The Impact of Methylarginine Metabolism on Disease Progression in Idiopathic Pulmonary Fibrosis

A thesis presented for the degree of Doctor of Medicine (Research) by

Harpreet Kaur Lota

Imperial College London

Medical Research Council
London Institute of Medical Sciences
Imperial College London
Hammersmith Hospital Campus
Du Cane Road
London W12 0NN
DECLARATION OF ORIGINALITY

I, Harpreet Kaur Lota, confirm that the work presented in this thesis is my own. Where information has been derived from external sources or data obtained by others, I can confirm this has been referenced in the text.

Harpreet Lota
February 2018
COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.
ABSTRACT

Idiopathic pulmonary fibrosis (IPF) is a complex, progressive and irreversible lung disease of unknown aetiology with a median survival of only three years following diagnosis. Signatures of increased nitric oxide (NO) and its generating enzyme, inducible nitric oxide synthase (iNOS), are observed in experimental models of idiopathic pulmonary fibrosis (IPF). Asymmetric dimethylarginine (ADMA) competitively inhibits iNOS, and is hydrolysed by dimethylarginine dimethylaminohydrolase (DDAH) 1 and 2. Previous evidence suggests that regulation of NO production via DDAH activity may play a role in IPF. Using both in vitro and human studies, this project investigated the relationship between the NO-ADMA-DDAH pathway and the progression of lung fibrosis; more specifically that variation in DDAH activity will impact on disease progression in IPF.

The in vitro study aimed to identify the mechanistic pathways linking DDAH activity to lung fibrosis in an experimental model of transforming growth factor-β (TGF-β) mediated epithelial to mesenchymal transition (EMT) in a human alveolar epithelial cell line. DDAH2 and iNOS were induced during TGF-β mediated EMT, indicating that DDAH may play a role in the transition towards the fibrotic phenotype. The human study aimed to identify the impact of DDAH single nucleotide polymorphisms (SNPs) with a known functional effect on ADMA levels in a normal cohort on longitudinal lung function decline and survival in a cohort of IPF patients. The results indicated that DDAH1 variant rs530006 (associated with lower ADMA levels in a normal cohort) was significantly associated with accelerated lung function decline and mortality in the IPF cohort. DDAH2 SNPs were not associated with lung function decline or mortality in the IPF cohort. Neither DDAH1 or DDAH2 SNPs were associated with lung function decline in a large cohort of patients with systemic sclerosis associated interstitial lung disease (SSc-ILD), supportive of the genetic distinction that is emerging between the two fibrotic lung diseases.

Overall, the evidence accumulated provides some novel insights into the potential role of DDAH in IPF. The complicated relationship between inherited genetic variants and environmental factors is a growing field in IPF, and DDAH biology may be one of the multiple pathways that leads to the fibrotic phenotype that we know as IPF.
ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisor James Leiper. Without his guidance, supervision, enthusiastic discussion and support of ideas, this project would not have taken shape. My colleagues in James' laboratory group were a talented team to work alongside and learn from, and I will always be grateful for the benchside skills they have taught me. Equally valuable was their company and sharing of ideas over many lively lunchtime discussions.

The members of the Interstitial Lung Disease Unit at the Royal Brompton Hospital have been invaluable for their collective inspiration and guidance over recent years. I am grateful for their support not only as collaborators and as individuals; Elizabeth Renzoni who ignited my passion for ILD as a Respiratory Registrar and supported my initial research endeavours; Athol Wells for his sound advice and discussions as my clinical mentor; and Gisela Lindahl and Carmel Stock who welcomed me into their lab, taught me the basics and supported my later work that formed a large part of the human study.

This thesis would not have been finished without my team at home; my father who patiently did the baby sitting; my brother who listened to my write-up worries; my husband who kept the house running; and last but not least, my baby boy for all the cuddles that kept me going whilst writing up on my maternity leave.

Finally, this thesis is dedicated to the memory of my dear Mum. She saw the early stages of this work and would have been so proud to see me achieve my MD(Res). I wouldn’t be the person that I am today without all the love, support, energy and laughter that she brought into my life.
# TABLE OF CONTENTS

DECLARATION OF ORIGINALITY .................................................................................................................. 1  
COPYRIGHT DECLARATION ............................................................................................................................. 2  
ABSTRACT .......................................................................................................................................................... 3  
ACKNOWLEDGEMENTS ................................................................................................................................. 4  
TABLE OF CONTENTS .................................................................................................................................. 5  
LIST OF FIGURES ........................................................................................................................................... 10  
LIST OF TABLES ............................................................................................................................................ 12  
ABBREVIATIONS ........................................................................................................................................... 13  
CHAPTER 1: INTRODUCTION ............................................................................................................................. 15 
  1.1 Idiopathic Pulmonary Fibrosis (IPF) ........................................................................................................ 15 
  1.1.1 Disease Overview .................................................................................................................................. 15 
  1.1.2 Epidemiology and Risk Factors ........................................................................................................... 15 
  1.1.3 Clinical Presentation and Diagnosis ..................................................................................................... 16 
  1.1.4 Markers of Disease Severity and Progression .................................................................................... 18 
  1.1.5 Biomarkers in IPF .................................................................................................................................. 20 
  1.1.6 Pathogenetic Mechanisms .................................................................................................................. 21 
  1.1.7 Epithelial to Mesenchymal Transition and IPF .................................................................................. 25 
  1.1.8 Genetics Basis of IPF .......................................................................................................................... 28 
  1.1.9 Current Treatment ................................................................................................................................ 30 
  1.2 The Nitric Oxide Signalling Pathway ....................................................................................................... 32 
  1.2.1 Nitric Oxide (NO) and Endogenous Production .................................................................................. 32 
  1.2.2 Asymmetric Dimethylarginine (ADMA) ............................................................................................. 33 
  1.2.3 Dimethylarginine Dimethylaminohydroxylase (DDAH) ................................................................... 34 
  1.2.4 NO-ADMA-DDAH Pathway and Disease States ................................................................................ 36 
  1.3 NO Signalling and Lung Fibrosis ............................................................................................................ 37 
  1.3.1 Role of NO in Wound Repair .......................................................................................................... 37 
  1.3.2 Overproduction of NO in IPF ............................................................................................................. 40 
  1.3.3 NO-ADMA-DDAH Pathway and IPF .................................................................................................. 42 
  1.3.4 NO-ADMA-DDAH Pathway and Epithelial to Mesenchymal Transition .......................................... 45 
  1.4 Human DDAH1 and DDAH2 Genes ....................................................................................................... 45 
  1.4.1 DDAH Genes ....................................................................................................................................... 45 
  1.4.2 Human DDAH Genetic Studies .......................................................................................................... 46 
  1.4.3 Rationale for a Targeted Genetic Association Study ....................................................................... 48
## CHAPTER 2: HYPOTHESIS AND STUDY AIDS ......................................................... 50

2.1 Hypothesis ........................................................................................................... 50
2.2 In Vitro Study ..................................................................................................... 50
2.3 Human Study ..................................................................................................... 50

## CHAPTER 3: MATERIALS AND METHODS ....................................................... 51

3.1 In Vitro Methods ................................................................................................. 51
  3.1.1 Cell Line ........................................................................................................ 51
  3.1.2 Cell Line Culture ........................................................................................... 51
  3.1.3 Cell Line Treatment ...................................................................................... 51
  3.1.4 Protein Harvesting from Cell Culture .......................................................... 51
  3.1.5 RNA Extraction from Cell Culture ............................................................... 52
  3.1.6 DNA Isolation from Human Whole Blood ................................................. 52
  3.1.7 Protein, RNA and DNA Quantification ........................................................ 52
3.2 Molecular Biology Methods .............................................................................. 52
  3.2.1 Reverse Transcription Quantitative PCR (RT-qPCR) .................................. 52
  3.2.2 Quantitative Polymerase Chain Reaction (qPCR) ....................................... 53
  3.2.3 Western Blotting ......................................................................................... 54
  3.2.4 KASP Genotyping Assay ........................................................................... 56
3.3 Biochemical Methods ........................................................................................ 58
  3.3.1 Methylarginine Quantification .................................................................... 58
  3.3.2 Nitric Oxide Quantification ........................................................................ 59
3.4 Human Study Methods ...................................................................................... 59
  3.4.1 Patient Recruitment and Diagnostic Criteria .............................................. 59
  3.4.2 Clinical Assessment ..................................................................................... 60
  3.4.3 Data Collection and Clinical End-Points .................................................... 61
  3.4.4 Statistical Methods ..................................................................................... 62

## CHAPTER 4: ESTABLISHMENT AND CHARACTERISATION OF AN IN VITRO MODEL OF EMT 64

4.1 Development of an Experimental Model of EMT ............................................. 64
  4.1.1 Introduction .................................................................................................. 64
  4.1.2 Protocols ..................................................................................................... 64
  4.1.3 Results ......................................................................................................... 65
  4.1.4 Discussion ................................................................................................... 68
4.2 Characterisation of the NO-ADMA-DDAH Pathway in EMT ............................. 70
  4.2.1 Introduction .................................................................................................. 70
  4.2.2 Protocols ..................................................................................................... 70
LIST OF FIGURES

