR-1 and FGFR-2) that upregulate pathways involved in angiogenesis once stimulated. For example, phosphatidylinositol 3-kinase (PI3K) and mitogen-activated protein kinase (MAPK) pathways. This results in endothelial cell proliferation and migration and pericyte recruitment. Mutations in FGFRs are some of the most prevalent mutations in human cancer, which has resulted in the development of FGF targeted therapy. For example, the multi-targeted antiangiogenic drug Cediranib. However, phase III studies have not been promising so far; addition of Cediranib did not improve survival in advanced NSCLC compared to carboplatin/paclitaxel alone. In addition, there was increased toxicity in the Cediranib arm resulting in an early termination of the trial.
PDGF is released from platelets in response to vascular damage. There are 5 PDGF ligands, which bind to PDGFRs on endothelial cells. Activation of these receptors results in angiogenesis. There is also evidence that PDGF activation results in increased cell proliferation and consequently may also directly result in tumourigenesis. In NSCLC increased PDGF expression in tumour tissue is associated with increased lymph node metastases and poor survival. PDGF is also targeted by some of the multi-targeted antiangiogenic therapies in development including Cediranib. Similarly, phase III trials have not been successful. For example Sorafenib and Motesanib have been shown to improve progression free survival in combination with chemotherapy compared to chemotherapy alone. However, these drugs have not improved overall survival and use of Motesanib results in increased adverse effects. Therefore, these agents are not used for lung cancer treatment presently.

1.5 The study of micro-RNA and messenger RNA expression in the pathogenesis of lung disease

1.5.1 The study of micro-RNAs in lung disease

Micro-RNAs (miRNAs) are small (around 20 nucleotides) non-coding ribonucleic acids (RNAs) that act to regulate gene expression. They do this via direct interaction between their 5’ region (seed sequence) and the 3’ region of their target messenger RNA (mRNA). The processing of miRNAs is outlined in figure 1.9.
**Figure 1.9: miRNA processing:** RNA polymerase II transcribes the miRNA gene to form pri-miRNA. Pri-miRNA is cleaved to pre-miRNA by Drosha/DGCR8 and exported to the cytoplasm by Exportin 5/RAN GTP. Dicer cleaves pre-miRNA to a double-stranded miRNA the active strand of which binds to Ago to form the RISC. RISC: miRNA-induced-silencing-complex.
MiRNAs are initially transcribed in the nucleus by RNA polymerase II to form pri-miRNA (100-1000 nucleotides). These are cleaved to form pre-miRNA (50-60 nucleotides) by the Drosha/DGCR8 enzymes in the nucleus. The pre-miRNAs are transported to the cytoplasm by Exportin 5/Ran GTP complex and are cleaved again by the Dicer protein to produce a mature miRNA duplex consisting of an active and passenger strand. The active strand associates with Ago proteins to form a miRNA-induced-silencing-complex (RISC) which can interact with target mRNA. miRNAs can suppress mRNA targets by either suppressing translation or causing mRNA degradation. However, more recently it has been shown that some miRNA/Ago complexes are transported into the nucleus and can bind to DNA promoter regions activating transcription and therefore can be positive regulators of gene expression as well as negative.

There is evidence that miRNAs are important in lung development. For example, Dicer-null mice have abnormal lung epithelium and airway branching. miRNAs may also be important in the regulation of inflammation in the lung and have been shown to alter expression in response to exogenous lipopolysaccharide (LPS) in animal models. Animal studies have also demonstrated altered lung miRNA expression in rats exposed to heavy smoke and those with emphysema in comparison to controls suggesting that miRNAs might influence the development of COPD.

Human studies also suggest a role for miRNAs in COPD pathogenesis. For example, let-7c is significantly suppressed in the sputum of patients with COPD compared to those without. Whole lung miRNA studies have shown that the miRNA
expression appears to differ between patients with and without COPD. (269, 270) For example, Ezzie et al demonstrated key differences in miRNA expression in pathways important to the development of COPD such as transforming growth factor-β and focal adhesion pathways. (269) However, very few studies have looked at the miRNA expression in COPD in individual cell types in the lung; this is important for two main reasons. Firstly, one cannot assume that miRNA changes in whole lung apply to each cell type, which makes it difficult to identify potential cellular pathogenic pathways in COPD. Secondly, miRNA signals from different cells might cancel each other out resulting in important miRNA targets being missed. Examples of miRNA studies focusing on single cell types include Schembri et al who demonstrated that 28 miRNAs were differentially expressed between smokers and non-smokers in bronchial epithelium. (271) This suggests that miRNA expression may become altered prior to the development of COPD. (271) Sato et al also showed that fibroblasts from patients with COPD have downregulated expression of miR-146a, which increased the COX-2 response and may therefore be important in the altered inflammation seen in COPD patients. (272) A third study demonstrated a low miR-218-5p level in lung tissue of COPD patients which was validated in Human Bronchial Epithelial Cells (HBECs) and related to airway obstruction. (273) This suggests that miR-218-5p may have a role in COPD pathogenesis. (273) However, as yet there are no reported studies on miRNA expression in pulmonary endothelial cells in COPD.

miRNA changes in COPD patients may precede the development of lung cancer. One study demonstrated significant differences in miRNA expression (in blood) in COPD and lung cancer patients compared to controls. (274) The miRNA signatures of COPD and lung cancer patients were more similar to each other than to controls.
suggesting that some miRNA changes in lung cancer are already present in COPD patients before cancer develops.\(^{(274)}\) Similarly, another study demonstrated that bronchoalveolar lavage (BAL) samples from people with COPD shared 5 deregulated miRNAs with samples from patients with adenocarcinoma and COPD.\(^{(275)}\) (These 5 miRNAs were not deregulated in patients with adenocarcinoma only suggesting they might be specific to adenocarcinoma development in COPD rather than normal lung.) It is possible that these deregulated miRNAs may have had a role in the development of cancer in the COPD patients because they represent epigenetic changes prior to the development of malignancy that persist in the setting of cancer. For example, miR-15b which was identified as upregulated in COPD and adenocarcinoma has previously been shown to be upregulated in NSCLC and is associated with a poor prognosis.\(^{(276)}\) Similarly, miR-365 which was identified as downregulated in COPD and cancer has been shown to target epithelial-mesenchymal transition.\(^{(277)}\)

In conclusion there is evidence that miRNA expression is altered in COPD patients and shares some similarity with miRNA changes in lung cancer. However, the evidence for this in specific cell types is lacking. Importantly there is no published data on the miRNA expression in pulmonary endothelial cells.

Despite a lack of studies looking at miRNA expression in pulmonary endothelial cells there are some studies suggesting miRNAs are involved in angiogenesis. This initially stemmed from murine model studies in which investigators produced Dicer knockout mice.\(^{(278)}\) Mouse embryos without Dicer demonstrated stunted growth and developmental defects.\(^{(278)}\) All Dicer knockout mice retarded phenotype died within 15 days of gestation suggesting this phenotype was not viable.\(^{(278)}\) Blood vessel formation in the embryos and yolk sacs was also severely compromised suggesting
Dicer (and thus miRNAs) could be key in angiogenesis. This was supported by \textit{in vitro} studies that used knockdowns of Dicer in Human umbilical vein endothelial cells (HUVECs). Knockdown of Dicer resulted in the alteration of key endothelial related genes such as VEGFR2 (Vascular Endothelial Growth Factor Receptor 2). Knockdown of Dicer also resulted in reduction of sprouting, tube forming ability and migration of endothelial cells suggesting a reduction in angiogenesis capability. Finally, mouse models with cell-specific inactivation of endothelial Dicer demonstrated reduced postnatal angiogenesis to several responses such as VEGF, tumours and wound healing.

There is also evidence that specific miRNAs directly affect angiogenesis. For example, Poliseno et al demonstrated that miR-221 and -222 were upregulated in HUVECs and appeared to reduce c-Kit protein levels which resulted in reduction of wound healing and tube formation in HUVECs. Interestingly, some miRNAs have been directly linked to angiogenesis in hypoxia. This may be particularly relevant in the setting of lung cancer and COPD, especially as inflamed tissues can become hypoxic. For example, miR-20a is downregulated by hypoxia and directly targets HIF-1α. Thus, reduction in miR-20a would increase HIF-1α and its target VEGF resulting in increased angiogenesis. MiR-424 has also been shown to be increased in response to hypoxia. This targets cullin-2 (CUL2), a scaffolding protein which is important in the assembly of the ubiquitin ligase system. This consequently stabilises HIF-1α and therefore promotes angiogenesis. This could perhaps be relevant to the increase in angiogenesis seen in cancer, but less so in the reduction of the vascular bed apparent in emphysema. However, other miRNAs induced in hypoxia might reduce HIF-1α.
activity. For example, miR-20b is induced in hypoxia and targets HIF-1α consequently leading to a reduction of VEGF.(286) Therefore, perhaps endothelial function depends on the balance of expression of different miRNAs and their cumulative effect on HIF-1α in hypoxia.

Mizuno et al also demonstrated that 2 other miRNAs (miR-199a-5p and -34a) might be involved in pulmonary endothelial function in COPD.(287) Both miRNAs were increased in lung tissue from COPD patients and negatively correlated with FEV1 percentage predicted.(287) Transfection of these miRNAs into human pulmonary endothelial cells (HPECs) resulted in reduced HIF-1α levels.(287) This suggests that increased levels of these miRNAs might reduce HIF-1α in the setting of COPD resulting in impairment of lung vessel homeostasis and loss of blood vessels. Consequently, these miRNAs could be important in the development of emphysema.

There is greater evidence that miRNAs are important in angiogenesis in lung cancer. For example, miR-494 enhances wound closure, migration and tube formation in HUVECS in vitro.(288) This is relevant in lung cancer as A549 (lung cancer cell line cells) can secrete and deliver miR-494 to HUVECs in vitro suggesting that a similar process could occur within the lung cancer microenvironment.(288) Similarly, miR-378 is upregulated in patients with NSCLC, which has metastasized.(289) In addition to contributing to tumour cell migration and invasion, miR-378 increases angiogenesis in implanted tumours in mouse models thereby suggesting that it may play a role in lung cancer metastasis formation.(289) Other studies have demonstrated certain miRNAs are suppressed in lung cancer, which has an effect on angiogenesis through lack of suppression of their targets. For example, miR-128 is reduced in the NSCLC tissues and appears to target VEGF-C.(290) Overexpression
of miR-128 leads to reduction of endothelial tube formation *in vitro* consequently suggesting that upregulation of VEGF and increased angiogenesis might occur in lung cancer through suppression of miR-128.(290) Another example of a similar mechanism is the downregulation of miR-206 in NSCLC.(291) This miRNA targets the protein 14-3-3ζ which, in turn, interacts with phosphorylated-signal transducers and activators of transcription 3 (p-STAT3) and activates HIF-1α.(291) Overexpression of miR-206 therefore inhibited HIF-1α activation and reduced tumour growth and angiogenesis in mouse models.(291) Downregulation of miR-206 could have the opposite effect and upregulate angiogenesis in NSCLC. This is consistent with data from NSCLC patients where low miR-206 expression was associated with poor survival.(291)

In summary, there is increasing evidence that miRNAs are involved in the regulation of angiogenesis particularly through the HIF-1α/VEGF pathway (figure 1.10). Although there are a few examples of miRNAs contributing to angiogenesis in lung cancer evidence of miRNAs influencing endothelial function in COPD is limited. This combined with the lack of effective treatments for COPD; means that the identification of new treatment targets such as miRNAs is warranted.
Figure 1.10: MiRNAs in the setting of hypoxia, NSCLC and COPD and their effect on angiogenesis through the HIF-1α/VEGF pathway. The central pathway represents the main HIF-1α/VEGF pathway that upregulates angiogenesis. The purple boxes represent processes or diseases where altered miRNA expression has been shown to affect the HIF-1α/VEGF pathway. The left side of the figure demonstrates miRNAs (in pale blue) that upregulate angiogenesis. Genes are in pale orange. The right side of the figure demonstrates miRNAs that downregulate angiogenesis. Blue upward arrows indicate that a miRNA/gene is upregulated. Red downward arrows indicate that a miRNA/gene is downregulated. CUL2 = cullin-2, p-STAT3 = phosphorylated-signal transducers and activators of transcription 3, HIF-1α= hypoxia inducible factor 1α, VEGF = vascular endothelial growth factor.
1.5.2 The study of messenger RNAs in lung disease

Studies investigating gene expression in lung tissue from COPD patients have been in place since the early 2000s. One study comparing gene expression in whole lung tissue from patients with and without emphysema demonstrated that genes involved in oxidative stress and inflammation were increased in emphysema patients. More surprisingly, genes involved in extracellular matrix (ECM) production were also increased. This suggests that there is an attempt to remodel connective tissue in patients with emphysema, but this is ultimately unsuccessful. In the same study angiogenesis related genes (such as endothelial cell growth factor 1) were downregulated and SERPINF1 (serpin family F member 1, a known inducer of endothelial cell apoptosis) was upregulated providing support for endothelial apoptosis in the development of emphysema. Similarly, other studies into whole lung gene expression have also demonstrated an increase in apoptosis-related genes in patients with emphysema.

A recent large study (n=581) into gene expression in COPD lung tissue provides further evidence that ECM genes are upregulated in COPD. Expression of these genes negatively correlated with lung function suggesting that expression of these genes increases as tissue destruction in COPD worsens. Campbell et al have also provided support for faulty tissue repair in COPD. They performed a study comparing gene expression in different areas of the same COPD lung. Genes involved in the transforming growth factor-β (TGF-β) pathway, actin organisation and integrin signalling were reduced in areas of worse emphysematous destruction.
It is likely that changes in gene expression are important in the development of COPD. For example, in a study comparing the expression of genes involved in inflammation, tissue remodelling and vessel maintenance between COPD patients, smokers and non-smoking controls there were significant differences between the groups, as well as large overlap between COPD patients and smokers, suggesting that gene expression changes in smokers precede the development of COPD. (296)

Early studies were generally based on expression of genes in whole lung tissue. More recently, gene expression studies have been performed on different cell types in COPD. For example, COPD gene signatures have been isolated from epithelial cells using brushing techniques at bronchoscopy. (297) These gene expression changes were also similar to changes seen in small airway epithelium in COPD patients suggesting that a ‘field of injury’ might be present in such patients throughout the bronchial epithelium. (297) Another study which looked at small airway epithelium has identified the Notch pathway as being of importance in COPD. (298) The Notch pathway is important in the control of epithelial differentiation in lung development. (298) Several Notch ligands and receptors are downregulated in COPD. (298) As the Notch pathway is important in lung development, perhaps the downregulation of this pathway in COPD is consistent with the faulty remodelling seen in COPD patients. Pierrou et al also looked at gene expression in the bronchial epithelium. (299) However, they identified genes relating to oxidative stress as being the most differentially expressed between patients with and without COPD. (299) Many of these changes were also seen in the epithelium from patients who were smokers (without COPD) suggesting that they may be important in the development of COPD. (299)
Another cell type which has been investigated in COPD is the fibroblast. Fibroblasts isolated from lung resections have been shown to express higher levels of the inflammatory marker serum amyloid A 2 (SAA2) in comparison to controls.(300) Fibroblasts isolated from COPD patients also respond to cigarette smoke extract (CSE) in vitro in a different manner to control fibroblasts in terms of gene expression in the TGF-β-Smad pathway.(301) As this pathway is important in ECM repair this change in response may be important in destruction to the ECM seen in emphysema. Perhaps this response in the TGF-β-Smad pathway represents a set of patients who are more susceptible to damage from cigarette smoke than controls.

Whole lung gene expression studies comparing patients with lung cancer to controls have also shown alteration in gene expression.(302) In addition differential gene expression studies have been performed on individual cell types in cancer-affected lung. For example, gene expression in epithelial cells from lung cancer patients shows alterations in chemokine signalling, cytokine receptors and cell adhesion pathways.(303) Pulmonary endothelial cells also show a difference in gene expression between tumour and normal tissue, such as matrix metalloproteinases.(304)

Studies have also looked at lung gene expression in lung cancer patients with and without COPD. For example, Boelens et al demonstrated that tumour tissue from patients with COPD was more likely to show a loss of 5q or a low expression of 5q genes than tumour tissue from patients without COPD.(169) However, expression of chemokine ligand 1 (CXCL14) (305) and cathepsin inhibitor Cystatin A (CSTA) (306) are upregulated in both COPD and lung cancer. Mutations in genes associated with lung cancer also seem to vary between patients with and without COPD. For
example, the *KRAS* (KRAS proto-oncogene, GTPase) (307), *EGFR* (Epidermal growth factor receptor) and *ALK* (ALK receptor tyrosine kinase) (308) mutations occur infrequently in COPD patients. Epigenetic control of gene expression may also be different between lung cancer patients with and without COPD. For example, Tessema et al demonstrated that methylation of microtubule-associated protein 1B (*MAP1B*), a cytoskeletal protein involved in cell motility, is more prevalent in lung tissue and lung tumours from patients with COPD. As the gene suppression is seen in COPD lung tissue in addition to the tumour tissue it is possible that this alteration in gene expression occurred prior to the development of lung cancer. In conclusion, there is evidence that gene expression in COPD lung is altered. Some changes in gene expression are also seen in lung cancer providing further evidence for shared pathogenesis. In a similar way to miRNA data, there have been a few studies investigating the gene expression in individual cell types in the lung including the endothelium in lung cancer. However, the expression of genes in the COPD endothelium has not been investigated yet.

1.5.3 *Pathway analysis of genomic data in the setting of lung cancer and COPD*

In recent years pathway analysis has become an important method for analysing genomic data. Microarray data produces huge lists of differentially expressed RNAs and it can be difficult to gain insight about underlying biology from the lists alone. Pathway analysis attempts to group genomic data into smaller sets of related RNAs. Therefore, rather than thousands of individual data points, the researcher can evaluate several hundred groups of data which simplifies the investigation. Also, by grouping data into pathways rather than individual RNAs it allows the researcher to gain an increased understanding about how differentially
expressed RNAs might be important in disease processes. (311) For example, Meng et al used pathway analysis to explore changes in long noncoding RNAs in sinonasal squamous cell carcinoma and identified multiple alterations in tumourigenic processes. (312)

Ingenuity pathway analysis (IPA; Qiagen) is a tool available as a web-based software application for this purpose. IPA studies data according to information available in the Ingenuity Knowledge Base, a manually curated database with data on genetic pathways, disease processes and pharmacology. (313) By using inbuilt algorithms IPA can provide the user with evidence on important biological mechanisms, pathways, regulators and functions relevant to the changes observed in microarray data. One can also predict the downstream effects of these pathways on other biological processes and diseases such as cancer.

IPA has been used multiple times in current literature in the analysis of respiratory disease including lung cancer and COPD. For example, He et al were interested in the development of lung cancer related to carbon nanotubes (CNTs). (314) They analysed gene expression in pulmonary epithelial cells exposed to CNTs and investigated the role of mesothelin (MSLN), a protein involved in carcinogenesis in lung cancer and mesothelioma. (314) By using IPA they were able to identify that MSLN was likely to act by inducing cyclin E (a protein involved in the cell cycle) expression, which increased cell proliferation. (314) They were able to support this hypothesis with further in vitro work. (314) Therefore, not only was a lung cancer target identified (MSLN), but the biological processes relevant to disease were also revealed allowing the group to gain more information on the precise function of the protein of interest. IPA has also been used to assess the effects of external stimuli on
cellular processes. For example, one group performed microarrays on lung epithelial cells exposed to cigarette smoke to determine the effects of smoke on cellular function.(315) Microarrays identified differentially expressed genes after exposure to smoke, but by utilising IPA the group were able to identify deregulation of important pathways such as the transforming growth factor beta-1 (TGF-β1) pathway.(315) Similarly, IPA has also been used to assess the cellular response to treatment for disease such as hesperidin (a glycoside flavonoid with anti-cancer effects).(316) When NSCLC cells were exposed to hesperidin there were significant changes in several important functions which could explain the anti-cancer effects such as changes in cell-cell interaction and signalling, cell-mediated immune response and the inflammatory response.(316)

IPA has also been used in the analysis of miRNAs in lung cancer. For example, Wang et al identified differentially expressed miRNAs in the blood of lung cancer patients in comparison to controls.(317) In this study expression of mir-675 was particularly dysregulated and Wang et al explored mir-675’s function further in IPA.(317) IPA demonstrated that mir-675 was important in cell death and survival pathways which could help explain its’ importance in the development of lung cancer.(317) Similarly, Ma et al looked at differences in miRNA expression in blood of patients with NSCLC and different epidermal growth factor receptor (EGFR) mutations.(318) The exon 19 deletion mutation (del19) has a favourable prognosis in comparison to single-point substitution mutation L858R in exon 21 (L858R).(318) The difference in miRNA expression seen between patients with the two different mutations was associated with alteration in several key pathways such as cell
survival, movement and proliferation which are key in the pathogenesis of lung cancer. (318)
In a similar fashion to lung cancer, COPD has been investigated using IPA. For example, an analysis of mRNA microarrays in peripheral blood mononuclear cells (PBMCs) has been conducted; IPA demonstrated that important immune system and inflammatory response pathways were altered in COPD patients. (319) There is also evidence that genetic pathways may change as COPD progresses. IPA of whole blood gene expression in patients admitted to intensive care in comparison to those admitted to a medical ward suggested that many upregulated genes were involved in neutrophil function. (320) Interestingly, the upregulated neutrophil pathways were associated with an increased risk of need for advanced respiratory support and bacterial infection. (320) This raises the possibility that monitoring of these pathways could help to identify patients with COPD with a worse prognosis. Other groups have also used IPA as a new way to interpret old data in COPD. For example, Kaneko et al performed a literature search to identify genes of importance in COPD and uploaded them into IPA. (321) This resulted in the identification of several pathways increased in COPD including those related to the immune response. (321)
Finally, IPA has also been used previously in the identification of pathways common to lung cancer and COPD. One study looking at protein expression in bronchoalveolar lavage (BAL) in patients with COPD, lung cancer or both diseases demonstrated that the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway is altered in both conditions. (322) NF-κB is a transcription factor that is important in inflammation and aberrant expression has previously been associated with carcinogenesis. (323, 324)
1.6 Isolation of Human Pulmonary Endothelial Cells (HPECs) for culture

The isolation of human endothelial cells for culture has been in place for the past 3-4 decades. (325, 326) Initially cells were primarily isolated by perfusing arteries and veins with collagenase and collecting flushed cells after digestion. (325, 326) However, the isolation of microvascular endothelial cells is more challenging and requires methods to select for these cells in order to avoid the contamination of cell cultures with other cells such as fibroblasts. (327-330) It has also been shown that endothelial cells from different organs in the body are inherently different from one another. (331, 332) This makes it difficult to use more easily available endothelial cell types like Human Umbilical Vein Endothelial Cells (HUVECs) when investigating markers identified in other endothelial cell isolates. (331) This prompted the development of techniques to isolate HPECs in the 1990s. (329, 330) Two publications on endothelial extraction were made around the same time: Carley et al and Hewett and Murray. (329, 330) Carley et al obtained lung tissue in a similar way to this study: by collecting tumour-free lung tissue removed at surgery for lung cancer. (329) The pleura was removed by scalpel (to reduce mesothelial cell contamination) and the tissue was minced and incubated with 0.1% collagenase for 1 hour. (329) The digested material was filtered, resuspended in media and plated on 1.5% gelatin coated plates. (329) After cells had been allowed to grow, the cells were removed from the plates and endothelial cells were selected for using fluorescent-activated cell sorting (FACS) by identifying endothelial cells with 1,1'-dioctadecyl-1,3,3',3'-tetramethyl-indocarbocyanine perchlorate-labeled acetylated low-density lipoprotein (Dil-Ac-LDL). (329) Cells isolated from the FACS sorting were then plated and allowed to grow to confluence. (329) These cells stained weakly for von
Willebrand Factor (vWF), immunostained for angiotensin converting enzyme (ACE; known to be expressed by lung endothelium) and bound *Ulex europaeus* agglutinin-1 (UEA-1, which binds to L-fucose residues on endothelial cells).(329) The cells also appeared to grow into tube like structures in Matrigel (a protein mixture secreted by Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells).(329)

Hewett and Murray obtained tissue from an alternate source: peripheral lung tissue from transplant surgery.(330) The initial part of their method was very similar to Carley et al. The pleura was removed and the lung was minced before digestion.(330) However, the lung was digested at room temperature for 16-20 hours in dispase prior to 15 minutes at 37°C in trypsin/ethylenediamine tetra-acetic acid (EDTA).(330) Again, the digested material was filtered and the cell solution was plated but on fibronectin plates.(330) After the cells reached confluence the cells were removed using trypsin and mixed with UEA-1-coated magnetic beads for 10 minutes at 4°C.(330) The beads were then resuspended in growth medium and plated on to 0.2% gelatin coated plates.(330) These cells formed characteristic ‘cobblestone’ morphology similar to HUVECs in culture and expressed typical endothelial antigens such as E-selectin.(330) The cells also formed tube like structures in matrigel.(330)

Due to the complex method in preparing HPECs many researchers have resorted to buying the cells commercially.(333-335) However, in recent years a few laboratories have continued to extract HPECs on site.(336-338) The methods used by these researchers are primarily based on the methods used by Hewett and Murray.(330) All of these techniques consist of mincing and digesting lung tissue prior to filtering and purifying with magnetic beads attached to different markers for endothelial cells such
as UEA-1, and CD31 (platelet and endothelial cell adhesion molecule 1).(336-338) Interestingly, these cells also have a cobblestone morphology and appear to resemble the cells from the Hewett and Murray paper more strongly than the paper from Carley et al.(329, 330, 336-338) One team also confirmed that this bead method could be used to extract HPECs from COPD tissue.(336) This, in addition to previous work performed in the Bicknell laboratory in the University of Birmingham for the isolation of HPECs for RNA work, resulted in the decision to base an extraction method using UEA-1 coated magnetic beads.(304)

1.7 Hypotheses of this thesis

This study was performed to investigate the possibility of shared pathogenesis between COPD and lung cancer in the context of the pulmonary endothelium. Epidemiological studies demonstrating an increased risk of lung cancer in patients with COPD (section 1.3) suggest common disease processes. Neutrophil transendothelial migration (TEM) is upregulated in COPD (section 1.4.1) and results in increased inflammatory cells in the lung. As inflammation is associated with cellular processes related to cancer such as cell proliferation (section 1.2.3) the upregulation of TEM in COPD could provide a mechanistic link between the two conditions.

Upregulation of TEM could be related to the endothelial dysfunction seen in COPD (section 1.4.1), but the pathogenesis of endothelial dysfunction remains unclear. Previous studies have demonstrated that miRNA and mRNA expression is abnormal in pulmonary COPD and lung cancer tissue, but endothelial specific studies are lacking. Therefore, this study was performed to identify alterations in microRNA and mRNA expression in COPD and lung cancer that could help to explain the
endothelial dysfunction seen in these conditions and that could potentially identify new shared treatment targets.

In order to investigate this hypothesis microRNA and mRNA expression was analysed using microarray. This method was chosen as previous studies had used this successfully in the context of whole lung genomic studies and in studies involving other pulmonary cell types. In particular this approach had been used with success investigating mRNA expression in pulmonary endothelial cells in lung cancer in the same laboratory previously.(304) An alternative method for performing this would be to use next generation sequencing (NGS) which has also been used to assess miRNA and mRNA expression in lung cancer.(339, 340) The main advantage of NGS is that it covers all areas of the transcriptome whereas microarray analysis uses probes, which provide limited coverage.(341) However, at the time of commencing this study there was limited access to NGS, which was also prohibitively expensive, and therefore the decision was made to continue with microarray analysis.

1.8 Aims

This project was undertaken to investigate new gene and miRNA targets in COPD and NSCLC lung tissue.

1) Characterise the miRNA and gene expression profile of COPD and NSCLC tissue.

There is evidence that gene and miRNA expression varies in lung tissue between normal lung and COPD and lung cancer.(269, 302) However, there is little data published on miRNA and gene expression in individual cell types. This is important as whole lung tissue gene/miRNA expression studies can
miss signals coming from individual cell types and it can be difficult to attribute results to specific cells. To address this further fresh tissue was isolated from non-COPD, COPD and NSCLC lung, which was used to extract macrophages and endothelial cells. Macrophages are inflammatory cells that phagocytose and kill micro-organisms. There are increased numbers present in COPD and they behave in an abnormal manner. (342) Endothelial cells may be of importance in COPD and lung cancer as inflammatory cells (such as neutrophils) must pass through endothelial tissue to reach the lung itself. (188) Due to time constraints endothelial cells only were chosen for further work. Endothelial cells were chosen primarily due to the lack of previous work in this area. The macrophage data was stored for further work.

2) Describe biological pathways of relevance to COPD pathogenesis within endothelial tissue, particularly those aspects shared with lung cancer

Gene and miRNA expression data produces huge datasets which can make the biological interpretation of such data difficult. (310) Therefore, to further interpret expression data a pathway analysis was performed using Ingenuity Pathway Analysis (IPA). This uses information on known molecular pathways to interpret expression data and to provide disease associations which allows the user to gain a deeper insight into their results.

3) Isolate and culture HPECS

Endothelial cells can vary between tissue types. (331) Consequently, it is ideal to investigate new endothelial targets using endothelial cells from the organ of interest. Thus, one aim of this thesis was to isolate and culture human
pulmonary endothelial cells (HPECs) for use for functional experiments to investigate new targets identified by the first aim.

4) Describe the function of miR-181b-3p in human endothelial tissue

The endothelial target that was most significantly differentially expressed between groups was taken forward for functional validation work. Several assays (such as matrigel and spheroid assays) were used to compare endothelial function in cells with and without an increase in expression of this target (miR-181b-3p) to gain an understanding of its' possible effects on angiogenesis.
CHAPTER TWO

METHODS
2.1 Patient data collection

2.1.1 Patient selection

45 patients undergoing thoracic surgery between December 2013 and November 2014 were included in the study. These patients were recruited from an existing tissue retrieval system: Midlands Lung Tissue Consortium (MLTC); (07/MRE08/42). Patients about to undergo thoracic surgery at Birmingham Heartlands Hospital (BHH) were consented by research staff during their pre-operative clinic appointments to donate excess lung tissue to research. Informed, written consent was obtained for each patient. An example of the consent form used is in Appendix 1. The majority of these patients were undergoing lobectomy or pneumonectomy for suspected lung cancer. However, patients were also recruited who were undergoing lung volume reduction surgery.

An additional 7 patients previously recruited in the same manner (by the Turner/Bicknell groups between July 2011 and April 2012) were also included in the analysis.

2.1.2 Clinical data

Clinical data was collected by the thoracic surgical teams at BHH as well as the thoracic research nursing team. Extraction of clinical data from BHH records was performed solely by me.

Baseline data including age, sex, body mass index (BMI), pack year history of smoking (calculated by multiplying number of cigarettes per day by number of years smoked and dividing by 20) and presence of chronic bronchitis was obtained from
the MLTC research database and computer databases at BHH. Copies of clinic and referral letters were available in BHH medical records. Pathological reports were reviewed to obtain tumour stage and type.

CT scans were performed either at BHH or referring hospitals. The images from CT scans not performed at BHH were transferred electronically to the BHH computer system. CT scan reports (where available) and images were examined for the presence of emphysema.

Lung function was performed either at BHH or referring hospitals. FEV1 was recorded for all patients in the research notes. Where available FVC, FEV1/FVC, total lung capacity (TLC), residual volume (RV), transfer factor for carbon monoxide (TLCO) and corrected gas transfer (KCO) were recorded using research notes and by contacting referring hospitals. FEV1 and FEV1/FVC values were compared to predicted values using standard reference equations for Caucasian adults (Appendix 2). This was used to calculate FEV1pp (percentage predicted) for all patients.

Patients were defined as having a diagnosis of COPD if they had a FEV1/FVC less than the lower limit of normal and/or presence of emphysema on the CT scan with a history of smoking.

A summary of information collected for each patient can be seen in table 2.1.
Table 2.1: Summary of information collected for each patient.

<table>
<thead>
<tr>
<th>Data collected</th>
<th>Type</th>
<th>Information used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Clinical information</td>
<td>Research database</td>
</tr>
<tr>
<td>Sex</td>
<td>Clinical information</td>
<td>Research database</td>
</tr>
<tr>
<td>Body mass index (BMI)</td>
<td>Clinical information</td>
<td>Research database</td>
</tr>
<tr>
<td>Pack year history of smoking</td>
<td>Clinical information</td>
<td>Research database/clinic letters</td>
</tr>
<tr>
<td>Chronic bronchitis</td>
<td>Clinical information</td>
<td>Clinic letters</td>
</tr>
<tr>
<td>Tumour stage</td>
<td>Pathology</td>
<td>Pathology reports</td>
</tr>
<tr>
<td>Tumour type</td>
<td>Pathology</td>
<td>Pathology reports</td>
</tr>
<tr>
<td>Presence of emphysema</td>
<td>Radiological</td>
<td>CT reports/images</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>Lung function</td>
<td>Computer databases and reference equations.</td>
</tr>
<tr>
<td>FEV1/FVC</td>
<td>Lung function</td>
<td>Computer databases and reference equations.</td>
</tr>
</tbody>
</table>

Clinical information for each patient was collected using multiple sources including the research database at BHH, computer records, clinic letters and pathology reports. Collected information for each patient was stored anonymously in a spread sheet format. FEV1pp = Forced expiratory volume in one second percentage predicted. FEV1/FVC = Forced expiratory volume in one second/Forced vital capacity.
2.2 Statistical analysis of clinical data

All statistical analysis was performed using IBM SPSS statistics version 22 (SPSS Inc, Chicago, USA). Patients with and without COPD were compared according to: age, sex, body mass index (BMI), smoking status, pack year history, FEV1pp and tumour stage. The Shapiro-Wilk test was used to determine the normality of each variable. Normally distributed variables are displayed as mean and standard error (SE). Non-parametric variables are displayed as median and interquartile range (IQR). Variables that were normally distributed were compared using t-tests. Variables not normally distributed were compared using the Mann-Whitney-U test. Categorical variables were compared using Chi-squared or Fisher’s exact tests depending on the number of patients in each group. Variables were determined as being statistically significantly different from one another using the cut off of \( p<0.05 \). Where experiments included different numbers of patients from the overall cohort, the above statistical analysis was repeated for each individual experiment.

2.3 Collection of lung tissue

Lung tissue was excised by surgical staff at BHH and kept in 0.9% saline. The rest of the lung tissue was transferred to pathology at BHH where it was examined by a pathologist. Where enough tissue was available tumour tissue was also taken and kept in 0.9% saline. Tissue samples were kept on ice and transferred to the University of Birmingham laboratories for further processing on the same day.

2.3.1 Extraction of macrophages

1 l of 0.9% saline was attached to a giving set and placed into a pressure bag. The giving set was connected to the needle from the inside of a 16 G cannula (after
removing the safety device with forceps). The lung sample was washed with saline by holding the sample with forceps and placing the needle into the lung tissue. The wash was collected into a bottle using a funnel. The lung was washed until the wash became clear. The wash was known as a bronchoalveolar lavage (BAL). The BAL was placed into 50 ml falcon tubes (as many as required to hold all of the BAL) and centrifuged at 40 g for 5 minutes to pellet the cells. The supernatants were removed and the cell pellets were collected to create a cell solution. 20 ml of cell solution was required. If necessary supernatant was used to make the cell solution up to 20 ml. If more than 20 ml of cell solution was collected this was placed into a second falcon tube. 12 ml of lymphoprep (Axis-Shield; 1114545) was added to a new falcon tube and 20 ml cell solution was overlaid onto the lymphoprep slowly using a pipette. This mixture was centrifuged at 100 g, acceleration 1, brake 0, temperature 20°C for 30 minutes. The interlayer was aspirated using a Pasteur pipette and placed into a new falcon tube. The volume of cell solution was made up to 50 ml with Dulbecco's Phosphate-buffered saline (DPBS); (Life Technologies (Gibco) 14190-094). This was centrifuged at 50 g for 10 minutes. The supernatant was discarded and the cells were resuspended in 30 ml DPBS. This was then centrifuged at 50 g for 5 minutes. The supernatant was discarded and the cells were resuspended in 10 ml DPBS. The cell solution was centrifuged at 50 g for 5 minutes and the supernatant was discarded. The cells were resuspended in 700 µl Trizol (Sigma TRI Reagent (Trizol); T9424) in an Eppendorf tube and stored at -80°C. This method has previously been validated by the Respiratory group at the University of Birmingham.(343)
Figure 2.1: Extraction of endothelial cells for future RNA work. A fresh lung sample is minced before being added to collagenase V for digestion. The digested sample is passed through a filter before being combined with ulex coated magnetic beads. The beads bind to endothelial cells before a magnet is used to isolate the beads from the rest of the solution. The beads are added to Trizol and stored at -80°C.
2.3.2 Extraction of endothelium

This method is summarized in figure 2.1. The lung tissue was divided into samples of approximately 3 cm in length. Larger pieces of lung are not digested as easily and produce a lower yield of endothelium. 25 mg collagenase (Collagenase type V; (C9263 Sigma-Aldrich, UK): 25 mg/ml in DMEM (41966-029 Life Technologies (Gibco)), 1 ml aliquots at -20°C) and 75 µl actinomycin (Actinomycin D; (A4262 Sigma-Aldrich, UK): 1 mg/ml in DMSO (Dimethyl sulfoxide; Sigma Aldrich; D2650); 75 µl aliquots at -20°C) were defrosted in a dry bath. 12.5 ml DMEM was added to a falcon tube. 25 mg of collagenase was added to the DMEM and vortexed to mix in fully. Finally, 75 µl actinomycin and 20 µl DNAse solution (RNase-Free DNAse Set; (79254 Qiagen)) were added to create ‘digestion solution’. One piece of lung was then added to 7.5 ml digestion solution, cut into fine pieces on a 10 mm plate and transferred to a falcon tube. 5 ml of digestion solution was used to wash the plate to obtain any remnants of tissue and this was transferred to the same falcon tube. The lung was then digested by placing the solution into a 37°C shaker for 1 hour. This was repeated for each 3 cm of lung tissue.

The following steps were repeated for each piece of lung. Streptavidin Dynabeads (Life Technologies (Invitrogen) M-280) were vortexed to resuspend them. For each sample 50 µl of beads was used for the first 1 g of tissue and 25 µl for each gram thereafter. Prior to use it is necessary to wash the beads with DPBS which is done using a magnet. The beads were placed into a falcon tube and the bead solution was made up to 500 µl with DPBS (Life Technologies (Gibco) 14190-094). The falcon tube containing the beads was placed into a magnet and inverted. The solution in the tube was removed before removing the falcon tube from the magnet. 500 µl of DPBS
was added to the beads and they were placed into the magnet again before the DPBS was removed as previously. The beads were resuspended in 250 µl of DPBS. Distilled water was added to Lectin from Ulex Europaeus (gorse, furze) (L8262 Sigma) to form a solution of 1 mg/ml. 12.5 µl of ulex solution was added to the beads and the tube was shaken to mix the beads with ulex. The bead solution was then placed into a 37°C shaker for 30 minutes in order to bind the ulex to the beads.

Following digestion the lung solution was filtered through a 100 µm filter (VWR International; 734-0004) before being passed through a similar 70 µm filter (VWR International; 734-0003). The lung solution was then centrifuged at 200 g for 5 minutes to pellet the cells. The supernatant was removed without disturbing the cell pellet and placed into a second falcon tube in order to prevent loss of cells. The cell pellet was resuspended in 5 ml DPBS and both the cell solution and supernatant were centrifuged again at 200 g for 5 minutes prior to removal of supernatants from both tubes, which were then discarded. The two cell pellets were combined in 250 µl of DPBS in an Eppendorf tube. The bead solution was then added and the resultant mixture blended on a wheel in a cold room for 30 minutes.

After mixing, the cell solution was transferred into a falcon tube labelled as ‘endothelial cells’, which was placed into a magnet and inverted. The solution was removed whilst the tube remained in the magnet and placed into another falcon tube labelled ‘lung bulk’. The ‘endothelial cells’ tube was then removed from the magnet and beads resuspended in 500 µl of DPBS. The ‘endothelial cells’ tube was placed into the magnet again and the process was repeated until the solution removed was clear. The ‘lung bulk’ tube was centrifuged at 200 g for 5 minutes to pellet the cells, prior to removal of supernatant and resuspension of the undisturbed cell pellets in
Figure 2.2: RNA extraction using the Qiagen miRNeasy kit. Chloroform is added to cell samples stored in Trizol and shaken. This is then centrifuged to separate the sample into 3 phases. The RNA-rich (aqueous phase) is extracted and combined with ethanol to bind the RNA. The sample is washed with a series of buffers before RNA is finally eluted into RNAse-free water and stored at -80°C.
700 µl Trizol (Sigma TRI Reagent (Trizol); T9424). The beads were also resuspended in 700 µl Trizol and stored separately at -80°C.

The entire method was repeated with tumour tissue to obtain tumour endothelium where available.

2.4 Preparation of samples for microarray analysis

2.4.1 RNA Extraction

This method is summarized in figure 2.2. RNA extraction was done using commercially available kits: Qiagen miRNeasy Mini Kit (217004) and the Qiagen RNase-Free DNase Set (79254). The kits include all the reagents and sample tubes required for RNA extraction. Prior to each extraction the bench was cleared and wiped with RNAseZap (ThermoFisher Scientific; AM9780) to limit RNAase exposure. All equipment was also wiped with RNAseZap. Only RNAse and DNAse free filter tips were used for the extraction.

Cell samples stored in Trizol were defrosted and vortexed for 1 minute. Samples were then placed on to the bench top at room temperature for 5 minutes. During this time the centrifuge was pre-cooled to 4°C. 140 µl chloroform (Sigma Chloroform minimum 99%; C2432) was added to each sample which was then shaken for 15 seconds. Each sample was then placed on the bench top for 2 minutes prior to centrifugation for 15 minutes at 12000 g and 4°C. This separates the sample into 3 phases: an upper, colourless phase containing RNA, a white interphase and a lower, red phase containing other organic material. For each sample the upper aqueous phase was transferred into 100 µl aliquots into a collection tube. 1.5 volumes of
100% ethanol (Sigma; E7023) was added to the aqueous phase before being directly transferred on to the RNeasy Mini spin column in one 2 ml collection tube. This was centrifuged at 8000 g for 15 seconds at room temperature. Flow-through was discarded, and 350 µl Buffer RWT was added to the spin column followed by centrifugation at 8000 g for 5 minutes. The flow-through was again discarded. 10 µl DNase I stock solution was added to 70 µl Buffer RDD (from the Qiagen RNase-Free DNase Set) and was mixed by gently inverting. The DNase I incubation mix was pipetted directly on to the spin column and placed on to the bench top for 15 minutes. It is important to pipette the incubation mix directly on to the spin column membrane to avoid the mix sticking to the column wall resulting in incomplete DNase digestion. 350 µl Buffer RWT was then pipetted to the spin column and centrifuged for 15 seconds at 8000 g and the flow-through was discarded. 500 µl of Buffer RPE was added to the spin column and centrifuged for 15 seconds at 8000 g and the flow through was discarded. Another 500 µl of Buffer RPE was added to the spin column and centrifuged for 2 minutes at 8000 g to dry the column membrane and to remove residual ethanol. The spin column was then transferred to a new collection tube and centrifuged at full speed for 1 minute to eliminate any possible carryover of Buffer RPE. The spin column was transferred to a new 1.5 ml collection tube and 30 µl of RNAse-free water was pipetted directly on to the spin column membrane. The column was then centrifuged at 8000 g for 1 minute. The eludate was then pipetted back on to the spin column which was centrifuged again at 8000 g for 1 minute. This acts to increase the final concentration of the RNA solution.

The concentration of each RNA sample was checked using a nanodrop (ThermoFisher Scientific Nanodrop 2000/2000c spectrophotometer). Prior to use the
nanodrop was cleaned and 1 µl RNAse-free water was used to blank set the nanodrop on the RNA setting. 1 µl of each sample was then added to the nanodrop to determine its' concentration in ng/µl. RNA samples were frozen at -80°C until they were required.

2.4.2 Quality control of RNA prior to microarray analysis

RNA samples with concentrations > 5 ng/µl were analysed by a core University service within Biosciences for their RNA Integrity Number (RIN). The RIN is calculated using the traces from electrophoresis of RNA, which are analysed using the Agilent 2100 Expert software to generate the RIN. RIN is a number in the range 1-10. 1 represents completely degraded RNA and 10 represents an intact RNA sample. Samples with RIN>5 were included in the microarray analyses. An example of electrophoresis and an individual sample trace is presented in figure 2.3.

2.4.3 Labelling RNA for mRNA microarray analysis

This was performed according to instructions in the Agilent One-Color Microarray-Based Gene Expression Analysis: Low Input Quick Amp Labeling protocol. 25 ng of RNA was used for each reaction. The amount of RNA solution containing 25 ng of RNA was transferred to a 1.5 ml microcentrifuge tube and the volume increased to 2.5 µl with RNAse free water. 2 µl of Spike mix (Agilent One-Color RNA Spike-in Kit; 5188-5282) was then added to each tube. Labelling was performed using Agilent Low-Input Quick Amp Labeling Kit, One-Color (5190-2305), which contains the required consumables. 0.8 µl of T7 Primer was added to 1 µl of nuclease-free water for each sample. The T7 Primer mix was added to the RNA sample prior to incubation at 65°C for 10 minutes and placed on ice for 5 minutes.
Figure 2.3: Quality control of extracted RNA. A: An example of electrophoresis of 10 RNA samples (and control). Intact RNA is visualised by sharp bands at 4000 and 2000 representing 28S and 18S rRNA bands respectively. B: An individual trace of a good quality RNA sample with RIN of 8.3.
The 5x First Strand Buffer was warmed at 80°C for 3 minutes and was vortexed to ensure resuspension. The cDNA master mix was prepared by adding 2 µl 5x First Strand Buffer, 1 µl 0.1 M DTT, 0.5 µl mM dNTP Mix and 1.2 µl Affinity Script RNase Block Mix per reaction. Each RNA sample was briefly spun in a microcentrifuge before 4.7 µl of cDNA master mix was added. This was then incubated at 40°C for 2 hours and 70°C for 15 minutes. The samples were then incubated on ice for 5 minutes. A Transcription master mix was created by mixing the following: 0.75 µl nuclease-free water, 3.2 µl 5x Transcription buffer, 0.6 µl 0.1 M DTT, 1 µl NTP Mix, 0.21 µl T7 RNA Polymerase Blend and 0.24 µl Cyanine 3-CTP. 6 µl of Transcription Master Mix was added to each sample before incubating the samples at 40°C for 2 hours.

The next step involved using the Qiagen RNeasy Mini Kit (74104) to purify the cRNA samples. 84 µl of nuclease-free water was added to the cRNA sample. 350 µl of Buffer RLT was then added to the sample and mixed using a pipette. 250 µl of ethanol was added to the mix, which was then transferred to an RNeasy Mini Spin Column in a 2 ml collection tube. The sample was spun in a centrifuge at 4°C for 30 seconds at 10 000 g, flow-through was discarded and 500 µl of Buffer RPE was added to the column. This was centrifuged at 4°C for 1 minute at 10 000 g. The collection tube was discarded and the spin column was transferred to a new collection tube (1.5 ml) for centrifugation at 4°C for 30 seconds at 10 000 g. The collection tube was discarded and replaced with a new collection tube (1.5 ml). 30 µl of RNase free water was added directly onto the spin column filter membrane. After 1 minute the sample was centrifuged at 4°C for 30 seconds at 10 000 g. The resultant cRNA sample was kept on ice.
After purification the sample was measured using a nanodrop with the Microarray Measurement setting (ThermoFisher Scientific Nanodrop 2000/2000c spectrophotometer). The nanodrop was cleaned prior to use and blank set with nuclease free water. RNA-40 was selected as the sample type. 1 µl of each sample was loaded on to the nanodrop. The following values were recorded: Cyanine 3 dye concentration (pmol/µl), RNA absorbance ratio (260nm/280nm) and cRNA concentration (ng/µl). These values were used to determine the yield and specific activity of each sample using the equations below:

\[
\text{YIELD: } \frac{(\text{Concentration of cRNA}) \times 30\mu l \text{ (elution volume)}}{1000} = \mu g \text{ of cRNA}
\]

\[
\text{SPECIFIC ACTIVITY: } \frac{\text{Concentration of Cy3} \times 1000}{\text{Concentration of cRNA}} = \text{pmol Cy3 per } \mu g \text{ cRNA}
\]

Only results with a yield >0.825 and a specific activity ≥ 6 were sent for microarray analysis.

2.4.4 Hybridization of mRNA for microarray

This was performed by me, under supervision, in the Biosciences department at the University of Birmingham according to instructions in the Agilent One-Color Microarray-Based Gene Expression Analysis: Low Input Quick Amp Labeling protocol.

Initially hybridization was performed using the Agilent Gene Expression Hybridization Kit (5188-5242). The following was mixed in a 1.5 ml nuclease-free microfuge tube:
600 ng labelled cRNA, 5 µl 10x Gene Expression Blocking Agent, 1 µl 25x fragmentation buffer and nuclease-free water to bring the total volume to 25 µl. This mix was incubated at 60°C for 30 minutes to fragment the RNA. Immediately after this the mixture was cooled on ice for one minute and 25 µl of 2x Hi-RPM Hybridization Buffer was added to stop the fragmentation reaction. The mixture was then centrifuged at 10,000 g for one minute to reduce bubble formation and was stored on ice.

A clean gasket (Agilent Hybridization Gasket slide kit; G2534-60014) was loaded into an Agilent SureHyb chamber base. The gasket was then loaded horizontally with 8 hybridized samples as in figure 2.4. 40 µl of sample was loaded into each well. An Agilent slide (SurePrint G3 Unrestricted GE 8X60K; G4858A) was placed active side down parallel to the gasket slide so that the ‘Agilent’ barcode faced down. The SureHyb chamber cover was clamped over the slides and hand-tightened. The chamber was rotated vertically to ensure that the bubbles in the slides were mobile (figure 2.5).

The chamber was then loaded into a hybridization chamber at 65°C for 17 hours. During this time the Gene Expression Wash Buffer 2 (Agilent Gene Expression Wash Pack; 5188-5327) was pre-warmed to 37°C.
Figure 2.4: Diagram of a loaded Agilent gasket. Eight hybridized RNA samples were loaded into the gasket horizontally. The gasket is read and the results are printed in the order above.

Figure 2.5: Magnified view of a RNA sample loaded into a gasket. The bubble seen should be freely mobile so that the sample can move freely within the gasket and therefore covers the entire Agilent slide.
After hybridization the array-gasket was removed and placed into Gene Expression Wash Buffer 1 (Agilent Gene Expression Wash Pack; 5188-5327). Whilst the array-gasket was submerged the array was separated from the gasket using forceps. The slide was removed from the wash buffer and placed into a slide rack also submerged in Gene Expression Wash Buffer 1, which was stirred with a magnetic stirrer for one minute. The slide rack was then removed and submerged into another dish containing Gene Expression Wash Buffer 2. This was stirred using a magnetic stirrer for one minute. The slide rack was removed slowly (over 5-10 seconds) from the wash buffer. The dry slide was then uploaded into a microarray scanner.