Figure 1. Mortality trends of IPF in the UK ................................................................. 16
Figure 2. High-resolution CT chest of the IPF lung ..................................................... 17
Figure 3. Surgical lung biopsy showing the UIP pattern ........................................... 17
Figure 4. Pathogenesis of IPF: proposed model of abnormal wound healing .......... 24
Figure 5. Important features and markers of EMT ....................................................... 25
Figure 6. TGF-β signalling via canonical and non-canonical pathways .................. 26
Figure 7. Overview of the NO-ADMA-DDAH pathway ............................................ 34
Figure 8. DDAH activity as a key determinant of NO bioavailability ....................... 35
Figure 9. Biphasic role of NO in regulating expression of iNOS via NF-κB ............... 38
Figure 10. Possible role of NO in normal wound healing following alveolar epithelial injury. 40
Figure 11. Example of a typical genotyping cluster plot .......................................... 58
Figure 12. Mechanism of KASP genotyping ................................................................. 57
Figure 13. Outline of EMT protocol used for in vitro experiments with A549 cell line .. 65
Figure 14. E-cadherin gene and protein expression in TGF-β mediated EMT .......... 65
Figure 15. α-SMA gene and protein expression in TGF-β mediated EMT ............... 66
Figure 16. Type 1 Collagen expression in TGF-β mediated EMT ......................... 66
Figure 17. MMP-1, MMP-2, MMP-7 and LOXL2 gene expression in TGF-β mediated EMT .... 67
Figure 18. TGF-β mediated Smad3 phosphorylation ................................................. 67
Figure 19. DDAH1 expression in TGF-β mediated EMT ........................................... 71
Figure 20. DDAH2 expression in TGF-β mediated EMT .......................................... 71
Figure 21. Methylarginine levels in TGF-β mediated EMT ....................................... 71
Figure 22. NOS gene expression in TGF-β mediated EMT ....................................... 72
Figure 23. iNOS protein expression in TGF-β mediated EMT ................................... 72
Figure 24. NOX levels in TGF-β mediated EMT ....................................................... 72
Figure 25. Effect of ADMA treatment on E-cadherin protein expression in EMT model ...... 77
Figure 26. Effect of ADMA treatment on α-SMA protein expression in EMT model ...... 77
Figure 27. Effect of ADMA treatment DDAH1 protein expression in EMT model .......... 78
Figure 28. Effect of ADMA treatment on DDAH2 protein expression in EMT model .... 78
Figure 29. Effect of ADMA treatment on iNOS protein expression in EMT model .......... 78
Figure 30. Effect of selective iNOS inhibition on E-cadherin protein expression in EMT ...... 83
Figure 31. Effect of selective iNOS inhibition on α-SMA protein expression in EMT ...... 83
Figure 32. Effect of selective iNOS inhibition on DDAH1 protein expression in EMT ...... 84
Figure 33. Effect of selective iNOS inhibition on DDAH2 protein expression in EMT ...... 84
Figure 34. Effect of selective iNOS inhibition on iNOS protein expression in EMT ...... 84
Figure 35. Summary of connections between NO, TGF-β and iNOS in the type 2 AEC ....... 87
Figure 36. Summary of the step-wise approach taken to select DDAH SNPs for analysis ...... 88
Figure 37. DDAH1 gene and SNPs ......................................................................... 89
Figure 38. DDAH2 gene and SNPs ......................................................................... 89
Figure 39. DDAH1 LD Plot ..................................................................................... 90
Figure 40. DDAH2 LD plot ..................................................................................... 91
Figure 41. Distribution of DDAH1 SNP genotypes amongst in IPF Cohort ............... 93
Figure 42. Distribution of DDAH2 genotypes in IPF cohort .................................... 94
Figure 43. DDAH1 SNP rs530006: ADMA concentrations by genotype ................. 96
Figure 44. DDAH1 SNP rs6576765: ADMA concentrations by genotype ................. 96
Figure 45. Baseline clinical data: age at presentation ................................................. 97
Figure 46. Baseline clinical data: gender ................................................................. 98
Figure 47. Baseline clinical data: smoking history .........................................................98
Figure 48. Baseline clinical data: summary of baseline pulmonary function .......................98
Figure 49. Clinical end point: time to first decline in FVC .............................................99
Figure 50. Clinical end point: length of survival from presentation .....................................99
Figure 51. Clinical end point: time to first event ................................................................99
Figure 52. rs530006 genotype distribution ......................................................................100
Figure 53. Baseline PFTs by rs530006 genotype ...............................................................100
Figure 54. Time to decline of FVC by carriage of rs530006 genotype ..............................101
Figure 55. Time to decline of FVC by carriage of rs530006 minor homozygote .................101
Figure 56. Length of survival by rs530006 genotype ......................................................103
Figure 57. Length of survival by carriage of rs530006 minor homozygote .......................103
Figure 58. Time to first event by rs530006 genotype .....................................................105
Figure 59. Time to first event by carriage of rs530005 minor homozygote .......................105
Figure 60. rs6576765 genotype distribution ..................................................................107
Figure 61. Baseline PFTs by rs6576765 genotype ............................................................107
Figure 62. Time to decline of FVC by carriage of rs6576765 genotype ............................108
Figure 63. Time to decline of FVC by carriage of rs6576765 major homozygote ..............108
Figure 64. Length of survival by rs6576765 genotype ....................................................110
Figure 65. Length of survival by carriage of rs6576765 major homozygote ......................110
Figure 66. Time to first event by rs6576765 genotype ...................................................112
Figure 67. Time to first event by carriage of rs6576765 major homozygote ......................112
Figure 68. rs805804 genotype distribution ....................................................................114
Figure 69. Baseline PFTs by rs804305 genotype ..............................................................114
Figure 70. Time to decline of FVC by carriage of rs805304 genotype .............................115
Figure 71. Length of survival by rs805304 genotype .....................................................115
Figure 72. rs805305 genotype distribution ....................................................................116
Figure 73. Baseline PFTs by rs805305 genotype ..............................................................116
Figure 74. Time to decline of FVC by carriage of rs805305 genotype .............................117
Figure 75. Length of survival by carriage of rs805305 genotype .....................................117
Figure 76. Baseline clinical data (SSc cohort): Age ..........................................................122
Figure 77. Baseline clinical data (SSc cohort): Gender ......................................................122
Figure 78. Baseline clinical data (SSc cohort): Smoking history ........................................122
Figure 79. Baseline clinical data (SSc cohort): Pulmonary function tests ..........................123
Figure 80. Clinical end point: time to decline of FVC (SSc Cohort) .................................123
Figure 81. rs530006 genotype distribution (SSc Cohort) .................................................124
Figure 82. Baseline PFTs by rs530006 genotype (SSc Cohort) ........................................124
Figure 83. TtD of FVC by carriage of rs530006 genotype (SSc Cohort) .........................125
Figure 84. TtD of FVC by carriage of rs530006 minor homozygote (SSc Cohort) ............125
Figure 85. rs6576765 genotype distribution (SSc cohort) ...............................................126
Figure 86. Baseline PFTs by rs6576765 genotype (SSc Cohort) .......................................126
Figure 87. TtD of FVC by carriage of rs6576765 genotype (SSc Cohort) ........................127
Figure 88. TtD of FVC by carriage of rs6576765 major homozygote (SSc Cohort) ..........127
Figure 89. rs805304 genotype distribution (SSc Cohort) ...............................................128
Figure 90. Baseline PFTs by rs805304 genotype (SSc Cohort) ........................................128
Figure 91. Time to decline of FVC by carriage of rs805304 genotype (SSc Cohort) ..........129
Figure 92. rs805305 genotype distribution (SSc Cohort) ...............................................129
Figure 93. Baseline PFTs by rs805305 genotype (SSc Cohort) ........................................130
Figure 94. Time to decline of FVC by carriage of rs805305 genotype (SSc Cohort) ..........130
LIST OF TABLES

Table 1. Diagnostic criteria for IPF ................................................................. 18
Table 2. Summary of common genetic variants linked to IPF .......................... 30
Table 3. Summary of negative phase II and phase III randomised controlled trials in IPF ...... 31
Table 4. Association of ADMA levels in examples of human diseases ......................... 37
Table 5. DDAH1 SNPs: associations with ADMA levels in normal cohorts and disease states 47
Table 6. DDAH2 SNPs: associations with ADMA levels in normal cohorts and disease states 47
Table 7. Reverse-transcription PCR protocol .................................................... 53
Table 8. Quantitative PCR protocol .................................................................. 54
Table 9. Western blotting protocol .................................................................... 55
Table 10. Primary antibodies and concentrations used .......................................... 55
Table 11. tSNPs genotyped in DDAH1 and DDAH2 genes ...................................... 92
Table 12. DDAH1 SNP genotype and allele frequency in IPF cohort ...................... 93
Table 13. DDAH2 SNP genotype and allele frequency in IPF ............................... 94
Table 14. rs530006 TtD of FVC: multivariate Cox proportional hazards analysis ...... 102
Table 15. rs530006 length of survival: multivariate Cox proportional hazards analysis .... 104
Table 16. rs530006 time to first event: multivariate Cox proportional hazards analysis ...... 106
Table 17. rs6576765 TtD of FVC: multivariate Cox proportional hazards analysis ........ 109
Table 18. rs6576765 length of survival: multivariate Cox proportional hazard analysis ...... 111
Table 19. rs6576765 time to first event: multivariate Cox proportional hazard analysis ..... 113
## ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-SMA</td>
<td>alpha-smooth muscle actin</td>
</tr>
<tr>
<td>ADMA</td>
<td>Asymmetric dimethylarginine</td>
</tr>
<tr>
<td>AEC</td>
<td>Alveolar epithelial cells</td>
</tr>
<tr>
<td>AGXT2</td>
<td>Alanine–glyoxylate aminotransferase</td>
</tr>
<tr>
<td>ANOVA</td>
<td>One-way analysis of variance</td>
</tr>
<tr>
<td>BALF</td>
<td>Bronchoalveolar lavage fluid</td>
</tr>
<tr>
<td>CEU</td>
<td>Utah residents with Northern and Western European ancestry in the International HapMap Project</td>
</tr>
<tr>
<td>CPI</td>
<td>Composite physiologic index</td>
</tr>
<tr>
<td>DDAH</td>
<td>Dimethylarginine Dimethylaminohydroxylase</td>
</tr>
<tr>
<td>EARSII</td>
<td>European Atherosclerosis Relatives Study II</td>
</tr>
<tr>
<td>EBC</td>
<td>Exhaled breath condensate</td>
</tr>
<tr>
<td>ECM</td>
<td>Extra-cellular matrix</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial to mesenchymal transition</td>
</tr>
<tr>
<td>eNOS</td>
<td>Endothelial NOS</td>
</tr>
<tr>
<td>FEV1</td>
<td>Forced expiratory volume in 1 second</td>
</tr>
<tr>
<td>FIP</td>
<td>Familial interstitial pneumonia</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence resonant energy transfer</td>
</tr>
<tr>
<td>FVC</td>
<td>Forced vital capacity</td>
</tr>
<tr>
<td>GER</td>
<td>Gastroesophageal reflux</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome wide association studies</td>
</tr>
<tr>
<td>HRCT</td>
<td>High-resolution computed tomography</td>
</tr>
<tr>
<td>HWE</td>
<td>Hardy–Weinberg equilibrium</td>
</tr>
<tr>
<td>IIPs</td>
<td>Idiopathic interstitial pneumonias</td>
</tr>
<tr>
<td>ILD</td>
<td>Interstitial lung disease</td>
</tr>
<tr>
<td>iNOS</td>
<td>Inducible NOS</td>
</tr>
<tr>
<td>IPF</td>
<td>Idiopathic pulmonary fibrosis</td>
</tr>
<tr>
<td>KASP</td>
<td>Kompetitive Allele Specific PCR</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage disequilibrium</td>
</tr>
<tr>
<td>L-NMMA</td>
<td>N-monomethyl-L-arginine</td>
</tr>
<tr>
<td>LOXL2</td>
<td>Lysyl oxidase homolog 2</td>
</tr>
<tr>
<td>MA</td>
<td>Methylarginine</td>
</tr>
<tr>
<td>MAF</td>
<td>Minor allele frequency</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix metalloproteinase</td>
</tr>
<tr>
<td>MUC5B</td>
<td>Mucin 5B</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>NF-κB</td>
<td>Nuclear factor kappa-light-chain-enhancer of activated B cells</td>
</tr>
<tr>
<td>nNOS</td>
<td>Neuronal NOS</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric oxide</td>
</tr>
<tr>
<td>NOS</td>
<td>Nitric oxide synthase</td>
</tr>
<tr>
<td>NOx</td>
<td>Nitrite (NO(_2)(^\cdot)) and nitrate (NO(_3)(^\cdot)) combined</td>
</tr>
<tr>
<td>NSIP</td>
<td>Non-specific interstitial pneumonia</td>
</tr>
<tr>
<td>O(_2)(^\cdot)</td>
<td>Superoxide</td>
</tr>
<tr>
<td>ONOO(_{\cdot})</td>
<td>Peroxynitrite</td>
</tr>
<tr>
<td>PF</td>
<td>Pulmonary fibrosis</td>
</tr>
<tr>
<td>PDGF</td>
<td>Platelet-derived growth factor</td>
</tr>
<tr>
<td>PFTs</td>
<td>Pulmonary function tests</td>
</tr>
<tr>
<td>PRMT</td>
<td>Protein arginine methyltransferases</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative Polymerase Chain Reaction</td>
</tr>
<tr>
<td>RBH</td>
<td>Royal Brompton Hospital</td>
</tr>
<tr>
<td>RNS</td>
<td>Reactive nitrogen species</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse Transcription Quantitative PCR</td>
</tr>
<tr>
<td>SDMA</td>
<td>Symmetric dimethylarginine</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSc</td>
<td>Systemic sclerosis</td>
</tr>
<tr>
<td>SSc-ILD</td>
<td>Systemic sclerosis associated interstitial lung disease</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transforming growth factor β</td>
</tr>
<tr>
<td>TLCO</td>
<td>Transfer capacity of the lung for carbon monoxide</td>
</tr>
<tr>
<td>UIP</td>
<td>Usual interstitial pneumonia</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 Idiopathic Pulmonary Fibrosis (IPF)

1.1.1 Disease Overview

Idiopathic pulmonary fibrosis (IPF) is a devastating, fibrotic lung disease that is the most progressive and the most prevalent of the idiopathic interstitial pneumonias (IIPs). IPF is defined as a specific form of chronic, progressive fibrosing interstitial pneumonia of unknown cause, occurring primarily in older adults, and limited to the lungs, and associated with the radiological and/or histopathologic pattern of usual interstitial pneumonia (UIP) [Raghu 2011]. The condition is typically characterised by progressive fibrosis with architectural destruction of the lung resulting in an irreversible loss of pulmonary function.

Despite advances in our understanding of the pathogenetic mechanisms of this disease and the recent development of IPF-specific therapies that may slow disease progression, IPF is almost invariably a fatal disorder with death after a median of 3 years following diagnosis [Ley 2011]. There remains a large unmet clinical and therapeutic need for patients with IPF.

1.1.2 Epidemiology and Risk Factors

The incidence of IPF is rising globally and is estimated at 3–9 cases per 100,000 per year in Europe and North America, [Hutchison 2015]. From a UK perspective, approximately 15,000 people have a diagnosis of IPF with an estimated annual incidence of 7.44 per 100,000 people. Approximately 5000 new cases are diagnosed each year and 5000 patients with IPF will die each year [Navaratnam 2011]. The burden of mortality from IPF is higher than for many cancers, and in the UK, more people will die each year from IPF than from ovarian cancer, lymphoma, leukaemia, mesothelioma or kidney cancer. Whether the observed rise in incidence represents a true increase or is due to earlier diagnosis, the median survival from diagnosis remains approximately 3 years, with a 5-year survival of 20-30% [Ley 2011].

The incidence increases with older age and presentation is typically in the sixth and seventh decades of life; patients aged less than 50 years are rare [Raghu 2006]. There is a higher predominance of the disease in men than in women (1.5 to 1.7:1) and the majority of patients have a history of cigarette smoking. Several factors have been
associated with an increased risk of developing the disease, namely cigarette smoking, environmental/occupational pollutants, microbial agents, chronic micro-aspiration secondary to gastroesophageal reflux, and genetic abnormalities [Raghu 2011]. The epidemiology indicates that the development of a familial form of IPF, familial interstitial pneumonia (FIP), is associated with older age, male sex and ever having smoked cigarettes [Steele 2005], indicating an interplay between environmental factors, and a genetic basis to IPF pathogenesis that is explored in later sections.

![Mortality trends of IPF in the UK](image)

**Figure 1. Mortality trends of IPF in the UK**
Estimated number of deaths from idiopathic pulmonary fibrosis clinical syndrome, age standardised to the 2008 population of England and Wales. ICD, International Classification of Diseases [Navaratnum 2011]

### 1.1.3 Clinical Presentation and Diagnosis
Progressive loss of the gas exchange surface area of the lung leads to worsening breathlessness and cough. A diagnosis of IPF should be considered in all adult patients with unexplained chronic exertional dyspnoea. The majority of IPF deaths occur directly from the progression of lung fibrosis and the development of respiratory failure rather than secondary to comorbid conditions [Ley 2011]. In view of its poor prognostic implications and the recent availability of novel therapeutic options, an early and accurate diagnosis is crucial in IPF.

IPF is characterised by a radiological and/or histopathological pattern of UIP which is central to the diagnosis. High-resolution computed tomography of the chest (HRCT chest) is an essential component of the diagnostic pathway and UIP is characterised by a predominantly peripheral, basal and subpleural distribution of reticular
abnormalities and clustered cystic airspaces known as ‘honeycombing’ with associated traction bronchiectasis and volume loss [Jacob 2015].

Figure 2. High-resolution CT chest of the IPF lung
Characteristic combination of peripheral, subpleural, and predominantly bibasilar reticular abnormalities with associated honeycomb change and traction bronchiectasis [Adapted from Spagnolo 2015(b)]

The histopathological hallmark of a UIP pattern on lung biopsy is a heterogeneous appearance at low magnification in which areas of fibrosis with scarring and honeycomb change alternate with areas of less affected or normal parenchyma [ATS/ERS Statement 2000]. Inflammation is usually minimal and fibrotic zones are composed mainly of dense collagen and areas of proliferating fibroblasts known as ‘fibroblastic foci’.