2.5 Microarray Analysis

All microarray analysis (other than Ingenuity Pathway Analysis) was performed using packages within the computer software R. R is available as a free to download software online (https://cran.r-project.org/mirrors.html) and is both a language and an environment for statistical computing and graphics.(345)

2.5.1 Quality analysis of microarray results

The Bioconductor package arrayqualitymetrics was used in R to determine the quality of each sample in the microarray experiments. Bioconductor is an online provider of freely accessible software for bioinformatics. It provides tools, which use the R software for the analysis of genomic data such as microarrays. (346) Arrayqualitymetrics produces a file containing information about the quality of each sample included in a microarray experiment. Results appear in the user’s current working directory and a summary can be viewed in the HTML page “index.html”. Arrayqualitymetrics highlights samples that are outliers according to 4 criteria:
distances between arrays, array intensity distributions, variance of arrays and array quality. (347) All text files (results) for each sample in each microarray experiment were saved together with a summary file (“targets.txt”) in the same folder. This consisted of the name of each sample text file and whether or not that sample was from a patient with or without COPD. An example of a typical “targets.txt” file is shown in figure 2.6.

To process array data in arrayqualitymetrics it is necessary to create an object consisting of the data called an ExpressionSet. This was done by processing the data initially using the programme limma (Linear Models for Microarray Data). Limma is a package designed for the analysis of gene expression microarray data and can be used to assess differential gene expression. It uses linear models to assess gene expression in multifactor experiments and borrows information across genes to provide a stable analysis for small sample sizes. (348) R was opened on the desktop and the Bioconductor package limma was installed using the following commands:

source("https://bioconductor.org/biocLite.R")

biocLite("limma")

Limma was opened using the command: Library(limma). The working directory was set using setwd("~/Desktop/wheremyfilesare"). The “targets.txt” file was read into limma (targets <- readTargets("targets.txt")) which was used to upload the data into an RGList object called x (x <- read.maimages(targets, path=".", source="agilent",green.only=TRUE)). To correct for background the background signal was then subtracted from the signal from each array (y <- backgroundCorrect(x, method="normexp", offset=16)). Array data was normalised
between arrays (y <- normalizeBetweenArrays(y, method="quantile")) and replicate spots were averaged (y.ave <- avereps(y, ID=y$genes$ProbeName)). This resulted in the production of an Elist object (y.ave). In order to upload this information into arrayqualitymetrics y.ave needed to be converted into an ExpressionSet. This was done using the R package Biobase. This is part of the Bioconductor project and is used in different packages to convert data into the ExpressionSet format.(349) Biobase was installed into R using the following commands:

source("https://bioconductor.org/biocLite.R")

biocLite("Biobase")

Biobase was opened using the command: library(Biobase). The Elist object y.ave was converted into an ExpressionSet using “eset <- ExpressionSet(assayData = assayDataNew(exprs = y.ave$E))”. Arrayqualitymetrics was installed into R using the following commands:

source("https://bioconductor.org/biocLite.R")

biocLite("arrayQualityMetrics")

Once arrayqualitymetrics was opened using “library(arrayQualityMetrics)” the ExpressionSet (eset) was run through arrayqualitymetrics using the following command: arrayQualityMetrics(expressionset = eset, outdir = "Report_for_eset", force = TRUE). A report (Report_for_eset) was formed in the current working directory. The file “index.html” in this report was reviewed and samples that were outliers in terms of quality control were excluded from further analysis. An example of a poor quality sample is seen in figure 2.7.
<table>
<thead>
<tr>
<th>SampleNumber</th>
<th>FileNmae</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>copd1 true.txt</td>
<td>copd</td>
</tr>
<tr>
<td>2</td>
<td>copd2 true.txt</td>
<td>copd</td>
</tr>
<tr>
<td>3</td>
<td>copd3 true.txt</td>
<td>copd</td>
</tr>
<tr>
<td>4</td>
<td>normal1 true.txt</td>
<td>normal</td>
</tr>
<tr>
<td>5</td>
<td>normal2 true.txt</td>
<td>normal</td>
</tr>
<tr>
<td>6</td>
<td>normal3 true.txt</td>
<td>normal</td>
</tr>
</tbody>
</table>

**Figure 2.6:** A “targets.txt” file used in arrayqualitymetrics. This is a file created in the program 'notepad' which numbers all samples within the microarray experiment and lists each sample name and patient condition (i.e COPD or normal).
Figure 2.7: Output from arrayqualitymetrics. This demonstrates that microarray sample 5 is an outlier in comparison to other microarrays performed based upon two criteria. A: Individual array quality plot. Samples of good quality concentrate around 0 on the axis. Sample 5’s intensity is spread out around the axis. B: Principle components analysis looking at distance between arrays. Sample 5 is far away from the other samples.
2.5.2 Assessment of differential gene expression using Limma

The text files for samples that passed the quality control in each microarray experiment were saved in a folder on the computer desktop. A summary file called “targets.txt” was created in the same folder consisting of the name of each sample text file and whether or not that sample was from a patient with or without COPD (as in figure 2.6).

R was opened on the desktop and the Bioconductor package limma was installed using the following commands:

source("https://bioconductor.org/biocLite.R")

biocLite("limma")

Limma was opened using the command: Library(limma). The working directory was set using setwd("~/Desktop/wheremyfilesare"). The “targets.txt” file was read into limma (targets <- readTargets("targets.txt")); which was used to upload the data into an RGList object called RG (RG <- read.maimages(targets, path=".", columns = list(G = "gMedianSignal", Gb = "gBGMedianSignal", R = "gProcessedSignal", Rb = "glIsPosAndSignif"), annotation = c("Row", "Col", "FeatureNum", "ControlType", "ProbeName")). To correct for background the background signal was then subtracted from the signal from each array (RG <- backgroundCorrect(RG, method="normexp", offset=16)). Array data was normalised between arrays (RG$G <- normalizeBetweenArrays(RG$G, method="quantile"); and green channel intensity values were log2 transformed (RG$G <- log2(RG$G)). The RGList object was converted into a MAList object for further manipulation: E <- new("MAList", list(targets=RG$targets, genes=RG$genes, source=RG$source, M=RG$Gb,
Replicate spots in the array data were averaged using: 
```r
E.avg <- avereps(E, ID=E$genes$ProbeName).
```
A design matrix was created for the linear modelling function in three steps:

```r
f <- factor(targets$Condition, levels = unique(targets$Condition))
design <- model.matrix(~0 + f)
colnames(design) <- levels(f)
```

The intensity values were applied to `lmfit` which fits the linear model for each gene (fit <- `lmFit(E.avg$A, design)`). A contrast matrix was set up to compare the two test groups (COPD and non-COPD): `contrast.matrix <- makeContrasts("copd-normal", levels=design)`. The matrix was applied to the modelled data and the statistics for the data were computed: `fit2 <- contrasts.fit(fit, contrast.matrix)`. This data was then outputted into the current working directory as follows:

```r
output <- topTable(fit2, adjust="BH", coef="copd-normal", genelist=E.avg$genes, number=1000000)
write.table(output, file="Treatment1_vs_Treatment2.txt", sep="\t", quote=FALSE)
```

This produced the text file “Treatment1_vs_Treatment2.txt” which was copied and pasted into an excel file for further analysis. The appearance of the “Treatment1_vs_Treatment2.txt” file can be seen in table 2.2. To obtain the gene name for each probe in the limma output file the function `VLOOKUP` was used comparing the list of probe names to a reference list of probes and gene names from agilent which was available to download online.(350)
Table 2.2: Output from limma.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Row</th>
<th>Col</th>
<th>FeatureNum</th>
<th>ControlType</th>
<th>ProbeName</th>
<th>logFC</th>
<th>AveExpr</th>
<th>t</th>
<th>P.Value</th>
<th>adj.P.Val</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_23_P424617</td>
<td>295</td>
<td>15</td>
<td>48231</td>
<td>0</td>
<td>A_23_P424617</td>
<td>1.89294</td>
<td>6.766007</td>
<td>5.68436</td>
<td>-0.000299</td>
<td>0.94705</td>
<td>-4.42954</td>
</tr>
<tr>
<td>A_33_P3300027</td>
<td>330</td>
<td>115</td>
<td>54071</td>
<td>0</td>
<td>A_33_P3300027</td>
<td>1.61983</td>
<td>7.198713</td>
<td>5.33265</td>
<td>-0.000472</td>
<td>0.94705</td>
<td>-4.43513</td>
</tr>
</tbody>
</table>

Each probe from the microarray experiment is listed in the first column. The other columns show the rest of the data extracted from limma. The logFC column gives the value of the contrast between the two patient groups. The AveExpr column is the average expression of the probe across all the arrays and channels in the experiment. P value lists whether or not the expression was significantly different between groups. Adj p val gives the p value adjusted for significance testing.
2.5.3 Assessment of false discovery rate (FDR)

The package LBE was used to determine the false discovery rate (FDR) values for the differential miRNA and gene expression analyses. FDR adjusts the test for significance to control for the expected proportion of incorrect rejections of null hypotheses in multiple tests.

The p values from a limma output were copied and pasted into a text file ("pvalues.txt") in a folder on the Desktop. R was opened on the Desktop and LBE was installed into R using the following commands:

source("https://bioconductor.org/biocLite.R")

biocLite("LBE")

LBE was opened in R (Library(LBE)) and the working directory was set to the folder containing “pvalues.txt”: setwd("~/Desktop/wheremyfileis"). The p value data was read into R (pvalues = read.table("pvalues.txt")) and the table was converted into a vector (unlist(pvalues, use.names = FALSE)). The LBE function was used to calculate the FDR based upon the p value data: LBE.res <- LBE(pvalues[[1]], a = NA, l = 0.05, ci.level = 0.95, qvalues = TRUE, plot.type = "none", FDR.level = 0.05, n.significant = NA). A text file of the results was created using the function: LBEwrite(LBE.res). This appeared in the current working directory and was copied and pasted into excel for further analysis.
2.5.4 Assessment of differential gene expression using Significance Analysis of Microarrays (SAM)

SAM is a statistical technique that finds genes that are significantly differentially expressed between groups in microarray experiments.(353) SAM calculates a statistic for each gene (d statistic) that measures the relationship between the expression of that gene and the different groups. SAM is installed via R using the following commands:

```
install.packages(c("samr", "matrixStats", "GSA", "shiny", "shinyFiles", "openxlsx"))
source("http://bioconductor.org/biocLite.R")
biocLite("impute")
```

SAM was started using the following commands in R:

```
library(shiny)
library(shinyFiles)
runGitHub("SAM", "MikeJSeo")
```

This brings up a browser window with a user interface in which it is possible to upload data.

For each array experiment the text files from samples passing the quality control analysis were opened and saved as excel files. The function VLOOKUP was used to extract the probe ID, gene ID and 'gProcessedsignal' (microarray signal) from each individual excel file (using Feature number as the control). Samples from patients without COPD were labelled as group 1 and those with COPD as group 2. This was used to create a SAM input file an example of which is shown in table 2.3.
Table 2.3: Example of a SAM input file.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE_BrightCorner</td>
<td>GE_BrightCorner</td>
<td>26160.82</td>
<td>24726.47</td>
<td>14630.95</td>
<td>24073.52</td>
</tr>
<tr>
<td>DarkCorner</td>
<td>DarkCorner</td>
<td>3.104816</td>
<td>3.44231</td>
<td>3.483712</td>
<td>2.878703</td>
</tr>
<tr>
<td>DarkCorner</td>
<td>DarkCorner</td>
<td>3.115946</td>
<td>3.405396</td>
<td>3.470299</td>
<td>38.07094</td>
</tr>
</tbody>
</table>

The first and second columns contain probe and gene names. The other columns contain the gProcessedSignal for each sample. ‘gProcessedSignal’ is the intensity with which each probe was recorded in the microarray. Samples are labelled into the two test groups (non-COPD (1) and COPD (2)).

Figure 2.8: An example SAM plot. This is a graphical representation of the statistics for each probe recorded in a microarray analysis. Probes that are upregulated in group 2 (COPD) are in red. Probes in green are downregulated in group 2.
The SAM input file was uploaded into the browser interface and results were limited to fold change 2. Fold change 2 was chosen as this represents a large biologically significant difference in expression between groups and has been previously used in other microarray studies.(354) SAM was run and an excel results file was produced containing a list of significant results and a plot of the statistics for each gene (figure 2.8). To obtain the gene name for each probe in the SAM output file the function VLOOKUP was used comparing the list of probe names to a reference list of probes and gene names from agilent which was available to download online.(350)

2.5.5 Ingenuity Pathway Analysis (IPA)

The data used for the IPA was drawn from the results of the SAM analysis performed in sections 3.3.2 (miRNA) and 3.4.2 (mRNA). Data was split into 4 different analyses according to the date that each microarray was performed. Data could not be directly combined owing to inherent significant differences between datasets most likely as a result of a variation in dye fluorescence between arrays. The results of the microarray analyses were limited to fold change 2 in order to restrict the IPA to only those genes with a significantly different expression between non-COPD and COPD. Therefore, 4 datasets were run through IPA:

- 2014 miRNA microarray data limited to probes significantly differentially expressed in 2015.
- 2015 miRNA microarray data limited to probes significantly differentially expressed in 2014.
- 2014 mRNA microarray data limited to probes significantly differentially expressed in 2016.
• 2016 mRNA microarray data limited to probes significantly differentially expressed in 2014.

Each array result was uploaded into IPA separately and a core analysis was performed. The core analysis compares the results of a microarray to existing knowledge in the Ingenuity Knowledge Base. This is a manually curated database consisting of millions of findings drawn from peer reviewed literature, public databases (such as miRBase) and experimental datasets.(355) Therefore, microarray results are compared to known relationships identified from multiple tissue types and disease processes. Significantly altered pathways are identified and Ingenuity places these into groups of related pathways (diseases and biological functions). For example, all pathways associated with cancer. The groups with the most significantly altered pathways are flagged up in the IPA. This process is outlined in figure 2.9. This can identify networks and pathways not previously identified in the tissue under study. Thus, a range of diseases and associations may be significant in the IPA including diseases not in the tissue of interest. This requires the researcher to examine whether or not these associations may have biological relevance.

IPA uses two main statistical methods to determine statistical significance of pathways under study:

1. P-value of overlap. This is used for all types of analysis in the IPA. The null hypothesis is that there is no overlap between the molecules in an array dataset and a particular known disease or pathway. Right tailed Fisher’s exact tests are used and the significance is set at p<0.05.

2. Z-score. This is used when determining the downstream effects of microarray data on certain processes and performs a correlation between what is known
and the researcher's data. A z-score $> 2$ predicts activation of a pathway and $< 2$ predicts inhibition. (355)

The results of the core analyses were reviewed and significantly altered pathways were identified for each subset. Firstly, a summary file was created which outlines the top canonical pathways, predicted upstream regulators, associated diseases, molecular pathways and biological networks associated with the microarray dataset. Each pathway analysis was opened in IPA and explored further. For example, significantly up- and down-regulated canonical pathways were opened in IPA. Where possible, IPA was used to predict the downstream effect of the microarray changes on effector molecules and disease processes. An example of how IPA was used to predict downstream biological effects is shown in figure 2.10.
Figure 2.9: The process by which Ingenuity identifies significant associations with diseases and biological functions. Microarray data is compared to known cellular processes in the Ingenuity knowledge base and significantly altered (p<0.05) processes are identified. Related processes are placed into groups (diseases and biological functions), which are highlighted in IPA.
Figure 2.10: Example of a canonical pathway significantly altered in a microarray analysis. Molecules highlighted in magenta were measured in the microarray analysis. Molecules in red or pink were upregulated in the analysis and molecules in green were downregulated (in COPD). Pathways in orange are predicted upregulated targets and pathways in blue are predicted to be downregulated.
2.6 Reverse transcription

2.6.1 Reverse transcription of RNA to cDNA for RNA samples of concentration >300 ng/µl

This method uses the Applied Biosystems High Capacity RNA-to-cDNA kit (4387406). This kit can be used for the conversion of 20 ng-2 µg RNA to cDNA. The kit consists of a buffer and an enzyme mix, which were initially thawed on ice. A 50 µl stock solution of RNA was created for each sample by diluting RNA to a concentration of 111 ng/µl. A multi-mix of buffer and enzyme was created by adding 10 µl 2X RT Buffer to 1 µl 20X RT Enzyme Mix per sample. 9 µl of RNA stock solution was added to 11 µl multi-mix to create 20 µl of reaction mix which was placed into a PCR tube. The tubes were then centrifuged at 1000 g for 1 minute to spin down the contents and to eliminate air bubbles.

The thermal cycler (TC-PLUS; Techne) was run for 60 minutes at 37°C, 5 minutes at 95°C then 10°C continuously. Samples were removed from the thermal cycler and stored at -15°C to -25°C.

2.6.2 Reverse transcription of RNA to cDNA for RNA samples of concentration <300 ng/µl

This method uses the Invitrogen SuperScript IV VILO Master Mix (11756050). The kit is able to convert 10 pg-25 µg RNA to cDNA. 35 µl of RNA stock solution was created for each RNA sample by diluting RNA with nuclease free water to a concentration of 1 ng/µl. 5 µl of stock solution was combined with 4 µl of SuperScript IV VILO Master Mix and 11 µl of nuclease-free water to create a 20 µl reaction mix which was placed into a PCR tube. The tubes were then centrifuged at 1000 g for 1 minute to spin down the contents and to eliminate air bubbles. The thermal cycler (TC-PLUS; Techne) was run for 10 minutes at 25°C, 20 minutes at 50°C, 5 minutes
at 85°C then 10°C continuously. Samples were removed from the thermal cycler and stored at -15°C to -25°C.

2.6.3 Reverse transcription of miRNA
This method uses the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems; 4366595). The kit components were initially thawed on ice. A Reverse Transcription master mix was created by mixing 0.15 µl 100 mM dNTPs, 1 µl MultiScribe Reverse Transcriptase (50 U/µl), 1.5 µl 10X Reverse Transcription Buffer, 0.19 µl RNase Inhibitor (20 U/µl) and 4.16 µl of nuclease-free water per sample. The master mix was mixed gently and centrifuged at 1000 g for 1 minute. The master mix was then kept on ice. 30 µl of RNA stock solution was created for each RNA sample by diluting RNA with nuclease free water to a concentration of 1 ng/µl. 5 µl of RNA stock solution was combined with 7 µl of master mix in a 0.2 ml reaction tube, was mixed gently, centrifuged at 1000 g for 1 minute and kept on ice. The 5X RT primer was thawed on ice prior to vortexing and centrifugation at 1000 g for 1 minute. 3 µl of 5X RT primer was added to the RNA/master mix solution. This was mixed by gently inverting prior to centrifugation at 1000 g for 1 minute. The final solution was kept on ice for at least five minutes prior to loading the thermal cycler. The thermal cycler (TC-PLUS; Techne) was run for 30 minutes at 16°C, 30 minutes at 42°C, 5 minutes at 85°C before being held at 10°C. The cDNA was stored at -15°C to -25°C.

2.7 Real-time Polymerase chain reaction (qPCR)
2.7.1 qPCR of mRNA
This was performed using the Universal Probe Library for qPCR (Roche; 04683633001). This consists of 90 hydrolysis probes that can be used with appropriate primers for qPCR. The mRNA of interest was initially searched for on the
Roche Universal Probe Library website. This gave the correct probe to use as well as sequences for the forward and reverse primer. The forward and reverse primer were ordered online via the Roche website. Primers used in this thesis are in table 2.4.

The cDNA to be used for the qPCR reaction was diluted four fold in order to produce 80 µl of diluted cDNA from each 20 µl reverse transcription reaction. A master mix of qPCR components was created from the following: 0.5 µl nuclease free water, 0.5 µl forward primer, 0.5 µl reverse primer, 10 µl EXPRESS qPCR Supermix Universal (Invitrogen; 11785200) and 0.5 µl probe for each reaction. The master mix was placed into a 1.5 ml Eppendorf tube, shaken to mix and centrifuged at 1000 g for 1 minute to displace the contents to the bottom of the tube. 12 µl of master mix was pipetted into wells on a 96-well plate. 8 µl of diluted cDNA was added to each master mix aliquot. The plate was centrifuged at 650 g for five minutes to drive the contents into the wells. The 96 well plate was transferred to a PCR machine (Applied Biosystems 7500 Fast) and was programmed to run as follows: step 1: 15 minutes at 96°C, step 2: 40 cycles of 15 seconds at 96°C, 30 seconds at 58°C and 30 seconds at 72°C, step 3: 3 minutes at 72°C.

2.7.2 qPCR of miRNA
This was performed using TaqMan microRNA assay kits (Applied Biosystems; 4427975). The TaqMan assay (20x) and cDNA samples were thawed on ice and resuspended by vortexing. The assay and cDNA were then centrifuged at 1000 g for one minute. The master mix (TaqMan Universal PCR Master Mix II, no UNG; Applied Biosystems, 4440040) was mixed by gently swirling the bottle. The qPCR reaction mix was prepared in a 1.5 ml Eppendorf tube for each sample. The samples were all
This table outlines the genes that were tested for in qPCR analyses performed in this study. The sequences of the primers used are given in addition to the probe that was used for each gene from the Roche universal probe library.
run in triplicate. The following were pipetted into an Eppendorf for each sample: 1.8 µl TaqMan Small RNA assay (20x), 2.4 µl cDNA, 18 µl master mix and 13.81 µl nuclease-free water to create a total volume of 36.01 µl. The Eppendorf was capped and inverted several times to mix before being centrifuged at 1000 g for one minute. 10 µl of qPCR reaction mix was transferred into a well of a 384-well plate. The plate was sealed and centrifuged at 650 g for one minute. The plate was inserted into a PCR machine (Roche Lightcycler) and the reaction volume was set at 10 µl. The PCR machine was programmed to run as follows: step one: 95°C for 10 minutes, step two: 40 cycles of 95°C for 15 seconds and 60°C for 60 seconds.

2.7.3 Analysis of qPCR results

All qPCR results were analysed using the delta Ct method. mRNA targets were normalised to the housekeeping gene Flotillin 2 to calculate the ∆Ct. (∆Ct =Ct of test mRNA – Ct of Flotillin 2.) Fold changes for mRNAs were calculated using the following formula:

$$\Delta \Delta \text{Ct} = \Delta \text{Ct}(\text{test group(sample 1,2,3,...)}) - \text{Mean } \Delta \text{Ct}(\text{control group})$$

Fold change = \(2^{-\Delta \Delta \text{Ct}}\)

To see if the difference between groups was statistically significant a one-tailed t-test was used to compare the mean ∆Ct in both groups. The Shapiro-Wilk test was used prior to this to confirm that the two groups were normally distributed.

MiRNA qPCR was performed in a similar manner but miRNA targets were normalised to the small RNA RNU48 rather than Flotillin 2.
Figure 2.11: Culture of Mouse Lung Endothelial Cells (MLECs). A fresh lung sample is initially minced before being added to collagenase I for digestion. The digested sample is passed through a filter before being centrifuged. The cell pellet is resuspended in medium and placed in a 10 cm plate. This is cultured until near confluence. The cell culture is then incubated with CD16/32 coated magnetic beads. The beads bind to macrophages, which are then removed (negative selection). The cells are replated and cultured until near confluence again. The cell culture is then incubated with CD102 coated beads, which bind to endothelial cells before a magnet is used to isolate the beads from the rest of the solution (positive selection). The beads are resuspended in medium and placed in a 10 cm plate. This is cultured until confluence before being split. The splitting is repeated until no beads remain.
2.8 Culture of Mouse Lung Endothelial Cells (MLECs)

2.8.1 Collagenase digestion of lung tissue

This method is outlined in figure 2.11. Mouse lung tissue was kindly donated by Dr Victoria Heath from the University of Birmingham. Strain C57BL6 (a non-inflammatory strain) was used and mice were culled by a schedule 1 method of culling (CO₂). Lungs were dissected out with scissors. Ethical approval for this mouse was covered by Home Office licence number PPL 70/8704. Prior to tissue digestion, MLEC/HPEC (human pulmonary endothelial cell) media and collagenase solution was prepared in a sterile environment in the tissue culture hood. MLEC/HPEC media was prepared by mixing the following: 400 ml DMEM-F12 with glutamine (11320033; Life Technologies (Gibco)), 45 mg heparin in 1 ml (Sigma; H3149), 5 ml Penicillin/streptomycin (Gibco; 15140-122), 1 ml gentamicin/amphotericin (R01510; Life Technologies (Gibco)), 250 µl Endothelial Cell Growth Supplement (E2759; Sigma Aldrich): 15 mg/ml in DMEM-F12 at 4°C and 100 ml Fetal Bovine Serum (Gibco; 10270106). The MLEC/HPEC media was sterilised by filtering the solution through a filter bottle. Collagenase solution was prepared by adding 0.1 g of type I collagenase (Gibco; 17100017) in 25 ml PBS. The solution is then incubated at 37°C for 1 hour. A further 25 ml PBS is added before filter sterilising with a 0.2 µm filter. PBS solution was created by mixing 5 PBS tablets (Sigma; P4417) with 1 l distilled water before autoclaving the solution. 0.1% gelatin solution was also created by adding 0.5 g gelatin (Fluka Biochemika; 04055) to 500 ml water followed by autoclaving the solution. Three 10 cm plates were prepared in the hood in order to clean the lung tissue (DMEM-F12, 70 % ethanol and MLEC/HPEC medium). Each plate consisted of 12 ml of solution. Tweezers were cleaned with 70% ethanol and
were used to hold and wash the lungs by placing the lungs in DMEM-F12 followed by 70% ethanol for 30 seconds. The lung was finally transferred to the MLEC/HPEC medium plate. To aid tissue digestion the mouse lungs were minced using a sterile scalpel until the lungs resembled pâté. The lungs were mixed with 10 ml of collagenase solution and incubated for 1 hour at 37°C in a falcon tube. The solution was mixed by shaking the sample by hand after 30 minutes. During this incubation period a 10 cm gelatin coated plate was prepared by placing 10 ml 0.1% gelatin on to a 10 cm plate and incubating this at 37°C. After the one hour digestion phase 10 ml MLEC/HPEC media was added to the lung solution to stop further digestion of tissue from occurring. The lung tissue was then placed into a petri dish and was passed in and out of a 20 ml syringe using a 19 gauge needle four times in order to limit clumping of cells. The solution was then filtered through a 70 µm filter into a falcon tube to remove cell debris. 20 ml MLEC/HPEC media was added to the lung solution prior to centrifugation at 200 g for 10 minutes. The supernatant was removed leaving 5 ml media at the bottom of the falcon tube. The cell pellet was resuspended and 10 ml MLEC/HPEC medium was added. The gelatin was removed from the 10 cm plate and the lung solution was added. This was then incubated at 37°C, 5% CO₂ overnight.

2.8.2 Negative selection of macrophages using magnetic beads

The lung culture was purified by the negative selection of macrophages on day two. Initially the MLEC/HPEC media was removed and the plate was washed twice using PBS. 5 ml of fresh MLEC/HPEC media was then added. The cells were then incubated at 4°C for 20 minutes to reduce the metabolism of the cells and prevent endocytosis of antibody. During this incubation step 10 µl of antibody to CD16/32
(ThermoFisher Scientific; Rat anti-mouse; 14-0161-82) was added to 3 ml PBS and stored on ice. After the 20 minute incubation step the MLEC/HPEC media was removed and the 3 ml antibody solution was added to the lung cells. The cells were incubated at $4^\circ$C for 30 minutes to allow the antibody time to bind to the macrophage cells. In order for the antibody to cover all the plate the solution was swirled around the plate after 15 minutes. During this phase the beads were prepared. 100 µl of magnetic dynabeads (ThermoFisher Scientific; 11035) were added to 1 ml PBS. The bead solution was added to a magnet and the solution was removed, effectively washing the beads. This was repeated two more times. After the third wash the beads were resuspended in 100 µl PBS. 50 µl of bead solution was then combined with 3 ml MLEC/HPEC medium. The antibody solution was removed from the plate and the plate was washed with PBS to remove traces of antibody solution. The 3 ml bead solution was added to the plate, which was incubated at $4^\circ$C for 30 minutes again with the plate swirled after 15 minutes. During this incubation phase a 10 cm gelatin coated plate was created by adding 10 ml 0.1% gelatin to a 10 cm plate and incubating this at $37^\circ$C, 5% CO$_2$. The bead solution was removed from the plate and this was washed three times with PBS to remove all traces of beads. The beads/cells were then removed from the plate by adding 2.5 ml 2x trypsin (Gibco; 15400054) and incubating the cells for two minutes at $37^\circ$C, 5% CO$_2$. Detachment of cells was observed under the microscope and 9 ml MLEC/HPEC media was added to inactivate the trypsin. The cells were transferred to a 15 ml falcon tube and placed into a magnetic holder for five minutes to give the beads time to attach to the magnet. The solution was removed and placed on to the prepared gelatin plate. The plate was placed into the incubator at $37^\circ$C, 5% CO$_2$. The cell culture was reviewed each day
and media was replaced every other day until large colonies of endothelial cells were visible.

2.8.3 Positive selection of endothelial cells by magnetic beads

Initially the medium was removed and replaced with 5 ml of fresh MLEC/HPEC medium. The plate was incubated at 4°C again for 15 minutes in order to reduce cell metabolism. During this time a 10 cm gelatin plate was prepared by adding 10 ml 0.1% gelatin to a 10 cm plate and incubating this at 37°C, 5% CO₂. The plate was labelled as ‘MLEC positive P1’. Anti-ICAM2 (CD102) antibody solution was prepared by adding 10 µl of anti-ICAM2 antibody (BD Pharmigen, Rat anti-mouse; 553326) to 3 ml PBS on ice. The medium was removed from the plate and replaced with the antibody solution. The plate was incubated at 4°C for 30 minutes, but the plate was swirled after 15 minutes in order to distribute the antibody around the plate. During this time the beads were washed as above and resuspended in 100 µl of PBS. 50 µl of beads was combined with 3 ml of PBS. After the incubation period the antibody solution was removed and the plate was washed with PBS to remove excess antibody. The bead solution was added to the plate, which was incubated at 4°C for 30 minutes, but swirled at 15 minutes to distribute the beads across the plate. After this time the bead solution was removed and the plate washed three times with PBS to ensure no traces of bead solution remained on the plate. The cells were removed using trypsin as in section 2.8.2 and 9 ml of media was added to neutralise the trypsin. The cell solution was transferred to a 15 ml falcon tube and placed into a magnetic holder for 5 minutes to give time for the beads to bind to the magnet. The solution was removed and discarded. The beads were resuspended in 10 ml MLEC/HPEC media and were added to the gelatin plate.
Figure 2.12: Culture of Human Pulmonary Endothelial Cells (HPECs). A fresh lung sample is minced before being added to collagenase V. The digested sample is passed through a filter before being overlaid on lymphoprep. The cells are centrifuged which leads to the erythrocytes concentrating at the bottom of the tube. All cell layers above the erythrocytes are removed and centrifuged. The cell pellet is resuspended in medium and placed in a 6 well plate. The cell culture is then incubated with ulex coated magnetic beads. The beads bind to endothelial cells before a magnet is used to isolate the beads. The beads are resuspended in medium and placed in a 6 well plate. The culture is split repeatedly until no beads remain.
2.9 Culture of Human Pulmonary Endothelial Cells (HPECS)

In order to have the ability to carry out functional validation work the culture of lung endothelial cells was attempted. This would ideally provide the opportunity to study the effects of altering a target of interest in the cell of interest rather than an alternative cell (such as Human Umbilical Vein Endothelial Cells (HUVECs)). The advantage of this is that alternative cell lines may not express the target of interest thus rendering functional validation in them impossible. This method was adapted from the endothelial extraction method in 2.3.2 and similar work on extracting pulmonary endothelial cells in mice.(357)

2.9.1 Collagenase digestion of lung tissue

Lung tissue was collected in the same manner as section 2.3. Three 100 mm plates were prepared. Plate 1: 12 ml DMEM-F12 (11320033; Life Technologies (Gibco)). Plate 2: 12 ml 70% ethanol. Plate 3: 12 ml MLEC/HPEC medium (section 2.8.1). Tweezers were placed in the ethanol plate before being used to transfer the lung sample to the DMEM-F12 plate for 30 seconds. The lung sample was then washed with ethanol for 30 seconds and transferred to the plate containing medium. The lung was divided into portions approximately 3x2 cm. One portion was minced on the lid of a 100 mm plate until it resembled pâté. The portion was added to a sterile 50 ml falcon tube with 12.5 ml collagenase solution and incubated for one hour in a 37°C shaker. Collagenase solution was created by adding 25 mg collagenase (Collagenase type V; (C9263 Sigma-Aldrich, UK): 25 mg/ml in DMEM, 1 ml aliquots at -20°C) to 12.5 ml DMEM-F12 (11320033; Life Technologies (Gibco)). The collagenase solution was filter sterilised using a 0.22 µm filter.
The other portions of lung were then minced and digested in the same fashion. Gelatin plates were then prepared by pipetting 3 ml 0.1% gelatin (section 2.8.1) into a well of a 6-well plate and incubating at 37°C (for at least 15 minutes). 10 ml medium was added to the first lung portion to stop digestion, which was passed in and out of a 10 ml pipette to disaggregate the cells. The lung solution was then passed through a 100 µm cell strainer (734-0004; VWR International) into a 50 ml falcon tube. 20 ml medium was added to the falcon tube, which was then centrifuged for 10 minutes at 200 g. The supernatant was removed by pouring into another falcon tube leaving 5 ml liquid at the bottom. The cell pellet was resuspended by pipetting gently and adding 20 ml medium. In order to remove the erythrocytes from solution (which are metabolically active and can limit endothelial growth) the cell solution was then overlaid on to 12 ml lymphoprep (1114545; Axis-Shield) and centrifuged at 100 g for 30 minutes (acceleration 1). All layers above the erythrocyte layer were removed with a sterile 10 ml pipette into another 50 ml falcon tube. The cell solution was centrifuged at 200 g for 10 minutes, and supernatant was removed with a sterile pipette. This process was repeated with the other lung portions. Two cell pellets (from two lung portions) were combined and resuspended in 5 ml medium in order to increase the number of endothelial cells on one plate. The gelatin was removed from the 6-well plate and the cell suspension was added to the plate and incubated at 37°C, 5% CO₂ overnight.

On day one the media was removed and the cells were washed four times with DMEM-F12. 3 ml medium was then added to the cells. The media was changed every other day until the cells grew to cover most of the plate, but were not confluent (to prevent cell clumping).
2.9.2 Positive selection of endothelial cells by magnetic beads

Firstly, Streptavidin Dynabeads (M-280; Life Technologies (Invitrogen)) were vortexed to resuspend the beads in solution. 5 µl of beads were combined with 200 µl PBS (section 2.8.1) in a 15 ml falcon tube. The beads were washed by placing the bead solution into a magnetic holder and pipetting off the solution whilst the tube was still in place. The beads were then resuspended in 200 µl PBS and placed into the magnetic holder again before the PBS was removed. The beads were resuspended in 1 ml MLEC/HPEC medium (section 2.8.1) and 12.5 µl of Ulex solution (section 2.3.2) was added to the beads. The bead solution was then placed in a 37°C shaker for 30 minutes so the Ulex could bind to the beads. During this time the cells were removed from the plate by using 2 ml 2x trypsin (Gibco; 15400054) and placing this in the incubator at 37°C, 5% CO₂ for five minutes. Detachment of the cells was observed under the microscope. 5 ml MLEC/HPEC media was added to the cells to neutralise the trypsin and the cells were transferred to a 15 ml falcon tube for centrifugation at 200 g for 5 minutes. The cell pellet was resuspended in the bead solution and vortexed very briefly to obtain a single cell suspension. The cells/beads were then placed on a wheel in the cold room at 4°C for 30 minutes to give the cells time to bind to the beads. The mixing of the cells on the wheel should encourage the proper mixing of cells and beads. During this time 2 wells of a 6 well plate were prepared by adding 0.1% gelatin (section 2.8.1) and placing in the incubator at 37°C, 5% CO₂. After the mixing in the cold room the cells were transferred into a 15 ml falcon tube, which was placed into a magnetic holder. Whilst the tube was still in the magnet the media was removed and combined with 2 ml fresh MLEC/HPEC media. This was then plated on to a plate labelled ‘HPEC negative P1’. The beads were
resuspended in 3 ml media and added to the other plate labelled ‘HPEC positive P1’. Both plates were incubated overnight at 37°C, 5% CO₂. Cells were cultured and split until enough beads were removed for flow cytometry.

2.10 Cell storage

Surplus cells used throughout this study were stored at -80°C. This was done by pelleting the cells in solution using centrifugation at 200 g for 5 minutes. The supernatant was removed and the cells were resuspended in 1 ml of freezing media and placed into a cryogenic storing tube (Thermofisher Scientific; 375418). Freezing media was created by adding 1 ml Dimethyl sulfoxide (DMSO) (Sigma Aldrich; D2650) to 9 ml of Fetal Bovine Serum (Gibco; 10270106). Cells were initially frozen in a ‘Mr Frosty’ freezing container (ThermoFisher Scientific; 5100-0001) in order to slowly cool the cells to prevent cellular damage prior to storing in boxes at -80°C.

Cells that were frozen were slowly defrosted using a water bath at 37°C under direct supervision. 10 ml of cell media (depending on cell type) was added to the cells which were centrifuged at 200 g for 5 minutes. The supernatant was removed to remove traces of DMSO prior to plating.

2.11 Flow cytometry

2.11.1 Flow cytometry to look for CD31 expression in cells isolated from mouse lung

This was performed at the University of Wolverhampton. Initially the MLECs were removed from the plate using 2.5 ml 2x trypsin (Gibco; 15400054) as in section 2.8.2. 400 000 cells were added to each of three tubes: control, CD102 (BD Pharmigen, Rat anti-mouse; 553326) and CD31 (BD Pharmigen, Rat anti-mouse; 550274). The
cells were centrifuged at 160 g for 5 minutes and the supernatant was discarded. The
cells were resuspended in 500 µl of 4 % Formaldehyde (Sigma; HT50-1-2) in order to
fix the cells and were left for 15 minutes at room temperature. During this time each
primary antibody solution was prepared on ice. 3 µl of antibody was added to 150 µl
FACS buffer (1:50 dilution). FACS buffer was created by mixing bovine serum
albumin (BSA) (Sigma; A2058) at a ratio of 0.2% with PBS (section 2.8.1). Sodium
azide (Sigma; 71289) is then added at a ratio of 0.02% with the PBS/BSA.
After the fixation phase 8 µl PBS was added to each of the three tubes which were
centrifuged at 1200 g for 5 minutes to remove the formaldehyde. After this 150 µl
FACS buffer was added to the control tube, 150 µl of CD102 solution was added to
the CD102 tube and 150 µl of CD31 solution was added to the CD31 tube. The tubes
were left at room temperature for 2 hours to allow the antibodies to bind. During this
time the secondary antibody (BD Pharmigen; anti-rat/FITC (fluorescein
isothiocyanate); 553881) was prepared by combining 7 µl of secondary antibody to
700 µl FACS buffer. To prevent degradation of the FITC the secondary antibody was
kept in the dark. After the 2 hour incubation period the cells were washed with 10 ml
PBS and were centrifuged at 1200 g for 5 minutes. The cells were resuspended with
200 µl secondary antibody solution per each tube. The cells were incubated for 1
hour in the dark to allow the secondary antibody time to bind. After this incubation
step the cells were washed again with 10 ml PBS and centrifuged at 1200 g for 5
minutes. Finally the cells were resuspended in 300 µl PBS and were analysed using
a flow cytometer (BD Biosciences; BD Accuri C6 plus) using the flow cytometer
software. The unstained cytogram was used initially and a gate was created by
drawing a plot around the main cell population. This gate was used to limit results for
the other cell conditions (i.e. only results falling within the gate were included in the histogram plot). The histogram for the unstained cells was compared to the histograms for the other cell conditions. Data was presented as percentage positive stained cells.

2.11.2 Flow cytometry to look for CD31 expression in cells isolated from the lung

To confirm whether or not the cells isolated using the method in 2.9 were endothelial cells flow cytometry to look at CD31 expression was performed. Prior to flow cytometry the cells to be used were washed with PBS (section 2.8.1) twice. Cells were then treated with 5 ml cell dissociation buffer (Sigma; C5789) and placed into an incubator at 37°C for 30 minutes. The cells were removed and placed into a 15 ml falcon tube before being centrifuged at 200 g for 5 minutes. This has the effect of removing the dissociation buffer. The cells were then resuspended in 5 ml of MLEC/HPEC medium (section 2.8.1) in order to count the cells. 10 µl of cell solution was removed and placed into a cell haemocytometer. Around 300 000 cells were used per condition in the flow cytometry experiment. Surplus cells were removed and frozen for storage at -80°C (section 2.10).

The cells were centrifuged at 200 g for 5 minutes and the supernatant was removed. The cells were then washed by resuspending them in 5 ml PBS and centrifuging at 200 g for 5 minutes. The supernatant was removed and cells were resuspended into PBS (100 µl per condition). 100 µl of cell solution was added to each eppendorf representing each condition in the experiment: unstained cells, secondary antibody only, secondary antibody and isotype antibody control and test (CD31 antibody and secondary antibody). A stock solution of CD31 antibody (Dako; JC70A) was made by
diluting the stock antibody solution in a ratio of 1:20 with FACS buffer (section 2.11.1) and vortexed to mix. 5 µl of CD31 antibody was added to the test eppendorfs and vortexed. In a similar fashion a stock solution of isotype control antibody (Sigma, Mouse IgG1 Negative Control Antibody; CBL610) was created by diluting the antibody solution at a ratio of 1:20 with FACS buffer and vortexed to mix. 5 µl of isotype antibody was added to the isotype antibody control eppendorfs and vortexed. All eppendorfs were transferred to ice and incubated for 1 hour. The eppendorfs were centrifuged at 2500 g for 5 minutes and the supernatant was removed in order to remove excess antibody. To further wash the cells the pellets were resuspended in 250 µl of FACS buffer, centrifuged at 2500 g for 5 minutes with the removal of the supernatant afterwards twice. Each cell pellet was then resuspended in 100 µl FACS buffer. The secondary antibody, anti-mouse/FITC (fluorescein isothiocyanate) (Sigma; F2883) was diluted in a ratio of 1:100 using FACS buffer. 10 µl of this was added to the eppendorfs: secondary only, secondary antibody, isotype antibody control and test. All eppendorfs were incubated on ice for 30 minutes. To protect the FITC from degradation this was performed in the dark. After this the eppendorfs were centrifuged at 2500 g for 5 minutes. To further wash the cells the pellets were resuspended in 250 µl of FACS buffer, centrifuged at 2500 g for 5 minutes with the removal of the supernatant afterwards twice. The cells were resuspended in 300 µl of FACS buffer and transferred into FACS tubes. The cells were run through the FACS machine (Becton-Dickinson Fluorescence activated cell sorter (FACSCalibur)) and the results processed using the online available software FloJo. The unstained cytogram was used initially and a gate was created by drawing a plot around the main cell population (figure 2.13). This gate was used to limit results for the other cell
conditions (i.e. only results falling within the gate were included in the histogram plot). The histogram for the unstained cells was compared to the histograms for the other cell conditions (figure 2.14). Data was presented as percentage positive stained cells.

2.11.3 Flow cytometry for cell cycle analysis

Analysis of cell cycle was achieved by fixing cells in alcohol and staining with propidium iodide. Propidium iodide binds to DNA proportionally to the amount of DNA in the cell. Therefore, cells in S phase fluoresce more brightly than G1 cells. Cells in the G2 phase are approximately twice as bright as G1. (358)

Firstly cells were washed using PBS (section 2.8.1) before being removed from the plate using 2x trypsin (section 2.9.3). Cells were added to 5 ml HUVEC media to inactivate the trypsin and were centrifuged at 200 g for 5 minutes to pellet the cells. HUVEC media was created by mixing the following in the sterile hood: 500 ml DMEM-F12 with glutamine (11320033; Life Technologies (Gibco)), 45 mg heparin in 1 ml (Sigma; H3149), 5 ml Penicillin/streptomycin (Gibco; 15140-122), 5 ml L-glutamine (Gibco; 25030-024), 1 ml bovine brain extract (created in Bicknell/Heath laboratory at University of Birmingham; in 1 ml aliquots at -20°C) and 50 ml Fetal Bovine Serum (Gibco; 10270106). HUVEC media was then sterilised by filtering the solution through a filter bottle.

The media was removed by suction and 1 ml 85 % ice-cold ethanol (VWR; 20821.330) was added to the cells and vortexted to mix. This fixes the cells. Cells were stored at 4°C until propidium iodide staining was performed but were used within 2 weeks. After fixation cells were centrifuged again at 200 g for 5 minutes and
Figure 2.13: A cytogram plot from the program ‘FloJo’: The dots represent individual cells. A gate has been drawn around the main cell population.

Figure 2.14: A histogram plot from flow cytometry: The plot compares unstained cells (green) to cells stained for CD31 (blue), secondary antibody only (orange) and isotype control (red). N=1
alcohol was removed. 1 ml propidium iodide staining solution was added to the cells which were vortexed to mix. Propidium iodide solution was created by mixing the following: 1 ml PBS, 10 µl Propidium Iodide (Invitrogen; P3566), 10 µl RNAse A (Qiagen; 1007885) and 10 µl 10 % Triton-X-100 (Sigma; X100). The propidium iodide solution was vortexed to mix and kept away from light.

The cells were transferred to an eppendorf tube and covered in foil to prevent light from damaging the propidium iodide stain. The cells were incubated at 37°C, 5 % CO₂ for 20 minutes to ensure staining of cells. Cells were centrifuged again to remove excess staining solution at 3800 g for 5 minutes. The supernatant was discarded and cells were resuspended in 300 µl FACS buffer. Cells were run through the CyAn flow cytometer (Beckman-Coulter, United Kingdom) and the results were analysed using FloJo. Two cytogram plots (figure 2.15) were used to isolate the desired population of cells and to identify single cells only. A histogram of cell counts was used to identify the proportion of cells within each phase of the cell cycle (figure 2.16).
Figure 2.15: Cytogram plots from the program 'FloJo' used for cell cycle analysis. A: An initial gate was drawn to identify the main cell population. B: A second gate was drawn to identify single cells within the main cell population. The single cell population was used for cell cycle analysis.
Figure 2.16: A histogram plot of cell counts for cell cycle analysis. The plot demonstrates the group of cells in G1 (blue), S (mustard) and G2 (green). The percentages of cells in each phase are listed next to the plot.
2.12 Cell culture for cellular functional work

All functional work was carried out using Human Umbilical Vein Cells (HUVECs) which were purchased through Lonza (C2519A). These are pooled donor cells. On arrival HUVECs were washed by adding them to 10 ml HUVEC media (section 2.11.3) and centrifuged at 200 g for 5 minutes to pellet the cells. The cell pellets were then resuspended in 10 ml HUVEC media and plated on to 0.1% gelatin coated (section 2.8.1) T75 flasks. Cells were cultured at 37°C, 5% CO₂. Once confluent the media was removed from the cells, which were then washed with 5 ml PBS (section 2.8.1). 2 ml 2x trypsin (Gibco; 15400054) was added to the cells which were placed into the incubator at 37°C, 5% CO₂. Cell detachment was observed at intervals under the microscope. 10 ml HUVEC media was added to the cells to neutralise the trypsin before centrifuging the cells at 200 g for 5 minutes. The cell pellet was resuspended in 30 ml of HUVEC media and divided equally into 3 different 100 mm 0.1% gelatin coated plates. The new plates were returned to the incubator at 37°C, 5% CO₂ until cells were confluent. Cells were passaged in this way until passage 3 or 4 and used for functional work at this point. Surplus cells were stored and defrosted for use as per section 2.10.

2.13 miRNA Inhibition and Overexpression

Transfection of miRNA mimics and inhibitors can be used to determine the effects of overexpression and inhibiting a miRNA on a cell. MiRNA mimics are synthetic double-stranded miRNAs that mimic naturally occurring miRNAs after transfection. MiRNA inhibitors are modified single-stranded RNAs that bind to and inhibit miRNAs after transfection. (359) Both mimics and inhibitors can be transfected using
lipofectamine as a transfection reagent. Nucleic acids have a negative charge which prevents them from crossing cell membranes (which are also negatively charged) thereby requiring another agent to assist miRNA mimics and inhibitors to cross the cell membrane. Lipofectamine (a liposome formulation) acts by forming a complex with the RNAs, which allows them to overcome the negative charge on the cell membrane and enter the cell. Once a miRNA or inhibitor is in a cell then functional work can be performed to see the effects of over-/under-expression of that particular miRNA.

2.13.1 Transfection of a miRNA inhibitor or mimic into HUVEC in a 6 well plate

On day one 5 ml 0.1% gelatin (section 2.8.1) was placed into a well of a 6 well plate. This was transferred into an incubator at 37°C, 5% CO₂ for at least 15 minutes so that the gelatin formed a coating on the plate. 2.5x10⁵ HUVEC were plated into the well in 5 ml HUVEC media (section 2.11.3) which resulted in a near confluent HUVEC culture. Cultures should be near confluence as cells require contact in order to grow. However, 100% confluence can result in contact inhibition making cells less likely to take up RNAs reducing the success of transfection. On day two transfection was performed. To maintain sterile conditions transfection was performed in a tissue culture hood after decontamination of the hood with 70% ethanol. To reduce the chance of contamination with RNases the hood was then cleaned with RNaseZap (ThermoFisher Scientific; AM9780). All pipettes and tube holders were also cleaned with RNaseZap. All plastic ware used during the transfection was single use and RNase/DNase free.
Mimics and inhibitors come as dried powder that needs to be dissolved into RNase free water prior to use. All mimics and inhibitors used were Qiagen miScript miRNA mimics or inhibitors (5 nmol). Prior to opening all mimics and inhibitors were centrifuged for one minute to drive the contents of the tubes down into the wells. Mimics and inhibitors were dissolved in 250 µl RNase free water and vortexed to mix to create stock solutions of 20 µM. Mimics and inhibitors were used at a final concentration of 10 nM. Therefore, the 20 µM stock solution was diluted in the first instance by combining 10 µl of 20 µM stock with 190 µl optimem (ThermoFisher Scientific; 31985070). This was vortexed to mix completely resulting in a 1000 nM stock solution. 10 µl of 1000 nM stock was combined with 160 µl optimem to create a mimic/inhibitor mix. 3 µl lipofectamine (Lipofectamine® RNAiMAX Transfection Reagent, ThermoFisher Scientific; 13778075) was also combined with 27 µl optimem and both mixtures were left to incubate at room temperature for 10 minutes. The 30 µl lipofectamine mix was then combined with the mimic/inhibitor mix, flicked to mix and left at room temperature for a further 10 minutes. During this time the cells were removed from the incubator, media was removed and cells were washed twice with 2 ml PBS (section 2.8.1) to ensure all traces of media were removed. The lipofectamine/mimic/inhibitor mix was then combined with 800 µl optimem before being transferred on to the cells. The cells were left in the incubator at 37°C, 5% CO₂ for 4 hours to allow time for transfection to occur. After this time the mimic/inhibitor was removed and replaced with 5 ml HUVEC media. Cells were left alone on day 3 and by day 4 they were ready for assays to be performed. All mimic/inhibitor experiments were also performed using the following controls:
• Negative siRNA (AllStars Negative Control siRNA, Qiagen; 1027280): Negative control for miRNA mimic. Final concentration 10 nM. This is a double-stranded scrambled non-coding short RNA sequence.