Figure 3. Surgical lung biopsy showing the UIP pattern
Characteristic abrupt juxtaposition of scarred lung with honeycombing (*) and nearly normal lung (+). Several pale fibroblastic foci are also seen (arrows). Hematoxylin– eosin stain, 20x. Usual interstitial pneumonia (UIP) [Adapted from Spagnolo 2015(b)]
IPF currently represents a diagnostic challenge and as set out in the official American Thoracic Society/European Respiratory Society/Japanese Respiratory Society/Latin America Thoracic Association (ATS/ERS/JRS/ALAT) 2011 guidelines [Raghu 2011], a multidisciplinary approach that integrates clinical, radiological and histopathological findings is the mainstay of clinical practice.

The diagnosis of IPF requires the following:

1. Exclusion of other known causes of ILD (e.g., domestic and occupational environmental exposures, connective tissue disease, and drug toxicity)

2. Presence of UIP pattern on HRCT in patients not subjected to surgical lung biopsy

3. Specific combinations of HRCT and surgical lung biopsy pattern in patients subjected to surgical lung biopsy

Table 1. Diagnostic criteria for IPF
Adapted from the ATS/ERS/JRS/ALAT 2011 guidelines [Rahu 2011]

There is no single test or biomarker that alone can diagnose IPF. Performing a surgical lung biopsy, recommended when the diagnosis cannot be made non-invasively, carries a mortality rate of 3–6% and is often inappropriate given the average age of onset, severity of disease at presentation and comorbidities of the individual patient [Wells 2016]. Currently, in roughly half of IPF patients, the ATS/ERS/JRS/ALAT 2011 diagnostic criteria for IPF are not satisfied, leaving clinicians to make the diagnosis on balance of probabilities. This is of particular relevance in the context of the selection of patients for novel IPF-specific anti-fibrotic therapies and for the cohort enrichment of clinical studies to ensure the inclusion of patients with a true diagnosis of IPF.

1.1.4 Markers of Disease Severity and Progression
The natural history of IPF is heterogeneous and unpredictable at the time of the diagnosis and phenotypes with a distinct clinical course have emerged [King 2011]. The majority of patients demonstrate a slow, gradual progression over many years. Some patients remain stable while others have an accelerated decline. Some patients may experience episodes of acute respiratory worsening known as ‘acute exacerbations’ which despite the misnomer, invariably are pre-terminal events. Worse prognosis is associated with older age (>70 years of age), smoking history, low body-mass index, severe physiological impairment, large radiological extent of disease, co-existing emphysema and pulmonary hypertension [Ley 2011].
The mainstay of clinical monitoring is the measurement of baseline and longitudinal pulmonary function tests (PFTs). Examples of the commonly used variables in clinical practice, each representing a different aspect of pulmonary function, are measures of lung volume (forced vital capacity, FVC), flow (forced expiratory volume in 1 second, FEV1) and gas exchange (transfer capacity of the lung for carbon monoxide, TLCO). Respiratory disorders typically affect these variables in predictable ways and spirometric measures of lung volume are of particular relevance to IPF, a disease characterised by a loss of lung volume and a resulting restrictive ventilatory defect.

Of the candidate variables, trends in FVC have consistently predicted mortality in IPF and FVC decline is generally viewed as the best marker of chronic disease progression [Wells 2013]. Although not a proven surrogate for mortality, a consensus exists that mortality is not a practicable primary end point for the demonstration of beneficial treatment effects and FVC trends are now also the preferred primary end point in IPF clinical trials. In addition, the time taken for the FVC to decline can be amalgamated with mortality in the evaluation of ‘progression-free survival’ as a composite end-point.

From the ATS/ERS/JRS/ALAT 2011 IPF guidelines, it was concluded that a relative change in FVC of 10% from baseline values constituted evidence of disease progression (in the absence of an alternative explanation such as a lower respiratory tract infection) [Raghu 2011]. A loss of ≥10% in FVC is a clinically meaningful event for an IPF patient as it translates into an increased rate of death of up to eight-fold in the following year [du Bois 2011]. An absolute decline of 10% represents a relatively
minor fall in mild disease but proportionally a much larger fall in severe disease [Wells 2013]. Measuring a relative decline of 10% in FVC deals with the confounding effect of measurement variation and the minimal clinically important cohort difference in FVC lies between 3% and 6% [du Bois 2011].

In a subset of IPF patients, the quantification of disease severity at baseline using pulmonary function tests alone may be confounded by factors such as emphysema and pulmonary hypertension and composite scoring systems have been developed to utilise both physiological and radiographic variables in an attempt to provide more accurate prognostic information. The composite physiologic index (CPI) is a measure that was derived by fitting pulmonary function tests (FEV1, FVC and TLCO) against disease extent on HRCT to predict the extent of disease on HRCT [Wells 2003]. Inclusion of the FEV1 and TLCO allowed for correction of the confounding effects of emphysema, and furthermore, the CPI was stronger predictor of mortality than individual measures of lung function alone. Despite its advantages, this composite approach has not been externally validated in any prospective clinical trials and its application in wider routine clinical practice is unknown and FVC remains the most widely used marker of disease progression in IPF.

1.1.5 Biomarkers in IPF

An interest in biomarkers has developed in IPF for a range of purposes including early diagnosis, disease monitoring, predicting future disease progression, determining response to therapy, and to serve as a surrogate endpoints for future clinical studies. Suitable biomarkers may also allow “cohort enrichment” and more accurate patient stratification in clinical studies and may thus be able to identify patient populations that respond better to therapeutic intervention than others. A number of potential serum biomarkers have been identified and the best described of these are matrix metalloproteinase (MMP)-7, chemokine (C-C motif) ligand (CCL)-18, Krebs von den Lungen (KL)-6 and more recently serum lysyl oxidase-like (LOXL)-2 [Maher 2014]. However, in each case these markers have typically been identified in pre-existing sample collections from single centres with blood drawn at a single time point, and are yet to prove to be effective for identifying individuals with progressive disease and may not map biological pathways for which potential IPF therapies exist. Although a more detailed review of biomarker discovery within IPF is outside the scope of this thesis, it is important to note that a biomarker to identify individuals with IPF at increased risk of disease progression remains elusive.
1.1.6 Pathogenetic Mechanisms

Understanding of the pathobiology of IPF has evolved greatly over the past two decades. Previously, it was widely held that a common pathogenetic sequence underpinned all fibrotic lung diseases, regardless of the aetiology; the “inflammatory fibrosis” hypothesis asserted that chronic inflammation resulted in lung injury and fibrogenesis which led to the end-stage fibrotic scar [Keogh 1982]. IPF was thought of as a chronic inflammatory condition predominantly based on parallels drawn with other interstitial lung disorders, for example, in systemic sclerosis, sarcoidosis and hypersensitivity pneumonitis, there is an identifiable inflammatory response, clinical and histological features of inflammatory change and a treatment response to immunosuppression [Selman 2002]. In IPF however, inflammation is not a prominent histopathological finding, clinical measurements of inflammation fail to correlate with disease stage, outcome does not improve despite potent anti-inflammatory or immunosuppressive therapy and experimental models have demonstrated that inflammation is not always required for the development of a fibrotic response [Selman 2001].

Although it is possible that inflammation accompanied the initial injury in the early phases of the disease, IPF patients respond poorly to anti-inflammatory therapies because there is little or no inflammation in advanced UIP [Thannickal 2004]. Consequently, use of the previous mainstay of IPF treatment, combined azathioprine and corticosteroid therapy, was “strongly recommended against” in the latest ATS/ERS/JRS/ALAT guidelines [Raghu 2015] and recent clinical trials have shifted their focus from anti-inflammatory and immunosuppressant compounds to molecules targeting fibrogenesis.

There is a growing body of evidence that points towards IPF as being an epithelial/fibroblastic disorder [Selman 2001]. The prevailing hypothesis is that IPF arises from an aberrant wound healing response occurring after repetitive alveolar cell injury in a genetically susceptible lung epithelium. Normal wound repair has distinct stages following epithelial damage; a coagulation phase, an inflammatory phase, a fibroblast migration/proliferation phase, and a remodeling phase where normal tissue architecture is restored [Wynn 2011]. In IPF, this process appears to be dysregulated, and transforming growth factor β (TGF-β) is considered to be the “master switch” in the fibrotic process [Willis 2007]. The absence of significant inflammation in the chronic disease process may be due to tissue responses that have successfully “switched off” classic inflammatory responses, but persist in a state of chronic repair [Thannickal
Current concepts suggest that alveolar epithelial cells (AEC) initiate the pathological process and drive the phenotypic changes seen among fibroblasts during the progression to end-stage fibrosis as follows:

1. Repetitive alveolar epithelial cell micro-injury and aberrant activation

Multiple microinjuries lead to marked epithelial injury and apoptosis of type 1 AECs and hyperplasia of type 2 AECs. This epithelial disruption induces the coagulation cascade and formation of a fibrin-rich provisional matrix (wound clot). In IPF however, there is an increased antifibrinolytic and hypercoagulable microenvironment; aberrantly activated AECs secrete tissue factor (TF) and plasminogen activator inhibitor 1 and 2 (PAI-1 and 2) which impede migration to repair the denuded wound surface and disrupt the normal degradation process of the provisional matrix [Selman 2002, Fernandez 2012(a)].