• Negative miRNA inhibitor (miScript Inhibitor Negative Control, Qiagen; 1027271): Negative control for miRNA inhibitor. Final concentration 10 nM. This is a single-stranded RNA sequence complementary to the negative siRNA above.

• Lipofectamine only. The lipofectamine mix (30 µl) was combined with 170 µl of optimem rather than mimic/inhibitor mix.

• Optimem only: 1000 µl optimem was added to the cells rather than lipofectamine/mimic/inhibitor mix.

Some mimic or inhibitor experiments were also performed using plates other than 6 well plates. The number of cells used and volumes of inhibitor, mimic and lipofectamine are listed for other plate types in table 2.5.
Table 2.5: The number of HUVEC plated and volumes of inhibitor, mimic and lipofectamine used for different plates in various mimic or inhibitor experiments.

<table>
<thead>
<tr>
<th>Plate used</th>
<th>Number of HUVEC per well</th>
<th>Inhibitor/mimic mix per well</th>
<th>Lipofectamine mix per well</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-well</td>
<td>$8 \times 10^5$</td>
<td>6.8 µl 1000 nM stock 108.2 µl optimem</td>
<td>2 µl lipofectamine 18 µl optimem</td>
</tr>
<tr>
<td>10cm or T25 flask</td>
<td>$1 \times 10^6$</td>
<td>2 µl 20 µM stock 678 µl optimem</td>
<td>12 µl lipofectamine 108 µl optimem</td>
</tr>
</tbody>
</table>

This table outlines the plates used for the miRNA inhibitor and mimic functional experiments (chapter 6.) For each type of plate used the number of HUVEC plated is listed. The volumes of inhibitor/mimic, lipofectamine and optimem used are given for each plate type.
2.14 Cell growth assay

This assay was performed to determine the effect of miRNA overexpression on cell growth. MiRNA mimic and the appropriate controls were transfected into HUVECs as per section 2.13 in T25 flasks. 4 hours after transfection the plates were washed with PBS (section 2.8.1) and 2x trypsin (Gibco; 15400054) was used to remove cells as per section 2.12. 5 ml HUVEC media (section 2.11.3) was added to each flask to neutralise the trypsin and cells were transferred into 15 ml falcon tubes. Cells were centrifuged at 200 g for 5 minutes to remove the trypsin. The supernatant was removed and the cells were resuspended in 5 ml HUVEC media. The cells were counted by a haemocytometer and diluted to a concentration of 12 500 cells per ml using HUVEC media. 1 ml of cell solution (from each condition) was added to 3 wells of 3 0.1% gelatin (section 2.8.1) coated 12 well plates and placed into an incubator at 37°C, 5% CO₂. The following day cells from 1 plate were washed with PBS and removed using 2x trypsin (as above). The cells were resuspended in 100 µl HUVEC media and counted with a haemocytometer. This was repeated the next day for another plate and on the third day for the following plate. The number of cells from days 1-3 were compared for each condition to gain an estimate of cell growth during this time.

2.15 Matrigel tube formation assay

Matrigel is a basement membrane protein mixture that was originally isolated from mouse sarcoma cells. (362) When endothelial cells are plated on to matrigel they assemble to form tubes in a mesh-like manner. (363) This allows analysis of the
differentiation phase of angiogenesis where endothelial cells initially migrate from existing blood vessels. (364)

Initially 1 ml PBS (section 2.8.1) was added to a well of a 12 well plate and removed by suction to wet the well. 70 µl of matrigel (Corning; 356234) was then placed immediately into the well and allowed to spread throughout the well by rocking the plate back and forth gently. The matrigel was then allowed to solidify by placing the plate into the incubator at 37°C, 5% CO₂ for 30 minutes. During this time endothelial cells from a 6 well plate were washed with PBS and removed using 2x trypsin (Gibco; 15400054) (section 2.12). Cells were placed into 5 ml of HUVEC media (section 2.11.3) to inactivate the trypsin and pelleted by centrifuging at 200 g for 5 minutes. The supernatant was removed and cells were resuspended in 1 ml HUVEC media and counted using a haemocytometer. Cells were diluted to a concentration of 140 000 cells per ml using HUVEC media and 1 ml of cell solution was added to the matrigel well. The plate was uploaded into an IncuCyte incubator (Essen BioScience, United States) at 37°C, 5% CO₂. This incubator contains a microscope which was programmed to take pictures at 6, 12, 18 and 24 hours using IncuCyteZoom2015A software (Essen BioScience, United States). The images were downloaded and analysed using the online available software ImageJ.

2.16 Scratch wound assay

This assay measures endothelial cell migration by timing how long endothelial cells take to close an induced wound in an endothelial monolayer.(365)

48 hours prior to commencing the assays a 96 well ImageLock plate (Essen Biosciences; 4379) was gelatin coated by placing 100 µl 0.1% gelatin (section 2.8.1)
into each well of the plate prior to incubating for at least 15 minutes at 37°C, 5% CO₂. 8000 endothelial cells (in 100 µl HUVEC media (section 2.11.3)) were placed into each well of the plate. The cells were placed into the incubator at 37°C, 5% CO₂ for 48 hours.

A ‘Woundmaker 96’ (Essen Biosciences; 4493) was used to create the wounds in this experiment. This is effectively a tool consisting of a metal plate and 96 metal pins which fits to the ImageLock plate. Initially the Woundmaker was placed into 70% ethanol to sterilise the pins. The pins were dried in the sterile field of a tissue culture hood prior to being placed on the ImageLock plate. The Woundmaker was moved across the ImageLock plate creating a linear wound in each well. The media was removed from each well using a multi-channel pipette prior to washing each well with 100 µl PBS (section 2.8.1) in order to removed dislodged cells. Finally 100 µl of HUVEC media was added to each well and the ImageLock plate was inserted into an Incucyte incubator at 37°C, 5% CO₂. The incubator was programmed to take images at 0, 6, 12, 18 and 24 hours using IncuCyteZoom2015A software. The images were downloaded and analysed using the online available software ImageJ.
Figure 2.17: Overview of the spheroid assay: Endothelial cells are mixed with methocellulose, placed into droplets, inverted and are incubated overnight (1). Endothelial cells form spheroids in the droplet (2). The spheroids are resuspended in collagen where they form sprouts (3).
2.17 Spheroid assay

2.17.1 Preparation of the spheroid assay

An overview of this assay is presented in figure 2.17. This assay effectively analyses endothelial sprouting from a collection of endothelial cells known as a spheroid. Initially, endothelial cells are labelled with a fluorescent dye, CFSE (Carboxyfluorescein succinimidyl ester), in order to identify them easily on microscopy. Cells were washed with PBS (section 2.8.1) and removed from plates by treating with 2x trypsin (Gibco; 15400-054) (section 2.12). 5 ml HUVEC media (section 2.11.3) was added to the cells to neutralise the trypsin before pelleting the cells by centrifuging at 200 g for 5 minutes. The supernatant was removed and cells were resuspended in 200 µl of PBS. 200 µl CFSE-PBS solution was added to the cells, which were then incubated in a water bath at 37°C for 10 minutes. CFSE-PBS solution was created by mixing 1 ml PBS and 1 µl CFSE (Invitrogen; C34570) in a sterile hood. 5 ml HUVEC media was added to the cells, which were pelleted by centrifuging at 200 g for 5 minutes. The supernatant was removed and cells were washed again with 5 ml HUVEC media by centrifuging at 200 g for 5 minutes.

The cells were counted using a haemocytometer. 1.12 ml of cells in HUVEC media (at a concentration of 1.25 x 10⁴ cells/ml) was added to 280 µl methocellulose solution. Methocellulose solution is created by placing 6g methylcellulose (Sigma; M0512) into a 500 ml bottle. A magnetic stir bar is added to the bottle which is autoclaved. 250 ml M199 (see below) is added and pre-heated to 65°C. This mixture is stirred for 30 minutes before adding a further 250 ml M199. The mixture is stirred in a cold room overnight. The following day the mixture is transferred into 50 ml falcon tubes and centrifuged at 3500 g at 4°C for 3 hours. 45 ml of the supernatant is
transferred into fresh tubes – this is 1.2% methocellulose. The M199 used in the methocellulose solution is created by mixing one vial of M199 powder (Sigma; M5017) with 1 l of water. 2 g of sodium bicarbonate (Sigma; S5761) is added before sterilising the solution by filtering through a filter bottle in the hood.

20 µl of cells/methocellulose was added to each well of a 60 well microplate (Nunc; 439225). The lid of the plate was attached and the plate was inverted. This creates hanging droplets. Spheroids should form within each droplet. The plate was incubated overnight at 37°C, 5 % CO₂. Spheroids were harvested from the plate by aspirating each drop using a p1000 tip and were placed into a 15 ml falcon tube. The spheroids were centrifuged at 160 g for 5 minutes and the supernatant was aspirated.

Collagen mix 1 was created by mixing the following on ice in a sterile hood: 1.37 ml collagen (Type I Rat tail, Temecula; 92590), 250 µl 10x DMEM (see below) and 880 µl sterile water. The 10x DMEM solution used in collagen mix 1 was made by adding one vial of DMEM 10X powder (Sigma; D5030) to 1 l of water and filter sterilizing in the hood using a filter bottle.

Collagen mix 2 was created by mixing the following on ice in the hood: 1.5 ml EBM2 media (Lonza; CC-4542 + CC-5036) and 1 ml methocellulose solution.

10 µl sodium hydroxide (Sigma-Aldrich; S2770) was added to collagen mix 1 prior to mixing collagen mixes 1 and 2 in a ratio of 1:1. 200 µl of the combined collagen mix was used to resuspend the spheroids which were then placed on to a well of a 24 well plate. The plate was incubated at 37°C, 5 % CO₂ for 10 minutes to allow the collagen to solidify. 100 µl of EBM2 media (endothelial cell growth media, Lonza; CC-4542 + CC-5036) was added to the well and the plate was placed back into
incubator at 37°C, 5 % CO₂ for 8 hours. It is important to use media with VEGF (vascular endothelial growth factor) during this step to ensure sprouting. After this time the media was aspirated and 100 µl of 4 % paraformaldehyde was added to the well. The plate was covered in foil and placed into the incubator at 37°C, 5 % CO₂ for 30 minutes to fix the cells. Cells were washed with 100 µl PBS which was aspirated before storing the cells in 1000 µl PBS. Foil was placed back around the plate which was stored at 4°C until the plate was ready to be viewed. Spheroids were imaged using a Zeiss 780 Zen confocal microscope.

2.17.2 Analysis of spheroids using ImageJ

Fluorescent spheroid images were uploaded into ImageJ and were analysed using the ‘Spheroid Analysis’ plugin developed at the University of Birmingham by Victoria Salisbury. The plugin works by modifying the original fluorescent images in several steps (figure 2.18). Initially a binary image is produced where the endothelial objects appear as white on a black background (A). The binary image is used to create a mask of the endothelial sprouts (excluding the spheroid itself as this is not required for analysis) (B). This mask generates two images, one demonstrating sprouts connected to the spheroid (C) and the second demonstrating sprouts not connected representing cells which had begun to migrate away from the spheroid mass (D). The mask is further altered using a second plugin (Skeletonize), which converts sprouts into one pixel in diameter (E). Finally a third plugin (Analyze Skeleton) assesses the length and number of sprouts and produces these results in a table. A final image demonstrating the analysed network over the original image is also produced (F). This allows the user to determine how accurate the analysis has been.
Figure 2.18: Analysis of an endothelial spheroid using the ‘Spheroid Analysis’ plugin in ImageJ. A: Original image showing a fluorescent spheroid with sprouting which is uploaded into ImageJ. Images B-G are created by analysing this image by running the ‘Spheroid Analysis’ plugin. B: Binary image. C: Mask of sprouts. D: Connected sprouts. E: Disconnected sprouts. F: Skeleton image of sprouts. G: Final image demonstrating the analysed network overlying the original fluorescent spheroid. Scale bar represents 200μm.
CHAPTER THREE

MICRORNA AND MESSENGER RNA

EXPRESSION IN LUNG ENDOTHELIUM
3.1 Hypothesis

There will be significant differences in pulmonary endothelial microRNA (miRNA) and messenger RNA (mRNA) expression by microarray between COPD and non-COPD tissue. Some of these differences will be also be present in pulmonary endothelial lung cancer tissue.

3.2 Aims of this chapter

This chapter will detail the analyses performed in investigating endothelial miRNA and mRNA expression by microarray in COPD, non-COPD and lung cancer.

Initial array results and possible targets will be described prior to validation work by qPCR. An overall conclusion for the chapter will then be presented.

3.3 miRNA expression in the lung endothelium

3.3.1 Subjects

Baseline demographics of the entire cohort of patients are shown in table 3.1. There were more patients in the 2015 array group (13 vs 8 patients). Patient groups did not differ in terms of any of the demographics listed. There was no significant difference in tumour stage in the 2 groups (table 3.2).

3.3.2 Combined analysis of 2014 and 2015 microarrays

As there were no significant differences between microarray groups the groups were combined into one analysis comparing COPD and non-COPD patients (tables 3.3 and 3.4). There was no difference between groups in terms of sex, age, BMI, current smoking status or tumour stage. FEV1pp was significantly lower in the COPD group
(p=0.026), as expected. Pack year history was greater in the COPD group and approached significance (p=0.061).

**Limma analysis**

Samples were analysed using Limma to perform differential miRNA expression analysis as per section 2.5.2. Background signal was corrected for and signal intensity was normalised between arrays prior to analysis. The package LBE was used to determine the false discovery rate (FDR) values for the differential miRNA analysis (section 2.5.3). 4 miRNAs were significantly differentially expressed between the 2 groups (p<0.05) and had a log fold change (logFC) >1: miR-4495, miR-4462, miR-3923, miR-892c-5p. miR-3923 is known to be expressed in the endothelium, but the other miRNAs are not known to be endothelial expressed. The FDR was not significantly different for any miRNA between groups. A summary of the array result can be found in table 3.5.

**Significance Analysis of Microarrays (SAM)**

SAM is an alternative tool to Limma for the analysis of microarray data. In addition to presenting data in a tabular format it also presents data in a graphical form and results can be limited by fold change. As only 4 miRNAs were significantly differentially expressed between COPD and non-COPD groups (using limma) when 2014 and 2015 groups were combined a SAM analysis was performed (section 2.5.4) comparing all 2014 miRNA microarray and all 2015 miRNA microarray results. The plot from this analysis is in figure 3.1.
Table 3.1: Patient demographics comparing 2014 to 2015 miRNA microarrays.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>2014</th>
<th>2015</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>21</td>
<td>8 (38.1%)</td>
<td>13 (61.9%)</td>
<td>0.387</td>
</tr>
<tr>
<td>Male patients</td>
<td>10 (47.6%)</td>
<td>5 (62.5%)</td>
<td>5 (38.5%)</td>
<td>0.659</td>
</tr>
<tr>
<td>COPD</td>
<td>11 (52.4%)</td>
<td>5 (62.5%)</td>
<td>6 (46.2%)</td>
<td>0.860</td>
</tr>
<tr>
<td>Age</td>
<td>67.00 (14.50)</td>
<td>66.00 (18.75)</td>
<td>67.00 (14.00)</td>
<td>0.374</td>
</tr>
<tr>
<td>BMI</td>
<td>26.19 (8.72)</td>
<td>27.44 (6.93)</td>
<td>21.78 (7.06)</td>
<td>0.472</td>
</tr>
<tr>
<td>Pack year history</td>
<td>29.76 (25.37)</td>
<td>35.00 (17.73)</td>
<td>26.54 (29.33)</td>
<td>0.505</td>
</tr>
<tr>
<td>Current smoker</td>
<td>2 (9.5%)</td>
<td>0</td>
<td>2 (15.4%)</td>
<td>0.505</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>85.09 (25.73)</td>
<td>83.08 (26.42)</td>
<td>85.09 (36.00)</td>
<td>0.340</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the 2014 and 2015 microarray analyses. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher’s exact test was used to identify significant differences between groups.

Table 3.2: Number of patients with each tumour stage in the 2014 and 2015 miRNA microarrays.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>2014</th>
<th>2015</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>2 (9.5%)</td>
<td>0</td>
<td>2 (15.4%)</td>
<td>0.523</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IA</td>
<td>4 (19.0%)</td>
<td>1 (12.5%)</td>
<td>3 (23.1%)</td>
<td>0.251</td>
</tr>
<tr>
<td>IB</td>
<td>4 (19.0%)</td>
<td>1 (12.5%)</td>
<td>3 (23.1%)</td>
<td>0.251</td>
</tr>
<tr>
<td>IIA</td>
<td>2 (9.5%)</td>
<td>2 (25%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IIB</td>
<td>3 (14.3%)</td>
<td>1 (12.5%)</td>
<td>2 (15.4%)</td>
<td>0.251</td>
</tr>
<tr>
<td>IIIA</td>
<td>5 (23.8%)</td>
<td>3 (37.5%)</td>
<td>2 (15.4%)</td>
<td>0.251</td>
</tr>
<tr>
<td>Lung metastases</td>
<td>1 (4.8%)</td>
<td>0</td>
<td>1 (7.7%)</td>
<td>0</td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the miRNA microarray analyses. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by year. Fisher’s exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in 2014 and 2015.
Table 3.3: Patient demographics comparing COPD to non-COPD patients in the combined miRNA microarray analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Non-COPD</th>
<th>COPD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>21</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Male patients</td>
<td>10 (47.6%)</td>
<td>5 (50.0%)</td>
<td>5 (45.5%)</td>
<td>1.000</td>
</tr>
<tr>
<td>Age</td>
<td>64.10 (11.33)</td>
<td>63.50 (12.07)</td>
<td>64.64 (11.17)</td>
<td>0.825</td>
</tr>
<tr>
<td>BMI</td>
<td>25.77 (5.44)</td>
<td>27.48 (6.57)</td>
<td>24.20 (3.85)</td>
<td>0.174</td>
</tr>
<tr>
<td>Pack year history</td>
<td>30.00 (43.50)</td>
<td>6.50 (42.50)</td>
<td>30.00 (30.00)</td>
<td>0.061</td>
</tr>
<tr>
<td>Current smoker</td>
<td>2 (9.5%)</td>
<td>1 (10.0%)</td>
<td>1 (9.1%)</td>
<td>1.000</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>85.09 (25.73)</td>
<td>94.27 (10.77)</td>
<td>74.79 (23.41)</td>
<td>0.026</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the combined microarray analyses. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher’s exact test was used to identify significant differences between groups. Significant p values (<0.05) are presented in bold.

Table 3.4: Number of patients with each tumour stage in the combined miRNA microarray analysis.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>Non-COPD</th>
<th>COPD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>2 (9.5%)</td>
<td>0</td>
<td>2 (18.2%)</td>
<td>0.407</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>4 (19.0%)</td>
<td>2 (20.0%)</td>
<td>2 (18.2%)</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>4 (19.0%)</td>
<td>2 (20.0%)</td>
<td>2 (18.2%)</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>2 (9.5%)</td>
<td>0</td>
<td>2 (18.2%)</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>3 (14.3%)</td>
<td>1 (10%)</td>
<td>2 (18.2%)</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>5 (23.8%)</td>
<td>4 (40.0%)</td>
<td>1 (9.1%)</td>
<td></td>
</tr>
<tr>
<td>Lung metastases</td>
<td>1 (4.8%)</td>
<td>1 (10.0%)</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the miRNA microarray analyses. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by group (COPD vs non-COPD). Fisher’s exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in patients with and without COPD.
Table 3.5: Limma analysis results from the combined miRNA microarray analysis

<table>
<thead>
<tr>
<th>COPD samples</th>
<th>non-COPD samples</th>
<th>Number of downregulated miRNAs (p&lt;0.05)</th>
<th>Number of upregulated miRNAs (p&lt;0.05)</th>
<th>Number of downregulated miRNAs with logFC&gt;1</th>
<th>Number of upregulated miRNAs with logFC&gt;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>10</td>
<td>120</td>
<td>90</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

This table outlines the number of patient samples with and without COPD included in the combined limma miRNA microarray analysis and the number of miRNAs significantly downregulated/upregulated in the COPD group. A moderated t-statistic is used to determine significance between groups.

Figure 3.1: SAM plot for the analysis comparing 2014 and 2015 miRNA microarrays. This is a graphical representation of the statistics for each probe. Probes that are upregulated in group 2 (2015) are in red. Probes in green are downregulated in group 2. The d-statistic is used to determine significance between groups.
181 probes were significantly differentially expressed with a fold change $\geq 2$ between the 2014 and 2015 miRNA microarrays. 61 probes were upregulated in the 2015 microarrays and 120 probes were downregulated. Most of the probes that were differentially expressed were control probes. This result suggests that the 2014 and 2015 miRNA microarrays were inherently significantly different from one another. This can occur due to differences in the fluorescence between arrays due to different consumables being used. As the dyes used in microarray analysis are photosensitive they can degrade over time resulting in different levels of fluorescence. Therefore, the decision was made not to combine the 2014 and 2015 miRNA microarrays in one analysis and they were analysed using SAM separately.

**Combining SAM analyses from 2014 and 2015 results**

In an attempt to combine results from both miRNA microarray analyses the following analysis was performed. The list of miRNAs significantly differentially expressed (between COPD and non-COPD) with a fold change 2 from the 2014 arrays was compared to the 2015 array data. miRNAs that were not significant in the 2014 arrays were deleted from the 2015 array data. This was used to create a SAM input file (table 2.3) for the 2015 arrays where only the probes relating to significantly differentially expressed miRNAs in 2014 microarrays were included. This shortened 2015 array data was run through SAM limiting the results to fold change 2. 165 probes were significantly (fold change 2) upregulated in COPD in this analysis. No probes were significantly downregulated. The SAM plot for this analysis and the corresponding heatmap is in figure 3.2.

The opposite analysis was also performed: the list of miRNAs significantly differentially expressed (between COPD and non-COPD) with a fold change 2 from
the 2015 arrays was compared to the 2014 array data. miRNAs that were not significant in the 2015 arrays were deleted from the 2014 array data. This was used to create a SAM input file (table 2.3) for the 2014 arrays where only the probes relating to significantly differentially expressed miRNAs in 2015 microarrays were included. This shortened 2014 array data was run through SAM limiting the results to fold change 2. 1696 probes were significantly (fold change 2) upregulated in COPD in this analysis. No probes were significantly downregulated. The SAM plot for this and the corresponding heatmap is in figure 3.2.

The two final SAM analyses (2015 data limited to significantly differentially expressed 2014 miRNAs and 2014 data limited to significantly differentially expressed 2015 miRNAs) were directly compared to one another using excel. Probes that were significantly differentially expressed in both analyses were identified. These probes related to 43 upregulated miRNAs that are in table 3.6.

3.3.3 Computer prediction of targets of identified miRNAs

The 43 miRNAs identified in the SAM analysis were uploaded into two computer prediction programs for miRNA targeting: TargetScan (368) and DIANA-Micro-T-CDS (369). Two programs were used in order to minimise missing targets. The results of the predictions for each miRNA were compared to the list of endothelial mRNA targets identified from the mRNA microarrays performed in section 3.4.2. 37 miRNAs were possible modifiers of the endothelial mRNAs of interest. A literature search was performed and 8 miRNAs were chosen for further validation work as these have been previously expressed in endothelium.(370-377) The miRNAs chosen for validation work are in table 3.7.
Figure 3.2: SAM plots and corresponding heatmaps for the miRNA microarray analysis using SAM. A: 2015 microarray data limited to probes significantly differentially expressed in 2014. B: 2014 microarray data limited to probes significantly differentially expressed in 2015. These are graphical representations of the statistics for each probe. Probes that are upregulated are in red and those that are downregulated are in green. SAM uses the $d$ statistic to determine significance between groups.
Table 3.6: Significantly upregulated miRNAs in COPD

<table>
<thead>
<tr>
<th>Significantly upregulated miRNAs</th>
<th>Significantly upregulated miRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-342-5p</td>
<td>hsa-miR-5589-5p</td>
</tr>
<tr>
<td>hsa-miR-423-3p</td>
<td>hsa-miR-6082</td>
</tr>
<tr>
<td>hsa-miR-466</td>
<td>hsa-miR-6721-5p</td>
</tr>
<tr>
<td>hsa-miR-582-3p</td>
<td>hsa-miR-181b-3p</td>
</tr>
<tr>
<td>hsa-miR-624-3p</td>
<td>hsa-miR-18b-3p</td>
</tr>
<tr>
<td>hsa-miR-660-3p</td>
<td>hsa-miR-193b-5p</td>
</tr>
<tr>
<td>hsa-miR-676-3p</td>
<td>hsa-miR-20b-3p</td>
</tr>
<tr>
<td>hsa-miR-3119</td>
<td>hsa-miR-374a-3p</td>
</tr>
<tr>
<td>hsa-miR-3126-3p</td>
<td>hsa-miR-378a-5p</td>
</tr>
<tr>
<td>hsa-miR-3160-5p</td>
<td>hsa-miR-4666a-3p</td>
</tr>
<tr>
<td>hsa-miR-3200-5p</td>
<td>hsa-miR-548b-5p</td>
</tr>
<tr>
<td>hsa-miR-3677-5p</td>
<td>hsa-miR-429</td>
</tr>
<tr>
<td>hsa-miR-4278</td>
<td>hsa-miR-23c</td>
</tr>
<tr>
<td>hsa-miR-4301</td>
<td>hsa-miR-4477a</td>
</tr>
<tr>
<td>hsa-miR-4712-5p</td>
<td>hsa-miR-4703-5p</td>
</tr>
<tr>
<td>hsa-miR-4720-3p</td>
<td>hsa-miR-4727-5p</td>
</tr>
<tr>
<td>hsa-miR-4733-3p</td>
<td>hsa-miR-4797-5p</td>
</tr>
<tr>
<td>hsa-miR-4757-5p</td>
<td>hsa-miR-5010-5p</td>
</tr>
<tr>
<td>hsa-miR-4771</td>
<td>hsa-miR-548aa</td>
</tr>
<tr>
<td>hsa-miR-4782-5p</td>
<td>hsa-miR-599</td>
</tr>
<tr>
<td>hsa-miR-4799-5p</td>
<td>hsa-miR-892c-5p</td>
</tr>
<tr>
<td>hsa-miR-5191</td>
<td></td>
</tr>
</tbody>
</table>

This table demonstrates the 43 miRNAs that were significantly upregulated in COPD vs non-COPD in both the SAM 2015 and 2014 miRNA microarray analyses.
Table 3.7: miRNAs selected for validation work

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Previous evidence of endothelial expression</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-181b-3p</td>
<td>Expressed in HUVEC <em>in vitro</em>. Downregulates NK-κB signalling.</td>
<td>(370)</td>
</tr>
<tr>
<td>hsa-miR-342-5p</td>
<td>Associated with endothelial dysfunction in obese children (n=70)</td>
<td>(371)</td>
</tr>
<tr>
<td>hsa-miR-18b-3p</td>
<td>Expressed in endothelial progenitor cells <em>in vitro</em>.</td>
<td>(372)</td>
</tr>
<tr>
<td>hsa-miR-193b-5p</td>
<td>Altered expression in chronic hypoxia model in mouse aortic endothelial cells.</td>
<td>(373)</td>
</tr>
<tr>
<td>hsa-miR-374a-3p</td>
<td>Controls pre-B-cell colony-enhancing factor/NAMPT expression in HPEC in response to mechanical stress <em>in vitro</em>.</td>
<td>(374)</td>
</tr>
<tr>
<td>hsa-miR-378a-5p</td>
<td>Detected in human corneal endothelial cells.</td>
<td>(375)</td>
</tr>
<tr>
<td>hsa-miR-429</td>
<td>Increased levels in human aortic endothelial cells results in suppression of Bcl-2 and apoptosis.</td>
<td>(376)</td>
</tr>
<tr>
<td>hsa-miR-23c</td>
<td>Enriched in endothelial cells and promotes angiogenesis in vitro by targeting Sprouty2 and Sema6A proteins.</td>
<td>(377)</td>
</tr>
</tbody>
</table>

This table outlines the 8 miRNAs that were chosen for validation work. These miRNAs were significantly upregulated in both the 2014 and 2015 SAM miRNA microarray analyses. All 8 miRNAs were predicted to target endothelial mRNAs of interest identified in the mRNA microarray analysis (section 3.4.2). The 8 miRNAs were chosen for further validation as they have been previously expressed in endothelium. Evidence for endothelial expression for each target is listed in the table with the corresponding reference.
3.3.4 qPCR validation of miRNA targets

qPCR validation of all potential miRNA targets was performed prior to moving on to functional validation. This was performed as due to the large number of genes tested in microarray false positive results can occur.(378) Prior to performing qPCR for miRNA targets it was necessary to confirm that pulmonary endothelium was enriched in the samples extracted using ulex-coated magnetic beads. (To ensure that endothelial extraction had been successful). To do this qPCR was performed to compare CD31 expression in the endothelial isolates (n=14) to the bulk remainder tissue (n=6). Figure 3.3 illustrates the results of this qPCR experiment. The expression of CD31 was significantly increased in the endothelial isolate by over 2.5 fold (p=0.012). It is worth noting that endothelial isolation using the same method has previously shown a greater enrichment for CD31 than seen in this experiment (eg 15-fold in lung tumour). The reduced fold increase seen here is likely due to the fact that the lung is highly vascular and the proportion of endothelial cells in the lung is high (30%). (304)

After endothelial enrichment was confirmed qPCR was performed to validate the expression of the potential miRNA targets in table 3.7. 4 COPD and 4 non-COPD samples were tested for expression of each of the miRNAs listed in table 3.7. Figure 3.4 illustrates the results of the qPCR experiments for each miRNA. The expression of all miRNAs was increased in COPD. However, only miR-181b-3p, miR-429 and miR-23c were significantly increased in COPD.

Significance testing between groups was performed using t-tests to compare mean delta Ct in COPD and non-COPD. Delta Ct was chosen as it is the outcome value directly influenced by external effects such as concentration and sampling.(379)
Figure 3.3: qPCR validation of endothelial enrichment. qPCR was used to determine the expression of $CD31$ in endothelial isolates in comparison to bulk tissue isolates. Flotillin 2 was used as the house-keeping gene to which the data was normalised. Expression of $CD31$ in the endothelial isolate was normalised to that of the bulk tissue sample. The double delta Ct method was used to compare the expression levels. Fold change was calculated from the delta Ct levels for $CD31$ expression in endothelial isolates ($n=14$) in comparison to mean delta Ct for bulk tissue isolates ($n=6$). Figure represents mean fold change; error bars being SEM (standard error of the mean) fold change. A t-test was used to determine significance between mean delta Ct values. * denotes that the p value reached significance ($p<0.05$).
Figure 3.4: qPCR validation of potential miRNA targets. qPCR was used to determine the expression of each miRNA target in COPD in comparison to non-COPD. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miRNA targets in COPD was normalised to that of non-COPD. The double delta Ct method was used to compare the expression levels. Fold change was calculated from the delta Ct levels for miRNA expression in COPD (n=4) in comparison to mean delta Ct for non-COPD (n=4). Figures represent mean fold changes; error bars being SEM fold change. A t-test was used to determine significance between mean delta Ct values. * denotes that the p value reached significance (p<0.05).
3.3.5 Validation of potential miRNA targets using qPCR in lung cancer

In order to identify potential miRNA targets in both COPD and lung cancer further miRNA qPCR experiments were performed using RNA isolated from lung tumour pulmonary endothelial cells (collected from patients in section 2.1.1). The expression of each miRNA in lung cancer (n=6) was compared to expression in non-COPD (n=4). MiR-181b-3p and -23c were not expressed in the lung tumour samples. However, miR-429 was significantly increased in the lung tumour samples by 9-fold. Figure 3.5 illustrates the results of the qPCR experiment for miR-429.

3.4 mRNA expression in the lung endothelium

3.4.1 Subjects

Baseline demographics of the entire cohort of patients are shown in table 3.8. Patient groups did not differ in terms of any of the demographics listed. There was no significant difference in tumour stage in the 2 groups (table 3.9).

3.4.2 Combined analysis of 2014 and 2016 microarrays

As there were no significant differences between the 2014 and 2016 microarray groups the two groups were combined into one analysis comparing COPD and non-COPD patients (tables 3.10 and 3.11). There was no difference between groups in terms of sex, age, BMI, current smoking status or tumour stage. FEV1pp was significantly lower in the COPD group (p=0.035). Pack year history was greater in the COPD group and approached significance (p=0.097).
Figure 3.5: qPCR validation of miR-429 in lung cancer. qPCR was used to determine the expression of each miR-429 in lung cancer in comparison to non-COPD. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-429 in lung cancer was normalised to that of non-COPD. The double delta Ct method was used to compare the expression levels. Expression was calculated from the delta Ct levels for miRNA expression in lung cancer (n=6) in comparison to mean delta Ct for non-COPD (n=4). Figures represent mean fold changes; error bars being SEM fold change. A t-test was used to determine significance between mean delta Ct values. * denotes that the p value reached significance (p<0.05).
Table 3.8: Patient demographics comparing 2014 to 2016 mRNA microarrays.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>2014</th>
<th>2016</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>14</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Male patients</td>
<td>6 (42.9%)</td>
<td>2 (28.6%)</td>
<td>4 (57.1%)</td>
<td>0.592</td>
</tr>
<tr>
<td>Age</td>
<td>64.14 (13.18)</td>
<td>65.00 (23.00)</td>
<td>68.00 (35.00)</td>
<td>1.000</td>
</tr>
<tr>
<td>BMI</td>
<td>22.37 (4.82)</td>
<td>26.53 (4.65)</td>
<td>24.20 (5.06)</td>
<td>0.389</td>
</tr>
<tr>
<td>Pack year history</td>
<td>29.29 (23.28)</td>
<td>40.00 (20.00)</td>
<td>19.00 (18.00)</td>
<td>0.259</td>
</tr>
<tr>
<td>Current smoker</td>
<td>3 (21.4%)</td>
<td>2 (28.6)</td>
<td>1 (14.3%)</td>
<td>1.000</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>85.38 (23.97)</td>
<td>90.38 (11.14)</td>
<td>80.39 (48.78)</td>
<td>0.458</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the 2014 and 2016 microarray analyses. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher's exact test was used to identify significant differences between groups.

Table 3.9: Number of patients in each group with each tumour stage.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>2014</th>
<th>2016</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.057</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIIB</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IIIA</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Lung metastases</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the mRNA microarray analyses. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by year. Fisher's exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in 2014 and 2016.
Table 3.10: Patient demographics comparing COPD to non-COPD patients in the combined mRNA microarray analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Non-COPD</th>
<th>COPD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>14</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Male patients</td>
<td>6 (42.9%)</td>
<td>3 (50%)</td>
<td>3 (37.5%)</td>
<td>1.000</td>
</tr>
<tr>
<td>Age</td>
<td>64.14 (13.18)</td>
<td>67.00 (15.99)</td>
<td>62.00 (11.31)</td>
<td>0.505</td>
</tr>
<tr>
<td>BMI</td>
<td>22.37 (4.82)</td>
<td>27.26 (5.39)</td>
<td>23.94 (4.13)</td>
<td>0.215</td>
</tr>
<tr>
<td>Pack year history</td>
<td>29.29 (23.28)</td>
<td>17.33 (16.29)</td>
<td>38.25 (24.55)</td>
<td>0.097</td>
</tr>
<tr>
<td>Current smoker</td>
<td>3 (21.4%)</td>
<td>1 (16.67%)</td>
<td>2 (25%)</td>
<td>1.000</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>85.38 (23.97)</td>
<td>100.49 (16.74)</td>
<td>74.05 (22.90)</td>
<td>0.035</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the combined microarray analyses. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher’s exact test was used to identify significant differences between groups. Significant p values (<0.05) are presented in bold.

Table 3.11: Number of patients in each group with each tumour stage.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>Non-COPD</th>
<th>COPD</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0.904</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IIIA</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Lung metastases</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the mRNA microarray analyses. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by group (COPD vs non-COPD). Fisher’s exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in patients with and without COPD.
**Limma analysis**

Samples were analysed using Limma to perform differential mRNA expression analysis as per section 2.5.2. 32 mRNAs were significantly differentially expressed between the 2 groups (p<0.05) and had a log fold change (logFC) >1: 20 downregulated and 12 upregulated. None of the mRNAs are known to be endothelial expressed. The FDR was not significantly different for any mRNA between groups. A summary of the array result can be found in table 3.12.

**Significance Analysis of Microarrays (SAM)**

In a similar way to the miRNA data a SAM analysis was performed (section 2.5.4) comparing all 2014 mRNA microarray and all 2016 mRNA microarray results. The plot from this analysis is in figure 3.6.

39154 probes were significantly differentially expressed with a fold change ≥ 2 between the 2014 and 2016 mRNA microarrays. 12695 probes were upregulated in the 2016 microarrays and 26459 probes were downregulated. This represents the vast majority of probes and included many control probes. This result suggests that the 2014 and 2016 mRNA microarrays were inherently significantly different from one another. Therefore, again, the decision was made not to combine the 2014 and 2016 mRNA microarrays in one analysis and they were analysed using SAM separately.
Table 3.12: Limma analysis results from the combined mRNA microarray analysis.

<table>
<thead>
<tr>
<th>COPD samples</th>
<th>non-COPD samples</th>
<th>Number of downregulated mRNAs (p&lt;0.05)</th>
<th>Number of upregulated mRNAs (p&lt;0.05)</th>
<th>Number of downregulated mRNAs with logFC&gt;1</th>
<th>Number of upregulated mRNAs with logFC&gt;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6</td>
<td>138</td>
<td>89</td>
<td>20</td>
<td>12</td>
</tr>
</tbody>
</table>

This table outlines the number of patient samples with and without COPD included in the combined limma mRNA microarray analysis and the number of mRNAs significantly downregulated/upregulated in the COPD group. A moderated t-statistic is used to determine significance between groups.

Figure 3.6: SAM plot for the analysis comparing 2014 and 2016 mRNA microarrays. This is a graphical representation of the statistics for each probe. Probes that are upregulated in group 2 (2016) are in red. Probes in green are downregulated in group 2. The d-statistic is used to determine significance between groups.
Combining SAM analyses from 2014 and 2016 results

In an attempt to combine results from both mRNA microarray analyses the following analysis was performed. The list of mRNAs significantly differentially expressed (between COPD and non-COPD) with a fold change 2 from the 2014 arrays was compared to the 2016 array data. mRNAs that were not significant in the 2014 arrays were deleted from the 2016 array data. This was used to create a SAM input file (table 2.3) for the 2016 arrays where only the probes relating to significantly differentially expressed mRNAs in 2014 microarrays were included. This shortened 2016 array data was run through SAM limiting the results to fold change 2. 1355 probes were significantly (fold change 2) upregulated in COPD in this analysis. 37799 probes were significantly downregulated. The SAM plot for this analysis and the corresponding heatmap are in figure 3.7.

The opposite analysis was also performed: the list of mRNAs significantly differentially expressed (between COPD and non-COPD) with a fold change 2 from the 2016 arrays was compared to the 2014 array data. mRNAs that were not significant in the 2016 arrays were deleted from the 2014 array data. This was used to create a SAM input file (table 2.3) for the 2014 arrays where only the probes relating to significantly differentially expressed mRNAs in 2016 microarrays were included. This shortened 2014 array data was run through SAM limiting the results to fold change 2. 2113 probes were significantly (fold change 2) upregulated in COPD in this analysis. No probes were significantly downregulated. The SAM plot for this analysis and the corresponding heatmap are in figure 3.7.

The two final SAM analyses (2016 data limited to significantly differentially expressed 2014 mRNAs and 2014 data limited to significantly differentially expressed 2016
mRNAs) were directly compared to one another using excel. Probes that were significantly differentially expressed in both analyses were identified. 2071 genes (including duplicates) were significantly positively upregulated in COPD in both analyses. This list was compared to a known list of endothelial expressed genes. [Personal communication with R. Bicknell.] 6 known endothelial genes were upregulated in COPD:

- **DLC1** (Rho GTPase activating protein).
- **HHIP** (Hedgehog interacting protein).
- **LTA4H** (Leukotriene A4 Hydrolase).
- **PPIL2** (Peptidlyprolyl isomerase 2).
- **TMEM154** (Transmembrane protein 154).
- **TP53** (Tumour suppressor 53).

These mRNAs were chosen for further validation work.
Figure 3.7: SAM plots and corresponding heatmaps for the mRNA microarray analysis using SAM. A: 2016 microarray data limited to probes significantly differentially expressed in 2014. B: 2014 microarray data limited to probes significantly differentially expressed in 2016. These are graphical representations of the statistics for each probe. Probes that are upregulated are in red and those that are downregulated are in green. SAM uses the d statistic to determine significance between groups.
3.4.3 **Tumour mRNA analysis**

In order to investigate whether there were any shared targets between COPD and lung cancer a further mRNA microarray was performed using three tumour endothelial samples. As the 2014 and 2016 non-COPD lung tissue endothelial arrays were inherently different from one another the decision was made to compare the three tumour endothelial samples to the non-COPD 2014 microarray data and the 2016 microarray data in separate analyses. A similar univariate analysis was performed (as in section 3.4.2) comparing the demographics of patients from the tumour microarray data to the non-COPD patients from the 2014 and 2016 microarray data (tables 3.13-3.16).

There were no significant differences in baseline demographics between non-COPD patients in either microarray analysis and patients in the tumour microarray analysis. There was also no significant difference in tumour stage between non-COPD patients and tumour patients.

**SAM analysis: 2014 non-COPD versus tumour**

A SAM input file (table 2.3) was created for the 2014 non-COPD and tumour data. This was run through SAM limiting results to fold change 2. 7261 probes were significantly downregulated in the tumour group. No probes were upregulated. The SAM plot for this analysis and the corresponding heatmap are in figure 3.8.
Table 3.13: Patient demographics comparing 2014 non-COPD patients to tumour tissue patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Non-COPD 2014</th>
<th>Tumour</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Male patients</td>
<td>1 (16.7%)</td>
<td>1 (33.3%)</td>
<td>0</td>
<td>1.000</td>
</tr>
<tr>
<td>Age</td>
<td>66.67 (12.34)</td>
<td>72.67 (14.57)</td>
<td>60.67 (7.77)</td>
<td>0.277</td>
</tr>
<tr>
<td>BMI</td>
<td>25.58 (4.69)</td>
<td>28.22 (4.13)</td>
<td>22.94 (4.12)</td>
<td>0.192</td>
</tr>
<tr>
<td>Pack year history</td>
<td>10.00 (48.75)</td>
<td>20.00</td>
<td>0</td>
<td>1.000</td>
</tr>
<tr>
<td>Current smoker</td>
<td>1 (16.7%)</td>
<td>0</td>
<td>1 (33.3%)</td>
<td>1.000</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>103.09 (33.59)</td>
<td>95.62 (13.12)</td>
<td>87.56 (33.81)</td>
<td>0.865</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the mRNA microarray analysis: non-COPD 2014 vs lung tumour. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher’s exact test was used to identify significant differences between groups.

Table 3.14: Number of patients in each group (non-COPD 2014 and tumour) with each tumour stage.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>Non-COPD 2014</th>
<th>Tumour</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.400</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIIA</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Lung metastases</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the mRNA microarray analysis: non-COPD 2014 vs lung tumour. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by group (non-COPD vs tumour). Fisher’s exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in each group.
Table 3.15: Patient demographics comparing 2016 non-COPD patients to tumour tissue patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Non-COPD 2016</th>
<th>Tumour</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Male patients</td>
<td>2</td>
<td>2 (66.7%)</td>
<td>0</td>
<td>0.400</td>
</tr>
<tr>
<td>Age</td>
<td>61.00 (12.51)</td>
<td>61.33 (18.18)</td>
<td>60.67 (7.77)</td>
<td>0.956</td>
</tr>
<tr>
<td>BMI</td>
<td>24.62 (5.59)</td>
<td>26.30 (7.26)</td>
<td>22.94 (4.12)</td>
<td>0.524</td>
</tr>
<tr>
<td>Pack year history</td>
<td>6.50 (42.00)</td>
<td>13.00</td>
<td>0</td>
<td>1.000</td>
</tr>
<tr>
<td>Current smoker</td>
<td>2 (33.3%)</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td>1.000</td>
</tr>
<tr>
<td>FEV1pp</td>
<td>96.45 (27.11)</td>
<td>105.36 (21.39)</td>
<td>87.56 (33.81)</td>
<td>0.484</td>
</tr>
</tbody>
</table>

A comparison of patient demographics for all patients included in the mRNA microarray analysis: non-COPD 2016 vs lung tumour. Continuous variables that are normally distributed are presented in standard font as mean (standard deviation). T-tests were used to identify significant differences between normally distributed groups. Continuous variables that are not normally distributed are presented in italics as median (interquartile range). The Mann-Whitney U test was used to identify differences between non-normally distributed groups. Categorical variables are presented with percentages and fisher’s exact test was used to identify significant differences between groups. Significant p values (<0.05) are presented in bold.

Table 3.16: Number of patients in each group (non-COPD 2016 and tumour) with each tumour stage.

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>All</th>
<th>Non-COPD 2016</th>
<th>Tumour</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cancer</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.400</td>
</tr>
<tr>
<td>0 (in situ)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IIIA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Lung metastases</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the total number of patients with each stage of cancer in the mRNA microarray analysis: non-COPD 2016 vs lung tumour. Percentages of patients with each tumour stage are listed in brackets. The number (and percentage) of patients with each stage is also listed by group (non-COPD vs tumour). Fisher’s exact test was used to determine if there was a significant difference between the proportion of patients with each stage of cancer in each group.
SAM analysis: 2016 non-COPD versus tumour

A SAM input file (table 2.3) was created for the 2016 non-COPD and tumour data. This was run through SAM limiting results to fold change 2. 25939 probes were significantly differentially expressed between the two groups (12646 were upregulated and 13293 were downregulated in the tumour data). The SAM plot for this analysis and the corresponding heatmap are in figure 3.8.

Comparison of tumour microarray analyses to COPD microarray analysis

To identify joint COPD and tumour targets the results of the tumour and COPD analyses were compared. Table 3.17 demonstrates whether the COPD targets from section 3.4.2 were also differentially expressed in lung tumour. None of the results for the COPD targets from the analysis: non-COPD 2014 vs lung tumour were consistent with results from the COPD vs non-COPD mRNA microarrays. However, in the non-COPD 2016 vs lung tumour analysis, DLC1, PPIL2 and TP53 were upregulated in lung tumour thus potentially being shared targets.

3.4.4 Validation of potential mRNA targets using qPCR

In order to validate the expression of the potential mRNA targets in table 3.17 qPCR (section 2.6.2) was performed using RNA isolated from endothelial samples collected from patients in section 2.1.1. qPCR was completed for the six potential mRNA targets comparing target expression in COPD (n=7) and non-COPD (n=7). The results of this analysis are shown in figure 3.9. The expression of LTA4H was increased by over two-fold in COPD approaching statistical significance (p=0.067). The expression of the other genes did not vary between groups. PPIL2 was not expressed in either group and therefore the data for this gene is not presented.
Figure 3.8: SAM plots and corresponding heatmaps for the mRNA lung tumour microarray analysis using SAM. A: non-COPD 2014 microarray data vs lung tumour. B: non-COPD 2016 microarray data vs lung tumour. These are graphical representations of the statistics for each probe. Probes that are upregulated are in red and those that are downregulated are in green. SAM uses the d statistic to determine significance between groups.
Table 3.17: A comparison of the results of the COPD mRNA microarray analysis and the two tumour mRNA microarray analyses for the 6 possible targets identified in the COPD analysis (section 3.4.2).

<table>
<thead>
<tr>
<th>COPD target gene</th>
<th>Direction in COPD (versus non-COPD, combined analysis)</th>
<th>Direction in Tumour (versus 2014 non-COPD)</th>
<th>Direction in Tumour (versus 2016 non-COPD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLC1</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
</tr>
<tr>
<td>HHIP</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>LTA4H</td>
<td>Up</td>
<td>Not significant</td>
<td>Not significant</td>
</tr>
<tr>
<td>PPIL2</td>
<td>Up</td>
<td>Not significant</td>
<td>Up</td>
</tr>
<tr>
<td>TMEM154</td>
<td>Up</td>
<td>Not significant</td>
<td>Down</td>
</tr>
<tr>
<td>TP53</td>
<td>Up</td>
<td>Not significant</td>
<td>Up</td>
</tr>
</tbody>
</table>

This table lists the 6 targets that were identified in section 3.4.2 as being significantly upregulated in COPD. The third column lists whether or not these targets were significantly altered in the non-COPD 2014 vs lung tumour mRNA microarray analysis. The fourth column lists whether or not these targets were significantly altered in the non-COPD 2014 vs lung tumour mRNA microarray analysis. The d statistic was used to determine significance between groups.
Figure 3.9: qPCR validation of potential mRNA targets. qPCR was used to determine the expression of each mRNA target in COPD in comparison to non-COPD. Flotilin 2 was used as the house-keeping gene to which the data was normalised. Expression of mRNA targets in COPD was normalised to that of non-COPD. The double delta Ct method was used to compare the expression levels. Fold change was calculated from the delta CT levels for mRNA expression in COPD (n=7) in comparison to mean delta Ct for non-COPD (n=7). Figures represent mean fold changes; error bars being SEM fold change. A t-test was used to determine significance between mean delta Ct values.
3.5 Discussion

3.5.1 miRNA expression in pulmonary endothelium

This study has identified eight miRNAs which are upregulated in COPD endothelial cells: miR-181b-3p, -342-5p, -18b-3p, 193b-5p, 374a-5p, 378a-5p, -429 and -23c. However, only miR-181b-3p, -429 and -23c reached statistical significance. These miRNAs were not identified in the previous COPD microarray studies, but this is not unexpected as results from this study are endothelial cell specific whereas other studies looked at whole lung or other cell types.