2. Profibrotic cytokine release by alveolar epithelial cells

The abnormally activated, hyperplastic type 2 AECs are the primary source for a diverse range of cytokines, growth factors and matrix metalloproteinases that induce a profibrotic environment. In IPF, hyperplastic type 2 AECs constitute the main site of the synthesis and activation of TGF-β1, the TGF-β isoform most closely related with the development of IPF which is elevated in the lungs of patients with IPF [Khalil 1996, Wolters 2014]. TGF-β1 is a pleiotrophic growth factor with chemotactic and proliferative properties that plays a pivotal role in driving the fibrotic process by inducing fibroblast-to-myofibroblast differentiation, epithelial to mesenchymal transition, and is the most potent inducer of extra-cellular matrix (ECM) production characterised to date [Scotton 2007, Fernandez 2012(b)].

3. Fibroblast activation and differentiation into myofibroblasts

AEC release of TGF-β1 promotes the migration, proliferation, and differentiation of fibroblasts into activated, invasive myofibroblasts, the key effector cell in fibrogenesis [Selman 2002]. These spindle-shaped cells are contractile, contain α-smooth muscle actin (SMA) stress fibers and localise to sites of active fibrosis, and are the primary cell type responsible for the synthesis and deposition of ECM [Fernandez 2012(a)]. It is proposed that the three main sources of activated myofibroblasts are derived from i) the recruitment and proliferation of resident lung fibroblasts, ii) the attraction of circulating fibrocytes (bone marrow-derived mesenchymal cells), and iii) from AECs via stimulation of epithelial to mesenchymal transition [Scotton 2007, King 2011].
4. Formation of fibroblastic foci and extracellular matrix deposition

The actively proliferating fibroblasts and myofibroblasts organise into distinct clusters named ‘fibroblastic foci,’ the pathologic hallmark of IPF. Representing microscopic zones of acute lung injury, these are characterised by highly active myofibroblasts in subepithelial regions, mostly in close proximity to injured or hyperplastic type 2 AECs or in the alveolar spaces [Selman 2002]. In an attempt to repair the damaged alveolus, the activated myofibroblasts secrete TGF-β1 which i) increases the transcriptional activation of collagen genes, via Smad signalling pathways, and collagen deposition and ii) induces an imbalance between matrix metalloproteinases and their inhibitors, thereby creating a microenvironment that favours exaggerated extracellular matrix accumulation [Fernandez 2012(b), King 2011].

5. Abnormal lung remodelling and destruction of the lung parenchyma

In addition to the central, synthetic role that the fibroblastic foci play, disruption of the basement membrane by matrix metalloproteinases allows the migration of fibroblasts and myofibroblasts deep into the alveolar spaces where proliferation and ECM deposition continues [King 2011]. Furthermore, signals responsible for myofibroblast apoptosis seem to be absent or delayed in IPF. An additional action of TGF-β1 release by the myofibroblasts is to promote type 2 AEC apoptosis, thereby preventing subsequent alveolar epithelial cell re-epithelialisation [Selman 2001]. Loss of basement membrane integrity, abnormalities of myofibroblast apoptosis, increased AEC apoptosis and the failure of AEC re-epithelialisation perpetuate progressive, abnormal lung remodeling and the subsequent architectural destruction of the lung parenchyma.
Figure 4: Pathogenesis of IPF: proposed model of abnormal wound healing

Top left: On-going micro-injury damages and activates alveolar epithelial cells which in turn induce an antifibrinolytic environment in the alveolar spaces, enhancing wound clot formation.

Bottom left: Abnormally activated alveolar epithelial cells release various cytokines and growth factors, including TGF-β1, and promote fibroblast migration, proliferation, and differentiation into invasive myofibroblasts. Subepithelial myofibroblasts and alveolar epithelial cells produce gelatinases that may increase basement membrane disruption and allow fibroblast–myofibroblast migration. Further recruitment of fibroblasts results from i) proliferation of resident mesenchymal cells, ii) attraction of circulating fibrocytes, and iii) stimulation of epithelial to mesenchymal transition.


Top right: An imbalance between interstitial collagenases and tissue inhibitors of metalloproteinases provokes the progressive deposit of extracellular matrix. Exaggerated extracellular matrix deposition results in the architectural destruction of the lung parenchyma. Signals responsible for myofibroblast apoptosis are absent or delayed, increasing cell survival. Myofibroblasts produce angiotensinogen that as angiotensin II provokes alveolar epithelial cell death, further impairing reepithelialisation.

Abbreviations: FGF-2 (fibroblast growth factor-2); MMP (metalloproteinase); PAI-1, PAI-2 (plasminogen activator inhibitor-1, -2); PDGF (platelet-derived growth factor); TGF-β (transforming growth factor-β); TIMP (tissue inhibitors of metalloproteinases); TNF-α (tumor necrosis factor-α); VEGF (vascular endothelial growth factor).

[Adapted from Selman 2001].
1.1.7 Epithelial to Mesenchymal Transition and IPF

Epithelial-mesenchymal transition (EMT) is a process in which fully differentiated epithelial cells undergo transdifferentiation to a mesenchymal phenotype giving rise to fibroblasts and myofibroblasts [Willis 2007]. Normal tissue homeostasis does not require EMT, rather, EMT is activated during conditions of remodelling and plays an integral role in three biological contexts: embryonic development, cancer progression and repair of injured tissue [Wolters 2014]. During EMT, epithelial cells undergo profound changes which can be broadly categorised into: loss of cell adhesion and apical-basal polarity; changes in transcriptional regulation and changes in cytoskeleton, which increase the motility of cells and enables the development of an invasive phenotype [Lamouille 2014].

Figure 5. Important features and markers of EMT
[Adapted from Bartis 2014]

EMT is thought to be a gradual process, meaning that there are intermediate cellular phenotypes during the transition. An epithelial cell undergoing EMT will express a set of different marker genes at various times during this process and it is challenging to identify with certainty epithelial cells undergoing EMT in complex pathological conditions [Bartis 2014]. EMT is characterised genetically by decreased expression of epithelial cell-associated genes, in particular, the gradual loss of E-cadherin is a universal feature of EMT, regardless of the initiating stimulus [Willis 2007]. However, the lack of marker specificity to define the mesenchymal phenotype makes it necessary to simultaneously evaluate a panel of mesenchymal markers in the context of EMT. A
frequently assessed marker in EMT is alpha-smooth muscle actin (α-SMA). Expression of α-SMA is characteristic for myofibroblasts which are laying down an excessive amount of ECM and thus fibrosis [Bartis 2014]. Other markers used to define the mesenchymal phenotype include N-cadherin, vimentin, fibroblast-specific protein-1 (also called S100A4), and type 1 collagen [Lee 2006, Bartis 2014].

![Diagram of TGF-β signalling pathways](image)

**Figure 6. TGF-β signalling via canonical and non-canonical pathways**

*Abbreviations:* Type II TGF-β receptor (RII), Type I TGF-β receptor (RI), phosphorylation (P), mitogen-activated protein kinases (MAPKs), phosphatidylinositol 3-kinase (PI3K), TNF receptor-associated factor 4/6 (TRAF4/6) and nuclear factor-κB (NF-κB). [Adapted from Costanza 2017]

Multiple pathways have been demonstrated to trigger EMT, however, TGF-β1 is considered to be the prototypical cytokine for the induction of EMT in fibrosis and mediates its effects predominantly via Smad-dependent (Smad2/3) pathways [Kasai 2005, Kolosova 2011]. TGF-β1 signals via two heterodimeric transmembrane receptors, the type II and type I (also known as activin receptor-like kinase-5, ALK5) receptors. TGF-β1 is secreted into the extracellular environment in its latent form and following cleavage and release from the latency-associated peptide/latent TGF-β1-binding protein complex, the active form of TGF-β1 binds to the type II receptor which then associates with the type 1 receptor in a heterodimer. The Smad anchor for receptor activation (SARA) site on the heterodimer modulates the formation of
complexes with Smad2 or Smad3. The type I receptor then phosphorylates Smad2 and Smad3 at serine residues which induces their association with Smad4 (co-Smad). The trimeric Smad complex then translocates into the nucleus and interacts with other transcription factors to regulate the transcription of TGF-β1 responsive genes including α-SMA and type 1 collagen by interacting with Smad-binding elements [Lamouille 2014, Fernandez 2012(b)].

The profibrotic effects of TGF-β1 signalling have been largely attributed to Smad3 signalling [Bonniaud 2004]. The involvement of the TGF-β1/Smad3 pathway in fibrotic EMT has been demonstrated in vivo in animal models of the eye and the kidney; in Smad3 null mice, EMT of lens epithelial cells was completely ameliorated [Saika 2004] and similarly in the kidney, Smad3 null mice were protected from tubulointerstitial fibrosis [Sato 2003].