There is very little evidence about the function of miR-181b-3p. However, recently this miRNA has been shown to regulate epithelial-mesenchymal transition (EMT) in breast cancer cells *in vitro*. miR-181b-3p appears to upregulate the expression of Snail, a protein of importance in inducing EMT. miR-181b-3p inhibitors reduce migration and mesenchymal markers in highly invasive breast cancer cells *in vitro*. Also, transfection of miR-181b-3p into breast cancer cells increases metastatic nodule formation in the lungs of mice. Therefore, it is possible that the upregulation of miR-181b-3p in the COPD lung may be a reason for the increased malignancy seen in this patient group. Further support for this hypothesis comes from the evidence that miR-181b-5p (a related miRNA from the same precursor as miR-181b-3p), is upregulated in lung squamous cell carcinoma and adenocarcinoma. miR-181b-5p targets and suppresses Ras association domain family member 1 (*RASSF1*), a known tumour suppressor. Another miR-181b-5p target of importance is programmed cell death 4 (*PCD4*), a DNA-binding protein tumour suppressor of importance in several cancer types including lung cancer. miR-181b promotes cell proliferation and migration and
suppresses apoptosis \textit{in vitro} in colorectal cancer (CRC) cells, potentially through the downregulation of \textit{PCD4}.\textsuperscript{(383)} There is evidence that miR-181b-5p may be important in EMT in the same way as miR-181b-3p: overexpression of miR-181b-5p in gastric cancer cells \textit{in vitro} appears to induce EMT.\textsuperscript{(385)} miR-181b also increases nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) activity by suppressing CYLD lysine deubiquitinase (\textit{CYLD}).\textsuperscript{(386, 387)} NF-κB is a ubiquitous transcription factor that plays an important role in the regulation of immune responses and inflammation;\textsuperscript{(323)} aberrant activation of NF-κB can also result in oncogenesis.\textsuperscript{(324)} By regulating \textit{CYLD}, miR-181b seems to be important in cell proliferation and resistance to apoptosis.\textsuperscript{(386)} Even transient expression of miR-181b-1 (stem loop precursor of miR-181b-3p and -5p) appears to be enough to upregulate NF-κB activity and transform cells to a malignant phenotype.\textsuperscript{(387)} The miR-181b- NF-κB pathway may also be of importance in pulmonary inflammation. For example, lipopolysaccharide (LPS) stimulation of human bronchial epithelial cells results in the upregulation of miR-181b.\textsuperscript{(388)} By overexpressing miR-181b in bronchial epithelial cells, p65 (a primary component of NF-κB) was increased in addition to interleukin (IL)-6 levels.\textsuperscript{(388)} This suggests that miR-181b acts as a pro-inflammatory factor by the upregulation of the NF-κB pathway \textit{in vitro}. There is other evidence that miR-181b-5p is important in inflammation. For example, miR-181b is upregulated in sepsis and its' suppression improved both bacterial clearance and mortality in mice models of sepsis.\textsuperscript{(389)} miR-181b is also increased in gingival tissue in periodontitis, an inflammatory dental condition associated with COPD.\textsuperscript{(390, 391)}

There is recent evidence that miR-181b may be involved in angiogenesis. Xu et al discovered that miR-181b is upregulated in hypoxic conditions in retinoblastoma
This team then overexpressed miR-181b in retinoblastoma cells before transferring the retinoblastoma culture medium to Human Umbilical Vein Endothelial Cells (HUVECs). Addition of the retinoblastoma medium increased capillary tube formation in the HUVECs. In contrast, when the same experiment was performed with suppressing miR-181b in retinoblastoma cells, capillary tube formation of HUVECs was reduced. This suggests that miR-181b enhanced the angiogenesis capability of HUVECs. Therefore, it is possible that upregulation of miR-181b seen in COPD endothelial cells in this study might lead to an increased capability of these cells to undergo angiogenesis which again, might contribute to the increased risk of lung cancer seen in such patients. In summary, there is evidence that miR-181b-3p may have a role in EMT and could potentially increase lung cancer risk in COPD patients. There is also evidence that miR-181b-3p’s precursor (miR-181b) and its’ related miRNA miR-181b-5p might increase cancer risk by the suppression of various tumour suppressors and by upregulating the NF-κB pathway. miR-181b also appears to be important in the upregulation of inflammation and therefore could potentially contribute to increased inflammation seen in COPD patients. Finally, miR-181b overexpression in endothelial cells appears to increase angiogenesis, an important mechanism involved in tumour progression (section 1.4.2).

However, the evidence for miR-181b’s involvement in malignancy, inflammation and angiogenesis is inconsistent. For example, Yang et al found reduced miR-181b expression in NSCLC tissue, which was associated with increased tumour size, higher cancer stage and reduced disease-free survival. miR-181b has also been shown to be downregulated in glioma and gastric cancer cell lines.
Sun et al have shown conflicting results to the above studies related to the NF-κB pathway. (370) In their study miR-181b appeared to suppress the NF-κB pathway and was suppressed (in aortic tissue) in response to delivery of proinflammatory stimuli in mice. (370) Furthermore, patients with sepsis had reduced circulating levels of miR-181b. (370) Similarly, miR-181b appears to be reduced in inflammatory plaques associated with atherosclerosis in mice and treatment with miR-181b microparticles reduces endothelial inflammation and atherosclerosis in mouse models. (396) Finally, miR-181b is also reduced in chick embryos in response to arsenic and its’ reduction seems to facilitate arsenic-associated angiogenesis. (397) Therefore, there is conflicting evidence about the role of miR-181b in tumorigenesis, inflammation and angiogenesis. Also, despite evidence linking miR-181b-3p to cancer this miRNA was not upregulated in lung cancer endothelial cells in this study. Most studies performed so far have only looked at miR-181b-3p’s precursor or related miRNAs rather than miR-181b-3p itself. Also, no studies of this miRNA have looked at its’ expression in COPD or pulmonary endothelial cells. Therefore, further functional validation work of miR-181b-3p is warranted.

MiR-429 has been associated with various types of cancers. However, in a similar way to miR-181b, miR-429 has been both up- and down-regulated in cancer. For example, it is increased in colorectal and ovarian cancers, but reduced in renal cancer. (398-400) However, miR-429 does appear to be more consistently increased in lung cancer. For example, Lang et al demonstrated that miR-429 was increased in lung cancer tissue and lung cancer cell lines. (401) Overexpression of miR-429 also resulted in increased cell proliferation, migration and invasion in A549 lung cancer cells. (401) It is possible that miR-429 may cause this change in cell behaviour by
targeting and suppressing known tumour-suppressors: phosphatase and tensin homolog (PTEN), a phosphatase involved in apoptosis, Ras association domain family member 8 (RASSF8), a protein that maintains adherent junction function and TIMP metalloproteinase inhibitor (TIMP2) which suppresses metastasis. (401) Interestingly, miR-429 has also been shown to target DLC1 (a gene identified in the mRNA microarray part of this study). MiR-429 suppresses DLC1 which increases cell proliferation in vitro. (402) Another possible target of importance in the increase cell proliferation seen with upregulation of miR-429 is p27kip1. (403) P27kip1 is a cyclin-dependent kinase inhibitor and is involved in the progression of the cell cycle. (403) In addition to miR-429 being upregulated in lung tumour versus normal lung miR-429 has also been shown to be upregulated in serum from patients with lung cancer compared to serum from patients with COPD. (404) Perhaps, therefore, it is possible that miR-429 is upregulated in COPD initially and further increases in expression of this miRNA contribute to the development of lung cancer in these patients. This is supported by results from the qPCR experiments in this study. MiR-429 was increased by 2-fold in COPD pulmonary endothelial cells, but 9-fold in lung cancer pulmonary endothelial cells (in comparison to non-COPD controls). Genetic variants in miR-200 family, of which miR-429 is a part, have also been shown to be related to lung cancer. The SNP in rs9660710 in the miR-200b/200a/429 cluster is significantly associated with lung cancer risk and is located in the regulatory elements in lung cancer cells. (405)

One mechanism of miR-429 upregulation that may be of particular importance in COPD is hypoxia. MiR-429 has been shown to be upregulated in hypoxic conditions in cardiomyocytes in vitro. (406) Downregulation of miR-429 in the same cell culture
also protected cells against apoptosis through the activation of notch1 (an intercellular signalling protein).(406) If miR-429 exerts similar effects on lung endothelial cells, it is possible that upregulation of miR-429 might result in endothelial apoptosis and consequent emphysema in COPD patients. Mir-429 is also upregulated in response to hypoxia in endothelial cells.(407) Bartoszewska et al demonstrated that miR-429 is upregulated by the hypoxia-inducible factor (HIF-)α in human umbilical vein endothelial cells.(407) Interestingly, miR-429 also downregulated HIF-α and therefore may act as a negative regulator of HIF-α during the cell response to hypoxia.(407) Another mechanism that may induce apoptosis in endothelial cells through upregulation of miR-429 is targeting Bcl-2, an anti-apoptotic protein.(376) In vivo studies using mouse models have demonstrated that miR-429 is upregulated in aortic endothelial cells of mice with atherosclerosis and that this is associated with endothelial apoptosis.(376) Again, if miR-429 has a similar function in pulmonary endothelial cells it is possible that this could result in pulmonary endothelial apoptosis and emphysema. Patients without emphysema are unlikely to be hypoxic and consequently this mechanism is most likely to be involved in the worsening of emphysema rather than emphysema initiation.

MiR-429 may also play an important role in pulmonary inflammation. Studies in rat pulmonary macrophages have shown that miR-429 is upregulated in response to lipopolysaccharide (LPS) stimulation.(408) The miR-429 upregulation results in the production of pro-inflammatory cytokines through the targeting of DUSP1 (dual-specificity phosphatase 1).(408) DUSP1 inactivated p38 MAPK (mitogen-activated protein kinase), an important intracellular signalling molecule which leads to the production of pro-inflammatory cytokines such as TNF-α (tumour necrosis factor α)
and IL-1β (interleukin-1β).(408) Thus, miR-429 effectively results in an increase in p38 MAPK activity and pro-inflammatory cytokine production. This may therefore have an important impact in COPD in which the level of pro-inflammatory cytokines are high.

In conclusion, miR-429 appears to have a role in pulmonary inflammation and endothelial apoptosis which are two important mechanisms in COPD pathogenesis. Also, miR-429 seems to be upregulated in lung cancer and targets known tumour suppressors. As miR-429 is increased in both COPD and lung cancer in this study it provides a potential common target for both diseases. Therefore, further functional validation of miR-429 in the pulmonary endothelium is required to establish its’ importance in these diseases further.

The available information on miR-23c is also limited. So far the only study in humans that looked at the expression of miR-23c was performed in prostate cancer cells. This showed an upregulation of miR-23c in prostate cancer chemoresistant cells in comparison to cells that were chemosensitive.(409) MiR-23c also appears to target the tumour suppressor PTEN in a similar way to miR-429 and may therefore play a role in tumourigenesis.(409) In support of this miR-23 may also increase cell proliferation; downregulation of miR-23 in vitro reduced the growth of lung cancer cells.(410) Other potential mechanisms of miR-23c action may come from studies looking at related miRNAs. For example, studies have demonstrated that miR-23 is upregulated in endothelial progenitor cells (EPCs) in patients with coronary artery disease compared to controls.(411) MiR-23 also targets VEGF – inhibition of miR-23 led to increased VEGF levels and angiogenic activities of EPCs in vitro.(411) Further support for the role of miR-23 in angiogenesis comes from studies of glioblastoma.
Microvascular proliferation is important in the pathogenesis of glioblastoma; miR-23 is downregulated in glioblastoma endothelial cells which results in the upregulation of two of its targets: \textit{ATP5A1} and \textit{ATP5B}, both ATP synthase proteins involved in cell metabolism.\textsuperscript{(412)} Both studies suggest that miR-23 inhibits angiogenesis through more than one mechanism. Upregulation of miR-23c could be a mechanism for the lack of angiogenesis seen in some COPD patients with emphysema, though evidence for this is conflicting. For example, one study which demonstrated that miR-23 was enriched in HUVECs also showed that the anti-angiogenic proteins Sprouty2 and Sema6A were targeted by miR-23 suggesting that miR-23 may be involved in the upregulation of angiogenesis instead.\textsuperscript{(377)}

MiR-23c may be increased in COPD as miR-23 is upregulated in response to hypoxia. MiR-23 also appears to target the apoptotic proteins caspase 7, \textit{BID} (BH\textsubscript{3} interacting-domain death agonist) and \textit{NIX-BNIP3L} (BCL2 interacting protein 3 like) which reduces apoptosis in response to hypoxia.\textsuperscript{(413)} This mechanism could potentially be important in tumourigenesis. Other studies support this and have shown that miR-23 also targets the apoptotic protein Fas thereby reducing apoptosis.\textsuperscript{(414, 415)}

In summary, there is little information regarding miR-23 although it may be involved in malignancy. Related miRNAs also appear to be involved in angiogenesis and apoptosis, but this evidence is conflicting. There is also evidence that miR-23a, a related miRNA, is in fact a tumour suppressor rather than 23c which promotes tumourigenesis.\textsuperscript{(416, 417)} Therefore, one has to approach extending results from related miRNAs to 23c function with caution. Its’ associations with cancer and the lack of previous investigations into its function make miR-23c a promising target to
investigate further. However, miR-23 did not seem to be expressed in lung cancer endothelial cells and therefore it is less likely to provide an endothelial mechanism for the common pathogenesis of COPD and lung cancer.

3.5.2 mRNA expression in pulmonary endothelium

This study identified 6 endothelial genes differentially expressed between COPD and non-COPD in microarray studies: DLC1, HHIP, LTA4H, PPIL2, TMEM154 and TP53. This was compared to the mRNA microarray of endothelium from tumours versus non-COPD lung tissue. All of these genes were differentially expressed between tumour and non-COPD with the exception of LTA4H, which suggested they might be possible shared markers of pathogenesis between COPD and lung cancer. Unfortunately, this was not supported by the qPCR validation studies. This is because each microarray contains thousands of genes and the probability of a false positive result increases with the number of hypotheses in a certain test. Thus, the likelihood of false positive results occurring in a microarray is huge. It is possible to limit the number of false positives using statistical tests for multiple testing (such as the false discovery rate; FDR) but as these methods are more stringent it can result in important genes being missed. Therefore, in this study, fold change was used to identify significantly differentially expressed genes with >2 used as a cut off (this suggests changes of biological relevance). This approach should identify the majority of genes of interest, but will also identify multiple false positives requiring the need for validation by qPCR. There was very little change in expression of TMEM154 and TP53 between COPD and non-COPD. PPIL2 was not obviously expressed in COPD or non-COPD pulmonary endothelium. DLC1, HHIP and LTA4H were increased in COPD with only LTA4H approaching significance (p=0.067).
tailed t-test was chosen in this instance as the microarray data provided a clear hypothesis for qPCR testing.

Leukotriene A4 hydrolase (LTA4H) is an enzyme that has been previously investigated in COPD. It has both pro- and anti-inflammatory actions. Its’ hydrolase function catalyses the conversion of leukotriene A4 (LTA4) to leukotriene B4 (LTB4), a neutrophil chemoattractant.(420) However, its’ aminopeptidase function degrades proline-glycine-proline (PGP), a protein with sequence homology to CXC chemokines and also acts as a neutrophil chemoattractant.(421) Therefore, activation of LTA4H may result in both the development and resolution of inflammation.(422) Previous studies have shown that LTA4H is increased in the lung tissue of murine COPD models and sputum from COPD patients.(422, 423) Interestingly, smoke appears to inhibit the aminopeptidase function of LTA4H, but not its’ hydrolase function.(422) Therefore, increased LTA4H in COPD patients would lead to an increase in both LTB4 and PGP.(422) This effect persists in COPD patients even after smoking cessation and therefore makes an intriguing target for the treatment of inflammation in COPD.(422) The difficulty with inhibiting LTA4H completely is the risk of a further reduction in aminopeptidase activity and accumulation of PGP. Murine models have shown that blockage of the aminopeptidase function of LTA4H results in an increase in neutrophilic inflammation and protease imbalance resulting in degradation of the extracellular matrix.(424) There are different ways of approaching this. Firstly, it is possible to augment the aminopeptidase function of LTA4H. For example, in a mouse COPD model the drug 4MDM (which selectively increases aminopeptidase activity) resulted in decreased PGP in BAL and reduced numbers of neutrophils in mouse lungs.(423) This acted to prevent emphysematous remodelling in the mouse
model. Alternatively, a new selective LTA4H hydrolase inhibitor, ARM1 has been
developed. (425) This would give the opportunity to selectively inhibit only the pro-
inflammatory functions of LTA4H.

In addition to inflammation, there is also evidence that LTA4H may be important in
endothelial apoptosis. Pulmonary hypertension (a condition that results in increased
blood pressure in pulmonary vessels) is associated with macrophage accumulation
near arterioles of the lung. (426) These macrophages express high levels of
LTA4H. (426) Macrophage-derived LTB4 was shown to induce apoptosis in
pulmonary artery endothelial cells (in vitro) by binding to its receptor BLT1 and
inhibiting the sphingosine kinase 1- endothelial nitric oxide synthase pathway. (426) If
LTA4H has similar effects on microvascular pulmonary endothelial cells it is possible
that it could induce emphysema through endothelial apoptosis.

There is less evidence available about LTA4H in lung cancer in comparison to
COPD. This would concur with the microarray findings that did not identify LTA4H as
an expressed target in lung cancer. However, LTA4H has been described in lepidic
lung carcinoma in situ. (427) This in addition to its known inflammatory role may mean
it could still potentially provide a link between the two conditions. None of the patients
in this study had lepidic carcinoma and therefore perhaps this is why LTA4H was not
differentially expressed between non-COPD and cancer groups.

In summary, LTA4H looks to be a potential important marker in COPD through both
inflammatory and apoptotic pathways. As it leads to the production of a
chemoattractant, LTB4, it is possible that upregulation of LTA4H might result in the
increase of transendothelial migration (TEM) of neutrophils seen in COPD patients.
3.5.3 Limitations

There are some limitations to this study. Firstly, it would be ideal to compare lung tissue from patients with COPD to tissue from patients with normal lungs. In this study only one patient without COPD did not have lung cancer. Therefore the miRNA and mRNA expression in non-COPD patients in this study could be different compared to patients without COPD or lung cancer. However, lung tissue is removed in this group of patients only rarely. For example patients with a suspicious nodule on CT scan that is confirmed as benign after removal. There were too few of these patients to provide an adequate control group for this study. Secondly, the COPD patients had only mild-moderate disease on average (mean FEV1pp 73%). In order to be fit for surgery all patients must have adequate lung function and thus most patients with severe COPD and lung cancer are not deemed suitable for surgery. I attempted to counter this selection bias by recruiting LVRS patients, however fewer procedures and smaller tissue volumes obtained in such patients limited the number I was able to recruit. Therefore it is possible that genetic changes in more severe cases of COPD might have been missed.

Samples collected by previous members of the Turner/Bicknell groups were also included in this study. All samples were collected in the same way and processed on the same type of microarray chip. However, it is possible that there may have been variability between how the samples were collected by different members of the team which could have possibly influenced results.

There were also only small numbers of patients in each microarray experiment. Power calculations performed in R using the sizepower package were performed to estimate the number of patients required in each group to achieve 90% power at
detecting difference between COPD and non-COPD, using data from prior published work on COPD.(269) 10 samples were required per group for miRNA microarrays and this was achieved. However, 35 samples were required per group for mRNA microarrays. This was not possible to achieve due to lack of sample availability and cost implications. This may be the main reason that there were more significantly differently expressed targets seen in the miRNA validation work compared to the mRNA validation work. Similar power calculations suggested that 6 samples were needed in each group to compare normal tissue and cancer. Only 3 samples were in each group. However, this analysis did manage to identify differentially expressed mRNAs with a significant FDR and so appeared to have adequate power.

3.6 Conclusions

To our knowledge this is the first study comparing miRNA and mRNA expression in COPD and non-COPD pulmonary endothelial cells. There are significant differences in the miRNA and mRNA expression between these two groups. Three known endothelial miRNAs were validated using qPCR as being upregulated in COPD: miR-181b-3p, -429 and -23c. miR-429 was also upregulated in lung cancer and provides the most promising target for shared pathogenesis between COPD and lung cancer.

One mRNA of interest was also upregulated in COPD: LTA4H. This gene appears to be important in inflammation, transendothelial migration and apoptosis and therefore targeting this gene may provide a new pathway to target in COPD. As LTA4H was not differentially expressed in non-COPD versus lung cancer in the microarray experiments it is less likely to provide a common target for the two pathologies.
CHAPTER FOUR
PATHWAY ANALYSIS OF MICRORNA AND MESSENGER RNA EXPRESSION DATA
4.1 Hypothesis

Intracellular pathways in pulmonary endothelial cells will differ significantly between COPD and non-COPD patients. Pathways relating to cancer development will be upregulated in COPD and will help to explain the increased risk of lung cancer development in patients with COPD.

4.2 Aims of this chapter

This chapter will outline the pathway analyses performed using microRNA (miRNA) and messenger RNA (mRNA) microarray data analysed in chapter 3.

4.3 Introduction

Pathway analysis of genomic data can provide a useful way of interpreting the biological significance of large datasets. Pathway analysis has been used successfully in the investigation of both COPD and lung cancer and identified shared biological pathways. Thus it was appropriate to use IPA upon the miRNA and mRNA microarray data described in chapter 3 to identify potential pathways involved in COPD and whether or not they could help to explain the increased risk of lung cancer in COPD patients.

4.4 Results

4.4.1 Diseases and biological functions associated with COPD.

Figures 4.1-4 demonstrate the top 10 diseases and biological functions associated with miRNA/mRNA expression in COPD endothelium for each of the 4 array analyses. Diseases and biological functions that are in the top 10 for all 4 analyses include: Cancer (the process most highly related to COPD in both mRNA arrays), Organismal injury and abnormalities and Reproductive system disease.
Figure 4.1: The top 10 diseases and biological functions altered in the 2014 miRNA microarray data. A threshold for significance is illustrated in the figure. Diseases and biological functions are listed in order of significance. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.

Figure 4.2: The top 10 diseases and biological functions altered in the 2015 miRNA microarray data. A threshold for significance is illustrated in the figure. Diseases and biological functions are listed in order of significance. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.3: The top 10 diseases and biological functions altered in the 2014 mRNA microarray data. A threshold for significance is illustrated in the figure. Diseases and biological functions are listed in order of significance. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.

Figure 4.4: The top 10 diseases and biological functions altered in the 2016 mRNA microarray data. A threshold for significance is illustrated in the figure. Diseases and biological functions are listed in order of significance. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
In order to explore the molecular and cellular processes altered in COPD in further detail all molecular and cellular processes significantly different in COPD were downloaded.

Important pathways that were independently identified in several of the subset analyses are listed in table 4.1. The results for significantly altered molecular and cellular processes are illustrated in figures 4.5-4.8.

In a similar fashion to molecular and cellular functions, disease associations were also investigated in further depth by downloading all disease associations that were deemed significant by IPA. The total number of diseases significantly associated in COPD for each analysis can be seen in table 4.2. The results for the diseases associated with COPD are illustrated in figures 4.9-4.12.

To identify whether or not the above disease associations were genuine, a literature search was performed to identify publicly available data on incidence and prevalence of co-morbid disease in COPD and to the co-morbidities present in the patients studied (table 4.3). The literature search was performed following a standard hierarchy. Initially systematic reviews were searched for each co-morbidity. Where this was not possible large cohort studies followed by case reports were used.
Table 4.1: Molecular and cellular processes altered in COPD in each microarray analysis.

<table>
<thead>
<tr>
<th>Category</th>
<th>2015 miRNA</th>
<th>2016 miRNA</th>
<th>2014 mRNA</th>
<th>2016 mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acid Metabolism</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Carbohydrate Metabolism</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Cell Cycle</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cell Death and Survival</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cell Morphology</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cell-mediated Immune Response</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell Signaling</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cell-mediated Immune Response</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell-To-Cell Signaling and Interaction</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Cellular Assembly and Organization</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cellular Compromise</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cellular Development</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cellular Function and Maintenance</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cellular Growth and Proliferation</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cellular Movement</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cellular Response to Therapeutics</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA Replication, Recombination, and Repair</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Drug Metabolism</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Free Radical Scavenging</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Expression</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hematopoiesis</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humoral Immune Response</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Immune Cell Trafficking</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Lipid Metabolism</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Molecular Transport</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Post-Translational Modification</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Synthesis</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Trafficking</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Molecule Biochemistry</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Vitamin and Mineral Metabolism</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

This table lists the molecular and cellular processes identified by IPA as being significantly altered in COPD. Each column represents an individual array analysis and ‘X’ indicates that the process was significantly altered in COPD. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.5: Molecular and cellular functions associated with differentially expressed miRNAs in COPD in the 2014 miRNA microarray data. The pie chart demonstrates the percentage of miRNAs (from the total number) involved in molecular and cellular functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.6: Molecular and cellular functions associated with differentially expressed miRNAs in COPD in the 2015 miRNA microarray data. The pie chart demonstrates the percentage of miRNAs (from the total number) involved in molecular and cellular functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.7: Molecular and cellular functions associated with differentially expressed mRNAs in COPD in the 2014 mRNA microarray data. The pie chart demonstrates the percentage of mRNAs (from the total number) involved in molecular and cellular functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.8: Molecular and cellular functions associated with differentially expressed mRNAs in COPD in the analysis: 2016 mRNA microarray data. The pie chart demonstrates the percentage of mRNAs (from the total number) involved in molecular and cellular functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Table 4.2: Diseases associated with COPD in each microarray analysis.

<table>
<thead>
<tr>
<th>Category</th>
<th>2015 miRNA</th>
<th>2016 miRNA</th>
<th>2014 mRNA</th>
<th>2016 mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Disease</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Behaviour</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cancer</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cardiovascular Disease</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Connective Tissue Disorders</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dermatological Diseases and Conditions</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Endocrine System Disorders</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal Disease</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Hematological Disease</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Hepatic System Disease</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Hereditary Disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunological Disease</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Infectious Diseases</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Inflammatory Disease</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Inflammatory Response</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Metabolic Disease</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Neurological Disease</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Nutritional Disease</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ophthalmic Disease</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Organismal Injury and Abnormalities</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Psychological Disorders</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Renal and Urological Disease</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Reproductive System Disease</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Respiratory Disease</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Skeletal and Muscular Disorders</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

This table lists the diseases identified by IPA as being significantly associated with COPD. Each column represents an individual array analysis and ‘X’ indicates that the disease was significantly associated with COPD. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.9: Diseases associated with differentially expressed miRNAs in COPD in the 2014 miRNA microarray data. The pie chart demonstrates the percentage of miRNAs (from the total number) involved in disease functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.10: Diseases associated with differentially expressed miRNAs in COPD in the 2015 miRNA microarray data. The pie chart demonstrates the percentage of miRNAs (from the total number) involved in disease functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.11: Diseases associated with differentially expressed mRNAs in COPD in the 2014 mRNA microarray data. The pie chart demonstrates the percentage of mRNAs (from the total number) involved in disease functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.12: Diseases associated with differentially expressed mRNAs in COPD in the 2016 mRNA microarray data. The pie chart demonstrates the percentage of mRNAs (from the total number) involved in disease functions significant according to IPA. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Table 4.3: Diseases identified as significant in IPA and their relation to known comorbidities of patients included in the study and the current literature

<table>
<thead>
<tr>
<th>Comorbidities identified in IPA</th>
<th>Comorbidities identified in patients included in the study</th>
<th>Comorbidities identified in the literature</th>
<th>Level of evidence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Disease</td>
<td></td>
<td>Not supported by current literature</td>
<td>Cohort</td>
<td>(428)</td>
</tr>
<tr>
<td>Behaviour</td>
<td>Heroin misuse (1)</td>
<td>Insomnia (OR 2.4)</td>
<td>Cross-sectional</td>
<td>(429)</td>
</tr>
<tr>
<td>Cancer</td>
<td>Lung (10) Lymphoma (1) Malignant myoepithelial carcinoma of left scapula (1)</td>
<td>Lung cancer (OR 2.11)</td>
<td>Systematic review</td>
<td>(430)</td>
</tr>
<tr>
<td>Cardiovascular Disease</td>
<td>Hypercholesterolaemia (4) Hypertension (4)</td>
<td>Cardiovascular disease (OR 2.46); Hypertension (OR 1.33)</td>
<td>Systematic review</td>
<td>(431)</td>
</tr>
<tr>
<td>Connective Tissue Disorders</td>
<td>Polyarthritis (1) Osteoarthritis (1)</td>
<td>COPD associated with Sjogren's syndrome, systemic sclerosis, rheumatoid arthritis</td>
<td>Cohort</td>
<td>(432) (433)</td>
</tr>
<tr>
<td>Dermatological Diseases and Conditions</td>
<td></td>
<td>OR of COPD in psoriasis (1.90)</td>
<td>Systematic review</td>
<td>(434)</td>
</tr>
<tr>
<td>Endocrine System Disorders</td>
<td>Diabetes (2) Hyperthyroidism (1) Hypothyroidism (1)</td>
<td>Diabetes (OR 1.36)</td>
<td>Systematic review</td>
<td>(431)</td>
</tr>
<tr>
<td>Gastrointestinal Disease</td>
<td>Gastroesophageal Reflux Disease (1)</td>
<td>Gastroesophageal reflux disease (Prevalence 30%); Colorectal polyps (OR 2.1)</td>
<td>Cohort</td>
<td>(435, 436)</td>
</tr>
<tr>
<td>Haematological Disease</td>
<td></td>
<td>Anaemia (Prevalence 6.2% to 46.3%)</td>
<td>Systematic review</td>
<td>(437)</td>
</tr>
<tr>
<td>Hepatic System Disease</td>
<td>OR of COPD in liver disease (2.1)</td>
<td>Cohort</td>
<td>(438)</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------</td>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Hereditary Disorder</td>
<td>Alpha-1-Antitrypsin Deficiency (Prevalence 0.63%)</td>
<td>Cohort</td>
<td>(439)</td>
<td></td>
</tr>
<tr>
<td>Immunological Disease</td>
<td>Human Immunodeficiency Virus (reduced CD4 counts) associated with reduced FEV1.</td>
<td>Cross-sectional</td>
<td>(440)</td>
<td></td>
</tr>
<tr>
<td>Infectious Diseases</td>
<td>OR for COPD with a history of TB (3.05)</td>
<td>Systematic review</td>
<td>(441)</td>
<td></td>
</tr>
<tr>
<td>Inflammatory Disease</td>
<td>CRP and IL-6 are related to FEV1/FVC. 30% of COPD patients show evidence of systemic inflammation</td>
<td>Cross-sectional</td>
<td>(66)</td>
<td></td>
</tr>
<tr>
<td>Metabolic Disease</td>
<td>Metabolic syndrome (Prevalence 30%). Reduced plasma leptin/adiponectin ratio in COPD.</td>
<td>Systematic review/cohort</td>
<td>(442, 443)</td>
<td></td>
</tr>
<tr>
<td>Neurological Disease</td>
<td>Childhood seizures (1)</td>
<td>Peripheral neuropathy (Prevalence 15%); Mild cognitive impairment (OR 1.87); Stroke (HR 1.24); Parkinson's disease (HR 1.37)</td>
<td>Cohort/cross-sectional</td>
<td>(444-447)</td>
</tr>
<tr>
<td>Nutritional Disease</td>
<td>Vitamin B12 deficiency (1)</td>
<td>Weight loss (OR 1.81)</td>
<td>Case-control</td>
<td>(448)</td>
</tr>
<tr>
<td>Ophthalmic Disease</td>
<td>Increased corneal endothelial vulnerability to intraocular surgical stress</td>
<td>Cohort</td>
<td>(449)</td>
<td></td>
</tr>
<tr>
<td>Psychological Disorders</td>
<td>Depression (OR 2.81); Anxiety (Prevalence 13-46%)</td>
<td>Systematic review</td>
<td>(450, 451)</td>
<td></td>
</tr>
<tr>
<td>Renal and Urological Disease</td>
<td>Renal stones (1)</td>
<td>Chronic Kidney Disease (OR 2.20)</td>
<td>Systematic review</td>
<td>(452)</td>
</tr>
<tr>
<td>Reproductive System Disease</td>
<td>Ovarian cyst (1)</td>
<td>Not supported by current literature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>-----------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory Disease</td>
<td>COPD (12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skeletal and Muscular Disorders</td>
<td>Myopathy (Prevalence 36.8%)</td>
<td>Cross-sectional (453)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All diseases significantly associated with COPD according to the IPA are listed in the first column. The second column lists whether or not these comorbidities were present in patients from the study with the number of patients in brackets. The third column lists associations (and if possible the prevalence and risk) of these comorbidities to COPD from the literature. OR = odds ratio, HR = hazards ratio.
4.4.2 Downstream disease analysis of miRNA and mRNA microarrays

Downstream prediction was performed for each of the microarray analyses focusing on the targets of interest identified in chapter 3: miR-181b-3p, miR-429, miR-23c and LTA4H. A key for the diagrams produced by Ingenuity can be seen in Appendix 3.

**Downstream analysis for 2015 miRNA microarrays**

There were no significant downstream effects predicted for the 2015 miRNA microarrays.

**Downstream analysis for 2014 miRNA microarrays**

MiR-181 and miR-429 were both associated with cell functions related to cancer. However, results were conflicting. For example, miR-181 was associated with other miRNAs predicting upregulation of neoplasia of tumour cell lines and miR-429 was predicted (with other miRNAs) to increase cell proliferation of carcinoma cell lines. Both associations would support a relationship between COPD and lung cancer. However, both miR-181 and miR-429 were associated with the downregulation of ‘cell proliferation of tumour cell lines’ and miR-429 with downregulation of ‘invasion of cell lines’. MiR-23 was associated with the upregulation of ‘cytokinesis of ventricular myocytes’ and the downregulation of ‘maturation of chondrocyte cell lines’. The figures for this analysis are in figures 4.13-18.
Figure 4.13: Downstream prediction of miRNA changes on cell proliferation of tumour cell lines in the 2014 miRNA microarray analysis. MiR-181 and -429 are involved in this pathway. Involved miRNAs are around the outside of the figure. All miRNAs are upregulated and are in red. The centre shape is blue suggesting that this function is downregulated in COPD.
Figure 4.14: Downstream prediction of miRNA changes on neoplasia of tumour cell lines in the 2014 miRNA microarray analysis. MiR-181 was involved in this pathway. All miRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
Figure 4.15: Downstream prediction of miRNA changes on cell proliferation of carcinoma cell lines in the 2014 miRNA microarray analysis. miR-429 is involved in this pathway. All miRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
Figure 4.16: Downstream prediction of miRNA changes on invasion of cell lines in the 2014 miRNA microarray analysis. MiR-429 is involved with this pathway. All miRNAs are upregulated and are in red. The centre shape is blue suggesting that this function is downregulated in COPD.
Figure 4.17: Downstream prediction of miRNA changes on cytokinesis of ventricular myocytes in the 2014 miRNA microarray analysis. MiR-23 is involved in this pathway. All miRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
Figure 4.18: Downstream prediction of miRNA changes on maturation of chondrocyte cell lines in the 2014 miRNA microarray analysis. MiR-23 is involved in this pathway. All miRNAs are upregulated and are in red. The centre shape is blue suggesting that this function is downregulated in COPD.
**Downstream analysis for 2016 mRNA microarrays**

LTA4H was involved in several pathways related to cancer including abdominal cancer, digestive organ tumour, intestinal cancer, intestinal tumour, neoplasia of epithelial tissue, renal lesion and tumourigenesis. All of these pathways involved very large numbers of genes and therefore the diagrams of the pathways in IPA could not accurately be reproduced here. Table 4.4 lists the overall downstream prediction for each pathway. The genes involved in each pathway in table 4.4 can be seen in Appendix 3. LTA4H also appeared to be involved in the downregulation of hypersensitivity reactions (figure 4.19). The full names of genes listed in figure 4.19 can be viewed in Appendix 3.

**Downstream analysis for 2014 mRNA microarrays**

Some pathways related to cancer were downregulated including digestive organ tumour and malignant neoplasm of the large intestine. These pathways involved very large numbers of genes and could not be reproduced easily in this thesis. Table 4.5 lists the overall downstream prediction for each pathway. The individual genes in each pathway can be seen in Appendix 3. LTA4H was also involved in the allergy pathway (figure 4.20), which was predicted to be downregulated in the same fashion as hypersensitivity reactions in the 2016 microarrays above. However, LTA4H was also shown to be involved in the upregulation of metabolism of eicosanoid, synthesis of fatty acid and synthesis of lipid. These pathways are illustrated in figures 4.21-4.23. The full names of genes contained within figures 4.21-4.23 can be viewed in Appendix 3.
Table 4.4: Downstream analysis predictions for diseases/cellular functions significant in the 2016 mRNA analysis.

<table>
<thead>
<tr>
<th>Disease or Function</th>
<th>Downstream Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumorigenes is of tissue</td>
<td>Down</td>
</tr>
<tr>
<td>Neoplasia of epithelial tissue</td>
<td>Down</td>
</tr>
<tr>
<td>Abdominal cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Digestive organ tumour</td>
<td>Down</td>
</tr>
<tr>
<td>Intestinal cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Intestinal tumour</td>
<td>Down</td>
</tr>
<tr>
<td>Renal lesion</td>
<td>Down</td>
</tr>
</tbody>
</table>

The first column of this table lists significant diseases and functions associated with COPD pulmonary endothelial tissue. The second column lists whether or not each disease or function was predicted to be up- or down-regulated. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.19: Downstream prediction of mRNA changes on hypersensitivity reactions in the 2016 mRNA microarray analysis. LTA4H is involved in this pathway. mRNAs that are upregulated are in red and those that are downregulated are in green. The centre shape is blue suggesting that this function is downregulated in COPD.
Table 4.5: Downstream analysis predictions for diseases/cellular functions significant in the 2014 mRNA analysis.

<table>
<thead>
<tr>
<th>Disease or Function</th>
<th>Downstream Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Malignant solid tumour</td>
<td>Down</td>
</tr>
<tr>
<td>Neoplasia of epithelial tissue</td>
<td>Down</td>
</tr>
<tr>
<td>Digestive system cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Abdominal cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Digestive organ tumour</td>
<td>Down</td>
</tr>
<tr>
<td>Large intestine neoplasm</td>
<td>Down</td>
</tr>
<tr>
<td>Malignant neoplasm of large intestine</td>
<td>Down</td>
</tr>
<tr>
<td>Urogenital cancer</td>
<td>Down</td>
</tr>
<tr>
<td>Genitourinary carcinoma</td>
<td>Down</td>
</tr>
<tr>
<td>Renal lesion</td>
<td>Down</td>
</tr>
</tbody>
</table>

The first column of this table lists significant diseases and functions associated with COPD pulmonary endothelial tissue. The second column lists whether or not each disease or function was predicted to be up- or down-regulated. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
Figure 4.20: Downstream prediction of mRNA changes on allergy in the 2014 mRNA microarray analysis. LTA4H is involved in this pathway. All mRNAs are upregulated and are in red. The centre shape is blue suggesting that this function is downregulated in COPD.
Figure 4.21: Downstream prediction of mRNA changes on metabolism of eicosanoid in the 2014 mRNA microarray analysis. LTA4H is involved in this pathway. All mRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
Figure 4.22: Downstream prediction of mRNA changes on synthesis of fatty acid in the 2014 mRNA microarray analysis. LTA4H is involved in this pathway. All mRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
Figure 4.23: Downstream prediction of mRNA changes on synthesis of lipid in the 2014 mRNA microarray analysis. LTA4H is involved in this pathway. All mRNAs are upregulated and are in red. The centre shape is orange suggesting that this function is upregulated in COPD.
4.4.3 miRNA and mRNA networks in COPD

Networks relating to miR-181, miR-429, miR-23 and LTA4H were identified. The full names of all genes included in the networks can be viewed in Appendix 3.

Network analysis for the 2015 miRNA microarrays

2 networks that were significant in this analysis also involved 2 of the possible targets: miR-181 and miR-23. MiR-181 was involved in network 1 (figure 4.24) and was associated with ‘Cancer’, ‘Dermatological Diseases and Conditions’ and ‘Organismal Injury and Abnormalities’. MiR-23 was involved in network 2 (figure 4.25) and was associated with ‘Cellular Development’, ‘Cellular Growth and Proliferation’ and ‘Connective Tissue Development and Function’.

Network analysis for the 2014 miRNA microarrays

2 networks that were significant in this analysis also involved 2 of the possible targets: miR-181 and miR-429. MiR-181 was involved in network 1 (figure 4.26) and was associated with ‘Organismal Injury and Abnormalities’, ‘Reproductive System Disease’ and ‘Cancer’. MiR-429 was involved in network 7 (figure 4.27) and was associated with ‘Connective Tissue Disorders’, ‘Developmental Disorder’ and ‘Hereditary Disorder’ functions.
Figure 4.24: A network identified in the 2015 miRNA microarrays. This is network 1 and it is associated with miR-181 (highlighted). MiR-181 is in red and is therefore upregulated. This appears to directly inhibit EPHB1 (EPH receptor B1), a tyrosine kinase which has been previously reduced in several malignancies.
Figure 4.25: A network identified in the 2015 miRNA microarrays. This is network 2 and it is associated with miR-23 (highlighted). MiR-23 is in red and is upregulated. This appears to be partially through indirect stimulation through reduced levels of hydrogen peroxide. MiR-23 directly inhibits 6 mRNAs in the pathway, which are all predicted to be downregulated: SERINC5 (Serine incorporator 5), DYDC2 (DPY30 domain-containing protein 2), YOD1 (Ubiquitin thioesterase OTU1), FGF2 (fibroblast growth factor 2), TGFB2 (transforming growth factor B2), GCC1 (GRIP and coiled-coil domain-containing protein 1).
Figure 4.26: A network identified in the 2014 miRNA microarrays. This is network 1 and it is associated with miR-181 (highlighted). MiR-181 is red and therefore is upregulated. This seems to be partly through indirect activation by Smad 2/3 (a transcription regulator) which is predicted to be upregulated. Another mechanism of upregulation appears to be through insulin which is reduced. Insulin acts indirectly on miR-181 leading to its inhibition.
Figure 4.27: A network identified in the 2014 miRNA microarrays. This is network 7 and it is associated with miR-429 (highlighted). MiR-429 is upregulated in red. This appears to directly inhibit EIF2S1 (eukaryotic translation initiation factor 2 subunit alpha) which is predicted to be downregulated. EIF2S1 is connected directly to 4 other mRNAs in the pathway. These include IMPACT and RAVER2 (ribonucleoprotein polypyrimidine tract protein-blinding 2) which are also both involved in regulation of mRNA translation. The final two mRNAs are NDRC (nexin-dynein regulatory complex) which is involved in cilia function and RAX (retina and anterior neural fold homeobox gene) which is involved in retinal development.
Network analysis for the 2016 mRNA microarrays

There were no networks involving LTA4H that were significant in the 2016 mRNA microarrays. However, network 22 was significant and involved miR-429. This can be viewed in figure 4.28. This network was associated with ‘Developmental Disorder’, ‘Hereditary Disorder’ and ‘Neurological Disease’ functions.

Network analysis for the 2014 mRNA microarrays

One network in the 2014 mRNA microarrays that was significant according to IPA involved LTA4H. This was network 4 (figure 4.29) and it was associated with ‘Respiratory System Development and Function’, ‘Cardiovascular Disease’ and ‘Cancer’. Therefore it is possible that through this pathway LTA4H might provide an important link between COPD and lung cancer.
Figure 4.28: A network identified in the 2016 mRNA microarrays. This is network 22 and it is associated with miR-429 (highlighted). MiR-429 is predicted to be upregulated and is in orange. This would be consistent with the miRNA array results. MiR-429 appears to inhibit 9 mRNAs that are predicted to be downregulated including KLF9 (Krueppel-like factor 9), which acts as a tumour suppressor and inhibits cellular proliferation. MiR-429 also targets CBX5 (Chromobox protein homolog 5) that appears to control angiogenic functions of endothelial progenitor cells. MiR-429 is also predicted to inhibit three mRNAs that are upregulated (in red). The yellow lines linking miR-429 to these mRNAs suggests that upregulation of miR-429 is inconsistent with upregulation of its target mRNAs.
Figure 4.29: A network identified in the 2014 mRNA microarrays. This is network 4 and it is associated with LTA4H which is upregulated. E2F1 (E2F transcription factor 1) which is also upregulated appears to be responsible for LTA4H activation as it is predicted to directly stimulate LTA4H. There is evidence that E2F1 is upregulated in inflammation and activates genes involved in the NF-κB pathway.
4.4.4 Canonical pathways in COPD

Significantly affected canonical pathways in the 2014 and 2016 mRNA microarray analyses were identified. LTA4H was identified in the ‘Eicosanoid Signalling’ pathway in both mRNA microarray analyses and the results of this are presented in figure 4.30.

The Eicosanoid signalling pathway from the 2016 arrays shows an upregulation of LTA4H and therefore an upregulation of the conversion of LTA4 to LTB4. LTB4 binds to its receptors promoting an increase in chemotaxis, cell proliferation and angiogenesis.

Other aspects of this pathway show upregulation of prostaglandins E2 (PGE2), F2 alpha (PGF2A), D2 (PGD2) and I2 (PGI2), which are associated with other disease processes of interest such as inflammation and cancer. Lipoxin A4 was also upregulated in the pathway, which results in vasodilation and was associated with inflammatory disease. Thromboxane A2 was reduced in this pathway, which was associated with reduced asthma risk. This result is not consistent with the other pathway results, which predict an increased asthma risk.

The Eicosanoid pathway from the 2014 arrays shows almost identical results to those from the 2016 array. However, unlike the results from 2016 the thromboxane A2 pathway was predicted to be upregulated, which is therefore more consistent with the rest of the pathway results.

The full names of genes included in the Eicosanoid signalling pathway can be viewed in Appendix 3.
Figure 4.30: Eicosanoid Signalling pathway according to results from the 2014 mRNA microarray analysis. This demonstrates that the highlighted proteins LTA4H, PLA2, PTGS, PDGS and PTGER3 have been measured as upregulated in this pathway. PLA2, PTGS and PDGS are involved in prostaglandin synthesis. PTGER3 is a prostaglandin receptor. The other proteins in this pathway are predicted to be upregulated (in orange). All disease and cellular functions controlled by this pathway (such as chemotaxis, cell proliferation and angiogenesis) are also predicted to be upregulated (in orange) other than ‘thrombosis’ which is coloured blue in the pathway.
4.5 Discussion

4.5.1 Diseases associated with COPD

Cancer

There were several interesting findings in the IPA. Firstly, the data from the four independent microarray analyses support the concept that COPD and lung cancer are related. Cancer is the process most highly related to COPD in both mRNA arrays and is in the top 10 associations for the miRNA arrays further supporting the hypothesis that cancer and COPD are related. Some of the biological functions associated with the arrays might explain why cancer is an important association such as inflammation. COPD is a disease driven by inflammation with increased levels of both systemic inflammatory cells and cytokines.(64) Supporting this concept, ‘inflammatory disease’ and the ‘inflammatory response’ were significant in all microarray datasets. This is also consistent with several other IPAs performed by different groups in COPD that all identify inflammatory pathways as being significantly altered.(319, 321, 322) These altered inflammatory pathways are of particular relevance in the setting of shared pathogenesis of COPD and lung cancer as malignancy risk appears to be increased in the setting of chronic inflammation.(149)

Several of the molecular and cellular processes were significantly altered in all four analyses, giving extra confidence to the conclusion that these pathways are important, since there was independent replication in more than one dataset. Pathways that were independently identified in several of the subset analyses included cellular growth and proliferation, cellular development, the cell cycle, cellular movement, cell death and survival, cell signalling and cell-cell interaction. Again,
several of these processes such as cell death and survival and cell-cell interaction have also been identified in another COPD IPA further supporting this study’s results. (321) All of these pathways would have the potential to be important in carcinogenesis and provide support for the hypothesis that carcinogenesis is more likely to occur in COPD lung. (454-456) There is existing evidence from other microarray studies in COPD that would support these findings. For example, Yang et al demonstrated that the NF (nuclear factor) – κB1 pathway is upregulated in COPD and this is associated with a reduction in apoptosis. (457) Similarly, Chen et al identified interleukin-6 (IL-6) as being upregulated in the JAK/STAT (Janus kinase/Signal transducer and Activator of Transcription) pathway that is associated with the promotion of inflammation and cell proliferation. (458) Finally, Wang et al identified down-regulation of the Wnt (wingless-type MMTV (mouse mammary tumor virus) integration site) pathway in COPD that would result in the dysregulation of cellular differentiation. (459)

Cell-cell signalling and interaction would also be of importance in inflammation and transendothelial migration (TEM), which in turn is a determinant of the level of inflammatory cells in the lung. There is evidence to support that the migration of neutrophils is upregulated in COPD and that this is related to disease severity. (231-233) Perhaps this could partly be due to the dysregulation of cell-cell interactions in the pulmonary endothelium. Furthermore, increased TEM could explain the increased inflammation seen in COPD (54) which is supported by these pathway results.

A downstream analysis was also performed in order to attempt to predict the direction of the effect the cellular and molecular processes would have on disease
processes. This downstream analysis was performed focusing on the miRNA and mRNA targets identified in chapter 3. In some cases these downstream predictions would support the hypothesis of underlying cellular changes in COPD resulting in an increased malignancy risk. For example, in the 2014 miRNA microarrays miR-181 was predicted (in a network with 6 other miRNAs) to upregulate neoplasia of tumour cell lines. MiR-429 was also predicted to increase cell proliferation of carcinoma cell lines. This would be consistent with previous data suggesting that overexpression of miR-181 and -429 is associated with lung cancer. However, the picture from the array results is not as simple as this. Both miR-429 and -181 appear to be involved in a network leading to the reduction of cell proliferation of tumour cell lines in the same arrays. In addition to this, miR-429 is also involved in a network resulting in reduction of invasion of cell lines. These findings would appear to contradict the initial two networks, which suggested an increased malignancy risk. As these networks occurred in the same array this cannot be explained by patient subsets. One possible explanation could be that the knowledge database that the IPA is based on is drawn from published literature and both miR-181 and miR-429 have also been associated with reduction in malignancy. For example, miR-181 has been shown to be reduced in NSCLC and miR-429 is reduced in renal cell carcinoma. Perhaps these networks reflect malignancy associations from tissues other than pulmonary endothelium that, by coincidence, contain RNAs significantly differentially expressed in these arrays. For example, the study showing reduced miR-181 in NSCLC used whole lung miRNA expression and could therefore reflect miRNA expression in other cell types meaning its findings would not be applicable to endothelial cells. This could potentially explain conflicting results but also
complicates the interpretation of the data. It is difficult to understand which networks might be relevant in the current setting and therefore, these particular networks fail to elucidate the shared pathogenic mechanisms between COPD and lung cancer. In theory conflicting results such as this could occur in many IPAs as the knowledge base reflects data from multiple cell types and diseases. However, a literature search could not readily identify examples of this and most publications on downstream analysis in IPA do not consist of conflicting results.(319, 321, 322, 460, 461)

LTA4H was also involved in multiple significant downstream analyses that involved cancer, but in every case the cancer association was negative. This complicates the association of these array results with malignancy further. It also appears to contradict current literature that associates an increase in LTA4H with various malignancies.(427, 462, 463) However, LTA4H was primarily associated with gastrointestinal cancers rather than lung cancers in the downstream analysis. Interestingly, in a previous COPD IPA, gastrointestinal cancer pathways were also significantly altered.(321) Despite this, it is probably difficult to extrapolate these particular results to the risk of lung cancer in COPD. Also, the IPA analysis does not distinguish between cell types during the analysis. Therefore, some of the associations reported in the results may be associations found in different cell types from different tissues which could partly explain why there are some contradictory results seen. Regardless, it is difficult to make firm conclusions on cancer risk based on these downstream results. It would be necessary to repeat the microarrays, perhaps with greater numbers, to further clarify the roles of miR-429, -181 and LTA4H on these networks in a COPD setting. It would also be useful to perform all array analyses at one time which would allow one to compare all array results with
one another rather than performing separate smaller analyses with less statistical power. Further work to explore cell function more with respect to targets would be useful, and is described in the future work section (section 7.5.1).