The lung epithelium is the key initiating component in the pathogenesis of IPF and there is accumulating in vitro and in vivo evidence that EMT contributes causally to IPF. It was first demonstrated that exposure of type 2 AECs to TGF-β1 induced EMT in vitro in a rat alveolar epithelial cell line (RLE-6TN) [Willis 2005]. A direct role for TGF-β mediated EMT in the lung in vivo was demonstrated by overexpression of active TGF-β1 using an adenoviral vector in triple transgenic mice in which type 2 AECs permanently express β-galactosidase [Kim 2006]. In this model, X-gal-positive cells (an indicator of whether a cell expresses a functional β-galactosidase) that expressed the myofibroblast marker α-SMA were identified in injured lungs demonstrating EMT in situ. One-third of cells identified as fibroblasts by the expression of vimentin were X-gal-positive, indicating an epithelial origin, and X-gal-positive cells accounted for most of the increase in vimentin-positive cells 21 days after instillation of TGF-β1 [Kim 2006].

Importantly, both of these research groups identified cells co-expressing epithelial markers and mesenchymal markers (specifically α-SMA) in AECs from IPF lungs. This finding was also recently confirmed in a study which additionally demonstrated elevated mRNA levels for type I collagen, α-SMA, and calponin 1, as well as the transcription factor SLUG which has been shown to promote EMT, in AECs isolated from IPF lungs, [Marmai 2011]. Taken together, this evidence suggests that a process of EMT is occurring in IPF.

The extent to which EMT contributes to IPF is under investigation and one ongoing controversy is whether epithelial cells acquire sufficient mesenchymal characteristics
that they can be classified as fibroblasts [Kage 2012]. In humans, the only data supporting this is the finding that fibroblasts isolated from IPF patients express the epithelial cell protein keratin 18 [Larsson 2008]. However, in murine models of bleomycin-induced pulmonary fibrosis, several lineage-tracing studies have shown that the transition of type 2 AECs accounts for 30–50% of cells expressing the mesenchymal proteins vimentin or fibroblast-specific protein-1, supporting the proposition that EMT is a potential source of myofibroblasts during fibrogenesis [Kim 2006, Tanjore 2009, Degryse 2010]. In contrast, one lineage-tracing study suggested no evidence for EMT, although the authors speculated that differences in the experimental setup could explain this [Rock 2011]. Resolution of this issue is complicated by the lack of definitive, cell specific markers for fibroblasts. Although the relative contribution of EMT to fibrosis in IPF remains to be determined, understanding the molecular interactions that lead to EMT will hopefully lead to the identification of novel therapeutic targets for fibrosis in the lung and elsewhere.

1.1.8 Genetics Basis of IPF

Evidence that IPF may have a genetic basis initially emerged from the clustering of pulmonary fibrosis within families. Termed as familial interstitial pneumonia (FIP), this rare subset accounts for less than 5% of all patients with IPF and is clinically and histologically indistinguishable from sporadic IPF [Raghu 2011]. Most affected families have an autosomal dominant vertical transmission pattern of inheritance with reduced penetrance. Rare genetic variants in eight genes are linked to the development of FIP (and are found in some cases of sporadic IPF), which broadly fall into two categories: three genes related to surfactant protein processing and trafficking (surfactant protein C, SFTPC; surfactant protein A2, SFTPA2; and ATP-binding cassette member A3, ABCA3) and five genes linked to telomere function (telomerase reverse transcriptase, TERT; human telomerase RNA component, hTR; dyskerin, DKC1; telomere interacting factor 2, TINF2; and regulator of telomere elongation helicase, RTEL1) [Kropski 2015].

In 2011, a genome-wide linkage study identified a locus on chromosome 11 that was significantly associated with the risk of developing IPF and subsequent resequencing of this region identified a common single nucleotide polymorphism (SNP) (rs35705950) in the promoter of the gene encoding mucin 5B (MUC5B) that was associated with a six- to eight-fold increased risk for IPF [Siebold 2011]. This association has since been validated in several independent cohorts [Noth 2013, Zhang 2011(a), Fingerlin 2013] and is strongly associated with both familial and sporadic IPF [Siebold 2011] but does not confer increased risk of systemic sclerosis associated ILD or sarcoidosis.
[Peljto 2012, Stock 2013]. Interestingly, the MUC5B promoter variant (the minor allele) that confers the greatest risk of IPF is associated with higher expression of MUC5B in the human lung and also associated with improved survival in IPF patients compared to non-carriers [Peljto 2013]. Previous animal studies have suggested that MUC5B regulates airway host defence but the precise mechanisms by which MUC5B influences fibrotic remodelling are uncertain at present [Kropski 2015].

In 2013, two large independent genome-wide association studies of IPF patients were conducted and together identified numerous genetic loci that confer IPF risk [Fingerlin 2013, Noth 2013]. The first evaluated 1616 IPF cases and 4683 controls; the previously reported association with MUC5B was confirmed and nine additional loci were significantly associated with IPF, including SNPs near TERT and hTR [Fingerlin 2013]. The second study comprised 1410 IPF cases and 2934 controls; five loci achieved genome-wide significance, including four SNPs on chromosome 11p15 (MUC5B and three SNPs within the Toll-interacting protein TOLLIP locus) and one on 17q21 [Noth 2013]. Linkage disequilibrium was reported to be low with MUC5B rs35705950 and suggested that TOLLIP may be an independent risk locus - IPF cases with the TOLLIP risk allele (the major allele) had decreased mortality compared to minor allele carriers [Kropski 2015].

Although these genetic variants have been identified as enhancing the risk of developing IPF, none have proven causal. Furthermore, the biological mechanisms through which these genetic variants contribute to disease pathogenesis are not fully understood and include alterations in host defence, DNA repair, cell senescence and epithelial barrier function in the lung [Mathai 2015]. The complex relationship between inherited genetic variants, how they interact with each other and with environmental factors, which leads to the phenotype of IPF needs further investigation. More importantly perhaps, the genetic variants associated with IPF may be useful in identifying prognostically distinct disease subtypes and allow earlier therapeutic intervention.
### Table 2. Summary of common genetic variants linked to IPF

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>SNP</th>
<th>IPF risk</th>
<th>IPF survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q14</td>
<td>IL1RN</td>
<td>rs608392</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs619598</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs2437988</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>3q26</td>
<td>hTR</td>
<td>rs6793225</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>4q13</td>
<td>IL8</td>
<td>rs4073</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs2227307</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>4q22</td>
<td>FAM13A</td>
<td>rs2609255</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>4q35</td>
<td>TLR3</td>
<td>rs3775291</td>
<td></td>
<td>Harmful</td>
</tr>
<tr>
<td>5p15</td>
<td>TERT</td>
<td>rs2736100</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>6p21</td>
<td>CDKN1A</td>
<td>rs2395655</td>
<td>Yes</td>
<td>Harmful</td>
</tr>
<tr>
<td>6p21</td>
<td>HLA-DRB1</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>6q24</td>
<td>DSP</td>
<td>rs2076295</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>7q22</td>
<td>Intergenic</td>
<td>rs6776443</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10q24</td>
<td>OBFC1</td>
<td>rs11191865</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>11p15</td>
<td>MUC5B</td>
<td>rs35705950</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MUC2</td>
<td>rs7934606</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs111521887</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs5743894</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOLLIP</td>
<td>rs2743890</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>13q34</td>
<td>ATP11A</td>
<td>rs1278768</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>14q21</td>
<td>MDGA2</td>
<td>rs7144383</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>16q14–15</td>
<td>Intergenic</td>
<td>rs2034450</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>17q13</td>
<td>TP53</td>
<td>rs12951053</td>
<td>No</td>
<td>Harmful</td>
</tr>
<tr>
<td></td>
<td>TP53</td>
<td>rs12602273</td>
<td>No</td>
<td>Harmful</td>
</tr>
<tr>
<td>17q21</td>
<td>MAPT</td>
<td>rs1981997</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>17q21</td>
<td>SPP2C</td>
<td>rs17690703</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>19q13</td>
<td>DPP9</td>
<td>rs12610495</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>19q13</td>
<td>TGFBI</td>
<td>rs1800470</td>
<td>No</td>
<td>Harmful</td>
</tr>
</tbody>
</table>

Abbreviations: interleukin-1 receptor alpha (IL1RN), human telomerase RNA component (hTR), interleukin 8 (IL8), family with sequence similarity 13 member A (FAM13A), toll-like receptor 3 (TLR3), telomerase reverse transcriptase (TERT), cyclin dependent kinase inhibitor 1A (CDKN1A), HLA class II histocompatibility antigen, DRB1 beta chain (HLA-DRB1), desmoplakin (DSP), oligonucleotide/oligosaccharide-binding fold containing 1 (OBFC1), mucin 5B (MUC5B), mucin 2 (MUC2), toll-interacting protein (TOLLIP), ATPase phospholipid transporting 11A (ATP11A), MAM domain containing glycosylphosphatidylinositol anchor 2 (MDGA2), tumour protein p53 (TP53), microtubule-associated protein tau (MAPT), signal peptide peptidase like 2C (SPP2C), dipeptidyl peptidase 9 (DPP9), transforming growth factor beta 1 (TGFBI) [Adapted from Kropski 2015].