IPA was also used to see if the miRNA targets and LTA4H were involved in any IPA networks. Genes of particular interest in these networks are listed in Appendix 3. Each network identified as significant included genes previously recognised as important in various malignancies, frequently lung cancer, further reinforcing the potential link between this condition and COPD. Several genes had also been previously altered in COPD. One particularly interesting find is the downregulation of EPHB1 (EPH receptor B1) which is directly inhibited by miR-181b in the 2015 miRNA microarrays (network 1). This tyrosine kinase receptor is involved in transducing signals for cell migration and under-expression has been associated with multiple cancer types. (464-470) This could theoretically provide a mechanism of how upregulation of miR-181b in COPD could predispose to cancer.

EPHB1 is indirectly linked to RHOA (ras homolog family member A) which again is predicted to be downregulated. RHOA is associated with pulmonary endothelial dysfunction in COPD and therefore may be of importance in this study. (226) However, previous studies have shown upregulation of RHOA in pulmonary endothelial dysfunction and consequently this result is not consistent with previous evidence. OGT (O-linked N-acetylglucosamine transferase), a gene connected to RHOA in network 1, may also have a role in COPD. Interestingly, OGT is also involved in the DNA damage response and could be of particular importance in COPD where oxidative stress is common. (79, 471) It is therefore possible that reduction in OGT expression as seen in this study could represent a defective
cellular response to oxidative stress. However, a review of the GEO database of microarray data from lung tissue does not show a difference in OGT expression between COPD and non-COPD. For example, an analysis of data uploaded by Spira et al does not show a statistically significant difference between OGT expression in normal lung and lung from patients with severe COPD. All the available data originates from whole lung, consequently, to establish the role of OGT in pulmonary endothelial cells would require further investigation using this specific cell type.

**Other diseases**

Ingenuity compares microarray results to existing knowledge on millions of findings including those from different tissue types to that under investigation. This can result in the IPA identifying associations with multiple diseases from different organs. This occurred in all four IPAs. The literature search showed that the IPA data was consistent with the vast majority of expected panel of comorbidities. Areas of particular commonality were cardiovascular disease, endocrine diseases and malignancy. It was notable that auditory disease and reproductive system disease which were both identified in IPA were not supported by current evidence in the literature about comorbidities in COPD. This might reflect associations between COPD-related endothelial dysfunction and these conditions not previously identified in the literature. Interestingly, endothelial dysfunction has been linked to reproductive system disease previously. For example, patients with erectile dysfunction have reduced FMD in comparison to controls. In females, endothelial dysfunction is associated with both polycystic ovary syndrome and pre-eclampsia in pregnancy. Perhaps an association with COPD and reproductive disease
has not been previously reported as patients with COPD are typically past reproductive age at presentation.

There is less evidence for endothelial dysfunction playing a role in auditory disease. However, in mouse models of otitis media HIF and VEGF appear to be upregulated suggesting that vascular proliferation may be important in the pathogenesis of this condition.(475) Therefore, the fact that these diseases were identified in array data from endothelial cells might reflect the systemic importance of endothelial dysfunction in COPD and its importance in the development of associated comorbidities.

Significant downstream analyses that were associated with disease processes other than cancer include downregulation of the ‘maturation of chondrocytes’ pathway that was associated with miR-23. There is some evidence for the degeneration of chondrocytes in bronchial cartilage in COPD and perhaps this pathway reflects this.(476) MiR-23 was also involved in the ‘cytokinesis of ventricular myocytes’, which was upregulated. This could reflect hypertrophic changes in ventricular myocytes occurring as part of pulmonary hypertension in response to hypoxia in COPD. There is recent evidence to suggest these changes are associated with alteration in miRNA expression.(477)

4.5.2 Pathways associated with COPD

Transforming growth factor (TGF)

‘Organismal injuries and abnormalities’ was a process that was significantly altered in COPD in all datasets. This is a general category and includes cellular functions occurring in several multicellular organisms including, primarily, humans, mice and rats. It involves injury mechanisms such as bleeding, oedema and haemorrhage, as well as functions involved in abnormal tissue repair (lesions, ulcers, scars and
wounds). [Personal communication with Dr Ruth Burton, Qiagen Discovery Informatics.] ‘Organismal injuries and abnormalities’ has also been associated with whole lung gene expression studies in COPD previously. (321) This broad category may well have some relevance in the setting of COPD in terms of abnormal tissue repair. Epithelial-mesenchymal transition, a process driven by transforming growth factor beta (TGF-β) has been identified as upregulated in COPD, particularly in the presence of airway fibrosis and provides a mechanistic link to lung cancer. (478) The transforming growth factor beta 2 (TGF-β2) pathway was involved in both the 2015 (network 2) and 2014 (network 1) miRNA microarrays (figure 4.31). TGF signalling has been previously shown to be important in the development of several cancers including non-small cell lung cancer. (479, 480) In the 2014 miRNA microarrays miR-181 is directly linked to Smad 2/3, which is predicted to be upregulated. The Smad proteins are important in the TGF-β pathway; TGF-β binds to its cell surface receptors, which results in the phosphorylation and activation of Smad 2/3 (figure 4.31). The Smad proteins act as regulators of transcription resulting in the increased expression of different genes. (481) Smad 3 in particular is important in the TGF-β mediated epithelial-mesenchymal transition pathway which is critical in carcinogenesis. (482) There is also evidence that the TGF-β pathway and EMT contributes to COPD pathogenesis: TGF-β and Smad proteins are increased in COPD and Smad levels appear to be related to the severity of airway obstruction. (483) The data presented here suggests that upregulation of this pathway could also act to predispose COPD patients to lung cancer.
Figure 4.31: The TGF-β pathway illustrated in a cell. TGF-β binds to the TGF-β receptor II. This recruits the TGF-β receptor I which is phosphorylated and activated. This in turn leads to the phosphorylation of Smad 2/3 which then bind to Smad 4. The Smad complex moves to the nucleus and stimulates the transcription of several genes. TGF-β = Transforming growth factor beta, TGFβRII = Transforming growth factor beta receptor II, TGFβRI = Transforming growth factor beta receptor I, P = Phosphate
Leukotriene A4 Hydrolase (LTA4H)

LTA4H was also associated with pathways other than those involved in malignancy. For example, it was associated with downregulation of hypersensitivity reactions and allergy in the mRNA microarrays. As previously discussed LTA4H results in the production of LTB4 (a neutrophil chemoattractant) and the destruction of PGP (another chemoattractant) and therefore has both pro- and anti-inflammatory effects (figure 4.32). These networks would suggest that the anti-inflammatory effects have a stronger role in COPD contradicting previous evidence that LTA4H is increased in COPD and has a mainly pro-inflammatory role. In support of this, the lines connecting LTA4H to the hypersensitivity and allergy pathways are yellow. This means that the level of LTA4H seen in the arrays is not consistent with the overall effect on the network. I.e. increased LTA4H should, in theory, be associated with an increase in hypersensitivity reactions or allergy not a decrease as seen in figures 4.19-20. Therefore, it is likely that the overall reduction seen in figures 4.19-20 is a result of the influence of genes other than LTA4H.

In the 2014 mRNA microarrays LTA4H is also associated with an upregulation in metabolism of eicosanoid and lipid synthesis pathways, which is consistent with its leukotriene metabolism function. This finding was further supported by the canonical pathway, ‘Eicosanoid signalling’, being upregulated (figure 4.30) in both the 2014 and 2016 mRNA microarrays. Five genes in this pathway were upregulated in the 2014 microarrays: LTA4H, PLA2 (phospholipase A2), PTGS (prostaglandin-endoperoxide synthase/cyclooxygenase), PTGDS (prostaglandin D2 synthase) and prostaglandin E receptor 3 (PTGER3). The same five mRNAs are upregulated in the 2016 mRNA microarrays and the enzyme TXS (thromboxane A synthase 1) was
significantly downregulated. Interestingly, most of these proteins have been previously associated with COPD. For example, LTA4H, PLA2 and PTGDS have been upregulated in sputum or broncho-alveolar lavage from patients with COPD.\textsuperscript{(422, 484, 485)} PTGS is increased in pulmonary fibroblasts isolated from patients with COPD and PTGER3 has been associated with cough.\textsuperscript{(486, 487)} This evidence increases the likelihood that ‘Eicosanoid signalling’ is a genuinely important in COPD. PTGDS levels are also associated with FEV1 levels suggesting that this pathway may have relevance in the development of COPD.\textsuperscript{(485)} Downstream prediction of these changes on ‘Eicosanoid signalling’ results in almost all pathways being upregulated with the exception of thromboxane synthesis in the 2016 microarrays.

The pathway controlled by LTA4H (conversion of LTA4 to LTB4) is predicted to upregulate several cellular processes that would be relevant to this study. Firstly, as discussed in chapter 3, an increase in LTB4 is associated with chemotaxis.\textsuperscript{(420)} This would clearly be important in the transendothelial migration of inflammatory cells in COPD. Secondly, this pathway also appears to increase cell proliferation. An increase in cell proliferation would be relevant in the development of cancer and an increase in LTA4H has been shown to be increased in cancers including lepidic lung cancer as above.\textsuperscript{(427, 462, 463)} There is evidence supporting increased cell proliferation in response to increased LTA4H. LTA4H is upregulated in chronic lymphocytic leukaemia cells with an aggressive phenotype rather than those with a lower, less aggressive proliferation rate.\textsuperscript{(488)} However, it is not clear that this cellular proliferation mechanism is relevant in endothelial cells as previous studies have demonstrated that LTB4 induces endothelial apoptosis, which would be consistent
with the development of emphysema. Similarly, IPA predicts that angiogenesis would be upregulated in response to increased LTA4H which appears to be inconsistent with its’ association with apoptosis. However, other \textit{in vitro} studies have shown that VEGF upregulates LTB4’s receptor (BLT2) in HUVECs and knockdown of BLT2 is associated with reduced angiogenesis. LTB4 has also been shown to be involved in angiogenesis in the setting of chronic inflammation that would be particularly relevant to COPD in which chronic inflammation occurs. The evidence for and against LTB4 being associated with increased angiogenesis is therefore contradictory and it is unclear what functional role LTB4 might play in angiogenesis in the setting of COPD.

Other genes predicted to be altered in this canonical pathway are explored in more detail in Appendix 3.
LTA4H is an enzyme with a hydrolase function that results in the production of LTB4 from LTA4. LTB4 is a neutrophil chemoattractant with pro-inflammatory effects. LTA4H also has an aminopeptidase function which degrades PGP. PGP is another neutrophil chemoattractant and therefore LTA4H has anti-inflammatory effects through its aminopeptidase function. Smoking inhibits the aminopeptidase function of LTA4H which consequently results in increased inflammation.

LTA4H = Leukotriene A4 Hydrolase, LTA4 = Leukotriene A4, LTB4 = Leukotriene B4, PGP = proline-glycine-proline
Angiogenesis factors were also present in several significantly altered networks. In the 2015 miRNA microarrays miR-23 is linked to fibroblast growth factor 2 (FGF2) which is associated with a worse prognosis in NSCLC. (491) The finding of miR-23 linking to FGF2 is of particular interest in endothelial cells as FGF2 is known to be important in angiogenesis, both in the normal and pathological setting. (492) This suggests that miR-23 may also have suppressive effects on angiogenesis through its interaction with FGF2 that could precipitate the development of emphysema. However, this is less likely to provide a mechanistic link between COPD and lung cancer as FGF tends to be increased in NSCLC. (491) In fact, anti-FGF agents are already in development for NSCLC such Cediranib. Unfortunately however, an initial phase III study has been disappointing with increased toxicity in the Cediranib arm with no increase in survival compared to traditional chemotherapy. (255) In the 2014 miRNA microarray VEGF is upregulated. As previously discussed this is the major growth factor involved in the angiogenesis pathway and therefore it is particularly relevant that this is increased in endothelial cells. It is known that VEGF is altered in COPD (210, 211) and is an important negative prognostic marker in lung cancer (243). Therefore, this supports the network in figure 4.25 being relevant in the relationship between the two conditions.

4.5.3 Limitations
Limitations described in section 3.5.3 also apply to the IPA as the same data was used. Briefly, most patients included in the study had lung cancer and so RNA expression in the non-COPD group could be different compared to patients without COPD or lung cancer. This could influence the pathways that emerged in the IPA.
Secondly, the COPD group had mild-moderate disease on average resulting in pathways associated with more severe cases of COPD being missed. However, due to surgery constraints these limitations are difficult to overcome as discussed in 3.5.3. Due to difficulties in combining microarray data (sections 3.3.2 and 3.4.2) each microarray experiment had to be analysed separately. This resulted in 4 separate smaller datasets rather than 2 larger ones for each type of RNA studied. Clearly this resulted in reduced statistical power. This would be possible to overcome in the future by processing all samples on the same day using the same consumables.

Analysis using Ingenuity also presents its own difficulties. For example, the Ingenuity knowledge base which is used in IPA includes associations in a wide range of tissue types and diseases. This is useful as a hypothesis generating tool as it may reveal new pathways of importance as yet unknown in the tissue of interest. However, it may also lead to apparent conflicting results thus complicating interpretation of data as seen in the downstream analysis in this study. MiR-429 and -181b-3p were associated with cellular mechanisms related to both increased and decreased risk of malignancy. This occurred in the same microarray and thus the same patient dataset. Therefore, one cannot conclude these results are due to differences in patients or array technique. It is difficult to elucidate which mechanisms might be relevant in the tissue type under study and so one may not be able to draw conclusions based on IPA alone. It is possible that larger microarray datasets with stronger statistical power would reduce the tendency to see conflicting outcomes in IPA. However, ultimately, clarification of the role of miRNAs/mRNAs in certain cell types and disease states would require functional cell based assays. This is described in further detail in future work (section 7.5.1).
4.6 Conclusions

IPA of microarray data comparing both mRNA and miRNA expression in COPD endothelial cells provides further support to existing evidence that COPD and lung cancer are linked. Cancer was the most significant disease process associated with the mRNA microarray results and was in the top ten significant disease processes for both miRNA microarray results. This is supported by the fact that several related cancer mechanisms such as cell growth and proliferation and cell death and survival were also significantly altered in COPD providing possible underlying reasons why patients with COPD might be at an increased risk of cancer development. Further in depth analysis has also highlighted several networks and pathways such as ‘Eicosanoid signalling’ which are altered in COPD and consist of multiple molecules associated with inflammation, angiogenesis and tumourigenesis suggesting that patients with COPD are at increased lung cancer risk.

However, it is impossible to draw firm conclusions from this analysis as the downstream analysis in particular highlighted several significant processes which were associated with reduced malignancy risk. It is likely that these results reflect disease processes occurring in different tissue types and may not be relevant to HPECs in COPD. Certainly, known epidemiological studies would support increased malignancy in COPD rather than the contrary.(11)

In conclusion therefore, it is necessary to validate the molecules contained in pathways that appear related to cancer risk in this study both quantitatively and functionally. This is beyond the scope of this thesis but such work would be useful to pursue once validation work on the miRNA and mRNA targets identified in chapter 3.
is complete in order to understand the effect of these targets on pulmonary endothelial cells further.
CHAPTER FIVE
ISOLATION AND CULTURE OF HUMAN PULMONARY ENDOTHELIAL CELLS (HPECs)
5.1 Hypothesis

It is possible to extract human pulmonary endothelial cells (HPECs) from donated lung tissue. These cells can be cultured and used for cellular functional validation work.

5.2 Aims of this chapter

This chapter will detail the development of a technique to isolate HPECs from donated lung tissue. This was attempted in order to allow for studying the effects of altering a target of interest in HPECs rather than an alternative cell (such as Human Umbilical Vein Endothelial Cells (HUVECs)). The advantage of this is that alternative cell lines may not express the target of interest. The decision to extract HPECs was made as lung tissue is readily available from the Midlands Lung Tissue Consortium (section 2.1.1). It is possible to buy HPECs, but these cells are both expensive and can be difficult to grow (personal communication with R. Bicknell). Extracting and culturing HPECs from fresh tissue could potentially provide an easily accessible and economically viable source of HPECs. There would also be the possibility of isolating HPECs from normal and diseased (eg COPD) lung and comparing cell behaviours.

5.3 Extraction of MLEC (mouse lung endothelial cells)

Before an extraction of human HPECs was attempted extraction was performed using mouse lung tissue. This was performed under the supervision of Angel Armesilla at the University of Wolverhampton. This was performed according to the method in section 2.8. After negative selection the cells were reviewed every day by Dr Armesilla and the medium was changed every couple of days. Figure 5.1 demonstrates the appearance of the culture after negative selection.
Figure 5.1: Appearance of mouse lung culture after negative selection: Colonies of endothelial cells were observed under the microscope. One of these colonies is highlighted at day one after negative selection. Only one MLEC extraction was performed and these images are representative of this extraction. Images were donated by the University of Wolverhampton without a scale bar.
After five days, large endothelial colonies were visible under the microscope so positive selection of endothelial cells was performed in order to purify the culture further. After positive selection the cells were reviewed by Dr Armesilla daily and media was replaced every other day. The appearance of the cells and beads at day one after positive selection is in figure 5.2. Once the cells reached confluence the cells were split into two plates. After 15 days enough of the magnetic beads had disappeared to perform flow cytometry.

To confirm that MLECs had been isolated flow cytometry was performed. The histogram for the control, CD31 and CD102 stained cells is shown in figure 5.3. The flow cytometry demonstrated that the MLECs positively stained for CD31 and CD102 suggesting a pure endothelial isolate. (The negative selection was discarded and therefore flow cytometry was not performed on this.)

5.4 Extraction of Human Pulmonary Endothelial Cells (HPECs)

This method was initially based upon the MLEC extraction method in section 2.8 and the endothelial extraction method in section 2.3.2 (which was used to collect cells for RNA extraction). The final extraction method is in section 2.9.
Figure 5.2: Day one after positive selection for CD31 in mouse lung cell culture. Individual endothelial cells and small endothelial cell cultures can be seen. Magnetic beads are also visible. Only one MLEC extraction was performed and these images are representative of this extraction. Images were donated by the University of Wolverhampton without a scale bar.
Figure 5.3: A histogram of flow cytometry performed on isolated MLECs. This was performed on the MLEC isolate from figures 5.1-2 (n=1). The blue histogram represents control cells (unstained), the orange histogram represents CD31 stained cells (75.4% positivity) and the pink histogram represents CD102 stained cells (93.4% positivity). As there is a shift in the histograms for CD31 and CD102 it shows that the cells positively bind these antibodies and suggests they are endothelial cells.
5.4.1 Development of the extraction method

Initial method (16.2.16): Method based on MLEC extraction using collagenase V for digestion

This initial method is outlined in Appendix 4. There were large amounts of sticky partially digested tissue after collagenase digestion using this method. This made it very difficult to disaggregate cells after digestion.

On day one the plate was reviewed under the microscope. Large amounts of erythrocytes were present and it was impossible to see any attached cells. Therefore, the media was removed and the plate was washed with DMEM-F12 with glutamine (11320033; Life Technologies (Gibco)) (10 ml) twice. 10 ml fresh MLEC/HPEC medium (section 2.8.1) was added to the plate, which was reviewed again under the microscope. Single cells were seen attached to the plate, but these were sparse. The cells were placed into the incubator again at 37°C, 5% CO₂ overnight.

On day two small colonies of endothelial cells were seen (figure 5.4) in addition to multiple erythrocytes. The medium was changed to fresh MLEC/HPEC medium (10 ml) and the cells were placed back into the incubator (37°C, 5% CO₂).

The plate was later checked on day five. Unfortunately no colonies of endothelial cells were visible at this point and so the experiment was abandoned.
Figure 5.4: Appearance of mixed lung cell culture on day two, extraction one (different areas of the same 10 cm plate). Multiple erythrocytes can be seen and attached cells are sparse. Small colonies of endothelial cells are seen and have been highlighted. Image taken with Leica microscope, objective: N PLAN 5x/0.12 PH0. Scale bar represents 1000μm.
Extraction 2 (3.3.16): Method with reduced volume of lung tissue, various digestion times and 6-well plates for culture

After the failure of the first extraction attempts were made in order to improve the yield of cultured endothelial cells. Firstly, the volume of lung digested at each extraction was reduced to 1 cm$^3$ in order to limit the development of large volumes of sticky tissue. It is likely that endothelial cells were trapped in this sticky tissue during attempt 1 and were therefore lost at the filtration stage. During this attempt four different time periods were used during the digestion phase in order to optimise digestion time: 30 minutes, 1 hour, 2 hours and 4 hours. These digestions were all performed using different pieces of the same lung. As the endothelial cells were sparse in attempt 1 a 6 well plate was used rather than a 10 cm plate for each digestion. Endothelial cells produce growth factors and therefore closer colonies of endothelial cells should encourage cell proliferation.(493)

During the extraction attempt large amounts of sticky tissue occurred for the digestions at 30 minutes and 1 hour. Therefore, again, it was difficult to pass this through a 20 ml syringe to disaggregate the cells. There was no sticky tissue seen after the 2 and 4 hour digestions and these solutions passed easily in and out of a 20 ml syringe.

The cells from each digestion were reviewed on day one. Again, multiple erythrocytes occluded the view of attached cells. Therefore, the cells were washed with 4 ml of DMEM-F12 with glutamine (11320033; Life Technologies (Gibco)) prior to the addition of 5 ml MLEC/HPEC media (section 2.8.1) to each well. Very few attached cells were seen on any of the wells. However, the well from the 1 hour digestion had the most attached cells present. The cells were left in the incubator at 37$^\circ$C, 5% CO$_2$ before review on day 4.
On day 4 no viable cells were seen in the wells from the 30 minutes, 2 hour and 4 hour digestions. Elongated individual cells were seen in the well from the 1 hour digestion. The media was replaced in the well from the 1 hour digestion only and the cells were incubated at 37°C, 5% CO₂. On day 10 no viable cells were seen and so the experiment was abandoned. It was concluded that 1 hour remained the optimal time for digestion.

**Extraction 3 (16.3.16): Method with altered means of cell disaggregation and removal of erythrocytes using bronchoalveolar lavage**

Due to the difficulty in passing the digested tissue in and out of a 20 ml syringe this step was removed from the protocol. However, to encourage cell disaggregation the digested tissue was passed in and out of a 10 ml pipette four times instead.

As multiple erythrocytes were present in the cultures from extractions 1 and 2 it was postulated that these cells might be competing with endothelial cells for nutrients in the culture thereby limiting endothelial cell growth. Therefore, the culture method was repeated with an attempt to remove erythrocytes. A section of lung tissue was flushed with 0.9 % saline using a cannula and pressure bag (as in section 2.3.1) in an attempt to flush out the erythrocytes prior to digestion.

The cells were reviewed on day one. There were no viable cells so this experiment was abandoned.
Extraction 4 (16.3.16): Method with removal of erythrocytes using lymphoprep

A second method of erythrocyte removal was attempted. After the lung had been digested and filtered the lung solution was overlaid on lymphoprep (Axis-Shield; 1114545). This method is outlined in Appendix 4.

The cells were reviewed on day one and multiple endothelial colonies were seen. These colonies were also seen on day two. Samples were reviewed on day 5. There had been a proliferation of the colonies (figure 5.5).

The cells were reviewed again on day seven and appeared similar. Cells were frozen in 90 % fetal calf serum (Gibco; 10270106) and 10 % dimethyl sulfoxide (Sigma; D2650) (FCS/DMSO) at -80°C. Cells were defrosted and replated weeks later but did not retain their endothelial appearance so the experiment was abandoned.

Extraction 5 (16.3.16): Method with digestion using collagenase I

As collagenase I (rather than V) was used in the MLEC extraction this was used during the digestion phase.

The cells were reviewed on day one and multiple endothelial colonies were seen. On day two multiple small cells and clumps of cells were seen. Samples were reviewed on day 5. The collagenase I well continued to show clumps of cells (figure 5.6).

In a similar fashion to extraction 4, on day 7 the cells were frozen and replated several weeks later but did not retain their endothelial appearance so the experiment was abandoned.

Extraction 6 (16.3.16): Method with fibronectin

Fibronectin was used to coat a well of a 6 well plate in an attempt to improve the number of adherent endothelial cells.

There were no viable cells on day one and so the experiment was abandoned.
Figure 5.5: Appearance of mixed lung cell culture on day five, extraction four. A: Image from extraction using lymphoprep. Primarily endothelial cells are seen with some clumps of other cells. B: Image from extraction using lymphoprep but higher magnification. (n=1)
Images taken with Leica microscope: objective of A: N PLAN 5x/0.12 PH0, objective of B: C PLAN 10X/0.22 PH1
Scale bar represents 1000μm.
Figure 5.6: Appearance of mixed lung cell culture on day five, extraction five. A: Image from extraction using collagenase I. Multiple small cells and clumps of cells are seen with possible endothelial cells. B: Image from extraction using collagenase I but higher magnification. (n=1)
Images taken with Leica microscope: objective of A/C: N PLAN 5x/0.12 PH0, objective of B/D: C PLAN 10X/0.22 PH1
Scale bar represents 1000μm.
Extraction 7 (12.4.16): Positive selection for endothelial cells attempted

This method is outlined in Appendix 4. Four pieces of lung were digested and were plated on 4 wells of a 6 well plate.

The cells were reviewed on day one and the four wells showed clusters of cells that were not obviously endothelial in nature.

On day 3 one well was confluent with primarily endothelial like cells and was split into 2 new wells. The second well showed primarily endothelial cells and was reaching confluence. The third well showed multiple endothelial colonies and the fourth well showed a few sparse single cells only.

On day 6 the cells were reviewed again. All 5 wells showed primarily endothelial cells at near confluence. Therefore positive selection for endothelial cells was attempted.

Positive selection of endothelial cells using mixed cell cultures from extraction 7

This method was based on the MLEC positive selection method in addition to positive selection performed when extracting HPECs for RNA isolation (section 2.3.2). The full method of this initial attempt can be reviewed in Appendix 4.

The cells were reviewed on day one after selection. There were adherent cells visible in addition to beads.

The cells were next reviewed on day three. Each well showed a proliferation of cells.

The cells from each well were removed after using 2x trypsin (section 2.9.2) and were replated on to 0.1% gelatin (section 2.8.1) coated 10 cm plates and left in the incubator at 37°C, 5% CO₂ overnight. The cells were reviewed on day four.

Unfortunately the cells seen were very sparse and it was likely that cells had been lost during the removal via trypsin. Media was changed on day 6 and 10. By day fourteen the cells had an elongated appearance similar to fibroblasts. It is likely that
the endothelial cells had died and a few contaminating cells had overtaken the cultures. The experiment was then abandoned.

**Extraction 8 (12.5.16): Method with removal of erythrocytes using red cell lysis buffer**

Two samples of lung were processed during this extraction. One sample was processed according to the method in extraction 4. The other sample was split into 2 digests: one digest was processed according to extraction 4. The second digest was performed using red cell lysis buffer (5830-100, Cambridge Biosciences). This was attempted to see if red cell lysis buffer provided a more efficient way at removing erythrocytes. The red cell lysis buffer method is outlined in Appendix 4.

All wells were reviewed on day one and showed clusters of cells, which were not obviously endothelial in nature. There were fragments of cells visible in the well where red cell lysis buffer had been used. These cells were reviewed and the media was replaced every other day. Unfortunately none of these digests produced enough endothelial cells to attempt positive selection. The experiment was therefore abandoned. As there were large amounts of debris seen in the wells from the red cell lysis buffer this method was not attempted again.

**Extraction 9 (4.8.16): Method combining multiple lung digests**

This extraction was performed using the method from extraction 4. The tissue was divided into three prior to digestion and 3 separate digests were performed. At the end of the lymphoprep phase all 3 cell pellets were combined and resuspended into 5 ml MLEC/HPEC media (section 2.8.1) and were plated into one well of a 0.1% gelatin (section 2.8.1) coated 6 well plate. This was performed in order to achieve a higher density of endothelial cells and to encourage endothelial cell growth.
On day one clusters of cells that were not obviously endothelial in nature were seen. On day four sparse endothelial cell clusters were seen and the media was changed. On day six the endothelial cell clusters were enlarging and the media was changed again. On day eight there were endothelial cell clusters, but also fibroblast like cells seen. The media was changed and cells were reviewed at day 12. At this point endothelial like cobblestone clusters were seen with areas of fibroblast like cells in sheets. Positive selection for endothelial cells was performed at this point.

**Positive selection of endothelial cells using mixed cell cultures from extraction 9 – reduction in number of streptavidin beads and mixing of cells/beads using a wheel**

In an attempt to optimise this step two key changes were implemented. Firstly, the number of beads was reduced from the previous positive selection attempt as it was postulated that the addition of too many beads previously may have prevented some of the endothelial cells from being able to adhere to the gelatin coated plate. The number of beads used was chosen after reviewing existing literature on the extraction of pulmonary endothelial cells.(494) Secondly, to attempt better mixing of cells with Ulex-coated beads, the cells were removed from the culture plate and resuspended in bead solution. The cells/beads were then placed on a wheel in the cold room at 4°C for 30 minutes to give the cells time to bind to the beads (rather than adding the beads to cells fixed on a plate). The mixing of the cells on the wheel should encourage the proper mixing of cells and beads. This method can be seen in detail in 2.9.2.

The cells were reviewed on day one after positive selection. The majority of cells were present in the negative plate without beads and appeared to be elongated in sheets. There were scanty cells seen with beads on the positive plate. The cells were
reviewed on the following day and had similar appearances. The negative plate was
confluent and was therefore split into 2 wells of a 6 well plate after removal with
trypsin (section 2.9.2).

On day three the positive plate showed only scanty cells. One of the negative plates
was left alone, but the other plate had media replaced with DMEM-F12 with
glutamine (11320033; Life Technologies (Gibco))/10% FCS (Gibco; 10270106) and
penicillin/streptomycin (Gibco; 15140-122) (no growth factors). This was done in an
attempt to see if any cells would start forming a cobblestone appearance in the
absence of growth factors. This is because sometimes primary endothelial cell
cultures can have a more elongated appearance in the presence of growth factors
(personal communication with R. Bicknell).

On day 6 the cells on the positive plate had started to proliferate and the media was
changed. The cells on the negative plates both retained a fibroblast like appearance
despite the removal of growth factors on one of these plates. The negative plates
were therefore discarded.

On day nine the cells on the positive plate had reached confluence. These cells did
not have a typical cobblestone appearance of endothelial cells as the cells extracted
by Hewett and Murray did. (330) The cells appeared more ‘spiky’ and elongated. The
cells were removed using trypsin (section 2.9.2) and split into 3 wells of a 0.1%
gelatin (section 2.8.1) coated 6 well plate.

On day 10 the 3 wells appeared near confluence. The appearance of these cells can
be seen in figure 5.7. 2 wells were removed using trypsin and frozen in FCS/DMSO
(as in extraction 4) at -80°C before being transferred to a liquid nitrogen freezer. The
third well was also removed using trypsin (section 2.9.2) and split into 3 wells of a 0.1% gelatin coated 6 well plate.

The cells were reviewed on day 14. The cells on the three wells appeared confluent and more fibroblast like. The media was replaced on two wells with MLEC/HPEC media (section 2.8.1). The media on the third well was replaced with media without growth factors (DMEM-F12/10% FCS and penicillin/streptomycin) to see if this would lead to the cells forming a more cobblestone like appearance. The cells were reviewed on day 16, but all three wells continued to have a fibroblast like appearance. The cells were removed using trypsin and were frozen at -80°C in FCS/DMSO before being transferred to a liquid nitrogen freezer.

5.4.2 Flow cytometry of positively selected lung cells

To determine whether or not HPECs had been isolated in extraction 9 flow cytometry for CD31 (characteristic of endothelial cells) was performed according to section 2.11.2. The following conditions were performed for both lung cells and HUVECs:

- Unstained cells
- Secondary antibody only (anti-mouse, FITC)
- Secondary antibody and isotype antibody control (mouse IgG)
- Test (CD31 antibody (mouse, anti-human) and secondary antibody)

The lung cells (from the positive selection from extraction 9) and commercially bought HUVECs (Cellworks, ZHC-2102) were plated on to 0.1% gelatin (section 2.8.1) coated T25 (lung cells) and T75 (HUVECs) flasks 3 days prior to flow cytometry. The HUVECs were used as a positive control. Analysis of the flow cytometry performed using the software ‘FloJo’ can be seen in figures 5.8-5.11. In both HUVECs and lung cells the unstained cytogram was used initially and a gate was created by drawing a
plot around the main cell population (figures 5.8 and 5.10). The gates were used to limit results for the other cell conditions. Interestingly, the cytogram plots for the lung cells and HUVECs looked different from one another. This suggested that they represent two different types of cell populations and thus the lung cells were unlikely to be endothelial in nature.

The histogram for the unstained cells was compared to the histograms for the other cell conditions (figures 5.9 and 5.11). The HUVECs stained for CD31 demonstrated a shift in the histogram (figure 5.9) suggesting that they stained positively for CD31 (as expected). Table 5.1 demonstrates that percentage positive cells was 95.9% in the CD31 stained group and median fluorescent intensity (MFI) was higher in this group in comparison to the control groups. However, there was no shift in the histogram for the lung cells stained for CD31. Table 5.2 demonstrates that percentage positive cells was 3.92% in the CD31 stained group and MFI was similar in this group in comparison to the control groups. This suggests that the lung cells did not stain for CD31 and therefore were not endothelial cells.
Figure 5.7: Appearance of lung cells on day 10 after positive selection using Ulex coated magnetic beads. These cells were positively selected from lung cells isolated in extraction 9. These wells consist of cells grown after one passage (on day 9) and were cultured in MLEC/HPEC media (section 2.8.1) on 0.1% gelatin (section 2.8.1) coated 6 well plates. All three wells were cultured in identical conditions and were passaged from the same original well of cells and are therefore replicates. Cells appear elongated and spiky. The multiple small dots visible are magnetic beads. A&B: Images from well 1. C&D: Images from well 2. E&F: Images from well 3. Images taken with Leica microscope: objective of A/C/E: N PLAN 5x/0.12 PH0, objective of B/D/F: C PLAN 10X/0.22 PH1. Scale bar represents 1000 µm.
Figure 5.8: A cytogram plot from the program ‘FloJo’ for HUVECs: The dots represent individual cells. A gate has been drawn around the main cell population. N=1

Figure 5.9: A histogram plot from flow cytometry performed on HUVECs: The plot compares unstained cells (green) to cells stained for secondary antibody only (orange), isotype control (red) and CD31 (blue). There is a shift in the CD31 stained population indicating positive staining for the cell marker in question. N=1
Figure 5.10: A cytogram plot from the program ‘FloJo’ for lung cells: The dots represent individual cells. A gate has been drawn around the main cell population. The dots are distributed in a different way to the HUVEC dots and therefore suggest that this population of cells is different to HUVECs. N=1

Figure 5.11: A histogram plot from flow cytometry performed on lung cells: The plot compares unstained cells (green) to cells stained for secondary antibody only (orange), isotype control (red) and CD31 (blue). There is no shift in the CD31 stained population indicating negative staining for the cell marker in question. N=1
### Table 5.1: Flow cytometry results for HUVECs.

<table>
<thead>
<tr>
<th></th>
<th>Percentage positive cells</th>
<th>MFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>HUVEC: CD31</td>
<td>95.9 %</td>
<td>8.26</td>
</tr>
<tr>
<td>HUVEC: isotype control</td>
<td>9.37 %</td>
<td>5.69</td>
</tr>
<tr>
<td>HUVEC: secondary antibody only</td>
<td>8.86 %</td>
<td>6.02</td>
</tr>
<tr>
<td>HUVEC: unstained</td>
<td>10.9 %</td>
<td>5.74</td>
</tr>
</tbody>
</table>

This table outlines the four different HUVEC groups analysed by flow cytometry and the percentage positive of each group that fluoresced. The MFI for each group is also listed. HUVEC = Human Umbilical Vein Endothelial Cells. MFI = Median Fluorescent Intensity. N=1

### Table 5.2: Flow cytometry results for lung cells.

<table>
<thead>
<tr>
<th></th>
<th>Percentage positive cells</th>
<th>MFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cells: CD31</td>
<td>3.92 %</td>
<td>3.92</td>
</tr>
<tr>
<td>Lung cells: isotype control</td>
<td>3.51 %</td>
<td>4.13</td>
</tr>
<tr>
<td>Lung cells: secondary antibody only</td>
<td>3.18 %</td>
<td>4.1</td>
</tr>
<tr>
<td>Lung cells: unstained</td>
<td>4.26 %</td>
<td>4.03</td>
</tr>
</tbody>
</table>

This table outlines the four different lung cell groups analysed by flow cytometry and the percentage positive of each group that fluoresced. The MFI for each group is also listed. MFI = Median Fluorescent Intensity. N=1
5.5 Discussion

After several attempts we were not successful in reliably extracting and culturing HPECs from human lung tissue. Cellular appearances from some of the extractions do demonstrate endothelial-like appearances (figures 5.4 and 5.5). This might suggest that endothelial cells were successfully grown in culture, but these early endothelial colonies did not survive over time. As the initial digested lung is cultured as a mixture of cell types it is likely that this included contaminating cells such as fibroblasts. It is known that contamination of endothelial cell cultures with fibroblasts results in the death of endothelial cultures where the contaminating cells effectively take over the cell culture. (328) It is important to note that the classical HPEC extraction papers by Carley et al and Hewett and Murray removed the pleura from the lung they used in extractions whereas this was not performed in this study. (329, 330) It is possible that mesothelial cell contamination could therefore have resulted in mesothelial cells overtaking early endothelial cell cultures in this study. Even though positive selection using Ulex-coated magnetic beads was used in extraction 9 to remove contaminating cells this selection may have occurred too late to obtain enough endothelial cells to achieve a successful culture. It is likely that the cells seen in figure 5.7 (after positive selection) were in fact derived from contaminating cells during positive selection rather than from cells that bound successfully to the Ulex beads. This hypothesis is supported by the fact that after positive selection only a few scanty cells were present in the positive selection culture whereas most cells were present in the negative selection culture (derived from cells that did not bind to the Ulex coated beads). Also, the morphology of the cells from the negative culture was similar to cells in the positive culture. Therefore, if the method was to be altered
again, it would be sensible to remove the pleura prior to mincing and digesting the lung tissue to minimise mesothelial cell contamination. A change of media could also possibly help to minimise contaminating cells. It is possible to buy commercially made endothelial cell media in which one can add different growth factors. By using these sorts of media it would be possible to tailor make media to avoid growth factors that would encourage contaminating cell growth (such as fibroblast growth factor, FGF) and to use growth factors that would encourage primarily endothelial cell growth (such as vascular endothelial growth factor, VEGF). Perhaps by discouraging the growth of contaminating cells in this way larger amounts of endothelial cells would have been able to proliferate in the extractions above. Finally, another way of limiting the amount of contaminating cells in the culture would be to remove contaminating cells from the culture manually. This could be achieved by using a cell scraper followed by washing away the loose cells. However, in order to do this accurately one would have to remove contaminating cells under the microscope. This would therefore present a challenge unless performed under a microscope in a sterile tissue culture hood. Removing the cells outside of such a hood could easily result in infection. Alternatively, one could remove the endothelial colonies manually using cloning rings and replate the colonies on to new wells. This would also have to be performed under microscope however.

A major impediment in the growth of HPECs was the low number of cells present in the initial isolate. As mentioned in the text, a large amount of sticky partially digested tissue was present after digestion with collagenase. It is very likely that this tissue may have consisted of endothelial cells. Therefore, when this tissue was filtered to achieve a single cell solution many of the lung endothelial cells would have been lost.
Indeed, this was not a problem during the MLEC extraction and one pair of mouse lungs provided enough MLECs to cover a whole 10 cm plate. However, several digestions of human lung tissue could not provide enough HPECs to successfully cover one well of a 6 well plate. To address this issue two methods were employed. Firstly, the cell digest was passed in and out of a 20 cm syringe via a needle in an attempt to disaggregate cells. However, the partially digested tissue made this impossible. Secondly, the digestion time was increased to 2 and 4 hours. Increasing the digestion time resulted in cell suspensions without partially digested tissue, but this resulted in poor cell yields presumably as a result of cellular damage. An alternative method to reduce the amount of partially digested tissue would be to use longer digestion times with a different enzyme such as neutral protease dispase. Dispase is more gentle on cell membranes so less likely to cause cell damage like collagenase at longer digestions and was used for over 16 hours in the extractions performed by Hewett and Murray.(330) A more thorough digest of tissue would hopefully reduce the amount of partially digested tissue and therefore reduce the number of endothelial cells removed at filtration.

Another possible way to attempt to increase the yield of endothelial cells would be to perform positive selection earlier. During this study positive selection in extraction 9 was performed after the initial cell cultures were approaching confluence. However, at this point other cell types were also visible. On day 6 after extraction 9 there were enlarging visible endothelial-like colonies. By day 8 fibroblast like cells were visible and positive selection was not performed until day 12. At this point the fibroblast like cells had formed sheets and were increasing in number. Perhaps, if the positive selection had been performed at day 6-8 the positive selection process may have
selected endothelial cells successfully. The total number of cells would have been fewer, but potentially the proportion of endothelial cells may have been higher increasing the final number of endothelial cells after positive selection. Other groups have used this approach with success. For example, Burg et al performed positive selection for HPECs after 5-7 days of mixed cell culture with the main purpose of preventing growth of contaminating cells.(338)

It is possible that patient selection might also be important in increasing the yield of HPECs. Perhaps the MLEC extraction worked well as it was performed on a young mouse. In theory one might expect younger cells to proliferate more effectively than older cells. Unfortunately I only had access to lung tissue from older patients. The youngest patient in this part of the study was 57 years old. However, as nearly all lobectomies and pneumonectomies are performed for the investigation and treatment of lung cancer it is unlikely that many patients would be below the age of 50. Tissue not suitable for lung transplant could provide another source of lung tissue potentially from younger patients. However, tissue like this is available only very infrequently and is not available at predictable times limiting its use for HPECs needed on a regular basis. It might be supposed that the healthier the donated tissue is, the better the extracted HPECs might be at proliferation. Also, as tissue from patients with emphysema often has reduced pulmonary vasculature one might expect endothelial extractions from tissue from COPD patients to be poor.(205) However, extraction 4 was performed using tissue from COPD lung and the initial cultures did have an endothelial like appearance (figure 5.5). Also, other groups have successfully grown HPECs from COPD lung.(336) Therefore, excluding patients with COPD from HPEC extraction attempts may not be a sensible strategy. One could also argue for the
extraction of HPECs from COPD and non-COPD lung which would allow the comparison of COPD and non-COPD HPECs in various functional work. Therefore, as yet, it is not clear which patients might provide lung tissue more amenable to HPEC extraction. Perhaps unknown genetic expression differences between patients might affect the success of HPEC isolation. However, this would be impossible to determine prior to every extraction and would be impractical. Other groups have also noticed that some HPEC extractions work better than others without obvious causes (Personal communication with Dr P Hewett.)

Finally, it is possible that the position of the lung used for endothelial extraction could influence the success of HPEC extraction. Hewett and Murray discovered that peripheral lung tissue (less than 1 cm from the lung edge) appeared to produce greater yields of HPEC than more proximal pieces of lung tissue (Personal communication with Dr P Hewett.)(330) However, the use of distal tissue has not always been specified by other groups that have successfully extracted HPECs. (336, 338) Also, in any case, the tissue used in this study was generally derived from peripheral pieces of lung as far away from the tumour as possible.

5.6 Conclusions

In summary this chapter presents the results of multiple attempts at extracting HPECs from human lung tissue obtained at lobectomy/pneumonectomy. Despite initial success at extracting similar cells from mouse tissue no successful cultures of HPECs were obtained. The results presented suggest that HPECs initially grew in mixed cell cultures, but the growth of these cells was limited by the more rapid growth of contaminating cells. Methods to overcome this could include attempts to limit mesothelial cell contamination (such as dissection of the pleura prior to digestion) or
attempts to increase overall cell yield (such as longer digestion times with alternative enzymes such as dispase). It is clear that several alternative strategies could be employed in further extraction attempts and it would be useful to pursue these strategies in the future. Unfortunately this was beyond the scope of this project due to time constraints.
CHAPTER SIX
IDENTIFYING THE ROLE OF MICRORNA TARGETS IN ANGIOGENESIS
6.1 Hypothesis

Increasing the expression of miR-181b-3p in endothelial cells will have effects on endothelial cell function that are relevant to the pathogenesis of COPD.

6.2 Aims of this chapter

This chapter outlines the functional work performed in human umbilical vein endothelial cells (HUVECs) to investigate the effect of miR-181b-3p on endothelial behaviour and function. Results for each assay performed will be outlined prior to a discussion and an overall conclusion for the chapter.

6.3 The role of miR-181b-3p in angiogenesis

To investigate whether or not miR-181b-3p influences endothelial function artificial mimics and inhibitors of miR-181b-3p were transfected into HUVECs and the altered expression of miR-181b-3p was confirmed using qPCR. Pooled HUVEC were purchased (Lonza) and used for all experiments in order to limit the effect of single nucleotide polymorphisms (SNPs) on endothelial function. HUVECs were grown and split in HUVEC media (section 2.11.3) and used at passage 3 or 4.

HUVECs were chosen for cellular functional work primarily as attempted cultures of HPECs failed (chapter 5) and therefore a ready supply of HPECs was not available. HPECs can be purchased, but these are very expensive and previous attempts at culturing them successfully in the Bicknell laboratory had proven difficult. HUVECs are relatively cheap and easy to culture (personal communication with R. Bicknell) and were therefore selected as the chosen cell for functional work.
6.3.1 Optimisation of miR-181b-3p mimic into HUVECs

An initial experiment was performed to transfect different concentrations of miR-181b-3p mimic at the following concentrations: 1 nM, 5 nM and 10 nM (figure 6.1). Each condition was repeated in quadruplicate. The expression of miR-181b-3p in each sample was then assessed by qPCR. All concentrations of mimic were successful at significantly increasing the expression of miR-181-3p (p<0.001). However, tukey’s test demonstrated that the expression of miR-181b-3p was significantly increased in cells transfected with 10 nM concentration of mimic in comparison to 1 nM (p<0.001) and 5 nM (p=0.021). Thus, 10 nM was used for further experiments.

To confirm that transfecting 10 nM miR-181b-3p mimic provided a stable increased level of miR-181b-3p expression in HUVECs three further experiments were performed confirming that miR-181b-3p expression was maintained up to three days after transfection (figure 6.2). Each experiment was repeated in quadruplicate. These experiments demonstrated that the expression of miR-181b-3p decreased over time but was still significantly higher in the transfected cells three days after transfection.
Figure 6.1: Optimisation of miR-181b-3p mimic concentration. The figure shows the relative expression of miR-181b-3p in HUVECs transfected with miR-181b-3p mimic at 1 nM, 5 nM and 10 nM in comparison to negative siRNA (control). qPCR was used to determine the expression of miR-181b-3p in each group. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-181b-3p in the mimic groups was normalised to that of the control group. The double delta Ct method was used to compare the expression levels. Expression was calculated from the delta Ct levels for miR-181b-3p expression in the mimic groups (n=4 per group) and the mean delta Ct for the control group (n=4). Figures represent mean fold changes; error bars being SEM fold change.

A one-way ANOVA test was used to determine significant differences between mean delta Ct in all groups (p<0.001). Tukey’s test was used post-hoc to look for significant differences between the mean delta Ct in the control group and each concentration of mimic. The tukey’s test also demonstrated that miR-181b-3p expression was significantly higher in cells transfected with 10 nM mimic in comparison to 1 nM (p<0.001) and 5 nM (p=0.021)

***= p<0.001 (Tukey’s test). (Negative siRNA = 1).
Figure 6.2: Relative expression of miR-181b following mimic transfection at 10nM concentration. qPCR was used to determine the expression of miR-181b-3p in each group. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-181b-3p in the mimic groups was normalised to that of the control groups. The double delta Ct method was used to compare the expression levels.

A: Mean relative expression of miR-181b-3p in HUVECs transfected with an artificial miR-181b-3p mimic 2 days post transfection. This represents the mean expression from 3 independent experiments. B: Mean relative expression of miR-181b-3p in HUVECs transfected with an artificial miR-181b-3p mimic 3 days post transfection. This represents the mean expression from 3 independent experiments.

A t-test was used to determine significant differences between the mean relative expression in the mimic group and the control group. n=3 in each group

***= p<0.001. Negative siRNA expression in both cases = 1
6.3.2 Optimisation of miR-181b-3p inhibitor into HUVECs

An initial experiment was performed to transfect different concentrations of miR-181b-3p inhibitor at the following concentrations: 1 nM, 10 nM and 50 nM (figure 6.3). The expression of miR-181b-3p in each sample was then assessed by qPCR. Only the 10 nM concentration resulted in a reduction of miR-181b-3p expression although the effect was small (23.4% reduction). To confirm this effect the experiment was repeated using the 10 nM concentration of inhibitor only repeating all conditions in quadruplicate (figure 6.4). This demonstrated a 53.9% reduction in miR-181b-3p expression but did not reach statistical significance.

In a similar way to miRNA mimic optimisation three further experiments were performed to confirm that inhibition of miR-181b-3p expression was maintained for three days after transfection (figure 6.5). Each experiment was repeated in quadruplicate. However, the miR-181b-3p was not successfully inhibited in any experiments. This is most likely because the background expression of miR-181b-3p was low in control cells (Ct values were primarily >35). Therefore, this suggests that miR-181b-3p expression could not be inhibited any further. MiR-181b-3p inhibitors were therefore not used in further experiments.
Figure 6.3: Optimisation of miR-181b-3p inhibitor concentration. The figure shows the relative expression of miR-181b-3p in HUVECs transfected with miR-181b-3p inhibitor at 1 nM, 10 nM and 50 nM concentration in comparison to a negative inhibitor (control). qPCR was used to determine the expression of miR-181b-3p in each group. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-181b-3p in the inhibitor groups was normalised to that of the control group. The double delta Ct method was used to compare the expression levels. Expression was calculated from the delta Ct levels for miR-181b-3p expression in the inhibitor groups (n=1 per group) and the control group (n=1). n=1 in each group so significance tests were not performed.
Figure 6.4: Inhibition of miR-181b-3p using an inhibitor at 10 nM. The figure shows the relative expression of miR-181b-3p in HUVECs transfected with miR-181b-3p inhibitor and negative inhibitor (control). qPCR was used to determine the expression of miR-181b-3p in each group. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-181b-3p in the inhibitor group was normalised to that of the control group. The double delta Ct method was used to compare the expression levels. Expression was calculated from the delta Ct levels for miR-181b-3p expression in the inhibitor group (n=4) and the mean delta Ct for the control group (n=4). Figures represent mean fold changes; error bars being SEM fold change. A t-test was used to determine significance between mean delta Ct values.
Figure 6.5: Relative expression of miR-181b following inhibitor transfection. qPCR was used to determine the expression of miR-181b-3p in each group. RNU48 was used as the house-keeping small RNA to which the data was normalised. Expression of miR-181b-3p in the inhibitor groups were normalised to that of the control groups. The double delta Ct method was used to compare the expression levels.