### 1.1.9 Current Treatment

The understanding of IPF pathobiology has substantially improved with time and novel compounds have been developed consequent to the increasing knowledge of the mechanisms underlying disease pathogenesis. This has changed the approach taken to treatment and recently, clinical trials have shifted their focus from anti-inflammatory and immunosuppressant compounds to molecules targeting components of the wound healing cascade and fibrogenesis [Spagnolo 2015(a)]. Despite this shift, the results of a number of high quality randomised, multicentre placebo controlled trials conducted in the last decade have been mostly disappointing.
<table>
<thead>
<tr>
<th>Study agent and/or treatment</th>
<th>Mechanism of action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azathioprine + prednisone + NAC (PANTHER)</td>
<td>Immunosuppressant + anti-inflammatory + antioxidant</td>
</tr>
<tr>
<td>Interferon gamma-1b</td>
<td>Anti-inflammatory, antifibrotic, and pro-Th1</td>
</tr>
<tr>
<td>Interferon gamma-1b (INSPIRE)</td>
<td>Anti-inflammatory, antifibrotic and pro-Th1</td>
</tr>
<tr>
<td>Sildenafil (STEP)</td>
<td>Phosphodiesterase 5 inhibitor and/or pulmonary vasodilator</td>
</tr>
<tr>
<td>Warfarin (ACE)</td>
<td>Anticoagulant</td>
</tr>
<tr>
<td>Bosentan (BUILD-1)</td>
<td>Dual endothelin receptor antagonist</td>
</tr>
<tr>
<td>Bosentan (BUILD-3)</td>
<td>Dual endothelin receptor antagonist</td>
</tr>
<tr>
<td>Ambrisentan (ARTEMIS)</td>
<td>Endothelin A receptor antagonist</td>
</tr>
<tr>
<td>Etanercept</td>
<td>TNF inhibitor</td>
</tr>
<tr>
<td>Everolimus</td>
<td>mTOR inhibitor - proliferation signal inhibitor</td>
</tr>
<tr>
<td>Macitentan (MUSIC)</td>
<td>Dual endothelin-receptor antagonist</td>
</tr>
<tr>
<td>Imatinib</td>
<td>PDGFR-α and -β Inhibitor</td>
</tr>
</tbody>
</table>

**Table 3. Summary of negative phase II and phase III randomised controlled trials in IPF**

Numerous clinical trials have demonstrated that inhibiting individual mediators or signalling pathways is largely ineffective in slowing the inexorable progression of IPF [Spagnolo 2015(a)]. This may be due to the plethora of mediators and signalling pathways in the wound healing process, many of which display considerable redundancy, that are likely to be involved in the pathogenesis of IPF [Maher 2007, Maher 2012]. Furthermore, the bleomycin mouse model, the most widely used model of experimentally-induced pulmonary fibrosis, recapitulates only partially the phenotype of progressive pulmonary fibrosis seen in IPF [Moelller 2008].

Only two compounds have emerged as efficacious from a series of multi-centre, double-blind and randomised, phase 3 clinical trials; on 15th October 2014, two drugs were approved for the treatment of IPF by the US Food and Drug Administration (FDA) – pirfenidone and nintedanib. Both reduce lung function decline and disease progression in patients with mild to moderate functional impairment and represent a huge breakthrough in the treatment of IPF but neither are curative. Notably, both are pleiotropic in their anti-fibrotic properties, reflective of the multiple mechanistic pathways of IPF, and treatment may necessitate a similarly multifaceted approach using combination regimens of antifibrotic agents in order to be effective [Wuyts 2014].

The proposed mechanism of action of the two compounds and the landmark phase 3 clinical trials establishing their effectiveness are only briefly summarised as a detailed
review is beyond the scope of this thesis. Pirfenidone, an orally administered pyridine, has been shown in both in vitro experiments and in vivo animal models of pulmonary fibrosis to exert anti-fibrotic, anti-inflammatory and anti-oxidant properties through down-regulation of profibrotic growth factors including platelet-derived growth factor (PDGF) and TGF-β; inhibition of inflammatory cytokine production and release; and reduction of lipid peroxidation and oxidative stress [Maher 2010, Borie 2016]. Safety and effectiveness of the drug were established in four phase III randomised clinical trials (Tanaguchi et al., CAPACITY 004, CAPACITY 006 and ASCEND trials). [Spagnolo 2015(b)]. Nintedanib is a potent intracellular inhibitor of fibroblast growth factor receptor (FGFR) 1, 2 and 3, PDGF receptor α and β, and VEGF receptor 1, 2 and 3 and it also inhibits members of the Src family of non-receptor tyrosine kinases [Hilberg 2008]. Two replicate randomised phase III trials (INPULSIS-1 and INPULSIS-2) demonstrated the safety and effectiveness of the drug [Spagnolo 2015(b)].

1.2 The Nitric Oxide Signalling Pathway

1.2.1 Nitric Oxide (NO) and Endogenous Production
Nitric oxide (NO) is a highly reactive, cellular signalling molecule and its diverse biological significance was recognised when the previously described endothelium-derived relaxing factor (EDRF) was identified as NO [Palmer 1987]. NO released from the endothelium was shown to locally activate soluble guanylate cyclase which catalyses the synthesis of cyclic guanosine monophosphate (cGMP), leading to smooth muscle arterial relaxation [Moncada 1991]. The NO/cGMP signalling pathway has since been implicated in a range of homeostatic and pathological processes, including regulation of vascular tone, neuronal transmission, immune responses, clotting cascades and fibrotic pathways [Moncada 1991]. NO can react with superoxide (O2⁻) to produce peroxynitrite (ONOO⁻), a potent oxidant and the key source of NO-mediated injury [Beckman 1996].

NO is endogenously synthesised from L-arginine by the nitric oxide synthase (NOS) family of enzymes [Knowles 1994]. Three NOS isoforms have been identified to date; neuronal NOS (nNOS), inducible NOS (iNOS), and endothelial NOS (eNOS). Each NOS is the product of a distinct gene and differs from the others in sub-cellular localisation, tissue distribution, and regulation [Alderton 2001].
1.2.2 Asymmetric Dimethylarginine (ADMA)

The endogenous production of NO is dependent on the availability of L-arginine, a common substrate shared by NOS and arginase, and on the activity of NOS which is regulated at many levels. A mechanism for the regulation of NOS is via endogenously produced inhibitors of NOS activity - the methylarginines (MAs). The MAs are naturally occurring amino acids that function as arginine analogues which bind to the active site of NOS and competitively inhibit enzymatic synthesis of NO, representing one of very few examples of an enzymatic pathway regulated by endogenous inhibitors [Vallance 1992].

Three MAs have been identified in mammals: asymmetric dimethylarginine (ADMA), N-monomethyl-L-arginine (L-NMMA) and symmetric dimethylarginine (SDMA). The asymmetric MAs, L-NMMA and ADMA, but not SDMA, are inhibitors of all three NOS isoforms [Vallance 1992]. The MAs are formed when arginine residues in proteins are methylated by the action of protein arginine methyltransferases (PRMTs). Subsequent proteolysis of these post-translationally modified proteins releases the free methylarginine residues into the cytoplasm [McBride 2001]. The MAs can move between the intracellular and extracellular compartments, resulting in variable local concentrations, and have been quantified in urine, plasma, cerebrospinal and bronchoalveolar lavage fluids [Bulau 2007]. Given its greater circulating concentrations, ADMA is considered to be the principal endogenously produced NOS inhibitor in vivo; although present in human plasma at a 10-fold lower concentration, L-NMMA can also inhibit NO production, while SDMA does not compete with arginine and its relevance is uncertain [Vallance 1992, Meyer 1997].

The endogenous production of ADMA has been proposed as an explanation for the ‘arginine paradox’; the observation that administering exogenous L-arginine in vivo or in vitro increases NO production by NOS despite intracellular concentrations of L-arginine that should saturate the enzyme [Tsikas 2000]. These findings are consistent with the presence of an endogenous competitive inhibitor at the active site of NOS. Thus, the degree of NOS inhibition that is exerted by the endogenous asymmetric methylarginines is dependent on the relative concentration of arginine and methylarginine at the active site of NOS [Leiper 2011] – the intracellular methylarginine:arginine ratio is therefore a critical determinant for NO availability.
1.2.3 Dimethylarginine Dimethylaminohydroxylase (DDAH)

Greater than 80% of the asymmetric MAs, ADMA and L-NMMA, are cleared by hydrolysis to citrulline and dimethylamines by the activity of dimethylarginine dimethylaminohydroxylases (DDAH) 1 and 2 [Ogawa 1987, Ogawa 1989]. The remainder is cleared by renal excretion and by a second minor pathway catalysed by alanine–glyoxylate aminotransferase (AGXT2) [Ogawa 1990]. Of note, SDMA is not hydrolysed by DDAHs and is mainly cleared by renal excretion and deamination by AGXT2.

To date, two isoforms of the DDAH enzymes have been identified in mammals, DDAH1 and DDAH2. The two isoforms have distinct tissue distributions with some relationship to NOS isoforms; DDAH1 is present at highest levels in the kidney, brain, liver and pancreas, whereas DDAH2 is found at high concentrations in the heart, kidney, placenta, lungs and immune tissues [Leiper 1999]. The expression pattern of DDAH1 overlaps more closely with the expression pattern of nNOS whereas DDAH2 expression has greater similarities with eNOS expression [Leiper 2011]. The finding that two isoforms of DDAH with different tissue distributions exist and that both proteins are highly conserved across mammalian species, with 98% similarity between murine
and human DDAH amino acid sequences, suggests that the regulation of methylarginine levels is of considerable biological importance [Leiper 1999].