A: Mean relative expression of miR-181b-3p in HUVECs transfected with an artificial inhibitor 2 days post transfection. This represents the mean expression from 3 independent experiments.

B: Mean relative expression of miR-181b-3p in HUVECs transfected with an artificial inhibitor 3 days post transfection. This represents the mean expression from 3 independent experiments.

A t-test was used to determine significant differences between the mean relative expression in the inhibitor group and the control group. n=3 in each group.
6.3.3 The effect of miR-181b-3p overexpression on cell growth

To determine whether or not miR-181-3p affects endothelial cell growth three separate experiments were performed. Briefly, HUVECs that had been transfected with miR-181b-3p mimic, negative siRNA, lipofectamine only or optimem only were plated into 3 wells (per condition) of 3 different 12 well plates. The number of cells per condition was counted 1-3 days post transfection (figure 6.6). There did not appear to be any difference in cell number in the miR-181b-3p, negative siRNA or lipofectamine conditions in any of the 3 experiments suggesting that miR-181b-3p overexpression does not have a significant effect on cell growth. However, cell number was significantly increased in the optimem only group which suggests that exposure to lipofectamine may impair cell growth.

6.3.4 The effect of miR-181b-3p overexpression on the cell cycle

To determine whether miR-181b-3p overexpression lead to any changes in the cell cycle, flow cytometry after propidium iodide staining was performed. HUVECs that had been transfected with miR-181b-3p mimic, negative siRNA, lipofectamine or optimem only were compared. This was performed in three separate experiments. Figure 6.7 demonstrates the proportion of cells in phases G1, S and G2 across all experiments. The proportion of cells in each phase did not differ between groups significantly suggesting that overexpression of miR-181b-3p did not alter the cell cycle. An example of the histogram plots of cell counts for cell cycle analysis from one experiment can be seen in figure 6.8.
Figure 6.6: Cell growth assay. This figure shows the number of cells day 1, 2 and 3 after transfection in cells transfected with miR-181b-3p mimic, negative siRNA, lipofectamine or optimem. Numbers represent the mean from three separate experiments. The Kruskal-Wallis test was used to determine significant differences between groups. N=3 in each group. Day 1; p = 0.046, day 2; p = 0.001, day 3; p < 0.001.

Figure 6.7: Cell cycle analysis according to flow cytometry. The figure shows the proportion of cells in G1, S and G2 phases of the cell cycle in cells transfected with miR-181b-3p mimic, negative siRNA, lipofectamine or optimem. The values represent the mean from three separate experiments. A one-way ANOVA test was used to determine significant differences between groups. Tukey’s test was used post-hoc to look for significant differences between individual groups. There were no significant differences detected. N=3 in each group.
Figure 6.8: Histogram plots of cell counts for cell cycle analysis. This represents a typical histogram plot from one experiment. A = HUVECs transfected with 181b-3p mimic. B = HUVECs transfected with a negative siRNA. C = HUVECs transfected with lipofectamine only. D = HUVECs exposed to optimem only (no transfection). There was no significant difference in the proportion of cells in each phase of the cell cycle compared between groups.
6.3.5 The effect of miR-181b-3p overexpression on HUVEC tube formation

To determine the effect of miR-181b-3p on HUVEC tube formation three separate matrigel assays were performed. HUVECs that had been transfected with miR-181b-3p mimic, negative siRNA, lipofectamine only or optimem only were seeded on to matrigel and images were taken of each condition at 6, 12, 18 and 24 hours. The images were uploaded and analysed by the program ImageJ. Specifically, number of nodes, number of junctions, number of meshes and total mesh area were assessed in each experiment (figures 6.9-6.12).

The four groups were significantly different in terms of number of nodes and number of junctions at each time point. A tukey’s test was performed post-hoc to compare individual groups which demonstrated that number of nodes and number of junctions were significantly reduced in HUVECs transfected with miR-181b-3p mimic compared to negative siRNA (p<0.001) at 6, 12 and 18 hours. The effect was lost at 24 hours.

The four groups were significantly different in terms of number of meshes at 6, 12 and 18 hours only, but were significantly different at all time points for total mesh area. A tukey’s test was performed post-hoc to compare individual groups which demonstrated that number of meshes and total mesh area were significantly reduced in HUVECs transfected with miR-181b-3p mimic compared to negative siRNA at 6 (p<0.001) and 12 (p<0.05) hours. The effect on number of meshes and total mesh area was lost at 18 hours.

Figure 6.13 demonstrates an example of images taken at 6, 12, 18 and 24 hours after HUVECs were added to matrigel in the miR-181b-3p mimic and negative siRNA conditions. Reduced tube formation in the mimic transfected HUVECs can be seen.
Figure 6.9: Number of nodes formed by HUVECs imbedded in matrigel. The figure shows the mean number of nodes present 6, 12, 18 and 24 hours after HUVECs were seeded on matrigel. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to look for significance between groups. *** = p <0.001 (ANOVA). Tukey’s test was used to compare groups which demonstrated that number of nodes was significantly reduced in the mimic group in comparison to the negative siRNA group at 6, 12 and 18 hours only (p<0.001 in all cases.)

Figure 6.10: Number of junctions formed by HUVECs imbedded in matrigel. The figure shows the number of junctions present 6, 12, 18 and 24 hours after HUVECs were seeded on matrigel. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to look for significance between groups. *** = p <0.001 (ANOVA). Tukey’s test was used to compare groups which demonstrated that number of junctions was significantly reduced in the mimic group in comparison to the negative siRNA group at 6, 12 and 18 hours only (p<0.001 in all cases.)
Figure 6.11: Number of meshes formed by HUVECs imbedded in matrigel. The figure shows the number of meshes present 6, 12, 18 and 24 hours after HUVECs were seeded on matrigel. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to look for significance between groups. *** = p<0.001 (ANOVA). Tukey’s test was used to compare groups which demonstrated that number of meshes was significantly reduced in the mimic group in comparison to the negative siRNA group at 6 (p<0.001) and 12 (p=0.028) hours.

Figure 6.12: Mesh area formed by HUVECs imbedded in matrigel. The figure shows the total mesh area at 6, 12, 18 and 24 hours after HUVECs were seeded on matrigel. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to look for significance between groups (p<0.001 at 6-18 hours, p=0.004 at 24 hours). *** = p <0.001, ** = p<0.01 (ANOVA). Tukey’s test was used to compare groups which demonstrated that total mesh area was significantly reduced in the mimic group in comparison to the negative siRNA group at 6 (p<0.001) and 12 and (p=0.042) hours.
Figure 6.13: An example of images from a matrigel experiment. A: HUVECs transfected with miR-181b-3p mimic. B: HUVECs transfected with negative siRNA. There is reduced tube formation seen in HUVECs transfected with the mimic. Scale bars represent 800μm.
6.3.6 The effect of miR-181b-3p overexpression on HUVEC wound healing

To assess the effect of miR-181b-3p overexpression on endothelial migration and wound healing three separate scratch wound experiments were performed. HUVECs that had been transfected with miR-181b-3p mimic, negative siRNA, lipofectamine only or optimem only were cultured in monolayers and a linear wound was created. The cultures were then imaged at 0, 6, 12, 18, 24 hours to determine the speed at which the wounds were closed. Each experiment was performed in triplicate. There was no significant difference between groups at 6, 18 and 24 hours although there was a trend for wound closure to be faster in the optimem only group (figure 6.14). Wound closure was significantly different between groups at the 12 hour time point. A tukey’s test was performed post-hoc to compare results between groups which demonstrated that the optimem only group had a smaller percentage wound area remaining compared to the lipofectamine group (p=0.015). This was the only significant difference detected in the post-hoc analysis.

As wound closure did not vary significantly between the miR-181b-3p mimic and negative siRNA groups these results suggest that miR-181b-3p does not have a significant effect on endothelial migration and wound closure. Figure 6.15 demonstrates an example of images taken at 0, 6, 12, 18 and 24 hours after a scratch wound was completed in the miR-181b-3p mimic and negative siRNA conditions. No obvious difference can be seen between groups.
Figure 6.14: Percentage wound area remaining during scratch wound assay. The average percentage wound remaining for each condition is recorded at 0, 6, 12, 18 and 24 hours. The data represents the mean from three separate experiments. An ANOVA test was used to look for significance between groups. $p=0.021$ at 12 hours. * = $p < 0.05$ (ANOVA). A tukey’s test demonstrated that the optimem only group had a smaller percentage wound area remaining compared to the lipofectamine group ($p=0.015$).
Figure 6.15: An example of images from a scratch wound experiment. A: HUVECs transfected with miR-181b-3p mimic. B: HUVECs transfected with negative siRNA. The remaining wound area is highlighted. There was no significant difference between groups. Scale bars represent 400 µm.
6.3.7 The effect of miR-181b-3p overexpression on HUVEC spheroid sprouting

To determine how miR-181b-3p overexpression in HUVECs alters endothelial sprouting three spheroid assays were performed. In each experiment HUVECs transfected with miR-181b-3p mimic, negative siRNA, lipofectamine only or optimem were used to create spheroids. Confocal microscopy was used to image the spheroids. Five spheroids were imaged per condition. The images were uploaded into ImageJ and were analysed using the ‘Spheroid Analysis’ plugin developed at the University of Birmingham by Victoria Salisbury. (366)

Number of sprouts, total sprout length and average sprout length were significantly different between groups (figures 6.16-18). Tukey’s tests were performed in post-hoc analyses for number of sprouts, total sprout length and average sprout length. These demonstrated that spheroids in the miR-181b-3p mimic group had reduced sprouting in comparison to the negative siRNA group suggesting that miR-181b-3p overexpression significantly reduced endothelial sprout formation. (Number of sprouts, p = 0.021; total sprout length, p = 0.003; average sprout length, p = 0.003.)

An example of a typical endothelial spheroid formed from HUVECs transfected with miR-181b-3p mimic and negative siRNA can be seen in figure 6.19. Reduced sprouting in the mimic transfected spheroid can be seen.
Figure 6.16: Number of endothelial sprouts from spheroids. The figure shows the total number of endothelial sprouts for each condition. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to determine significant differences between groups. *** = p <0.001 (ANOVA). A tukey’s test demonstrated that there was a reduced number of sprouts in the mimic group in comparison to the negative siRNA group (p=0.021).

Figure 6.17: Total length of endothelial sprouts from spheroids. The figure shows the total endothelial sprout length for each condition. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to determine significant differences between groups. *** = p <0.001 (ANOVA). A tukey’s test demonstrated that there was a reduced total sprout length in the mimic group in comparison to the negative siRNA group (p=0.003).
Figure 6.18: Average length of endothelial sprouts from spheroids. The figure shows the average endothelial sprout length for each condition. The numbers represent the mean from 3 separate experiments. An ANOVA test was used to determine significant differences between groups \( (p=0.003) \). ** = \( p < 0.01 \) (ANOVA). A tukey’s test demonstrated that there was a reduced average sprout length in the mimic group in comparison to the negative siRNA group \( (p=0.003) \).
Figure 6.19: An example of an endothelial spheroid assay. A: Spheroid created from HUVECs transfected with miR-181b-3p. B: Spheroid created from HUVECs transfected with negative siRNA. There is reduced sprouting seen in the spheroid transfected with the miR-181b-3p mimic. Spheroids were imaged using a Zeiss 780 Zen confocal microscope. Scale bar represents 200μm.
6.3.8 Identification of miR-181b-3p targets

As the above results suggest an inhibitory role for miR-181b-3p in angiogenesis an attempt was made to identify possible miR-181b-3p gene targets that could be responsible for this effect. Computer prediction programs were used in section 3.3.3 to help to identify targets. Two predicted targets were identified in the top 100 list of targets that looked particularly promising: \textit{ELTD1} (adhesion G protein-coupled receptor L4) and \textit{IL8} (interleukin 8). These have both been previously associated with an upregulation of angiogenesis.\cite{497, 498} Therefore, qPCR was performed to look for the expression of \textit{ELTD1} and \textit{IL8} in cells that have been transfected with miR-181b-3p mimic (n=4) and compare this to expression in cells that had been transfected with a negative siRNA (n=4). Neither gene was significantly suppressed in the samples transfected with miR-181b-3p (figure 6.20). This suggests that either the target prediction was false and miR-181b-3p is acting through suppression of other genes, or miR-181b-3p targets \textit{ELTD1} and \textit{IL8} at a translational level. To identify this western blotting of ELTD1 and IL8 could be performed to identify reduced protein levels.
Figure 6.20: Mean relative expression of *ELTD1* and *IL8* in cells transfected with miR-181b-3p mimic and negative siRNA. qPCR was used to determine the expression of each mRNA target in HUVECs transfected with miR-181b-3p mimic in comparison to HUVECs transfected with a negative siRNA (control). *Flotillin 2* was used as the house-keeping gene to which the data was normalised. Expression of mRNA targets in HUVECs transfected with the mimic was normalised to that of HUVECs transfected with the negative siRNA. The double delta Ct method was used to compare the expression levels. The charts represent the mean expression from four separate experiments. A t-test was used to determine significant differences between the 2 groups. No significant difference between groups was found.
6.4 Discussion

6.4.1 The evidence for miR-181b-3p’s involvement in angiogenesis

In summary the results above suggest that miR-181b-3p overexpression significantly impairs tube formation and sprouting of endothelial cells *in vitro*. This does not appear to be through alteration of cell growth or changes to the cell cycle. Matrigel assays measure several angiogenic processes including adhesion, migration, cell alignment and formation of tubes.(499) The assay primarily assesses the differentiation stage of angiogenesis where cells form vessel sprouts.(500) This is supported by results from the spheroid assay in which endothelial sprouting was significantly reduced by miR-181b-3p overexpression. Therefore, the results of this study suggest that increased miR-181b-3p levels in endothelial cells impair angiogenesis. The spheroid assay possibly provides more conclusive evidence than the matrigel assay that this is the case. It is known that endothelial cells become undifferentiated in culture and therefore may function differently from cells *in vivo*. (501) However, when endothelial cells form spheroids the outer cells become differentiated and more closely resemble cells found in the vasculature.(502) They also produce capillary sprouts in collagen which have lumen and are lined by a flattened layer of endothelial cells rather like capillaries *in vivo*. (503) Therefore it is likely that spheroid assays more accurately reflect the process of angiogenesis. Consequently the results of the spheroid assay in this study strongly support the case for impaired angiogenesis in the context of miR-181b-3p overexpression.

Wound closure however does not appear to be influenced by miR-181b-3p. This is unexpected due to the consistent reduction in tube formation and endothelial sprouting seen when miR-181b-3p is overexpressed. However, scratch wound
assays primarily assess cell migration and proliferation, but do not assess the same number of functions as the matrigel and spheroid assays. Therefore, this could suggest that the reduction in angiogenesis is not primarily as a result of changes in cell migration; instead it could be due to cell alignment and endothelial sprout formation. The lack of change seen in cell proliferation in the scratch wound assay supports the cell growth experiments, which were non-significant between miR-181b-3p overexpression and control siRNA groups. Development of the pulmonary vasculature appears to be key in the development of alveolar structures in the developing lung. More recently it has also been postulated that the pulmonary microvascular circulation is important in maintenance of alveolar structures by producing factors termed ‘angiocrines’. Preserving the lung vasculature could therefore promote growth and preserve the architecture of alveoli. One possible angiocrine is retinoic acid. Retinoic acid stimulates tube formation in lung endothelial cells in vitro suggesting it enhances angiogenesis. Patients with emphysema also have increased levels of CYP26A1 (an enzyme that degrades retinoic acid) in the endothelium. This suggests that patients with emphysema may have reduced pulmonary endothelial retinoic acid levels. A reduction of retinoic acid could potentially result in dysregulated maintenance and repair of alveoli resulting in alveolar damage and emphysema. In support of this, murine models have demonstrated that retinoic acid can enhance lung growth both after pneumonectomy and in the setting of drug-induced disrupted alveolar development in adult mice and rats. These findings also suggest that the lung may have more intrinsic regenerative ability than previously thought further supporting the concept of lung regeneration as a treatment strategy for emphysema.
Retinoic acid agonists have been used in patients with alpha 1 antitrypsin deficiency and moderate-severe COPD (REPAIR trial). However, this trial did not demonstrate an improvement in lung density on CT scanning. This could mean that retinoic acid is unlikely to improve COPD in human disease, but as the study was only carried out over one year and COPD progresses slowly, the study may simply have not been long enough to detect a treatment effect. The study also focused on patients with more severe disease and it is possible that too much tissue destruction existed in these patients to demonstrate an effect. Perhaps retinoic acid would have been more effective in patients with milder disease with more residual lung tissue.

As previously discussed, there is little known information on miR-181b-3p’s function other than it appears to upregulate epithelial-mesenchymal transition in breast cancer cells. However, some studies do support the results of this study as they show evidence for a reduction in angiogenesis secondary to miRNAs related to miR-181b-3p. For example, miR-181b (the precursor of miR-181b-3p) is reduced in chick embryos when arsenic is added which is associated with a reduction in angiogenesis. In mouse models, injection of glioma cells expressing miR-181b mixed with matrigel resulted in reduced angiogenic responses in the matrigel plugs. This was not seen when the glioma cells expressed a negative control miRNA suggesting miR-181b was responsible for the reduced angiogenic response. This effect appeared to be through the suppression of insulin-like growth factor 1 (IGF-1), a tyrosine kinase receptor that regulates multiple pathways including cell proliferation. Results from one other study appear to conflict this however. As previously mentioned, one group identified that miR-181b was increased in retinoblastoma cells and enhanced angiogenesis through targeting
Programmed cell death-10 (PDCD10) and GATA binding protein 6 (GATA6). (392) There are no other studies looking at angiogenesis in relation to miR-181b-3p as yet and thus it is difficult to draw firm conclusions as a result of the evidence available.

6.4.2 Limitations

The main limitation of this study is the relatively small number of assays performed. Both the matrigel and scratch wound assay are relatively simple and test a restricted number of functions associated with angiogenesis. It has been previously argued that the tubes formed in matrigel assays do not have lumens typical of capillaries and therefore are limited in assessing angiogenesis function. (513) (Although this has been disputed by other groups who have identified such lumens under microscopy. (514)) Nevertheless, as the matrigel assay involves seeding endothelial cells on to a matrix only and does not involve input from other cell types it is unlikely to fully recreate the angiogenic process. This limitation has been partly assuaged by the use of the spheroid assay which produces capillary-like networks with lumen that more accurately reflect capillaries in vivo. (503) However, complex assays which include support from other cell types (co-culture assays) may be more accurate still. (515) Therefore, by limiting the assays performed to simple tests the degree of involvement of miR-181b-3p in angiogenesis function is not yet certain. However, due to time constraints it was not possible to perform co-cultures during the course of the study. These assays are considered in more depth in section 7.5.1.

The other main limitation is that HUVECs were used for the functional work rather than HPECs. This was primarily due to the ease of availability of HUVECs in comparison to HPECs which proved difficult to isolate (chapter 5) and are expensive to purchase. It could therefore be argued that one cannot assume that HPECs would
behave in the same fashion if the level of miR-181b-3p expression in these cells was increased in the same manner as HUVECs. Indeed, studies have shown that although glycoprotein expression in different microvascular endothelial cells is similar between organ types there are key differences which, in theory, could influence cell function.(331, 332) If time permitted it would therefore have been prudent to repeat studies in HPECs.

6.5 Conclusions

In summary this chapter presents evidence that miR-181b-3p expression can be increased using transfection in HUVECs with success and without influencing the proportion of cells in each phase of the cell cycle or cell growth. Preliminary studies suggest that 181b-3p overexpression reduces tube formation and sprouting in HUVECs and could therefore limit angiogenesis. Previous results (chapter 3) demonstrate that miR-181b-3p is increased in HPECs in patients with COPD. An increase in miR-181b-3p in this setting could therefore impair HPEC function. This might be important in the development of emphysema as other work supports the notion that the lung vasculature is key in the structural maintenance of alveoli. Correcting miR-181b-3p expression levels could therefore be a possible route for treating emphysema by promoting support of alveolar structure and regeneration. However, further work would be required to support or refute this hypothesis.
CHAPTER 7

FINAL DISCUSSION AND FUTURE WORK
7.1 Summary of project findings

7.1.1 Key project findings

This project has demonstrated that messenger RNA (mRNA) and microRNA (miRNA) expression in human pulmonary endothelial cells (HPECs) is significantly altered in patients with COPD. Four novel endothelial targets were also validated using qPCR as upregulated in COPD, one of which was also upregulated in lung cancer potentially providing a shared endothelial target between the two conditions. Pathway analysis further supported the hypothesis that COPD and lung cancer are related, as several pathways were upregulated in COPD with mechanistic links to lung cancer. For example, the transforming-growth factor-β (TGF-β) pathway that promotes epithelial-mesenchymal transition. (146) Finally, one endothelial target, miR-181b-3p, was also validated functionally using endothelial assays. Upregulation of miR-181b-3p reduced tube formation and sprouting in endothelial cells suggesting that this miRNA suppresses angiogenesis. Upregulation of miR-181b-3p may thus contribute to the endothelial dysfunction seen in COPD. (218) HPECs are important in the structural maintenance of alveoli (506) and therefore upregulation of miR-181b-3p may lead to alveolar damage promoting the development of emphysema.

This study has importance for two main reasons. Firstly, it is the only known study that has investigated miRNA and mRNA expression in HPECs in COPD and lung cancer. The majority of previous work in this area has used whole lung samples, which may miss signals from individual cell types. It is also difficult to attribute miRNA or mRNA expression findings to distinct cell types. Thus, the identified targets from this study are more likely to represent genuine endothelial targets as they were identified from this cell type alone. Secondly, this study is important, as it has
identified novel COPD targets. Current treatment for COPD manages symptoms only and does not alter the course of the disease. Therefore, identification of new possible treatment targets is essential and miR-181b-3p may provide a new mechanism of treatment through alveolar regeneration and repair.

7.1.2 Summary of the analysis of microRNA and messenger RNA expression in lung endothelium in COPD and lung cancer.

The analysis was carried out by completing microarrays for both miRNA and mRNA expression in isolated pulmonary endothelial cells from patients with and without COPD. Where possible, lung tumour tissue was also collected to assess expression in lung cancer. Subsequent analysis using the program SAM (Significance Analysis of Microarrays) demonstrated significant differences in both miRNA and mRNA expression in patients with and without COPD. Fourteen endothelial targets were chosen for further validation with qPCR which resulted in the identification of four new targets: miR-23c, -181b-3p, -429 and LTA4H. MiR-429 was then validated as upregulated in lung cancer providing a possible shared target between the two conditions.

All miRNA targets have little published data about them and are novel targets in the setting of COPD. However, LTA4H has been previously shown to be elevated in COPD and has potential roles in inflammation and transendothelial migration.

7.1.3 Summary of findings from Ingenuity Pathway Analysis

Microarray results comparing COPD to non-COPD patients were uploaded into the program Ingenuity Pathway Analysis (IPA). This identified several disease associations with the results. Interestingly, cancer was the most significant disease process associated with the mRNA IPA results and the third most significant disease process associated with miRNA results. This further supports the hypothesis that
cancer and COPD are related. In addition, multiple cancer-related cellular mechanisms were dysregulated in COPD such as cell growth and proliferation providing evidence as to why the two conditions might be associated. To investigate the function of the targets identified above further in-depth pathway downstream analysis was performed. This linked targets to disease relevant pathways such as Eicosanoid signalling and the transforming growth factor-β pathway.

7.1.4 Summary of findings from Human Pulmonary Endothelial Cell extractions.

Multiple attempts at isolating HPECs from human lung tissue were performed using modifications of the isolation of endothelial cells technique for RNA extraction (section 2.3.2). This technique used the mincing of lung tissue and digestion by collagenase followed by filtration and plating of the cell solution. Purification of the HPEC extraction was attempted using Ulex europaeus agglutinin-1 coated magnetic beads. However, no modifications of this technique resulted in reliable enough cultures of HPECs to use said cells in endothelial functional work. This mainly appeared to result from rapid growth of contaminating cells that overcame the growth of HPECs.

7.1.5 Summary of findings from functional studies investigating the role of miR-181b-3p in angiogenesis.

Artificial miR-181b-3p mimics were transfected into HUVECs using lipofectamine. This resulted in the increased expression of miR-181b-3p over at least three days. Transfected cells were used in functional experiments to determine the effect of miR-181b-3p overexpression on angiogenesis. Cell growth and the cell cycle did not seem to be altered by transfection. Matrigel studies showed reduced tube formation
in HUVECs transfected with miR-181b-3p mimics. Spheroid studies also showed reduced endothelial sprouting in these cells. These results suggest that miR-181b-3p overexpression impairs angiogenesis. Scratch wound adhesion was unaffected by the addition of a miR-181b-3p mimic.

**7.2 Possible mechanisms to explain miR-181b-3p's role in angiogenesis**

MiRNAs generally act to suppress target mRNAs either at the translational level or result in mRNA degradation. Therefore, it is likely that miR-181b-3p's effects on angiogenesis are a result of target suppression of angiogenesis-related genes. MiR-181b-3p targets were searched for using two computer programs (TargetScan (368) and DIANA-Micro-T-CDS (369)). Two genes within the top 100 targets have been previously related to angiogenesis: *ELTD1* (adhesion G protein-coupled receptor L4) and *IL8* (interleukin 8). Both of these targets appear to upregulate angiogenesis. Consequently, it could be postulated that miR-181b-3p overexpression in COPD results in a reduction in angiogenesis secondary to *ELTD1* and *IL8* suppression. To investigate this, a qPCR was performed to look for a reduction in *ELTD1* and *IL8* expression in HUVECs transfected with miR-181b-3p mimic (section 6.3.8). No reduction was seen suggesting that *ELTD1* and *IL8* were not true targets or miR-181b-3p inhibits the expression of *ELTD1* and *IL8* at a translational level. This could be investigated further using Western Blotting and confirmed with luciferase reporter assays (section 7.5.2). If one looks further down the list of potential targets then other possible angiogenesis genes can also be identified. Firstly, vascular endothelial growth factor C (*VEGFC*) was on the predicted miR-181b-3p target list albeit at position 4050. VEGF is a key mediator in angiogenesis and it is possible that this may be a particularly relevant pathway to
explain miR-181b-3p’s function. Other possible targets that could possibly explain reduced tube formation secondary to miR-181b-3p overexpression include tubulin related genes such as TBCA (tubulin cofactor A) and TUBB1 (tubulin beta 1). Tubulin forms an important part of the cytoskeleton and thus disruption to tubulin related genes could impair endothelial tube formation. (518, 519)

Further clues to miR-181b-3p’s function can be found in the IPA (chapter 4). MiR-181 was associated with network 1 (figure 4.24) where it appears to directly inhibit EPHB1 (EPH receptor B1), a tyrosine kinase which has previously been associated with angiogenesis and neovascularization. (520) This, in addition with EPHB1’s associations with increased cell motility, could explain why miR-181b-3p results in reduced tube formation in matrigel studies. (521) Additionally, another member of this network, RHOA (ras homolog family member A, a GTPase), also contributes to cell motility by contributing to actin-myosin contractility. (522) This could also possibly help to explain the difference seen in endothelial tube formation and sprouting with miR-181b-3p overexpression although miR-181 and RHOA are not directly linked in the network.

Another network identified in figure 4.26 also contains miR-181. However, the miRNA is not linked to any targets in the network and so this particular figure does not help to shed light on miR-181 function per se. However, it may help to explain why miR-181 is upregulated. In this network miR-181 appears to be directly activated by Smad 2/3 which is a regulator of transcription. (481) Smad is known to be upregulated in COPD and appears to be associated with smoking. (483) Perhaps increased smoke exposure in COPD patients induces Smad which consequently results in miR-181b-3p activation.
In summary, there are several pathways, which may be relevant in the context of miR-181-3p overexpression and angiogenesis. However, it would be necessary to validate these pathways first before any conclusions about miR-181b-3p’s function can be drawn. Without further exploration of miR-181b-3p’s functional pathways it would limit the use of this miRNA as a potential treatment target as it would be difficult to predict unwanted effects.

7.3 The potential of miR-181b-3p as a treatment target

Recently, miRNAs have been increasingly investigated as treatment targets. In fact, a miRNA inhibitor (Miravirsen) has been used successfully in phase 2 trials in the treatment of Hepatitis C. (523) This inhibitor binds to miR-122 to form a stable heteroduplex thus impairing its function. (523) MiR-122 is required for the propagation of the Hepatitis C virus and Miravirsen successfully reduced Hepatitis C titres in a phase 2 trial. (523) There are no miRNA treatment studies in COPD as yet, but animal models do provide some evidence that miRNA based therapy may be of use in pulmonary disease. For example, in a lung cancer mouse model systemic delivery of a miR-34a mimic led to a reduction in tumour area. (524)

As technology is already in place to develop therapeutic miRNA inhibitors it is possible that such an inhibitor could be developed for miR-181b-3p, which could be trialled in animal models in the first instance. However, there are considerations that would need to be taken into account before employing such a strategy. For example, COPD is an umbrella term for multiple clinical phenotypes. It has already been discussed (section 1.4.1) that the endothelium appears to behave differently in the settings of chronic bronchitis and emphysema. (210, 211) Therefore, it is possible that a drug that inhibits miR-181b-3p and enhances angiogenesis in the context of
emphysema would not be appropriate to use in patients with chronic bronchitis with increased VEGF expression. This leads to another consideration: targeting of therapy. It is conceivable that inhibition of miR-181b-3p might have undesirable effects on other pulmonary cell types. To minimise this would require consideration of which route (local versus systemic) and method (i.e. what vector) would be most appropriate. (525) This may be particularly difficult in endothelial cells as a systemic delivery would most likely be required. MiRNAs can also target multiple mRNAs involved in many different cellular pathways; the computer prediction programs predicted thousands of targets for miR-181b-3p. Therefore, as well as unwanted effects in other cell types, it is quite possible that miR-181b-3p inhibition could also have unwanted effects in endothelial cells too. Consequently, a thorough investigation of miR-181b-3p and its targets would be warranted before developing drugs to inhibit the miRNA. Finally, it is unlikely that miR-181b-3p will be an effective target to treat emphysema on its own. COPD is a complex disease involving multiple cells and inflammatory pathways and several miRNAs have previously been identified as being up or down regulated in COPD. (525) Hence, to be successful in treating COPD, investigating and targeting multiple abnormal miRNAs would probably be the treatment strategy most likely to succeed.

To summarize, miRNAs present new possible treatment targets and miRNA inhibitors have been used in previous phase 2 human trials suggesting that targeting miR-181b-3p may be a valid strategy in COPD. However, as miRNAs have varied targets and COPD has a heterogeneous nature, miR-181b-3p’s function needs to be explored in further detail prior to developing drugs for its’ inhibition.
7.4 Conclusions

This study has identified significant differences in both miRNA and mRNA expression in human pulmonary endothelial cells (HPECs) between patients with and without COPD. Fourteen potential targets known to be expressed in endothelial tissue were assessed further, four of which were validated by qPCR. One target, miR-429, was also upregulated in lung cancer HPECs. Another target, miR-181b-3p, has been validated functionally and appears to reduce endothelial tube formation and sprouting suggesting an inhibitory role in angiogenesis.

Previous work suggests that there is an abnormal amount of endothelial tissue in COPD which behaves in an atypical fashion. Furthermore, there is evidence that vasculature has an important supportive role in pulmonary tissue promoting the growth and preserving the structure of the alveoli. Faulty pulmonary vascular tissue could therefore result in the dysregulation of alveolar maintenance and repair resulting in alveolar damage and emphysema. Consequently, targeting and normalizing the levels of the aforementioned endothelial targets could provide a new mechanism of treatment for COPD which is focused on repair and regeneration of lung tissue.

Patients with COPD are at an increased risk of lung cancer despite adjusting for smoking history. Thus, the identification of miR-429 as a target increased in both COPD and lung cancer provides an exciting new potential shared target between the two conditions. This target needs to be functionally validated but targeting and modifying this miRNA could provide a new treatment route in both conditions. In
COPD patients targeting miR-429 may perhaps offer a possibility of treating an at risk patient group before the development of lung cancer.

7.4.1 Reflections on the study hypothesis

This study was undertaken to explore the hypothesis that lung cancer and COPD might have a common pathogenesis. The endothelium plays an important role in transendothelial migration (TEM), the process by which inflammatory cells such as the neutrophil cross the endothelial barrier into the lung. As inflammation is also associated with cancer, it is possible that dysregulation of the endothelium resulting in upregulation of TEM and inflammation could promote the development of cancer in the COPD lung. This study supports that miRNA and mRNA expression is significantly altered in COPD. The fact that one of the identified endothelial targets in this study (miR-181b-3p) alters endothelial function implies that the miRNA and mRNA changes seen in COPD pulmonary endothelium may have functional effects. MiR-429, another identified target, was upregulated in COPD and upregulated to a further extent in lung cancer. This suggests that changes in the COPD pulmonary endothelium might represent a pre-malignant state that can be targeted.

However, the functional studies from chapter 6 indicate that miR-181b-3p reduces angiogenesis. This could potentially result in emphysema (through lack of support of alveolar epithelial cells) but does not fit with current knowledge about lung cancer in which angiogenesis is upregulated. Therefore, miR-181b-3p is unlikely to represent a shared endothelial target. As miR-429 is upregulated in both COPD and lung cancer it is a more likely shared target but further work would need to be done to validate miR-429 before this hypothesis could be proven. For example,
transendothelial migration studies (section 7.5.4) would be necessary to support the hypothesis that miR-429 would increase TEM (and consequently inflammation increasing cancer risk). In conclusion therefore, this work supports the theory that the pulmonary endothelium behaves in a dysfunctional manner in COPD and has identified potential targets that could contribute to this. However, a shared COPD/lung cancer target has not been functionally validated in this study. Thus, further work (section 7.5) would be required to fully explore the hypothesis of this thesis (section 1.7). In particular, transendothelial migration studies are required to determine whether or not the identified targets influence TEM and inflammation seen in COPD, which could predispose the lung to malignant changes.

7.5 Future work

Identified future work can be divided into four sections. Firstly, the further investigation of RNA targets identified in this study using more complex endothelial functional assays. Secondly, the identification of miRNA targets to recognize potential treatment pathways. Thirdly, transendothelial assays to investigate the role of identified targets in the interaction of endothelial and inflammatory cells. Lastly, the development of drugs and in vitro and in vivo investigation of treating RNA targets.

7.5.1 Endothelial functional assays

**MiR-181b-3p**

In the first instance it would be useful to further investigate miR-181b-3p’s effects on angiogenesis by means of more complex angiogenesis assays, such as co-culture assays. These assays co-culture endothelial cells with other human cells in a matrix. The main role of supportive cells in these assays is to secrete extracellular matrix and
growth factors (526) resulting in endothelial cells forming anastomosing tubules over
the course of 1-2 weeks.(527) Matrigel assays look at cell-cell assembly rather than
true endothelial migration and sprouting angiogenesis, hence this assay is more
reflective of the in vivo situation.(526) MiR-181b-3p appears to reduce endothelial tube
formation, thus it would be interesting to see whether or not angiogenesis is reduced in
this assay.

Another co-culture of interest uses endothelial cells with pulmonary epithelial cells of
bronchial or alveolar origin.(528) The epithelial cells start to form bronchiolar or
alveolar-like structures in vitro depending on their cell of origin.(528) These cultures
have primarily been used to look at epithelial cell development, but it would be
intriguing to see whether or not alteration of the supporting endothelial cells with a miR-
181b-3p mimic would result in alteration of epithelial cell structures.(528) As current
evidence suggests that pulmonary vasculature is key in the development and
maintenance of lung tissue one might expect to see malformation of epithelial
structures in these assays.(505, 506)

**MiR-429, miR-23c and LTA4H**

It would be ideal to perform similar functional assays (matrigel, scratch wound spheroid
and co-culture assays) with these other three validated targets. The miRNA targets
could be investigated in an identical manner to miR-181b-3p using commercially
bought miRNA mimics or inhibitors transfected into endothelial cells using
lipofectamine. Overexpression of LTA4H could be achieved using lentivirus
transfection.(365) Underexpression could be performed in a similar fashion to the
miRNA work using commercially available siRNA.(304) Overexpression/underexpression would require confirmation at the RNA level using
quantitative PCR (qPCR) and, in the case of LTA4H, at the protein level using Western Blot analysis.

7.5.2 MiRNA target identification

MiRNAs function by targeting and (usually) suppressing the function of mRNAs.(368) Therefore, after identifying miRNAs that appear to have a role in endothelial function the next critical step would be to identify which mRNAs are being targeted in order to have these effects. The easiest way to do this is via computer prediction programs such TargetScan discussed in Chapter 3.(368) However, these programs all have different algorithms producing different results and can produce several false positives.(529) Therefore, it is necessary to confirm computer prediction results in the laboratory setting before one can confidently identify a miRNA target.

One way of looking for miRNA targets would be to perform gene expression analysis (such as with a microarray, Chapter 3) in endothelium in which the miRNA had been over-expressed. This technique has been used successfully to identify miRNA targets in other cell types.(516) The results of this gene expression analysis could be compared to similar results from endothelium transfected with a negative control siRNA (short inhibitory RNA). Genes that were significantly downregulated could likely be targets of the miRNA. However, there are two limitations to this approach. Firstly, miRNAs can regulate transcriptional levels of their targets, detectable by microarray screening. However, miRNAs often work by translational blocking and therefore this type of targeting would not be visible by microarray screening.(517) Secondly, this approach identifies both indirect and direct mRNA targets and cannot identify the miRNA binding sites.(530) Therefore, although this approach would
support findings from computer prediction programs the results would still need to be validated in some way.

One example of a laboratory method able to confirm mRNA-miRNA binding is through use of a luciferase-reporter assay. This technique involves the insertion of mRNA target sequences into a firefly luciferase reporter vector, which is inserted into a cell. A miRNA is then transfected into the cell and, if the miRNA binds to the putative mRNA sequence, there is a reduction in luciferase activity. This is a useful way of confirming specific mRNA-miRNA interactions, and could validate potential targets identified in computer predictions or via gene expression screening. However, this method cannot identify multiple targets at once and is labour intensive.

In order to overcome difficulties with luciferase-reporter assays immunoprecipitation methods have been developed. These involve immunoprecipitation of the RNA-Inhibiting-Silencing-Complex (RISC) in which miRNAs bind to mRNA. It is then possible to identify the targeted mRNAs using microarray analysis. In more recent years this technique has been further advanced by stabilizing the mRNA-miRNA interactions initially using crosslinking with ultraviolet radiation (high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation; HITS-CLIP). Ultraviolet cross-linking can also be further stabilized by incubating photoactivatable nucleosides with cells in culture (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation; PAR-CLIP). These techniques allow the researcher to accurately look for mRNA-miRNA interactions on a large scale. Unfortunately, the techniques also require specialist equipment and are far more costly in comparison to simpler reporter assays, which limit them for wide range use.
Another technique for the potential to identify multiple miRNA targets at once utilizes streptavidin beads in a similar fashion to the pulmonary endothelial cell extraction in section 2.3.2. MiRNA is biotinylated and miRNA/mRNA duplexes are captured (after cell lysis) using streptavidin beads. mRNA targets can then be identified by microarray analysis. This method is advantageous has it does not require irradiation prior to cell analysis. However, in a similar manner to immunoprecipitation technologies microarray is required which drastically increases cost in comparison to reporter methods.

Finally, proteomic techniques have been developed to identify miRNA targets such as SILAC (stable isotope labelling with amino acids in cell culture). These help to screen for changes in protein expression in relation to introduction of a miRNA. These methods have the advantage that they look at protein rather than gene expression and therefore identify expression changes post translation. However, currently their depth of coverage is not as great as gene expression approaches which would result in missing potential targets. These approaches may also identify indirect targets and therefore further validation with a luciferase reporter assay would be necessary.

7.5.3 Immunohistochemistry

After identification of miRNA targets a way of validating these in humans on the protein level would be to use immunohistochemistry on formalin fixed paraffin embedded sections. By staining both healthy lung and COPD lung sections one could identify whether or not the protein staining for the targets varied between COPD and non-COPD. It would also be possible to use endothelial markers (such as PECAM-1) to identify target protein expression in lung endothelium.
7.5.4 Neutrophil Transendothelial Migration studies

As previously discussed, transendothelial migration (TEM) may be a way that the endothelium influences the development of COPD due to the fact that neutrophils play an important role in the inflammatory response in COPD.\(^{(54)}\) In order to reach the lung tissue neutrophils must bind to, and migrate through, the endothelium.\(^{(188)}\) Therefore, to fully investigate the role of the above mentioned targets in COPD it would be necessary to perform TEM assays after overexpression/underexpression of miRNAs or mRNAs of interest in endothelial cells (as above). The altered endothelium could then be used in two types of co-culture assay, after isolation of neutrophils from blood using density gradient centrifugation with Percoll.\(^{(537)}\) Firstly, static co-cultures could be used to monitor neutrophil interaction with cultured endothelium. Static co-cultures involve adding isolated neutrophils to the surface of a confluent endothelial culture.\(^{(538)}\) Phase-contrast video-microscopy is then used to identify the proportion of neutrophils that have successfully transmigrated across the endothelial layer.\(^{(538)}\) These assays would give an idea of whether or not the above targets affect TEM. However, in order to monitor TEM in real time flow co-cultures would be required. This requires mounting microslides containing confluent endothelium on a phase-contrast, video microscope in a Perspex chamber.\(^{(539)}\) Isolated neutrophils are then drawn through the microslides using a syringe pump.\(^{(539)}\) A video recording can be performed of neutrophil transmigration across the HPEC layer.\(^{(539)}\) By comparing the results of these assays to the same assays using endothelium exposed to negative siRNA one can identify the influence of the targets on TEM. I have been successful in gaining grant funding to carry out this work from the British Lung Foundation (ERN_17-0008).
7.5.5 In vitro drug response

After identifying functional changes in endothelium with the alteration of a target of interest it would be necessary to identify how treatment of these alterations affects endothelial function. For example, there are already drugs that target LTA4H (section 3.5.2) such as 4MDM (which selectively increases LTA4H aminopeptidase activity) or ARM1 (a selective LTA4H hydrolase inhibitor).(423, 425) One could perform similar endothelial functional studies to above with increased endothelial LTA4H expression with and without these drugs in order to identify whether or not these medications could target the effects of increased LTA4H on endothelial cells. It may also be interesting to identify drug responses to airway cells on an ‘airway-on-a-chip’ co-culture in vitro. These co-cultures consist of airway epithelium with supportive endothelium that experiences fluid flow (540) and allow measurement of inflammatory markers in response to drug treatments in the co-culture. If 4MDM or ARM1 resulted in reduced inflammation on the airway-on-a-chip this may suggest that targeting LTA4H would be an effective strategy to reduce pulmonary inflammation in COPD.

As the other validated RNAs are all miRNAs and therefore likely have multiple mRNA targets they are less likely to be useful as drug targets. This is a result of an increased chance of unwanted side effects. Therefore, it would be more useful to identify possible drug targets from identified miRNA targets rather than using the miRNAs themselves.

7.5.6 In vivo drug response

In a similar way to above, treatment of LTA4H/miRNA targets could be tested in animal models of COPD. It is possible to induce emphysema in mice using chronic smoke exposure.(423) This would provide an ideal model to test whether or not treatment of the targets influences COPD development. Advantages of using mouse models
include their relatively low cost and the large availability of gene sequences and antibodies. (541) It is also possible to produce genetically modified mice to investigate the role of certain genes in COPD. (541) However, there are key differences between mouse and human pulmonary anatomy, which should be taken into account when interpreting mouse model results. For example, mice have fewer submucosal glands, less airway branching and do not have respiratory bronchioles. (542) This means caution needs to be applied when using these models as small airway remodeling and increased mucus production are important in COPD pathogenesis. (543) Smoke induced models of COPD also induce relatively mild disease (GOLD stage I/II) rather than severe COPD. (544) This suggests that effective interventions seen in models may not be effective in more severe cases of COPD in humans.

7.5.7 Macrophages

Finally, the macrophage targets were not followed up in this thesis. This would be a larger undertaking, involving all the steps included in chapters 3, 4 and 6, plus elements analogous to that described here for the endothelium (future work section). One of the miRNAs associated with COPD appears to target iron regulation genes, which we have recently shown to be relevant to malignancy (545), and aligns to genetic associations of COPD (546, 547) hence this pathway might be of particular interest for studies of cancer-COPD shared pathogenesis. This is likely to be the subject of further grant applications by our group.
References


ClinicalTrials.gov [https://clinicaltrials.gov/]


359. Qiagen. Guidelines for miRNA mimic and miRNA inhibitor experiments. For miRNA research. 2015.

Salisbury V. High Resolution Imaging and Analysis of Endothelial Tubulogenesis and Blood Vessel Formation: University of Birmingham; 2017.


Yuan JS, Reed A, Chen F, Stewart CN, Jr. Statistical analysis of real-time PCR data. BMC bioinformatics. 2006;7:85.


Pre-treatment serum levels of bFGF and VEGF and its clinical significance in endometrial carcinoma. Gynecol Oncol. 2013;128(3):454-60.


600. Pasieka Z, Stepien H, Komorowski J, Kolomecki K, Kuzdak K. Evaluation of the levels of bFGF, VEGF, sICAM-1, and sVCAM-1 in serum of patients with thyroid cancer. Recent Results Cancer Res. 2003;162:189-94.


651. Maeng YS, Kwon JY, Kim EK, Kwon YG. Heterochromatin Protein 1 Alpha (HP1alpha: CBX5) is a Key Regulator in Differentiation of Endothelial Progenitor Cells to Endothelial Cells. Stem Cells. 2015;33(5):1512-22.


Park J, Park SY, Kim JH. Leukotriene B4 receptor-2 contributes to chemoresistance of SK-OV-3 ovarian cancer cells through activation of signal transducer and activator of transcription-3-linked cascade. Biochim Biophys Acta. 2016;1863(2):236-43.


Hay A, Wood S, Olson D, Slater DM. Labour is associated with decreased expression of the PGF2alpha receptor (PTGFR) and a novel PTGFR splice variant in human myometrium but not decidua. Mol Hum Reprod. 2010;16(10):752-60.


Keating GL, Reid HM, Eivers SB, Mulvaney EP, Kinsella BT. Transcriptional regulation of the human thromboxane A2 receptor gene by Wilms' tumor (WT)1 and


<table>
<thead>
<tr>
<th>Cytokine</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>Function</td>
</tr>
<tr>
<td>G protein coupled receptor</td>
<td>Growth Factor</td>
</tr>
<tr>
<td>Ion channel</td>
<td>Kinase</td>
</tr>
<tr>
<td>Ligand-dependent Nuclear Receptor</td>
<td>Mature microRNA</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>
Peptidase

Phosphatase

Transcription Regulator

Translation Regulator

Transmembrane Receptor

Transporter

Non-targeting interactions

Activation, causation, expression, localization, membership, modification, molecular cleavage, phosphorylation, protein-DNA interactions, protein-RNA interactions, regulation of binding, transcription

Inhibition, ubiquitination

Inhibits and acts on

Leads to
Processing yields

miRNA targeting

Translocation

Reaction

Enzyme catalysis

Direct interaction

Indirect interaction

Increased measurement

Decreased measurement

Predicted activation

Predicted inhibition

Leads to activation

Leads to inhibition

Findings inconsistent with state of downstream molecule
Effect not predicted

List of gene names

<table>
<thead>
<tr>
<th>IPA</th>
<th>Figure/Table</th>
<th>Gene symbol</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016 mRNA</td>
<td>4.19: Hypersensitivity reaction</td>
<td>ADIPOQ</td>
<td>Adiponectin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BGN</td>
<td>Biglycan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAMP</td>
<td>Cathelicidin Antimicrobial Peptide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCL3L3</td>
<td>C-C motif chemokine ligand 3 like 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCNA2</td>
<td>Cyclin A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHRM5</td>
<td>Cholinergic Receptor, Muscarinic 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DCD</td>
<td>Dermcidin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DCN</td>
<td>Decorin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DHFR</td>
<td>Dihydrofolate reductase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2F1</td>
<td>E2F transcription factor 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCER1A</td>
<td>Fc fragment of IgE receptor Ia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCGR2B</td>
<td>Fc fragment of IgG receptor IIb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GPR34</td>
<td>G protein-coupled receptor 34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HHIP</td>
<td>Hedgehog inhibitory protein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HMGCS2</td>
<td>3-hydroxy-3-methylglutaryl-CoA synthase 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HPGDS</td>
<td>Haematopoietic prostaglandin D synthase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HRH3</td>
<td>Histamine receptor H3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IGKV1D-8</td>
<td>Immunoglobulin kappa variable 1D-8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL1A</td>
<td>Interleukin 1 alpha</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KLK4</td>
<td>Kallikrein related peptidase 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LCE3D</td>
<td>Late cornified envelope 3D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LTA4H</td>
<td>Leukotriene A4 Hydrolase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MKI67</td>
<td>Marker of proliferation Ki-67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MME</td>
<td>Membrane metalloendopeptidase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NLRP12</td>
<td>NLR family pyrin domain containing 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIP5K1B</td>
<td>Phosphatidylinositol-4-phosphate 5-kinase type 1 beta</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PTGS1</td>
<td>Prostaglandin-endoperoxide synthase 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S100A7</td>
<td>S100 calcium binding protein A7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SYK</td>
<td>Spleen associated tyrosine kinase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGM6</td>
<td>Transglutaminase 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TP53</td>
<td>Tumour protein 53</td>
</tr>
<tr>
<td>2014 mRNA</td>
<td>4.20: Allergy</td>
<td>ADIPOQ</td>
<td>Adiponectin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AQP3</td>
<td>Aquaporin 3</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1S</td>
<td>Complement C1S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>Complement C3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAMP</td>
<td>Cathelicidin Antimicrobial Peptide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNA2</td>
<td>Cyclin A2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCD</td>
<td>Dermcidin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESR1</td>
<td>Oestrogen receptor 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCER1A</td>
<td>Fc fragment of IgE receptor la</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCGR2B</td>
<td>Fc fragment of IgG receptor Iib</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMGCS2</td>
<td>3-hydroxy-3-methylglutaryl-CoA synthase 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPGDS</td>
<td>Haematopoietic prostaglandin D synthase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGKV1D-8</td>
<td>Immunoglobulin kappa variable 1D-8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLK4</td>
<td>Kallikrein related peptidase 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LCE3D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTA4H</td>
<td>Leukotriene A4 Hydrolase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MK167</td>
<td>Marker Of Proliferation Ki-67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MME</td>
<td>Membrane metalloendopeptidase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIP5K1B</td>
<td>Phosphatidylinositol-4-phosphate 5-kinase type 1 beta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTGER3</td>
<td>Prostaglandin E Receptor 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTGS1</td>
<td>Prostaglandin-endoperoxide synthase 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S100A7</td>
<td>S100 calcium binding protein A7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYK</td>
<td>Spleen associated tyrosine kinase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGM6</td>
<td>Transglutaminase 6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**4.21: Metabolism of Eicosanoid**

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADIPOQ</td>
<td>Adiponectin</td>
</tr>
<tr>
<td>CADM1</td>
<td>Cell adhesion molecule 1</td>
</tr>
<tr>
<td>CAMP</td>
<td>Cathelicidin Antimicrobial Peptide</td>
</tr>
<tr>
<td>CYP1A2</td>
<td>Cytochrome P450 family 1 subfamily A member 2</td>
</tr>
<tr>
<td>ERBB4</td>
<td>Erb-b2 receptor tyrosine kinase 4</td>
</tr>
<tr>
<td>ESR2</td>
<td>Estrogen receptor 2</td>
</tr>
<tr>
<td>FCER1A</td>
<td>Fc fragment of IgE receptor la</td>
</tr>
<tr>
<td>FCGR2B</td>
<td>Fc fragment of IgG receptor Iib</td>
</tr>
<tr>
<td>FGFR3</td>
<td>Fibroblast growth factor receptor 3</td>
</tr>
<tr>
<td>HPGDS</td>
<td>Haematopoietic prostaglandin D synthase</td>
</tr>
<tr>
<td>IGF1</td>
<td>Insulin like growth factor 1</td>
</tr>
<tr>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
</tr>
<tr>
<td>LTA4H</td>
<td>Leukotriene A4 hydrolase</td>
</tr>
<tr>
<td>PLA2G2A</td>
<td>Phospholipase A2 group IIA</td>
</tr>
<tr>
<td>PLA2G4F</td>
<td>Phospholipase A2 group IVF</td>
</tr>
<tr>
<td>PTGS1</td>
<td>Prostaglandin-endoperoxide synthase</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>SYK</td>
<td>Spleen associated tyrosine kinase</td>
</tr>
<tr>
<td>TACR1</td>
<td>Tachykinin receptor 1</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumour protein 53</td>
</tr>
</tbody>
</table>

### 4.22: Synthesis of fatty acid

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACSM3</td>
<td>Acyl-CoA synthetase medium-chain family member 3</td>
</tr>
<tr>
<td>ADIPOQ</td>
<td>Adiponectin</td>
</tr>
<tr>
<td>AWAT2</td>
<td>Acyl-CoA wax alcohol acyltransferase 2</td>
</tr>
<tr>
<td>CADM1</td>
<td>Cell adhesion molecule 1</td>
</tr>
<tr>
<td>CAMP</td>
<td>Cathelicidin Antimicrobial Peptide</td>
</tr>
<tr>
<td>ERBB4</td>
<td>Erb-b2 receptor tyrosine kinase 4</td>
</tr>
<tr>
<td>ESR2</td>
<td>Estrogen receptor 2</td>
</tr>
<tr>
<td>FCER1A</td>
<td>Fc fragment of IgE receptor Ia</td>
</tr>
<tr>
<td>FCGR2B</td>
<td>Fc fragment of IgG receptor IIb</td>
</tr>
<tr>
<td>FGFR3</td>
<td>Fibroblast growth factor receptor 3</td>
</tr>
<tr>
<td>HPGDS</td>
<td>Haematopoietic prostaglandin D synthase</td>
</tr>
<tr>
<td>IGF1</td>
<td>Insulin like growth factor 1</td>
</tr>
<tr>
<td>INS</td>
<td>Insulin</td>
</tr>
<tr>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
</tr>
<tr>
<td>LTA4H</td>
<td>Leukotriene A4 hydrolase</td>
</tr>
<tr>
<td>PLA2G2A</td>
<td>Phospholipase A2 group II A</td>
</tr>
<tr>
<td>PLA2G4F</td>
<td>Phospholipase A2 group IVF</td>
</tr>
<tr>
<td>PTGS1</td>
<td>Prostaglandin-endoperoxide synthase 1</td>
</tr>
<tr>
<td>RGN</td>
<td>Regucalcin</td>
</tr>
<tr>
<td>SCD5</td>
<td>Stearoyl-CoA desaturase 5</td>
</tr>
<tr>
<td>SYK</td>
<td>Spleen associated tyrosine kinase</td>
</tr>
<tr>
<td>TACR1</td>
<td>Tachykinin receptor 1</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumour protein 53</td>
</tr>
</tbody>
</table>

### 4.23: Synthesis of lipid

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCG5</td>
<td>ATP binding cassette subfamily G member 5</td>
</tr>
<tr>
<td>ACSM3</td>
<td>Acyl-CoA synthetase medium-chain family member 3</td>
</tr>
<tr>
<td>ADIPOQ</td>
<td>Adiponectin</td>
</tr>
<tr>
<td>AGTR2</td>
<td>Angiotensin II receptor type 2</td>
</tr>
<tr>
<td>AWAT2</td>
<td>Acyl-CoA wax alcohol acyltransferase 2</td>
</tr>
<tr>
<td>C3</td>
<td>Complement C3</td>
</tr>
<tr>
<td>CADM1</td>
<td>Cell adhesion molecule 1</td>
</tr>
<tr>
<td>CAMP</td>
<td>Cathelicidin Antimicrobial Peptide</td>
</tr>
<tr>
<td>DGAT2</td>
<td>Diacylglycerol O-acyltransferase 2</td>
</tr>
<tr>
<td>ERBB4</td>
<td>Erb-b2 receptor tyrosine kinase 4</td>
</tr>
<tr>
<td>ESR1</td>
<td>Estrogen receptor 1</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name and Description</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>ESR2</td>
<td>Estrogen receptor 2</td>
</tr>
<tr>
<td>ETV1</td>
<td>ETS variant 1</td>
</tr>
<tr>
<td>FCER1A</td>
<td>Fc fragment of IgE receptor Ia</td>
</tr>
<tr>
<td>FCGR2B</td>
<td>Fc fragment of IgG receptor Iib</td>
</tr>
<tr>
<td>FGFR3</td>
<td>Fibroblast growth factor 3</td>
</tr>
<tr>
<td>FGFR3</td>
<td>Fibroblast growth factor receptor 3</td>
</tr>
<tr>
<td>FSHR</td>
<td>Follicle stimulating hormone receptor</td>
</tr>
<tr>
<td>GPAM</td>
<td>Glycerol-3-phosphate acyltransferase, mitochondrial</td>
</tr>
<tr>
<td>GPAT2</td>
<td>Glycerol-3-phosphate acyltransferase 2, mitochondrial</td>
</tr>
<tr>
<td>HPGDS</td>
<td>Haematopoietic prostaglandin D synthase</td>
</tr>
<tr>
<td>IGF1</td>
<td>Insulin like growth factor 1</td>
</tr>
<tr>
<td>INS</td>
<td>Insulin</td>
</tr>
<tr>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
</tr>
<tr>
<td>LPCAT2</td>
<td>Lysophosphatidylcholine acyltransferase 2</td>
</tr>
<tr>
<td>LTA4H</td>
<td>Leukotriene A4 hydrolase</td>
</tr>
<tr>
<td>MC5R</td>
<td>Melanocortin 5 receptor</td>
</tr>
<tr>
<td>P2RX1</td>
<td>Purinergic receptor P2X 1</td>
</tr>
<tr>
<td>PIP5K1B</td>
<td>Phosphatidylinositol-4-phosphate 5-kinase type 1 beta [Homo sapiens]</td>
</tr>
<tr>
<td>PLA2G2A</td>
<td>Phospholipase A2 group IIa</td>
</tr>
<tr>
<td>PLA2G4F</td>
<td>Phospholipase A2 group IVF</td>
</tr>
<tr>
<td>PNPLA3</td>
<td>Patatin like phospholipase domain containing 3</td>
</tr>
<tr>
<td>PTGS1</td>
<td>Prostaglandin-endoperoxide synthase 1</td>
</tr>
<tr>
<td>RGN</td>
<td>Regucalcin</td>
</tr>
<tr>
<td>RLBP1</td>
<td>Retinaldehyde binding protein 1</td>
</tr>
<tr>
<td>SCD5</td>
<td>Stearoyl-CoA desaturase 5</td>
</tr>
<tr>
<td>SLC22A2</td>
<td>Solute carrier family 22 member 2</td>
</tr>
<tr>
<td>SYK</td>
<td>Spleen associated tyrosine kinase</td>
</tr>
<tr>
<td>TACR1</td>
<td>Tachykinin receptor 1</td>
</tr>
<tr>
<td>TIRAP</td>
<td>TIR domain containing adaptor protein</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumour protein 53</td>
</tr>
<tr>
<td>2015 miRNA</td>
<td>4.24: Network 1</td>
</tr>
<tr>
<td>DICER1</td>
<td>Dicer 1, ribonuclease III</td>
</tr>
<tr>
<td>EPHB1</td>
<td>EPH receptor B1</td>
</tr>
<tr>
<td>FOS</td>
<td>Fos proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td>HMGCR</td>
<td>3-hydroxy-3-methylglutaryl-coenzyme A reductase</td>
</tr>
<tr>
<td>KDR</td>
<td>Kinase insert domain receptor</td>
</tr>
<tr>
<td>KRBOX4</td>
<td>KRAB box domain containing 4</td>
</tr>
<tr>
<td>2014 miRNA</td>
<td>Figure 4.26: Network 1</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 2016 mRNA  | Figure 4.27: Network 7 | BTF3L4 | Basic transcription factor 3 like 4 |

Table of Genes:

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGT</td>
<td>O-linked N-acetylglucosamine (GlcNAc) transferase</td>
</tr>
<tr>
<td>PRDX6</td>
<td>Peroxiredoxin 6</td>
</tr>
<tr>
<td>PRKCA</td>
<td>Protein kinase C alpha</td>
</tr>
<tr>
<td>PRKG1</td>
<td>Protein kinase, cGMP-dependent, type I</td>
</tr>
<tr>
<td>REL</td>
<td>REL proto-oncogene, NF-kB subunit</td>
</tr>
<tr>
<td>RHOA</td>
<td>Ras homolog family member A</td>
</tr>
<tr>
<td>SCN11A</td>
<td>Sodium voltage-gated channel alpha subunit 11</td>
</tr>
<tr>
<td>SLC35E2B</td>
<td>Solute carrier family 35 member E2B</td>
</tr>
<tr>
<td>SUFU</td>
<td>SUFU negative regulator of hedgehog signaling</td>
</tr>
<tr>
<td>THAP3</td>
<td>THAP domain containing 3</td>
</tr>
<tr>
<td>ZNF385B</td>
<td>Zinc finger protein 385B</td>
</tr>
<tr>
<td>DYDC2</td>
<td>DPY30 domain containing 2</td>
</tr>
<tr>
<td>EI24</td>
<td>EI24, autophagy associated transmembrane protein</td>
</tr>
<tr>
<td>EPOR</td>
<td>Erythropoietin receptor</td>
</tr>
<tr>
<td>FGF2</td>
<td>Fibroblast growth factor 2</td>
</tr>
<tr>
<td>GCC1</td>
<td>GRIP and coiled-coil domain containing 1</td>
</tr>
<tr>
<td>GUF1</td>
<td>GUF1 homolog, GTPase</td>
</tr>
<tr>
<td>HBEGF</td>
<td>Heparin binding EGF like growth factor</td>
</tr>
<tr>
<td>MSX2</td>
<td>MSH homeobox 2</td>
</tr>
<tr>
<td>NCL</td>
<td>Nucleolin</td>
</tr>
<tr>
<td>NR3C1</td>
<td>Nuclear receptor subfamily 3 group C member 1</td>
</tr>
<tr>
<td>PDGFRA</td>
<td>Platelet derived growth factor receptor alpha</td>
</tr>
<tr>
<td>RB1</td>
<td>RB transcriptional corepressor 1</td>
</tr>
<tr>
<td>RPS15</td>
<td>Ribosomal protein S15</td>
</tr>
<tr>
<td>SEPP1</td>
<td>Selenoprotein P</td>
</tr>
<tr>
<td>TGFB2</td>
<td>Transforming growth factor beta 2</td>
</tr>
<tr>
<td>YOD1</td>
<td>YOD1 deubiquitinase</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Description</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
</tr>
<tr>
<td>C11orf95</td>
<td>Chromosome 11 open reading frame 95</td>
</tr>
<tr>
<td>CCND2</td>
<td>Cyclin D2</td>
</tr>
<tr>
<td>CCDC30</td>
<td>Coiled-coil domain containing 30</td>
</tr>
<tr>
<td>CGNL1</td>
<td>Cingulin like 1</td>
</tr>
<tr>
<td>EIF2S1</td>
<td>Eukaryotic translation initiation factor 2 subunit alpha</td>
</tr>
<tr>
<td>FAM149B1</td>
<td>Family with sequence similarity 149 member B1</td>
</tr>
<tr>
<td>IMPACT</td>
<td>Impact RWD domain protein</td>
</tr>
<tr>
<td>JAKMIP3</td>
<td>Janus kinase and microtubule interacting protein 3</td>
</tr>
<tr>
<td>OR2W5</td>
<td>Olfactory receptor family 2 subfamily W member 5</td>
</tr>
<tr>
<td>OGFOD3</td>
<td>2-oxoglutarate and iron dependent oxygenase domain containing 3</td>
</tr>
<tr>
<td>NDRG1</td>
<td>N-myc downstream regulated 1</td>
</tr>
<tr>
<td>PEX7</td>
<td>Peroxisomal biogenesis factor 7</td>
</tr>
<tr>
<td>PTER</td>
<td>Phosphotriesterase related</td>
</tr>
<tr>
<td>RAB37</td>
<td>RAB37, member RAS oncogene family</td>
</tr>
<tr>
<td>RAVER2</td>
<td>Ribonucleoprotein, PTB binding 2</td>
</tr>
<tr>
<td>2016 mRNA</td>
<td>Retina and anterior neural fold homeobox</td>
</tr>
<tr>
<td>SSTR1</td>
<td>Somatostatin receptor 1</td>
</tr>
<tr>
<td>ZNF566</td>
<td>Zinc finger protein 566</td>
</tr>
</tbody>
</table>

Figure 4.28: Network 22

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Gene Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFMD</td>
<td>Arylformamidase</td>
</tr>
<tr>
<td>ARMC8</td>
<td>Armadillo repeat containing 8</td>
</tr>
<tr>
<td>CBX5</td>
<td>Chromobox 5</td>
</tr>
<tr>
<td>CUL3</td>
<td>Cullin 3</td>
</tr>
<tr>
<td>EGLN3</td>
<td>Egl-9 family hypoxia inducible factor 3</td>
</tr>
<tr>
<td>ERAL1</td>
<td>Era like 12S mitochondrial rRNA chaperone 1</td>
</tr>
<tr>
<td>HOOK1</td>
<td>Hook microtubule tethering protein 1</td>
</tr>
<tr>
<td>HS3ST1</td>
<td>Heparan sulfate (glucosamine) 3-O-sulfotransferase 1</td>
</tr>
<tr>
<td>HS3ST2</td>
<td>Heparan sulfate (glucosamine) 3-O-sulfotransferase 2</td>
</tr>
<tr>
<td>IPO8</td>
<td>Importin 8</td>
</tr>
<tr>
<td>KCTD16</td>
<td>Potassium channel tetramerization domain containing 16</td>
</tr>
<tr>
<td>KIRREL3</td>
<td>Kirre like nephrin family adhesion molecule 3</td>
</tr>
<tr>
<td>KLF9</td>
<td>Kruppel like factor 9</td>
</tr>
<tr>
<td>KLHL10</td>
<td>Kelch like family member 10</td>
</tr>
<tr>
<td>NACC1</td>
<td>Nucleus accumbens associated 1</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>PALM2</td>
<td>Paralemmin 2</td>
</tr>
<tr>
<td>PLPPR4</td>
<td>Phospholipid phosphatase related 4</td>
</tr>
<tr>
<td>SLC45A3</td>
<td>Solute carrier family 45 member 3</td>
</tr>
<tr>
<td>TLL2</td>
<td>Tolloid like 2</td>
</tr>
<tr>
<td>TMEM178A</td>
<td>Transmembrane protein 178A</td>
</tr>
<tr>
<td>TMIGD3</td>
<td>Transmembrane and immunoglobulin domain containing 3</td>
</tr>
<tr>
<td>TMPO</td>
<td>Thymopoietin</td>
</tr>
<tr>
<td>VASH2</td>
<td>Vasohibin 2</td>
</tr>
<tr>
<td>WDFY4</td>
<td>WDFY family member 4</td>
</tr>
<tr>
<td>WDR91</td>
<td>WD repeat domain 91</td>
</tr>
<tr>
<td>ZNF236</td>
<td>Zinc finger protein 236</td>
</tr>
<tr>
<td>ZNF585B</td>
<td>Zinc finger protein 585B</td>
</tr>
<tr>
<td>ZNF705A</td>
<td>Zinc finger protein 705A</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>Aldehyde dehydrogenase 3 family member A1</td>
</tr>
<tr>
<td>APC</td>
<td>Anaphase promoting complex (complex)</td>
</tr>
<tr>
<td>ASTN2</td>
<td>Astrotactin 2</td>
</tr>
<tr>
<td>CCDC80</td>
<td>Coiled-coil domain containing 80</td>
</tr>
<tr>
<td>CCNA2</td>
<td>Cyclin A2</td>
</tr>
<tr>
<td>CCNB1</td>
<td>Cyclin B1</td>
</tr>
<tr>
<td>Cdc2</td>
<td>Cyclin dependent kinase 1</td>
</tr>
<tr>
<td>CEP55</td>
<td>Centrosomal protein 55</td>
</tr>
<tr>
<td>CHEK2</td>
<td>Checkpoint kinase 2</td>
</tr>
<tr>
<td>DLGAP5</td>
<td>DLG associated protein 5</td>
</tr>
<tr>
<td>E2f</td>
<td>E2F transcription factor</td>
</tr>
<tr>
<td>E2F1</td>
<td>E2F transcription factor 1</td>
</tr>
<tr>
<td>GEN1</td>
<td>GEN1, Holliday junction 5' flap endonuclease</td>
</tr>
<tr>
<td>GINS2</td>
<td>GINS complex subunit 2</td>
</tr>
<tr>
<td>GSTA4</td>
<td>Glutathione S-transferase alpha 4</td>
</tr>
<tr>
<td>LRRC3</td>
<td>Leucine rich repeat containing 3</td>
</tr>
<tr>
<td>LTA4H</td>
<td>Leukotriene A4 Hydrolase</td>
</tr>
<tr>
<td>MBTPS2</td>
<td>Membrane bound transcription factor peptidase, site 2</td>
</tr>
<tr>
<td>PAX2</td>
<td>Paired box 2</td>
</tr>
<tr>
<td>PHF21B</td>
<td>PHD finger protein 21B</td>
</tr>
<tr>
<td>PLPPR4</td>
<td>Phospholipid phosphatase related 4</td>
</tr>
<tr>
<td>PTPRU</td>
<td>Protein tyrosine phosphatase, receptor type U</td>
</tr>
<tr>
<td>RASSF2</td>
<td>Ras association domain family member 2</td>
</tr>
<tr>
<td>Rb</td>
<td>RB transcriptional corepressor</td>
</tr>
<tr>
<td>Srebp</td>
<td>Sterol regulatory element binding</td>
</tr>
</tbody>
</table>

**2014 mRNA**

**Figure 4.29: Network 4**
<table>
<thead>
<tr>
<th>Figure 4.30: Eicosanoid signalling pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>THBS2</td>
</tr>
<tr>
<td>TP53</td>
</tr>
<tr>
<td>TWIST2</td>
</tr>
<tr>
<td>ZNF385B</td>
</tr>
<tr>
<td>12-HETE</td>
</tr>
<tr>
<td>15-S-HETE</td>
</tr>
<tr>
<td>ALOX5</td>
</tr>
<tr>
<td>ALOX5AP</td>
</tr>
<tr>
<td>ALOX12</td>
</tr>
<tr>
<td>ALOX15</td>
</tr>
<tr>
<td>CYSLTR</td>
</tr>
<tr>
<td>DPEP</td>
</tr>
<tr>
<td>GGT1</td>
</tr>
<tr>
<td>LTB4R</td>
</tr>
<tr>
<td>LTA4H</td>
</tr>
<tr>
<td>LTA4R</td>
</tr>
<tr>
<td>LTC4S</td>
</tr>
<tr>
<td>PGD2</td>
</tr>
<tr>
<td>PGDS</td>
</tr>
<tr>
<td>PGE2</td>
</tr>
<tr>
<td>PGES</td>
</tr>
<tr>
<td>PGF2a</td>
</tr>
<tr>
<td>PGFS</td>
</tr>
<tr>
<td>PGH2</td>
</tr>
<tr>
<td>PGI2</td>
</tr>
<tr>
<td>PGIS</td>
</tr>
<tr>
<td>PLA2</td>
</tr>
<tr>
<td>PTGDR</td>
</tr>
<tr>
<td>PTGER1</td>
</tr>
<tr>
<td>PTGER2</td>
</tr>
<tr>
<td>PTGER3</td>
</tr>
<tr>
<td>PTGER4</td>
</tr>
<tr>
<td>PTGFR</td>
</tr>
<tr>
<td>PTGIR</td>
</tr>
<tr>
<td>PTGS</td>
</tr>
<tr>
<td>TBXA2R</td>
</tr>
<tr>
<td>TXS</td>
</tr>
</tbody>
</table>

This table lists gene names included in various figures in chapter 4. The table lists the IPA involved, the figure, gene symbols and gene names included in each figure.
### Downstream analysis for 2016 mRNA microarrays

<table>
<thead>
<tr>
<th>Disease or Function</th>
<th>Downstream Prediction</th>
<th>Genes Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumorigenesis of tissue</td>
<td>Down</td>
<td>ABCA6, ABCC6, ABCC9, ABCG5, ABHD12B, ABHD17A, ACAN, ACPP, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMTS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHFE1, ADIPOQ, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP1, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKR30A, ANKR53, ANP32C, AOX1, APOA5, APOB, ARHGAP39, ASPHD1, ASPN, ASTN2, ATCAV1, ATG9A, ATP1A4, ATP2B2, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCAT1, BCL11A, BDKRB1, BGN, BIR7, BNC2, BOLL, BPINF1, BRD3, BRSK2, BSPH1, BTK, C12orf56, C17orf51, C19orf81, C1orf186, C1S, C20orf196, C22orf46, C3, C3orf22, C6orf10, C7orf25, C9orf152, CA16, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACN54, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASP8, CASQ1, CATSPER4, CBLN4, CCL4, CCDC141, CCDC183, CCDC33, CCDC54, CCDC80, CCNA2, CCNB1, CD160, CD177, CD19, CD1C, CD5L, CDH18, CDH9, CDKL1, CDKL2, CEACAM18, CEACAM7, CEACAM8, CENPE, CENPF, CEP152, CEP55, CETFN1, CHEK2, CHRM5, CHRNA1, CHRNA5, CHRNAB3, CHTOP, CLIP2, CLC, CLDN4, CLEC4F, CLEC6A, CLEC11A, CLMN, CLMP, CNP1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COP7B, COQ2, CP2A, CPNE6, CPNE9, CPXM1, CPXM2, CR2, CRNN, CSN2, CTNNBP2, CUX2, CXCR2, CXorf67, CYP1A2, CYP2S1, DAAM2, DBH, DCBLD2, DCD, DCN, DDX10, DEFB128, DGAT2, DGKG, DHFR, DIO2, DIP2C, DIS3L2, DLD1, DLGAP5, DLX4, DNH17, DNL12, DNACD6, DOK3, DPCR1, DPY19L3, DRGX, DSC2, DSCR3, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, EML6, ENPP6, ERSB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM65C, FBNL1, FBXO24, FCAMR, FCAR, FCER1A, FCER1G, FCGR2B, FCRL2, FCRL3, FERMT1, FEEDER, FGFR2, FGFR3, FMO5, FOXS2, FOXL1, FSHR, FZD5, GAB4, GABRB3, GAD, GALT9, GAPT, GCKR, GEN1, GFPT1, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPA33, GPAM, GPMAP, GPC6, GPR149, GPR21, GPR24, GPR45, GPRC5D, GREB1L, GRIK3, GRIN2A, GRIN2C, GRIP1, GRK1, GRK3, GRM1, GSTA4, GUCA1B, GUCY2D, HAPLN4, HAS1, HDCC2, HES6, HHIP, HLA-DQA2, HMGA2, HMGC51, HMGS2, HMGN5, HPGDS, HRH3, HS3ST2, HSD11B1, HSD3B2, ID4, IGF1, IGSF3, IL1A, INHBE, ITGB6, ITGBL1, ITLN1, IWS1, JAKMIP2, KCNA6, KCNE1, KCNH8, KCNJ1</td>
</tr>
<tr>
<td>Neoplasia of epithelial tissue</td>
<td>Down</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>ABCA6, ABCC6, ABCC9, ABCG5, ABHD12B, ABHD17A, ACAN, ACPP, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMTS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHEF1, ADIPOQ, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP1, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, APOA5, AQP4, ARHGAP39, ASPH1, ASPL, ASTN2, ATCAY, ATCAY, ATG9A, ATP1A4, ATP2B2, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCAT1, BCL11A, BDKRB1, BGN, BIRC7, BNC2, BOLL, BPIFB1, BRD3, BRSK2, BSHP1, BTK, C12orf56, C17orf51, C19orf81, C1orf186, C1S, C20orf196, C22orf46, C3, C3orf22, C6orf10, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG4, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASP8, CASQ1, CATSPER4, CBLN4, CCDC141, CCDC183, CCDC33, CCDC54, CCDC80, CCNA2, CCNB1, CD160, CD177, CD19, CD1C, CD5L, CDH18, CDK12, CEACAM18, CEACAM7, CEACAM8, CENPE, CENPF, CEP152, CEP5, CETN1, CHEK2, CHRM5, CHRNA1, CHRNA5, CHRNBD3, CHTOP, CILP2, CLC, CLDNL4, CLEC4F, CLEC6A, CLEC6L1, CLECL1, CLMN, CLMP, CNPY1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COPS7B, COQ2, CPA2, CPN6, CPNE9, CPXM1, CPXM2, CR2, CRN, CSN2, CTTNBP2, CUX2, CXCR2, CXorf67, CYPA12, CYP2S1, DAAM2, DBH, DCHBD2, DCD, DCPN, DXD10, DEFB128, DGAT2, DGKG, DHFR, DIO2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, DNH17, DNA12, DNAJC6, DOK3, DPCR1, DPY10L3, DRG5, DSC2, DSCR3, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, EM6L, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ET10, F10, F11, FAM107A, FAM129C, FAM65C, FBXL1, FBXO24, FCAMR, FCAR, FCER1A, FGR2B, FCRL2, FCRL3, FERMT1, FETUB, FG6, FGFR9, FGFR3, FM05, FOXF2, FOXL1, FSHR, FZD5, GAP4, GABRQ, GABRR3, GAK, GALNT9, GAN, GCKR, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5, KCNQ2, KCTD16, KHDC1, KIF2B, KIRREL2, KIRREL3, KIT, KLHL10, KLHL29, KLHL38, KLK2, KLK4, KNL1, KPRP, KRT34, KRT36, KRTAP19-1, KSR2, LHGDH, LACRT, LAMA2, LAMA3, LAMP3, LCE1C, LCE3D, LCE3E, LCN1, LGI3, LILRA4, LINC00675, LMX1A, LOC102724788, PRODH, LPCA72, LRRC10, LRRQ3, LRRK2, LTA4H, LUM, LYN, MAGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAP3K2, MARCH1, MAS1L, MAST4, MBTPS2, MC5R, MCRIP2, MEST, MIA2, MICALCL, MKI67, MME, MMP25, MORN1, MPP2, MS4A5, MTRR, MUC12, MUC22, MUC5B, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYADML2, MYBPC2, MYL3, MYH2, MYO3A, MZT2B, NAP1L4, NEDF, NEF, NIN, NINL, NKKD2, NKKX3-1, NLRP12, NLRP5, NLRP8, NLRP9, NPSR1, NRP3, NRXN1, NRXN2, NT5C1A, NTRK1, NTRK2, NuMbl, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, OC90, OADMM, OLFML2A, OPN3, OR10G3, OR10G9, OR12D2, OR1D5, OR1N1, OR2A1/OR2A2, OR2D3, OR2M7, OR2T1, OR4C6, OR4Q3, OR51A7, OR51G1, OR52N5, OR56B4, OR7E5P, OR7G2, OR8A1, ORC6, OTOG, OTOGL, P2RX1, P2RX7, PACSIN3, PAD2, PAMR1, PAX2, PCDC18, PDE4DIP, PDGFRA, PDIK1L, PDZD2, PFKB2, PGK2, PHEX, PHF21B, PID1, PIK3R3, PIPIK1B, PLA2G2A, PLA2G4F, PLCH1, PLEKHO1, PLK1, PLXNA4, PMFBP1, PNPLA3, PPBP, PPAT4A, PPIL2, PPIP5K1, PPP2R2C, PRDM15, PRDM16, PRLA1, PRKCG, PRKCH, PRKQ, PRR30, PRSS38, PS11, PSM5D, PTER, PTGER3, PTGS1, PTGRH, PTPRU, Rab17, Rab27B, Rab41, Rab44, RabNP2, Rap1GAP, RASGRF1, RA5, RAB8B, RBMP20, REG3A, RGN, RIMK1A, RIMS3, RLB1P1, ROD1, ROX2, ROS1, RP1L1, RP S6K1L1, SARDH, SCD5, SCN3A, SCN7A, SCN1D, SCTR, SERPINA12, SERPINA3, SERPINB11, SERTD1B, SFRP1, SFRP2, SFRP5, SFTPA1, SFTPA2, SH3RF3, SHASA2, SIGLEC12, SIRPB1, SIRPB2, SIX3, SIX6, SLC11A1, SLC17A2, SLC22A11, SLC22A2, SLC22A20, SLC24A2, SLC24A3, SLC24A4, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC4A8, SLC6A12, SLC7A14, SLC8A1, SLC9A5, SMPD3B, SOPCS2, SOX3, SP8, SPACA5/SPAC5B, SPATA22, SPINT3, SP OCD1, SPO11, SPRY3, SRMS, SSTD2, SSTR9, STRC, SYCP1, SYK, SYNJ2, TACR1, TAL2, TBL3, TCOX1, TCF7, TCPO1L, TGM6, TH, THBS1, THBS2, TLR10, TM4SF5, TM6SF2, TMC2, TMEM151B, TMEM154, TMEM184A, TMEM257, TMEM266, TMEM87A, TMEM97, TNFRSF13B, TNIK, TONS1, TOP2A, TP53, TP53TGF, TPO, TPRX1, TPSD1, TRABD2A, TRIP13, TRPM3, TSHB, TSPER, TSSK1B, TTC24, TTC39A, TUBB2A, TUBB8, UBL4B, UBXN11, UCHL1, UGT3A1, UMODL1, UNC80, UPB1, UROS, USP17L2 (includes others), USP49, VASH2, VSIG1, VSIG10, VSIG10L, VSTM4, WAA5A, VWDE, WBSR17, WDFY4, W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Name</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>FDC12, WFDC5, WIF1, WT1, XIRP2, YME1L, ZAN, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF49 1, ZNF568, ZNF585B, ZNF625, ZNF705A</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Abdominal cancer</strong></td>
<td><strong>Down</strong></td>
<td></td>
</tr>
<tr>
<td>ABCA6, ABCCC6, ABCG5, ABHD12B, ACAN, ACPP, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMTS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHFE1, ADIPOQ, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP1, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, APOA5, AQP3, AQP4, ARHGAP39, ASN, ASTN2, ATCAY, ATG9A, ATP1A4, ATP2B2, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCAT1, BCL11A, BDKRB1, BGN, BIRC7, BNC2, BOLL, BPIFB1, BRD3, BRSK2, BSPH1, BTK, C12orf56, C17orf51, C1orf186, C1S, C20orf196, C15S, C22orf46, C3, C3orf22, C6orf10, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG4, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASP8, CASQ1, CATSPER4, CBLN4, CCDC141, CCDC183, CCDC33, CCDC80, CCNA2, CCNB1, CD160, CD177, CD19, CD1c, CD5L, CDH18, CDH9, CDKL2, CEACAM18, CEACAM7, CECAAM8, CENPE, CENPF, CEP152, CEP55, CETN1, CHEK2, CHRM5, CHRNA5, CHRNB3, CHTOP, CILP2, CLC, CLDN4, CLEC4F, CLEC6A, CLMN, CLMP, CNPY1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COPS7B, COQ2, CPA2, CPNE6, CPNE9, CXM1, CFXM2, CR2, CRN, CSN2, CTNBIP2, CUX2, CXXCR2, CXYorf67, CYP1A2, DAAM2, DBH, DBCDLD2, DCCN, DDX10, DGAT2, DKG, DHFR, DIO2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, Dnah17, Dnai2, DNAJC6, DOK3, DPCR1, DPY19L3, DRGX, DSC2, DSCR3, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ET1V1, F10, F11, FA107A, FAM129C, FAM65C, FBLN1, FBXO24, FCAMR, FCR, FCER1A, FCGR2B, FCRL2, FCRL3, FCRFB, FERMT1, FETUB, FGF16, FGF6, FGF9, FGF13, FMO5, FOXF2, FOX11, FSHR, FXYD3, FZD5, GAB4, GABRQ, GABRR3, GAK, GALNT9, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPA33, GPAM, GPAT2, GPC6, GPR149, GPR21, GPR34, GPR45, GPRC5D, GREB1, GRK3, GRIN2A, GRIN2C, GRIP1, GRK1, GRK3, GRM1, GSTA4, GUCA1B, GUCY2D, HAPLN4, HAS1, HDCC, HES6, HHIP, HLA-DQ2, HMGA2, HMGC1S, HMGS2, HMGN5, HPGDS, HRRH3, HS3ST2, HSD11B1, HSD3B2, ID4, IGF1, IGSF3, IL1A, INHBE, INS, ITGB6, ITGBL1, ITLN1, IWS1, JAKMIP2, KCNA6, KCNE1, KCNH8, KCNJ1, KCNO2, KCTD16, KHDC1, KIF2B, KIRREL2, KIRREL3, KIT, KLH10, KLHL29, KLHL38, KLK2, KLK4, KNL1, KPRP, KRT34, KRT36, KRTAP19-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Symbols</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1, KSR2, L2HGDH, LACRT, LAMA2, LAMA3, LAMP3, LCE1C, LCE3E, LCN1, LGI3, LILRA4, LINC00675, LMX1A, LOC102724788/PRODH, LPCAT2, LRRC10, LRRQ3, LRRK2, LTA4H, LUM, LYNX1, MAGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAP3K2, MARCH1, MAS1L, MAST4, MBTPS2, MCF5R, MCRIP2, MEST, MIA2, MICALCL, MKI67, MME, MMP25, MORN1, MPP2, MS4A5, MTRR, MUC12, MUC22, MUC5B, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYADM1L, MYBPC2, MYH2, MYO3A, MZT2B, NAP1L4, NBR2, NEBL, NGEF, NIN, NINL, NLD3, NKKD2, NKKX3-1, NLRP12, NLRP5, NLRP8, NLRP9, NPSR1, NRP3, NRPX1, NRPX2, NT5C1A, NTRK1, NTRK2, NUB1, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, OC90, ODAK, OLFML2A, OPN3, OR10G3, OR10G9, OR1N1, OR2A1/OR2A42, OR2D3, OR2M7, OR4C6, OR4Q3, OR51A7, OR51G1, OR52N5, OR56B4, OR7G2, OR8A1, ORC6, OTOG, OTOGL, P2RX1, P2RX7, PACSIN3, PADI2, PAMR1, PAX1, PCDHB1, PDE4DIP, PGDFRA, PDK1, PDZD2, PFKFB2, PGK2, PHEX, PHF21B, PID1, PIK3R3, PLP2, PLA2G2A, PLA2G4F, PLCH1, PLEKHO1, PLK1, PLXNA4, PMFB1, PNPLA3, PPBP, PPFIA4, PPII2, PPI5K1, PPR2R2C, PRDM15, PRDM16, PRIMA1, PRKCG, PRKC, PRK2, PRR27, PR30, PRR5, PRRX1, PRSS38, PSLG11, PSM5D5, PTER, PTGER3, PTGS1, PTGRH, PTPT3, PTB17, RAB41, RAB44, RARB2, RAP1GAP, RASGRF1, RASSF2, RBBP8NL, RBM20, RBMY1A1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(includes others), REG3A, RGN, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RPS6KL1, SARDH, SCD5, SCN3A, SCN7A, SCNN1D, SCTR, SERPINA12, SERPINA3, SERPINB11, SETD1B, SFRP1, SFRP5, SFTPA1, SFTPA2, SH3RF3, SHISA2, SIGLEC12, SIRPB1, SIRPB2, SIX3, SIX6, SLC11A1, SLC17A2, SLC22A11, SLC22A2, SLC22A20, SLC24A2, SLC24A3, SLC24A4, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC4A8, SLC6A12, SLC7A4, SLC8A1, SLC9A5, SMDP3B, SORCS2, SOX30, SP8, SPPA/SPSA5, SPPA/SPSA5B, SPATA22, SPOCD1, SPON1, SRMS, SSC5D, SXX3, STAR8D9, STRC, SYCP1, SYK, SYNJ2, TACR1, TAL2, TBL3, TBX6, TBXAS1, TCF7, TIP10L, TGM6, TH, THBS1, THBS2, TRIL, TM4SF5, TM6SF2, TMCI, TEM151B, TEM154, TEM158A, TEM26B, TEM75A, TEM97, TNFRSF13B, TNK1, TONSL, TOP2A, TP53, TP53TG5, TPO, TPRX1, TPS1D1, TRABD2A, TRIP13, TRPM3, TSPEAR, TSSK1B, TTC24, TUBB2A, TUBB8, TWIST2, UBL4B, UBXN11, UCHL1, UGT3A1, UMODL1, UNC80, UPB1, UROS, USP17L2 (includes others), USP49, VASH2, VSIG1, VSIG10, VSIG10L, VSTM4, VWA5A, VVDE, WBSCR17, WDFY4, WDC12, WDFC5, WIFI1, WT1, WT1-AS, XIRP2, YME1L1, ZAN, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF585B, ZN
Digestive organ
tumour

<table>
<thead>
<tr>
<th>Down</th>
</tr>
</thead>
</table>
| ABCA6, ABCC6, ABCC9, ABCG5, ACAN, ACPP, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMTS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHFE1, ADIPOQ, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP1, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, APOA5, AQP3, AQP4, ARHGAP39, ASPN, ASTN2, ATCAY, ATG9A, ATP1A4, ATP2B2, ATP2C1, ATP8A2, ATRNL1, AZGP1, B3GNT8, BACH2, BANK1, BCAT1, BCL11A, BDKRB1, BGN, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, C12orf156, C17orf51, C1orf186, C1S, C20orf196, C22orf46, C3, C3orf22, C6orf10, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG4, CACNG7, CADM1, CALB2, CALCA, CAMP, CASP8, CASQ1, CASK, CASKIP4, CCDC141, CCDC183, CCDC33, CCDC54, CCDC80, CCL19, CCNA2, CCNB1, CD160, CD177, CD1C, CDH18, CDKL2, CEACAM18, CEACAM8, CENPE, CENPF, CEP152, CEP55, CHEK2, CHRM5, CHRNA1, CHRNA5, CHRN3B, CHTOP, CIPL2, CLC, CLDN4, CLEC4F, CLEC6A, CLEC8L1, CLMN, CLMP, COL1A1, COL8A2, COE2C12, COMP, COP5B, COQ2, CPA2, CPXM1, CPXM2, CR2, CRNN, CSN2, CTNBP2, CUX2, CXCR2, CXorf67, CYP1A2, DAAM2, DBH, DCBLD2, DCN, DDX10, DGAT2, DGKG, DHFR, DIO2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, DNAH17, DNAI2, DNAJC6, DOK3, DPCR1, DPY19L3, DRG, DSC2, E2F1, E2F7, EBF1, EDAR, EGF, EGFLAM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM65C, FBNL1, FBXO24, FCAMR, FCAR, FERMT1, FERMT2, FGF6, FGFR3, FMO5, FOX2, FOXL1, FSHR, FYXD3, FXD5, GAB4, GABRAQ, GAK, GALNT9, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPAM, GPT2, GPCG6, GPR149, GPR21, GPR34, GPR45, GPRC5D, GREB1L, GRIK3, GRIN2A, GRIN2C, GRIP1, GRK1, GRM1, GSTA4, GUCA1B, HAPLN4, HAS1, HHC, HHC2, HES6, HISP, HLA-
| DQA2, HMGA2, HMGCS1, HMGCSD2, HMGCS5, HOTAIR, HPGDS, HPR3, HSST2, HSD3B2, ID4, IGF1, IGF3, IGF5, IGF7, IL1A, INHBE, INS, ITGB6, ITGB1, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNQ1, KCNJ8, KCNJ15, KCNN2, KCTD16, KIF2B, KIRREL2, KIRREL3, KIT, KLHL29, KLHL38, KLRK1, KLRK2, KLK4, KLK9, KNRK1, KPRP, KRT34, KRT36, KRTAP19-1, KSR2, L2HGDH, LACRT, LAMA2, LAMA3, LAMP3, LCE1C, LCE3E, LGI3, LILRA4, LINC00675, LMX1A, LOC102724788, PRODH, LRPC12, LRRC10, LRRK3, LRRK2, LTAV4, LVN, MAEGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAP3K2, MARCH1, MAS1L, MAST4, MBTPS2, MC5R, MCF2L,
<table>
<thead>
<tr>
<th>Gene Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCRIP2, MEST, MIA2, MICALCL, MKI67, MME, MMP25, MORN1, MMP2, MS4A5, MTRR, MUC12, MUC22, MUC5B, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYADML2, MYBPC2, MYCL, MYH2, MYO3 A, NAPL4, NEBL, NGEF, NIN, NINL, NKKD2, NKKX3-1, NLRP12, NLRP5, NLRP8, NLRP9, NPSR1, NRNX1, NRNX2, NT5C1A, NTRK1, NTRK2, NUMBL, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, OC90, ODAH, OLFML2A, OPN3, OR10G9, OR1N1, OR2D3, OR2M7, OR4C6, OR4Q3, OR51A7, OR51G1, OR52N5, OR56B4, OR7G2, OR8A1, ORC6, OTOG, OTQGL, P2RX1, P2RX7, PACSIN3, PADI2, PAMR1, PAX2, PCHBD1, PDE4DIP, PDEG FRA, PDZD2, PFKFB2, PGK2, PHF21B, PID1, PIK3R3, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLE KHO1, PLK1, PLXNA4, PMFBB1, PNPLA3, PPFAA, PPI1L2, PPI5K1, PPR2R2C, PRDM15, PRDM16, PRIMA1, PRKCG, PRKCR, PRR27, PRR30, PRR5, PRRX1, PRSS38, PSC1, PSM5D, PTER, PTG S1, PTPRH, PTPRU, RAB17, RAB27B, RANBP2, RAP1GAP, RASGFRF1, RASSF2, RB8B8NL, RBM 20, RBYMA1 (includes others), REG3A, RGN, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RPS6KL1, SARDH, SC5D, SC7A, SCNN1D, SCT, SERPINA12, SERPINA3, SERPINB11, SETD1B, SFRP1, SFRP5, SFPA1, SH3F3D, SHISA2, SIGLEC12, SIRPB2, SIX3, SIX6, SLC11A1, SLC17A2, SLC22A11, SLC22A2, SLC22A4, SLC22A5, SLC24A3, SLC24A4, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC6A12, SLC7A14, SLC8A1, SLC9A5, SLC9A9, SORCS2, SOX30, SPATA22, SPOCD1, SPON1, SRMS, SSCP5D, SSSX3, STARD9, STRC, SYCP1, SYK, SYNJ2, TACR1, TAL2, TBL3, TBX6, TBXAS1, TCF7, TCP10L, TGM6, TH, THBS1, THBS2, TLR10, TM4SF5, TM6SF2, TMC2, TME M151B, TMEM154, TMEM184A, TMEM97, TNFRSF13B, TNK1, TONSL, TOP2A, TP53, TP53T5G, TPO, TRPV1, TRPS1, TRABD2A, TRIP13, TRPML1, TSPEAR, TSSK1B, TTC24, TUBB2A, TUBB8, UB XN11, UCHL1, UGT3A1, UMODL1, UCN80, UBP1, UROS, USP49, VASH2, VSIG10, VSIG10L, VST M4, VWA5A, VVDE, WBSCR17, WDFY4, WDFD12, WDFD5, WT1, XIRP2, YME1L1, ZAN, ZMYND1 9, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF585B, ZNF625, ZNF705A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA6, ABCC6, ABCC9, ABCG5, ACAN, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHFE1, ADIPQO, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKRD30A, ANKRD53, AOX1, APOA5, AQP3, ARHGA39, ASTN2, ATCA, ATP1A4, ATP2B2, ATP2C1, AT RNL1, B3GNT8, BACH2, BANK1, BCA1, BCL11A, BDKRB1, BGN, BIRC7, BNC2, BOLL, BRD3, BR</td>
</tr>
</tbody>
</table>

Intestinal cancer  Down  ABCC6, ABCC9, ABCG5, ACAN, ACSM3, ACTC1, ADAM12, ADAM21, ADAM30, ADAM33, ADAMS16, ADAMTS17, ADAMTS7, ADARB2, ADCY2, ADHFE1, ADIPQO, ADORA1, AGBL1, AGMAT, AGTR2, AIF1, AIM1L, AK5, AKAP17A, AKR1B10, ALDH3A1, ANGPT4, ANK2, ANK3, ANKRD30A, ANKRD53, AOX1, APOA5, AQP3, ARHGA39, ASTN2, ATCA, ATP1A4, ATP2B2, ATP2C1, AT RNL1, B3GNT8, BACH2, BANK1, BCA1, BCL11A, BDKRB1, BGN, BIRC7, BNC2, BOLL, BRD3, BR
<table>
<thead>
<tr>
<th>Gene Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK2, BTK, C1orf51, C1orf186, C1S, C2orf46, C3, C7orf25, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CADM1, CAMP, CASP8, CATSPER4, CCDC141, CCDC33, CCDC54, CCDC80, CCNA2, CCNB1, CD160, CD177, CD1C, CDH18, CDKL2, CEACAM18, CEACAM7, CEACAM8, CENPE, CENPF, CEP152, CEP55, CHEK2, CHRM5, CHRNA1, CHRNA5, CLC, CLDN4, CLEC4F, CLEC6A, CLMN, CLMP, COL16A1, COL8A2, COLEC12, COMP, COPS7B, COQ2, CPA2, CPXM1, CR2, CRNN, CTNNBP2, CUX2, CXCR2, CYP1A2, DBH, DCBLD2, DCN, DDX10, DGKG, DHFR, DIO2, DI P2C, DIS3L2, DLC1, DLGAP5, DLX4, D納H17, DNAI2, DOK3, DPCR1, DRGX, E2F1, E2F7, EBF1, E2F4, EGFL6, EGFLAM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM65C, FBLN1, FBXO24, FCAMR, FCAR, FCGR2B, FCRL2, FCRL3, FERMT1, FETUB, FGF6, FGFR3, FMO5, FOXF2, FOXI1, FSFR, FZD5, GAB4, GAK, GALNT9, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPAM, GPC6, GPR149, GPR21, GPR34, GPR45, GREB1L, GRIK3, GRIN2A, GRIN2C, GRP1, GRK1, GRM1, GSTA4, GUCA1B, GUCY2D, HAPLN4, HAS1, HHIP, HMGS1, HMGS2, HMGN5, HPGDS, HRH3, HSD3B2, IG1, IGSF3, IL1A, INHBE, INS, ITGB6, ITGB1L, ITL1N, JAKMIP2, KCNA6, KCNE1, KCNH8, KCNJ15, KCNQ2, KCTD16, KIF2B, KIRREL2, KIRREL3, KIT, KLHL29, KLHL38, KLK2, KPRP, KRT34, KRT36, KSR2, L2HGHDH, LAMA2, LAMA3, LAMC3, LCE1C, LCE3E, LILRA4, LINC00675, LOC102724788, PRODH, LPCAT2, LRR1Q3, LRRK2, LAT4H, LYNX1, MAGEC1, MAGEE2, MAMD2, MAML3, MAP1B, MAP3K2, MARCH1, MAST4, MBT2S2, MC5R, MCF2L, MCRIP2, MEST, MIA2, MKI67, MME, MMP25, MS4A5, MTRR, MUC12, MUC5B, MUC6, MUC7, MUSK, MXRA5, MXR, MYBPC2, MYH2, MYO3A, NBR2, NEBL, NGEF, NIN, NINL1, NDK2, NKX3-1, NLRP12, NLRP5, NLRP8, NLRP9, NPSR1, NRXN1, NRXN2, NT5C1A, NTRK1, NTRK2, NUTM2F/NUTM2G, NXPH3, NYAP2, OC90, ODAM, OLFML2A, OPN3, OR10G9, OR1N1, OR2D3, OR2M7, OR51A7, OR52N5, OR56B4, OR7G2, ORC6, OTOL, P2RX1, P2RX7, PASC1N3, PAD2, PAMR1, PAX2, PCDHB1, PDE4DIP, PDGFRA, PDZD2, PFKFB2, PGK2, PHF21B, PID1, PIK3R3, PIK5B1, PLA2G2A, PLA2G4F, PLCH1, PLK1, PLXNA4, PMFBP1, PNPLA3, PPFIA4, PPR2R2C, PRDM15, PRDM16, PRKCG, PRKCCZ, PRR30, PRRX1, PRSS38, PSD5D1, PTGS1, PTPRH, PTPRU, RAB17, RANBP2, RAP1GAP, RASGRF1, RBBP8NL, RBM20, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RP56KL1, SARDH, SCD5, SCN3A, SCN7A, SCN11D, SCTR, SERPINA12, SERPINA3, SERPINB11, SETD1B, SFPR1, SFPR5, SH3RF3, SHISA2, SIGLEC12, SIX3, SLC11A1, SLC22A11, SLC22A20, SLC24A2, SLC24A3, SLC24A4, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC6...</td>
</tr>
<tr>
<td>A12,SLC7A14,SLC8A1,SLC9A5,SORCS2,SOX30,SP8,SPATA22,SPOCD1,SPON1,SRMS,SSC5D,SSX3,STRC,SYCP1,SYK,SYNJ2,TACR1,TAL2,TBL3,TBX6,TBXAS1,TCF7,TCP10L,TGM6,THBS1,THBS2,TLR10,TM2C,TMEM151B,TMEM184A,TMEM97,TNFRSF13B,TNIK,TONG,TOP2A,TP53,TP53TGST,TPO,TRABD2A,TRIP13,TRPM3,TPSEAR,TSSK1B,TUBB2A,TUBB8,UBXN11,UCHL1,UGT3A1,UNC80,UPB1,UROS,VASH2,VSIG10,VSTM4,VWA5A,VWDE,WBSCR17,WDFY4,WT1,XIRP2,YME1L1,ZAN,ZMYND19,ZNF236,ZNF365,ZNF385B,ZNF408B,ZNF568,ZNF585B,ZNF625</td>
</tr>
<tr>
<td>Renal lesion</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>ABCA6, ACAN, ADAMTS7, ADCY2, ADIPOQ, ADORA1, AGBL1, AGTR2, ANK2, ANKRD30A, ARHGAP39, BDKRB1, C1S, C3, CAMD1, CATSPER4, CENPE, CENPF, CHEK2, CLDN4, CPXM2, CRN, D CN, DGAT2, DIS3L2, DLGAP5, DNAH17, DPCR1, DSCR3, ERBB4, ERCC5, ESR1, FCAMR, FCA R, GFR3, GABRB3, GABRA2, GRIK3, IGF1, IGSF3, INS, KCNJ1, KIT, KLK2, KLK4, LAMA2, LRA RK2, LTA4H, MAST4, MKI67, MME, MUC12, MUC5B, MUC6, MUSK, MYO3A, NGEF, NLRP12, NTRK 2, PAD2, PDE4DIP, PDGFRB, PTPB, PPP2, PRKRU, RANBP2, RAP1GAP, RGN, RP1L1, SRP1, S LC39A8, SLC47A1, SLC4A8, SYNJ2, TBL3, TCF7, TH, THBS1, TMEM266, TNFRSF13B, TNIK, TOP 2A, TP53, TPSD1, TUBB2A, UNC80, WIF1, WT1, WT1-AS, XIRP2, ZAN</td>
</tr>
</tbody>
</table>

The table lists the disease or function significantly altered in COPD, the direction it is altered in and the genes involved in the disease/function. The p-value of overlap is used to determine significance between groups with a cut-off of <0.05.
### Downstream analysis for 2014 mRNA microarrays

<table>
<thead>
<tr>
<th>Disease or Function</th>
<th>Downstream Prediction</th>
<th>Genes Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Down</td>
<td>ABCA6, ABCC6, ABCC9, ABCG5, ACPP, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, ADCY2, ADGRF1, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, AQP3, AQP4, ASPHD1, ASPN, ATSTN2, ATG9A, ATP1A4, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BDKRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, BTN1A1, C15orf32, C17orf51, C19orf81, C1orf111, C1orf186, C1QTNF9B-AS1, C1S, C2orf46, C3, C7orf25, C7orf65, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CA4NB4, CACNG6, CACNG8, CADM1, CALB2, CALCA, CAMP, CASQ1, CCDC54, CCDC80, CCL19, CCNA2, CCNB1, CD160, CD177, CD19, CD1B, CD1C, CD5L, CDH9, CDKL2, CEACAM18, CEACAM8, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRN3, CILP2, CLCNKA, CLEC6A, CLEC7A, CLDN1, CLM, CLMP, CNPY1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COQ2, CPA2, CPNE9, CPX1M, CPXM2, CR2, CRNN, CTNBP2, CUX2, CXCRI, CXorf67, CYP1A2, CYP2S1, CYP4A22, DAAO, DCBLD2, DCD, DCN, DEFB128, DGAT2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, DNA12, DNN1, DPC1, DRGX, DSC2, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, ELF5, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERIC4H, ESPN, ESR1, ESR2, ET1, EVA1A, F10, F11, FAM107A, FAM129C, FAM195A, FAM209B, FBGN1, FBP1, FCER1A, FCGR2B, FCRL1, FCRL2, FCRL3, FCRL6, FERMT1, FGF16, FGF21, FGF9, FGF10, FOS1, FOS2, FSHR, FZD5, GABRG2, GALT1, GCKR, GEN1, GFTP2, GFRRA2, GINS2, GLI2, GNM1, GOLGA8K (includes others), GAPA33, GPM, GAPA2, GPC6, GPR149, GPR15, GPR34, GPR45, GPRC5D, GREB1L, GRIK3, GRIN2A, GRK1, GRK3, GSTA4, GUCA1B, GUCY2D, HAS1, HRCTR1, HEDDC2, HES6, HHIP, HLA-A, DQA2, HMGA2, HMGCS1, HMGCS2, HMGN5, HPGDS, HSD3B2, ID4, IGF1, IGF1R, IGF3, INS, ITGB6, ITGBL1, ITLN1, JAK2, KCNA6, KCNE1, KCNH8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KLK4, KPRP, KRT34, KRTAP19-1, L1TD1, L2HGDH, LAMA2, LAMP3, LCE2B, LCE3D, LCE3E, LGI3, LILRA4, LINC00675, LOC102724788/PRODH, LRPCAT2, LRRK2, LRTOMT, LTA4H, LUM, LYNX1, MAGEC1, MAGEE2, MAMDC2, M</td>
</tr>
<tr>
<td>Gene Symbols</td>
<td>Malignant solid tumour</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>AML3, MAP1B, MAP3K2, MARCH1, MAS1L, MAST4, MBTPS2, MC5R, MCF2L, MEST, MFSD2A, MA2, MICALCL, MKI67, MMD2, MME, MORN1, MS4A5, MTRR, MUC12, MUC22, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKD2, NKKX3-1, NLRP12, NLRP5, NLRP8, NPSR1, NRXN2, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, ODA5, OLFML2A, OPN3, OR10G9, OR12D2, OR1D5, OR1N1, OR2A1/OR2A42, OR2D3, OR4C6, OR4Q3, OR51G1, OR52I1, OR56B4, OR7G2, OR8A1, ORC6, OTOG, P2RX1, PACSIN3, PADI2, PAMR1, PAX2, PDGFRA, PDIK1L, PFKFB2, PGA5 (includes others), PGK2, PHF21B, PID1, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLEKHO1, PLXNA4, PMFBP1, PNLPL3, PPFBP, PPFIA4, PPI2L, PPPIP5K1, PPP1R14C, PRKCG, PRM1, PRR15L, PRR30, PRRX1, PRSS38, PSG11, PTER, PTGER3, PTGS1, PTPRH, PTPRU, RAB17, RAB27B, RAB44, RAP1GAP, RASGRF1, RASSF2, RBM20, RBM1A1 (includes others), REG3A, RFTN1, RGN, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RPS6KL1, RUNDC3A, S100A7, SBSN, SCD5, SCN7A, SCNN1D, SCTR, SETD1B, SFRP1, SFRP2, SFRP5, SFTPA1, SFTPA2, SHC3, SHISA2, SIGLEC12, SIRPB1, SIRPB2, SIS3, SLC11A1, SLC17A2, SLC22A2, SLC22A20, SLC24A2, SLC24A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC4A8, SLC5A10, SLC6A12, SLC9A5, SORCS2, SOX30, SPAC5A5/SPAC5B, SPINT3, SPOCD1, SPP1, SPRY3, SRMS, SSC5D, SSX3, STMN1, STRC, STRN, SUN5, SYCP1, SYK, TACR1, TCP10L, TGM6, TH, THBS2, TIRAP, TLR10, TM4SF5, TMEM154, TMEM236, TMEM257, TMEM266, TME8, TMN5, TNFRSF13B, TNK1, TONSLL, TOP2A, TOP3, TPO, TPSD1, TRPM3, TSHB, TSSK1B, TTC39A, TUBB2A, TUBB8, TWIST2, UBXN11, UCHL1, UGTL3A1, UMODL1, UROS, USP49, VASH2, VSG10L, VSTM4, VW5A5A, WWDE, WDF4, WFD5C5, WIF1, WT1, WT1AS, XIP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625</td>
<td>Down</td>
<td>ABCA6, ABCB6, ABCC9, ABCG5, ACPP, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS72, ADCY2, ADGRF1, ADHFE1, ADHDQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, AQP3, AQP4, ASPHD1, ASPN, ASTN2, ATG9A, ATP1A4, ATP2C1, ATP8A2, ATRLN1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BDGRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, BTN1A1, C15orf32, C17orf51, C19orf81, C1orf111, C1orf186, C1S, C22orf46, C3, C7orf25, C7orf65, C9orf152, C6A, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASC5, CASQ1, CCDC54, CCDC80, CCL19, CCNA2, CCNB1, CD160, CD177, CD19, CD1B, CD1C, CD5L, CDH9, CDKL2, CE</td>
</tr>
</tbody>
</table>
ACAM18, CEACAM8, CENPE, CEP55, CHEK2, CHRNA5, CHRNB3, CILP2, CLCNKA, CLEC6A, CLEC1L, CLMN, CLMP, CNPY1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COQ2, CPA2, CPNE9, CPXM1, CPXM2, CR2, CRNN, CTNNBP2, CX2, CXCR2, CXorf67, CYP1A2, CYP2S1, CYP4A22, DAAM2, DCBLD2, DCD, DCN, DEFB128, DGAT2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, DNAI2, DNAJC6, DPCR1, DRGX, DSC2, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, ELF5, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM195A, FAM209B, FBLN1, FBPI, FCER1A, FGFR2B, FCRL1, FCRL2, FCRL3, FCRLB, FERMT1, FGFR16, FGFR21, FGFR9, FGFR3, FM05, FOXI1, FSHR, FZD5, GABRQ, GALNT9, GAPT, GCKR, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPA33, GPAM, GPAT2, GPC6, GPR149, GPR34, GPR45, GPRC5D, GREB1L, GRIK3, GRIN2A, GRK1, GRK3, GRIA4, GUCA1B, GUCY2D, HAS1, HCRR1, HDCC2, HES6, HHIP, HLA-DQA2, HMGA2, HMGCS1, HMGCS2, HMGN5, HPGD5, HSD3B2, ID4, IGF1, IGF3, INS, ITGB6, ITGB5, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNH8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLG3, KLK4, KPRP, KRT34, KRTAP19-1, L1TD1, L2HGHD, LAMA2, LAMP3, LCE3D, LCE3E, LG13, LILRA4, LINCO0675, LOC102724788/P, RODH, LPAT2, LRRK2, LRTOMT, LT4A4H, LUM, LYNX1, MAGEC1, MAGEE1, MAMDC2, MAML3, MAP1B, MAP3K2, MARCH1, MAS1L, MAST4, MBTPS2, MC5R, MCF2L, MEST, MFS2A, MIA2, MIKCI, MKI67, MMD2, MME, MORN1, MS4A5, MTRR, MUC12, MUC22, MUC6, MUC7, MUSK, MXR5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKD2, NKKX3-1, NLRP12, NLRP5, NLRP8, NPSR1, NRXN2, NUTM2A, NUTM2B, NUTM2F, NUTM2G, NXPH3, NYAP2, ODAM, OFMIL2A, OPN3, OR10G9, OR12D2, OR1D5, OR1N1, OR2A1, OR2A42, OR2D3, OR4C6, OR4Q3, OR51G1, OR52I1, OR56B4, OR7G2, OR8A1, ORC6, OTG2, P2RX1, PACSIN3, PADI2, PAMR1, PAX2, PDGFRA, PDIK1L, PFKFB2, PGA5 (includes others), PGK2, PHF21B, PID1, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLEKHO1, PLXNA4, PMLF1, PNPLA3, PPBP, PPFIA4, PPIL2, PPPI5K1, PPP1R14C, PRKCG, PRM1, PRR15L, PRR30, PRRX1, PRSS38, PSG11, PTER, PTGER3, PTGS1, PTPRH, PTTRU, RAB17, RAB27B, RAB44, RAP1GAP, RASGRF1, RASSF2, RBM20, RBM15Y1A1 (includes others), REG3A, RFTN1, RGN, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RPL1L1, RPS6KL1, RUND3A, S100A7, SBSN, SCD5, SCN7A, SCN1D, SCTR, SETD1B, SFRP1, SFRP2, SFRP5, SFTA1, SFTPA2, SHC3, SHISA2, SIGLEC12, SIRPB1, SIRPB2, SIX3, SLC11A1, SLC17A2, SLC22A2,
SLC22A20, SLC24A2, SLC24A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC4A8, SLC5A10, SLC6A12, SLC9A5, SORCS2, SOX30, SPAC5A5/SPAC5A5B, SPINT3, SPOCD1, SPON1, SPRY3, SRMS, SSC5D, SSX3, STMN1, STRC, STRN, SUN5, SYCP1, SYK, TACR1, TCP10L, TGM6, TH, THBS2, TIRAP, TLR10, TM4SF5, TMEM154, TMEM257, TMEM266, TMEM87A, TME M97, TNFRSF13B, TNK1, TONSL, TOP2A, TP53, TPO, TPSD1, TRPM3, TSHB, TSSK1B, TTC39A, TUBB2A, TUBB3, TWIST2, UCHL1, UGT3A1, UMODL1, UROS, USP49, VASH2, VSIG10L, VSTM4, VWAA5A, VWDE, WDF5C, WDF6, WIF1, WT1, WT1-AS, XIRP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625

Neoplasia of epithelial tissue

Down

ABC6A, ABCC6, ABCG5, ACPP, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, ADCY2, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKR30A, ANKR53, ANP32C, AOX1, AQP4, ASPHDI, ASN2, ATG9A, ATP1A4, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BKRB1, BIR7, BNC2, BOLL, BRD3, BRSK2, BTK, C17orf51, C19orf81, C1orf186, C1S, C22orf46, C3, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASC5, CASQ1, CCDC54, CCDC80, CCNA2, CCNB1, CD160, CD177, CD19, CD1C, CD5L, CDH9, CDKL2, CEACAM18, CEACAM8, CENPE, CEP55, CHEK2, CHRNA5, CHRNB3, CLP2, CLCNKA, CLEC6A, CLEC1L, CLMN, CLMP, CNPY1, COL16A1, COL1A2, COL7A2, COLEC12, COMP, COQ2, CPA2, CPNE9, CPXM1, CPXM2, CR2, CRN, CTNBP2, CUX2, CXCRL2, CXorf67, CYP1A2, CYP2S1, CYPA422, DAAM2, DCBLD2, DCD, DNCB1, DFF48, DGAT2, DIP2C, DIS3L2, DLC1, DLPAP5, DLX4, DNAI2, DNAJC6, DPCR1, DRGX, DSC2, E2F1, E2F7, EBF1, EDAR, EGF, EGFAM, EML6, ENPP6, EBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM195A, FBLN1, FCR1A, FGR2B, FCR2C, FCR3, FERMT1, FGFR9, FGFR3, FMO5, FOXI1, FSHR, FZD5, GABRG, GALNT9, GAPT, GCKR, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPA33, GPAM, GPAT2, GP6C, GPR149, GPR34, GPR45, GPRC5D, GREB1L, GRK3, GRIN2A, GRK3, GSTA4, GUC1A1B, GUC2D, HAS1, HDCC2, HES6, HHIP, HLA-DQA2, HMG2A, HMGCS1, HMGCS2, HMGN5, HPGDS, HSD3B2, ID4, IGF1, ITGB6, ITGBL1, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNE2, KCNH8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KLK4, KPRR, KRT34, KRTAP19-1, L1TD1, L2HGDH, LAMA2, LAMP3, LCE3D, LCE3E, LGI3, LILRA4, LINCO00675, LOC102724788/P
<table>
<thead>
<tr>
<th>Digestive system cancer</th>
<th>Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA6, ABCC6, ABCC9, ABCC5, ACP5, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, ADCY2, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT2, ANKRD30A, ANKRD53, ANP32C, AOX1, AQP3, AQP4, ASPP1, ASTN2, ATG9A, ATP1A4, ATP2C1, ATP5A2, ATRN1, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BDKRB1, BIRC7, BNC2, BOL, BRD3, BRSK2, BTK, C17orf51, C1orf186, C15S, C22orf46, C3, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG7, CADM1, CALB2, CALCA, CAMP, CASCA5, CASQ1, CDC45, CDDC80, CCL19, CCNA2, CCNB1, CD160, CD177, CD1C, CDH9, CDFK2, CEACAM18, CEACAM8, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRN8, CLIP2, CLCNKA, CLCN6A, CLD1, CLMN, CLMP, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COQ2, CPA2, CPXM1, CPXM2</td>
<td></td>
</tr>
<tr>
<td>Abdominal cancer</td>
<td>Down</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
</tr>
</tbody>
</table>
| ABCA6, ABCB6, ABCC9, ABCC5, ACPP, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, ADCY2, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, AQP3, AQP4, ASPN, ASTN2, ATG9A, ATP1A4, ATP2C1, ATP8A2, ATRNL1, AWAT2, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BDKRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, C17orf51, C1orf186, C1S, C22orf46, C3, C7orf25, C9orf152, CA6, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CACNG7, CACNG8, CADM1, CALB2, CALCA, CAMP, CASC5, CASO1, CCDC54, CCDC80, CCNA2, CCNB1, CD160, CD177, CD19, CD1C, CD5L, CDH9, CDKL2, CEACAM18, CEACAM8, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRNB3, CILP2, CLCNKA, CLEC6A, CLMN, CLMP, CNPY1, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COQ2, CPA2, CPNE9, CPXM1, CPXM2, CR2, CRNN, CTTNB2, CUX2, CXCR2, CXorf67, CYP1A2, CYP4A22, DAAM2, DCBLD2, DCN, DGAT2, DIP2C, DIS3L2, DLC1, DLGAP5, DLX4, DNAI2, DNAJC6, DPCR1, DRGX, DSC2, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM195A, FBLN1, FCER1A, FCRGR2B, FCRL2, FCRL3, FCRBL, FERMT1, FGFR16, FGFR9, FGFR3, FMO5, FOXL1, FSHR, FZD5, GABRQ, GALT9, GEN1, GFPT2, GFRAR2, GINS2, GLI2, GNMT, Golga8K (includes others), GPA33, GPAM, GPAT2, GPC6, GPR149, GPR34, GPR45, GPRC5D, GREB1L, GRIK3, GRIK4, GRK1, GRK3, GSTA4, GUCA1B, GUCY2D, HAS1, HCCD2, HES6, HHIP, HLA-DQA2, HMGA2, HMQCS1, HMQCS2, HMGN5, HPDGS, HD3B2, ID4, IGFl, INS, ITGB6, ITGB1L, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNE2, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KLLK1, KPRP, KRT34, KRTAP19-1, L1TDL1, LH2GDH, LAMA2, LAMP3, LCE3E, LGI3, LINRA4, LINC00675, LOC102724788, PRODH, LPCAT2, LRRK2, LTA4H, LUM, LYNX1, MAGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAPK32, MARCH1, MAS1L, MAST4, MBTPS2, MC5R, MCF2L, MEIST, MFSD2A, MIA2, MICALCL, MKI67, MMET, MORN1, MS4A5, MTRR, MUC12, MUC22, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKD2, NKX3-1, NLRP12, NLRP5, NPSR1, NRXN2, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, ODAM, OLFML2A, OPN3, OR10G9, OR1N1, OR2A1/OR2A42, OR2D3, OR4C6, OR4Q3, OR51G1, OR56B4, OR7G2, OR8A1, ORC6, OTOG, P2RX1, PACSIN3, PADI2, PAMR1, PAX2, PDGFRA, ...
<table>
<thead>
<tr>
<th>Digestive organ tumour</th>
<th>Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA6, ABCC6, ABCG5, ACPP, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS2, ADCY2, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP1, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKRD30A, ANKRD53, ANP32C, AOX1, AQP3, AQP4, ASPN, ASTN2, ATG9A, ATP1A4, A TP2C1, ATP8A2, ATRN1, AZGP1, B3GNT8, BACH2, BANK1, BCL11A, BDKRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, C1orf51, C1orf186, C1S, C22orf46, C3, C7orf25, C9orf152, CA6, CACNA1C, CACNA2D2, CACNB4, CACNG7, CADM1, CALB2, CALCA, CAMP, CASC5, CASQ1, Ccdc54, CCDC80, CCL19, CCNA2, CCB1, CD160, CD177, CD1C, CDH9, CDK2, CEACAM18, CEACAM8, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRNB3, CLIP2, CLCNKA, CLEC6A, CLE CL1, CLMN, CLMP, COL16A1, COL1A1, COL8A2, COLEC12, COMP, COQ2, CPA2, CPXM1, CPXM2, CR2, CRN, CTCNB2, CUX2, CXCR2, CXorf67, CYP1A2, CYP4A22, DAAM2, DCBLD2, DCN, D GAT2, DIP2C, DIS3L2, DLCL1, DLGAP5, DLX4, DNAI2, DNAJC6, DPCR1, DRGX, DSC2, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESPN, ESR1, ES R2, ET1, F10, F11, FAM107A, FAM129C, FAM195A, FBLN1, FCR1A, FCGR2B, FCR2, FCR3, FCR8, FERMT1, FGF3, FMO5, FOXL1, FSHR, FZD5, GABRQ, GALNT9, GEN1, GFT P2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), REG3A, RFTN1, RGN, RIMKLA, RIMS3, RLB1, ROR1, ROR2, ROS1, RPL1L, RPS6KL1, R UNDC3A, SBSN, SCD5, SCN7A, SCN1D, SCTR, SETD1B, SF3BP1, SFRP5, SFTA1, SFTA2, SHTS, SIRPB1, SIRPB2, SIX3, SLC11A1, SLC17A2, SLC22A1, SLC22A2, SLC23A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC4A8, SLC6A12, SLC9 A5, SORCS2, SOX30, SPAC5/SPAC5A5B, SPOCD1, SPO1N, SRMS, SSC5D, SSX3, STMN1, STR C, SYCP1, SYK, TCP10L, TGM6, TH, THBS2, TLR10, TM4SF5, TMEM154, TMEM266, TME M87A, TMEM97, TNNFRSF13B, TNK1, TOSNO, TOP2A, TP53, TPO, TPSD1, TRPM3, TSSK1B, TUBB 2A, TUBB8, TWIST2, UBXN11, UCHL1, UGT3A1, UMODL1, UROS, USP49, VASH2, VSG10L, VST M4, VWA5A, WDFE, WDFY4, WDHC5, WIF1, WT1, WT1- AS, XRIP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625</td>
<td></td>
</tr>
<tr>
<td>others(), GPAM, GPAT2, GPC6, GPR149, GPR34, GPR45, GPRC5D, GREB1L, GRIK3, GRIN2A, GRK1, GSTA4, GUCA1B, GUCY2D, HAS1, HDDC2, HES6, HHIP, HLA-DQA2, HMGCS1, HMGCS2, HMGN5, HPGD, HSD3B2, ID4, IGF1, INS, ITGB6, ITGBL1, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNJ8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KLK4, KPRP, KRT34, KRTAP19-1, L1TD1, L2HGDH, LAMA2, LAMP3, LCE3E, LGI3, LILRA4, LINC00675, LOC102724788/PRODH, L21CAT, LRRK2, LTA4H, LYNX1, MAGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAP3K2, MARC-H1, MAS1L, MAST4, MBTPS2, MC5R, MCF2L, MEST, MFS2DA, MIA2, MICALCL, MKI67, MME, MORN1, MS4A5, MTRR, MUC12, MUC22, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKO2, NKK3-1, NLRP12, NLRP5, NLRP8, NPSR1, NRXN2, NUTM2A/NUTM2B, NUTM2F/NUTM2G, NXPH3, NYAP2, ODA2, OFM1L2A, OPN3, OR10G9, OR1N1, OR2D3, OR4C6, OR4Q3, OR51G1, OR56B4, OR7G2, OR8A1, ORC6, OTOG, P2RX1, PACSIN3, PADI2, PARM1, PAX2, PDGFRA, PFKFB2, PAG5 (includes others), PGK2, PHF21B, PID1, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLEKHO1, PLXNA4, PMFB1, PNPLA3, PPFIA4, PPL2, PPIP5K1, PRKCG, PRR30, PRRX1, PRSS38, PSON1, PTER, PTGS1, PTPRH, PTPRJ, RAB17, RAB27B, RAP1GAP, RASGRF1, RASSF2, RBM20, RBMY1A1 (includes others), REG3A, RFTN1, RGN, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RPS6KL1, RUND3A, SBSN, SCD5, SCN7A, SCN9A, SCTR, SERT1B, SFPR1, SFPR5, SFTP1A, SHISA2, SIGLEC12, SIRPB2, SIX3, SLC11A1, SLC17A2, SLC22A2, SLC22A20, SLC24A2, SLC24A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC6A12, SLC9A5, SORCS2, SOX30, SPOCD1, SPON1, SRMS, SSC5D, SSX3, STMN1, STRC, SYCP1, SYK, TACR1, TCP10L, TGM6, TH, THBS2, TLR10, TM4SF5, TMEM154, TMEM97, TNSFRSF13B, TNK1, TNSL1, TOP2A, TP53, TPO, TRPS1, TSSK1B, TUBB2A, TUBB8, UBXN11, UCHL1, UGT3A1, UMODL1, UROS, USP49, VASH2, VSIG10L, VSTM4, VWA5A, VVDE, WDFY4, WFDC5, WT1, XIRP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625</td>
<td></td>
</tr>
</tbody>
</table>

| Large intestine neoplasm | Down | ABCA6, ABCC6, ABCC9, ABCC5, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, Ancy2, ADHF1E, ADIPOQ, AGMAT, AGTR2, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKR D30A, ANKRD53, AOX1, AQP3, ASTN2, ATP1A4, ATP2C1, ATRNL1, B3GNT8, BACH2, BANK1, BC L11A, BDKRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, C17orf51, C1orf186, C1S, C22orf46, C3, |
| C7orf25, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CADM1, CAMP, CCDC54, CCDC80, CCN A2, CCNB1, CD160, CD177, CD1C, CDH9, CDKL2, CEACAM18, CEACAM8, CENPE, CEPF, CEP 55, CHEK2, CHRNA5, CLEC6A, CLMN, CLMP, COL16A1, COLEC12, COMP, COQ2, CPA2, CPXM1, CR2, CRNN, CTTNPBP2, CUX2, CXCR2, CXorf67, CYPA12, DCBLD2, DCN, DGAT2, DIP2C, DIS3L2, DL1C, DLGAP5, DLX4, DNAI2, DPCR1, DRGX, E2F1, E2F7, EBF1, EDAR, EGFL6, EGFLAM, EML6, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FA M195A, FBLN1, FCGR2B, FCRL2, FCRL3, FERMT1, FGFR3, FMO5, FOXI1, FSHR, FZD5, GALNT9, GEN1, GFPT2, GFRA2, GINS2, GLI2, GNMT, GOLGA8K (includes others), GPAM, GPC6, GPR149, GPR34, GPR45, GREG1L, GRK3, GRIN2A, GRK1, GFN4A, GUCA 1B, GUCY2D, HAS1, HHIP, HMGCS1, HMGCS2, HMGN5, HPGDS, HSD3B2, IGF1, INS, ITGB6, ITG BL1, ITLN1, JAKMIP2, KCNA6, KCNE1, KCNH8, KCNJ15, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KLK4, KPRP, KRT34, L2HDH, LAMA2, LAMP3, LCE3E, LILRA4, LINC00675, LOC102724788/PRO DH, LPCAT2, LRRK2, LTA4H, LYNX1, MAGEC1, MAGEE2, MAMDC2, MAML3, MAP1B, MAP3K2, MARCH1, MAST4, MBTPS2, MC5R, MCF2L, MEST, MFSD2A, MIA2, MKI67, MME, MS4A5, MTRR, M UC12, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKD2, NKX3- 1, NLRP12, NLRP5, NLRP8, NSR1, NRXN2, NUTM2F, NUTM2G, NXPH3, NYAP2, ODAM, OLFM12, OPA3, OR10G9, OR1N1, OR2D3, OR65B4, OR7G2, ORC6, P2RX1, PACSIN3, PAD2, PAMR1, PA2X1, PDGFA, PFKFB2, PKG2, PHF21B, PIP2D, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLXNA4, PMFBP1, PPLA3, PPFA2, PRKCG, PRR30, PRRX1, PRSS38, PTGS1, PRPT1, RAB17, RAP1GAP, RASGRF1, RBM20, RFTN1, RIMKLA, RIMS3, RLBP1, ROR1, ROR2, ROS1, RP1L1, RP S6KL1, RUND3A, SCDD5, SCN7A, SCNN1D, SCTR, SETD1B, SFRP1, SFRP5, SHISA2, SIGLEC12, SIX3, SLC11A1, SLC22A20, SLC24A2, SLC24A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC6A12, SLC9A5, SORCS2, SOX30, SPOCD1, SPOC1, SPON1, SRMS, SSC5D, SSX3, STRC, SYCP1, SYK, TACR1, TCP10L, TGM6, THEB2, TLR10, TMEM97, TNFRSF13B, TNK1, TONS, TOP2A, TP53, TPO, TRPM3, TSSK1B, TUBB2A, TUBB8, UBXN11, UCHL1, UGT3A1, UROS, VASH 2, VSTM4, VWA5A, VWDE, WDFY4, WT1, XIRP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625 |

| Malignant neoplasm of large intestine | Down | ABCA6, ABCC6, ABCG9, ABCC5, ACSM3, ACTC1, ADAM12, ADAM30, ADAMTS17, ADAMTS7, ADCA2, ADHFE1, ADIPOQ, AGMAT, AGTR2, AKAP17A, ALDH3A1, AMER2, ANGPT4, ANK3, ANKR D30A, ANKRD53, AOX1, AQP3, ASTN2, ATP1A4, ATP2C1, ATRNL1, B3GNT8, BACH2, BANK1, BC |
L11A, BDKRB1, BIRC7, BNC2, BOLL, BRD3, BRSK2, BTK, C17orf51, C1orf186, C1S, C22orf46, C3, C7orf25, CACNA1C, CACNA1G, CACNA2D2, CACNB4, CADM1, CAMP, CDC20, CCDC80, CCNA2, CCNB1, CD160, CD177, CD1C, CDH9, CDKL2, CEACAM18, CEACAM8, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CLC6A, CLMN, CLMP, COL16A1, COLEC12, COMP, COQ2, CPA2, CPGM1, CR2, CRNN, CTTNBP2, CUX2, CXCR2, CXorf67, CYF1A2, D2B, DCBLD2, DCN, DIP2C, DIS3L2, DCL1, DLGAP5, DLX5, DNA2, DPR1, DRG1, E2F1, E2F7, EBF1, EDAR, EGFL1, EGFAM, EML2, ENPP6, ERBB4, ERCC4, ERCC5, ERICH4, ESR1, ESR2, ETV1, F10, F11, FAM107A, FAM129C, FAM195A, FBLN1, FGCR2B, FCRL2, FCRL3, FERMT1, FGFR3, FMO5, FOX11, FSHR, FZD5, GALNT9, GEN1, GFPT2, GFA, GINS2, GLI2, GNMT, GOLGA8(K includes others), GPAM, GPC6, GPR149, GPR34, GPR45, GREB1L, GRIK3, GRIN2A, GRK1, GSTA4, GUCA1B, GUCY2D, HAS1, HHIP, HMGCS1, HMGCS2, HMGN5, HPGDS, HSD3B2, IGFI, INS, ITGB6, ITGB8, ITP1, ITLN1, JAKMIP2, KCNA6, KCNE1, KCN1, KCN8, KCN15, KIF2B, KIRREL3, KIT, KLHL29, KLHL38, KPR, KRT34, L2GHDH, LAMA2, LAMP3, LCE3E, LILRA4, LINC00675, LOC102724788, PRODH1, LPPL2, LRRK2, LTA4H, LYNX1, MAGEC1, MAGEE2, MAMD2, MAML3, MAP1B, MAP3K2, MARCH1, MAST4, MBTPS2, MC5R, MCF2L, MEST, MFS2D2, MIA2, MII67, MME, MS4A5, MTRR, MUC12, MUC6, MUC7, MUSK, MXRA5, MXRA8, MYO3A, NEBL, NGEF, NIN, NINL, NKD2, NKX3-1, NLRP12, NLRP5, NLRP8, NPSR1, NNXR2, NUTM2F, NUTM2G, NXP3, NYAP2, OSM, OLFML2A, OSMN, OR10G9, OR1N1, OR2D3, OR56B4, OR7G2, ORC6, P2RX1, PANS3, PAD2, PARM1, PAX2, PDGFRA, PFKFB2, PGK2, PHF21B, PID1, PIP5K1B, PLA2G2A, PLA2G4F, PLCH1, PLXNA4, PMF1P1, PNPLA3, PPF1A4, PRKCG, PRR30, PRX1, PRSS38, PTGS1, PTPRH, PTGPR, RAB17, RAP1GAP, RASGRF1, RBM20, RFTN1, RIMKLA, RIMS3, RBP1, ROR1, ROR2, ROS1, RP1L1, RP5, R6K1, RUNDC3A, SCD5, SCN7A, SCNN1D, SCTR, SETD1B, SFRP1, SFRP5, SHISA2, SIGLEC12, SIX3, SLC11A1, SLC22A20, SLC24A2, SLC24A3, SLC26A3, SLC28A1, SLC39A8, SLC45A3, SLC46A2, SLC47A1, SLC6A12, SLC9A5, SORCS2, SOX30, SPOC1, SPON1, SRRMS, SSCD5, SSX3, SSTR, SYCP1, SYK, TAC1R1, TCP10L, TGM6, THBS2, TLR10, TMEM97, TNFRSF13B, TNIK, TONSL, TOP2A, TP53, TPO, TRPM3, TSSK1B, TUBB2A, TUBB8, UBXN11, UCHL1, UGT3A1, UROS, VASH2, VSTM4, VWA5A, VWDE, WDFY4, WT1, XIRP2, ZMYND19, ZNF236, ZNF365, ZNF385B, ZNF491, ZNF568, ZNF625

Urogenital cancer Down ABCA6, ABCB6, ABCG1, ACPP, ACSM3, ACTC1, ADAMTS17, ADAMTS7, ADcy2, ADIPOQ, AGTR2, ALDH3A1, ANK3, ANKRD38A, ASPN, ASTN2, ATG9A, ATP2C1, ATP8A2, ATRNL1, AW
<p>| AT2, AZGP1, BANK1, BCL11A, BIRC7, BNC2, C1S, C3, CACNA1C, CACNA1G, CACNG8, CADM1, CALB2, CALCA, CASC5, CASQ1, CCDC80, CCNA2, CCNB1, CD19, CD5L, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRN3, CLMN, CNPY1, COL1A1, COLEC12, CPNE9, CPXM1, CPXM2, CRN, CTNBP2, CUX2, CXCR2, CYP1A2, CYP4A22, DCC, DGAT2, DIS3L2, DLC1, DLGAP5, DLX4, DNAJC6, DPCR1, E2F1, E2F7, EGFLAM, EML6, ERBB4, ERCC5, ESP1, ESP2, ETV1, F10, F11, FAM107A, FAM129C, FBLN1, FCRL2, FCRL3, FGF16, FGF9, FGFR3, FOXI1, FSHR, GABRQ, GEN1, GLI2, GPA33, GPAT2, GRK3, GRK3, HLA-DQA2, HMGAA2, HMGAN5, HSD3B2, IGF1, INS, ITLN1, JAKMIP2, KCHN8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLF4, KRT34, L1TD1, L2HGDH, LAMA2, LAMP3, LGI3, LILRA4, LPCAT2, LRRK2, LTA4H, LUM, MAGEC1, MAML3, MAP1B, MAP3K2, MAS1L, MAST4, MBTPS2, MC5R, MFSD2A, MICALCL, MKI67, MME, MUC12, MUC22, MUC6, MUC7, MUSK, MXRA8, MYO3A, NGEF, NINL, NLRP12, NLRP5, NUTM2A/NUTM2B, NYAP2, OLFM2A, OR10G9, OR2A1/OR2A42, OR2D3, OR7G2, PADI2, PAX2, PDGFRα, PDK1L1, PFKFB2, PGK2, PHF21B, PIP5K1B, PLA2G2A, PLXNA4, PMFBP1, PNPLA3, PPBP, PPFIA4, PPI2, PRKCG, PRSS38, PTGER3, PTGS1, PTPRU, RAB44, RAP1GAP, RASSF2, RBM20, RGN, RIMKLA, ROS1, RP1L1, SCN7A, SCN11D, SFRP1, SFTPA2, SIGLEC12, SIRP1, SIRPB1, SLC11A1, SLC22A2, SLC22A3, SLC24A3, SLC39A8, SLC45A3, SLC47A1, SLC4A8, SLC9A5, SORCS2, SPACA5/SPACA5B, SPOCD1, SRMS, SXS3, STMN1, STRC, TAC1R1, THBS2, TM4SF5, TMEM266, TMEM87A, TNFRSF13B, TNIK, TOP2A, TP53, TPSD1, TRPM3, TSSK1B, TUBB2A, TWIST2, UCHL1, USP49, WDFY4, WIF1, WT1, WT1-AS, XIRP2, ZNF236, ZNF365, ZNF385B, ZNF491 |
| Genitourinary carcinoma | Down | ABCA6, ABCC6, ABCC9, ABCG5, ACPP, ACSM3, ACTC1, ADAMTS17, ADAMTS7, ADCY2, ADIPQ, AGTR2, ANK3, ANKR3D30A, ASPN, ASTN2, ATG9A, ATP2C1, ATP8A2, ATRN1L1, AWAT2, BANK1, BCL11A, BIRC7, BNC2, C1S, C3, CACNA1C, CACNA1G, CACNG8, CADM1, CALB2, CASC5, CASQ1, CCDC80, CD19, CD5L, CENPE, CENPF, CEP55, CHEK2, CHRNA5, CHRN3, CLMN, CNPY1, COL1A1, COLEC12, CPNE9, CPXM1, CPXM2, CRN, CTNBP2, CUX2, CXCR2, CYP1A2, CYP4A22, DCC, DGAT2, DIS3L2, DLC1, DLGAP5, DNAJC6, DPCR1, E2F7, EGFLAM, EML6, ERBB4, ERCC5, ESP1, ESP2, ETV1, F10, F11, FAM107A, FAM129C, FCRL2, FCRL3, FGF16, FGF9, FGFR3, FOXI1, FSHR, GABRQ, GEN1, GLI2, GPA33, GPAT2, GRK3, GRK3, HLA-DQA2, HMGAA2, HMGAN5, HSD3B2, IGF1, ITLN1, JAKMIP2, KCHN8, KCNJ15, KCNK9, KIF2B, KIRREL3, KIT, KLHL29, KLF4, KRT34, L1TD1, L2HGDH, LAMA2, LAMP3, LGI3, LILRA4, LPCAT2, LRRK2, LTA4H, LUM, MAGEC1, MAML3, MAP1B, MAP3K2, MAS1L, MAST4, MBTPS2, MC5R, MFSD2A, MICALCL, MKI67, MME, MUC12, MUC22, MUC6, MUC7, MUSK, MXRA8, MYO3A, NGEF, NINL, NLRP12, NLRP5, NUTM2A/NUTM2B, NYAP2, OLFM2A, OR10G9, OR2A1/OR2A42, OR2D3, OR7G2, PADI2, PAX2, PDGFRα, PDK1L1, PFKFB2, PGK2, PHF21B, PIP5K1B, PLA2G2A, PLXNA4, PMFBP1, PNPLA3, PPBP, PPFIA4, PPI2, PRKCG, PRSS38, PTGER3, PTGS1, PTPRU, RAB44, RAP1GAP, RASSF2, RBM20, RGN, RIMKLA, ROS1, RP1L1, SCN7A, SCN11D, SFRP1, SFTPA2, SIGLEC12, SIRP1, SIRPB1, SLC11A1, SLC22A2, SLC22A3, SLC24A3, SLC39A8, SLC45A3, SLC47A1, SLC4A8, SLC9A5, SORCS2, SPACA5/SPACA5B, SPOCD1, SRMS, SXS3, STMN1, STRC, TAC1R1, THBS2, TM4SF5, TMEM266, TMEM87A, TNFRSF13B, TNIK, TOP2A, TP53, TPSD1, TRPM3, TSSK1B, TUBB2A, TWIST2, UCHL1, USP49, WDFY4, WIF1, WT1, WT1-AS, XIRP2, ZNF236, ZNF365, ZNF385B, ZNF491 |</p>
<table>
<thead>
<tr>
<th>Renal lesion</th>
<th>Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA6, ADAMTS7, ADCY2, ADIPOQ, AGTR2, ANKRD30A, BDKRB1, C1S, C3, CADM1, CENPE, CENPF, CHEK2, CPXM2, CRNN, CYP4A22, DCN, DGAT2, DIS3L2, DLGAP5, DPCR1, ERBB4, ERCC5, ESR1, FGBR3, GABRF2, GPRAT2, GRIK3, IGFR1, INS, KCNJ15, KIT, KLL4, L1TD1, LAMA2, LRRK2, LTA4H, MAST4, MKI67, MME, MUC12, MUC6, MUSK, MYO3A, NGEF, NLRP12, PADI2, PDGFRB, PPRP, PRL2, RAP1GAP, RGN, RP1L1, SFRP1, SLC39A8, SLC47A1, SLC4A8, STMN1, TH, TMEM266, TNFRSF13B, TNK, TP53, TPSD1, TUBB2A, WIF1, WT1, XIRP2</td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Microarray</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>EPHB1</td>
<td>2015 miRNA</td>
</tr>
<tr>
<td>FOS</td>
<td>2015 miRNA</td>
</tr>
<tr>
<td>Gene</td>
<td>miRNA Year</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>OGT (O-linked N-acetylglucosamine transferase)</td>
<td>2015</td>
</tr>
<tr>
<td>REL</td>
<td>2015</td>
</tr>
<tr>
<td>RHOA (ras homolog family member A)</td>
<td>2015</td>
</tr>
<tr>
<td>Gene</td>
<td>Year</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>FGF2</td>
<td>2015</td>
</tr>
<tr>
<td>TGF-β2</td>
<td>2015</td>
</tr>
</tbody>
</table>

**FGF2** (Fibroblast growth factor 2)

- **miRNA**: 2
- **Effect**: Down
- **Function**: Promotes cellular growth and differentiation including angiogenesis
- **Expression**: Increased expression in bronchial epithelium in COPD
- **Diseases**: Astrocytoma, Bladder cancer, Breast cancer, Colorectal cancer, Endometrial cancer, Head and neck cancer, Hepatocellular carcinoma, Lung cancer, Melanoma, Oesophageal cancer, Ovarian cancer, Pancreatic cancer, Prostate cancer, Renal cell carcinoma

**TGF-β2** (Transforming growth factor beta 2)

- **miRNA**: 2
- **Effect**: Down
- **Function**: Regulates gene transcription. Promotes epithelial-mesenchymal transition
- **SNP**: SNP associated with severe COPD
- **Diseases**: Breast cancer, Cervical cancer, Colon cancer, Gastric cancer, Glioma, Hepatocellular carcinoma, Lung cancer, Melanoma, Neuroblastoma, Oesophageal cancer, Ovarian cancer, Pancreatic cancer, Prostate cancer, Renal cell carcinoma

(492, 582-600)
<table>
<thead>
<tr>
<th>miRNA</th>
<th>YOD1</th>
<th>2015 miRNA</th>
<th>2</th>
<th>Down</th>
<th>Deubiquitinase enzyme</th>
<th>Cervical cancer</th>
<th>(615, 616)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smad 2/3</td>
<td>2014 miRNA</td>
<td>1</td>
<td>Up</td>
<td>Regulates transcription</td>
<td>Increased in COPD related to the severity of airway obstruction</td>
<td>Bladder cancer, Breast cancer, Colorectal cancer, Hepatocellular carcinoma, Gastric cancer, Ovarian tumours, Prostate cancer</td>
<td>(481, 483, 617-623)</td>
</tr>
<tr>
<td>VEGF</td>
<td>2014 miRNA</td>
<td>1</td>
<td>Up</td>
<td>Major growth factor involved in angiogenesis</td>
<td>Increased in chronic bronchitis. Reduced in emphysema.</td>
<td>Bladder cancer, Breast cancer, Cervical cancer, Colorectal cancer, Endometrial carcinoma, Gastric cancer, <strong>Lung cancer</strong>, Lymphoma, Oesophageal carcinoma, Osteosarcoma, Ovarian cancer, Pancreatic cancer, Prostate cancer, Renal cell carcinoma</td>
<td>(211, 243, 624-636)</td>
</tr>
<tr>
<td>Gene</td>
<td>Year</td>
<td>miRNA Type</td>
<td>Score</td>
<td>Regulation</td>
<td>mRNA/Protein Function</td>
<td>Disease/Condition</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>------------</td>
<td>-------</td>
<td>------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>EIF2S1</td>
<td>2014</td>
<td>miRNA</td>
<td>7</td>
<td>Down</td>
<td>Regulates protein synthesis</td>
<td>Brain tumours, Breast cancer, Leukaemia, <strong>Lung cancer</strong>, Melanoma, Prostate cancer, Renal cell carcinoma, Thyroid carcinoma</td>
<td></td>
</tr>
<tr>
<td>IMPACT</td>
<td>2014</td>
<td>miRNA</td>
<td>7</td>
<td>Down</td>
<td>Regulates mRNA translation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAVER</td>
<td>2014</td>
<td>miRNA</td>
<td>7</td>
<td>Down</td>
<td>Regulates mRNA translation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBX5</td>
<td>2016</td>
<td>mRNA</td>
<td>22</td>
<td>Down</td>
<td>Organises chromatin. Promotes cell proliferation. Controls the differentiation and angiogenic function of endothelial progenitor cells.</td>
<td>Breast cancer, <strong>Lung cancer</strong></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Year</td>
<td>RNA Type</td>
<td>Regulation</td>
<td>Function</td>
<td>COPD Associations</td>
<td>Cancer Associations</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>----------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>VASH2 (Vashohi bin-2)</td>
<td>2016</td>
<td>mRNA</td>
<td>Up</td>
<td>Promotes angiogenesis, cell proliferation and epithelial-mesenchymal transition.</td>
<td>Endometrial cancer, Breast cancer, Hepatocellular carcinoma, Ovarian carcinoma, Pancreatic cancer</td>
<td>(656-661)</td>
<td></td>
</tr>
<tr>
<td>E2F1 (E2F transcription factor 1)</td>
<td>2014</td>
<td>mRNA</td>
<td>Up</td>
<td>Promotes transcription of genes including the NF-κB pathway and cell proliferation.</td>
<td>Bladder cancer, Breast cancer, Colorectal cancer, Gastric cancer, Lung cancer, Oesophageal cancer, Ovarian cancer, Prostate cancer, Renal cell cancer, Testicular cancer</td>
<td>(662-673)</td>
<td></td>
</tr>
</tbody>
</table>

The table lists each gene, the microarray analysis it was identified in, whether it was up or downregulated in COPD, its function and COPD or cancer associations.
**Genes altered in the canonical pathway ‘Eicosanoid signalling’**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression level</th>
<th>Function</th>
<th>COPD associations</th>
<th>Cancer associations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYSLTR (Cysteinyl leukotriene receptor)</td>
<td>Up (predicted)</td>
<td>Receptor for cysteinylated leukotrienes. Signalling via CYSLTR is associated with bronchoconstriction and eosinophilia.</td>
<td>Increased levels in bronchial cells in COPD. Increased in exacerbations. Associated with emphysema in animal models.</td>
<td>Gastric cancer</td>
<td>(674-678)</td>
</tr>
<tr>
<td>PLA2 (Phospholipase A2)</td>
<td>Up</td>
<td>Arachidonic acid production from phospholipids.</td>
<td>Increased activity in BAL fluid.</td>
<td>Astrocytoma, Colorectal cancer, Endometrial cancer, Gastric cancer, Pancreatic cancer, Prostate cancer</td>
<td>(484, 700-706)</td>
</tr>
<tr>
<td>------------------------</td>
<td>----</td>
<td>-----------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>PTGDR (Prostaglandin D receptor)</td>
<td>Up (predicted)</td>
<td>Activation results in cytokine release and macrophage migration. Increases migratory capacity and survival of neutrophils in vitro.</td>
<td>Increased neutrophils and reduced lung compliance after PTGDR activation.</td>
<td>Bladder cancer, Cervical cancer, Colorectal cancer, Lung cancer, Lymphoma</td>
<td>(707-712)</td>
</tr>
<tr>
<td>PTGDS (Prostaglandin D synthase)</td>
<td>Up</td>
<td>Synthesises PGD2. Urinary biomarker for lupus nephritis. Increased levels in preterm births. Associated with coronary vasoconstriction. Induces apoptosis and prevents cell cycle progression.</td>
<td>Increased levels in moderate vs mild COPD. Increased PGD2 levels in BAL from COPD patients. Levels related to FEV1.</td>
<td>Lung cancer, Medulloblastoma, Meningioma, Ovarian cancer, Testicular cancer</td>
<td>(485, 713-720)</td>
</tr>
</tbody>
</table>
| **PTGER2**  
(Prostaglandin E Receptor 2)  | Up (predicted) | Reduces fibroblast proliferation. Increases HIF-1α and VEGF expression. Induces cirrhosis association immunosuppression. Increases cell invasion and migration. Induces IL-8. | Increased levels in COPD fibroblasts resulting in increased VEGF and IL-8 production. | Breast cancer, Colorectal cancer, Hepatocellular carcinoma, **Lung cancer**, Gastric cancer, Neuroblastoma, Oesophageal cancer, Prostate cancer, Squamous cell carcinoma | (730-741) |
| **PTGER3**  
| **PTGER4**  
(Prostaglandin E receptor 4)  | Up (predicted) | Increases the production of IL-6, IL-8, M-CSF and VEGF. Increases IL-23 and induces Th17 cells. Induces bone remodelling. Increases angiogenesis. Increases PGE2 (positive feedback). Reduces TEM in HPEC. Involved in the cellular response to oxidative stress. Inhibits platelet | | Breast cancer, Cervical cancer, Colon cancer, Endometrial carcinoma, Gallbladder carcinoma, Gastric cancer, Head and neck cancer, **Lung cancer**, Melanoma, Prostate cancer, Renal cell carcinoma | (753-771) |
| (PTGIR) Prostaglandin I (prostacyclin) receptor | Up (predicted) | Increases angiogenesis. PTGI binds to PPARy (protective in lung cancer). | Reduction seen in pulmonary endothelial cells in emphysema. | Breast cancer, **Lung cancer** (772-777) |
| (PTGFR) Prostaglandin F receptor | Up (predicted) | Induces bronchoconstriction and airway plasma exudation. Reduced in myometrium during parturition. Regulates endothelial network formation. Previous TEM in endothelium | Bronchoconstriction | Breast cancer, **Lung cancer**, Renal Cell cancer (778-784) |
| PTGS (Prostaglandin synthase) | Up | Production of prostaglandin H2 from arachidonic acid | Higher levels in fibroblasts cultured from COPD lung. | Bladder cancer, Colorectal cancer, **Lung cancer** (486, 785-787) |
| TXS (Thromboxane synthase) | Down (2016 arrays only) | Converts PGH2 to TXA2. TXA2 is a vasoconstrictor that promotes platelet aggregation and is important in angiogenesis. Inhibition leads to lung cancer cell apoptosis. | Bladder cancer, **Lung cancer** (797-800) |

The table lists each gene, whether it was up or downregulated in COPD, its function and COPD or cancer associations. LTA4 = Leukotriene A4, LTB4 = Leukotriene B4, EPCs = Endothelial Progenitor Cells, BAL = Bronchoalveolar lavage, PGD2 = Prostaglandin D2, FEV1 = Forced Expiratory Volume in 1 second, MMP2 = Matrix Metalloproteinase 2, IL = Interleukin, NF-Kb = Nuclear factor kappa-light-chain-enhancer of Activated B cells, HIF-1α = Hypoxia Inducible Factor 1α, VEGF = Vascular Endothelial Growth Factor, M-CSF = Macrophage Colony Stimulating Factor, Th17 = T helper 17 cells, PPARγ = Peroxisome Proliferator Activated Receptor Gamma, TEM= Tumour Endothelial Marker, COX-2 = Cyclo-oxygenase 2, PGH2 = Prostaglandin H2, TXA2 = Thromboxane A2