Via the regulation of endogenous NOS inhibition by the asymmetric MAs, DDAHs represent a potential mechanism for in vivo control of NO bioavailability. Specific DDAH inhibition, using either a pharmacological approach (synthetic pharmacological DDAH1 inhibitor L-257) or a genetic approach (DDAH1 heterozygous knockout mice), increases the intracellular concentration of ADMA enough to inhibit NOS activity, suggesting that DDAH contributes to the control of NO in vivo [Leiper 2007, Nandi 2012].

Heterozygous deletion of DDAH1 results in a 40–45% reduction in total DDAH activity in several tissues in vivo, suggesting that DDAH1 is the primary isoform responsible for metabolising ADMA [Leiper 2007]. The dominance of DDAH1 activity in vivo and the difficulty in the purification of active recombinant DDAH2 protein make a direct comparison of respective isoform activity difficult [Leiper 2011] and additionally, the protein-protein interactions between DDAH1 and DDAH2 are unknown at present.

Thus DDAH activity is a key determinant of the bioavailability of NO. Furthermore, it appears that this signalling pathway has been highly conserved through evolution adding weight to its biological significance [Tran 2000]. The NO-ADMA-DDAH pathway thus offers a potential endogenous mechanism to achieve tissue-specific regulation of NO production. Pharmacological modification of DDAH has therefore been proposed as a mechanism for manipulating endogenous ADMA concentrations and regulating the production of nitric oxide in situations where alterations in nitric oxide signalling have been shown to contribute to pathophysiology [Leiper 2011].

![Diagram of the NO-ADMA-DDAH pathway](image)

**Figure 8.** DDAH activity as a key determinant of NO bioavailability
1.2.4 NO-ADMA-DDAH Pathway and Disease States

It is widely acknowledged that NO can play both a pathological and protective role, and human disease states may be associated with excessive or under-availability of NO. ADMA has been implicated in the pathogenesis of a variety of diseases - in the vast majority there is an association between increased plasma ADMA concentrations and disease pathology - implying that raised ADMA levels may reduce NO bioavailability and accelerate disease progression. The first observations emerged from patients with end-stage renal failure who demonstrated elevated plasma ADMA concentrations and reduced NO production [Vallance 1992]. Elevated ADMA concentrations have also been associated with major cardiovascular risk factors, including hypertension and hypercholesterolaemia and importantly, are an independent predictor of increased mortality and morbidity in patients with cardiovascular disease [Boger 2009].

More recently in chronic renal failure, it has emerged that a DDAH1 gene variant with increased renal DDAH1 mRNA transcription was associated with lower plasma ADMA levels and a steeper rate of renal function decline [Caplin 2010]. Additionally, proximal tubule–specific DDAH1 knockout mice demonstrated tubular cell accumulation of ADMA and lower NO concentrations, but unaltered plasma ADMA concentrations and were protected from markers of renal fibrosis (reduced kidney tissue mass, collagen deposition, and profibrotic cytokine expression) in two independent models of renal injury [Tomlinson 2015]. This indicates that reduced renal-specific ADMA metabolism via DDAH may be protective against progressive renal damage and that the measurement of circulating ADMA levels alone may be an imprecise marker of renal methylarginine metabolism [Tomlinson 2015].

The reduction of endogenous ADMA via the activation of DDAH could thus be used as a strategy for treating disorders in which ADMA levels are increased. The converse could be applicable to diseases in which the over production of NO has been implicated in disease progression, such as inflammatory diseases, cancer and sepsis [Leiper 2011]. Fewer studies have examined the association between reduced ADMA levels and diseases in which excessive NO production contributes to pathology. However, the therapeutic potential of modulating endogenous ADMA concentrations by pharmacologically manipulating DDAH activity remains an attractive therapeutic approach in the treatment of these diseases.
Human diseases associated with higher ADMA levels

- Asthma (sputum)
- Renal failure
- Peripheral vascular disease
- Pulmonary hypertension
- Septic shock
- Atherosclerotic coronary disease
- Diabetes mellitus
- Preeclampsia
- Alzheimer disease
- Connective tissue disease
- Liver disease
- Hyperthyroidism
- Hypothyroidism
- Stroke

Human diseases associated with lower ADMA levels

- Diabetes mellitus
- Alzheimer disease (cerebrospinal fluid)

Table 4. Association of ADMA levels in examples of human diseases
Plasma levels of ADMA unless stated [Adapted from Caplin 2012]

1.3 NO Signalling and Lung Fibrosis

1.3.1 Role of NO in Wound Repair

NO is required for normal wound healing, although the exact mechanisms of action by which NO modulates the process are not fully understood [Witte 2002]. Underproduction of NO has been linked to impaired wound healing, however, overproduction at the wrong time-point, may be just as detrimental [Schwentker 2002]. This is particularly relevant in the context of IPF - a disorder characterised by runaway wound healing that is central to its pathogenesis.

Wound healing involves many cell types including epithelial cells, inflammatory cells, fibroblasts, endothelial cells and platelets; all of these cells are capable of producing high levels of NO in response to cytokines, growth factors and inflammatory stimuli through the activity of iNOS (or lower levels constitutively, via eNOS) [Schwentker 2003]. The expression of iNOS in mammalian cells is governed predominantly by the transcription factor, NF-kB, and importantly, NO appears to have a biphasic role which enables it to both up- and down-regulate the expression iNOS via NF-kB depending on the local concentration of NO, thus creating a feedback loop [Connelly 2001].
The dual effects of NO on i) iNOS transcription and ii) cellular cytokine release appear to be involved in modulating key phases of the wound healing response (Figure 10):

1. Inflammatory Phase
Up-regulation of iNOS expression occurs early in wound repair, and may peak as early as 48 hours based on the finding of elevated levels of nitrite and citrulline, end products of NOS enzymatic activity, in wound fluid during the first 3 days of wound healing [Albina 1990]. A key role for iNOS-derived NO in the inflammatory phase is the promotion of cytokine expression, including the activation of latent TGF-β1 [Vodovotz 1999] and production of VEGF [Papapetropoulos 1997]

2. Proliferative Phase (and Angiogenesis)
NO has varied effects on proliferation and differentiation in different cell types. NO directly stimulates the proliferation of human lung fibroblasts [Romanska 2002] and indirectly via the activation of latent TGF-β1 [Vodovotz 1999]. NO-mediated iNOS transcription, probably via NF-kB, and subsequent NO release appears to be an early event (within 3 hours) in the proliferative response of human lung fibroblasts [Romanska 2002].

Although NO can lead to the activation of latent TGF-β1 [Vodovotz 1999], TGF-β1 can reduce the production of NO via suppression of iNOS by destabilising iNOS mRNA [Vodovotz 1997]. By way of a negative feedback loop, iNOS is downregulated in fibroblasts in the later stages of wound healing, possibly associated with a switch from proliferative to synthetic phenotype [Romanska 2002].

NO stimulates the proliferation of endothelial cells, protects against apoptosis and mediates the production of vascular endothelial growth factor (VEGF), the most potent
angiogenic factor active during wound healing [Nissen 1998, Schwentker 2002]. Furthermore, VEGF increases NO production via up-regulation of eNOS [Hood 1998].

3. Remodelling Phase
In animal studies and in vitro, the link between NO synthesis and wound collagen deposition has been well described, however, the clear-cut enhancement of collagen synthesis or gene expression has not been found [Witte 2002]. Treatment with NO donors, dietary arginine, or iNOS overexpression increases the collagen content of experimental wounds [Shi 2002, Schwentker 2002] and similarly, NOS inhibition has been found to decrease collagen and granulation tissue formation in experimental burn wounds [Ackay 2000]. In vitro, both wound-derived and normal skin-derived fibroblasts produce increased collagen following NO donor treatment and decreased collagen after NOS inhibition [Schaffer 1997]. This appears to be primarily due to posttranslational enhancement of collagen synthesis, not de novo transcription of the relevant collagen genes [Witte 2000].

Collagen accumulation is tightly regulated through the activity of collagenases and their inhibitors, tissue inhibitors of metalloproteinases (TIMP). A potential mechanism of posttranslational collagen regulation by NO is via inhibition of the collagenolytic pathway and enhancement of collagen accumulation [Witte 2000]. In murine pulmonary fibroblast culture, exogenous NO induces the expression of TIMP-1, as well as TGF-β1, type I collagen and the phosphorylation of Smad2 [Hsu 2006]. An alternate mechanism is via regulation of protein kinase C (PKC) activity by inhibiting PKC activity, NO could down-regulate PKC-related collagen synthesis in fibroblasts [Clementi 1995].
Figure 10. Possible role of NO in normal wound healing following alveolar epithelial injury

Of note, dietary supplementation of L-arginine, the substrate for NOS, has been shown to enhance collagen deposition and wound strength in humans and animal models [Kirk 1993]. Some of this effect may be due to increased production of L-ornithine - a precursor of L-proline during collagen synthesis. L-ornithine is generated from L-arginine through the action of arginase, an enzyme that probably directly competes with NOS for the substrate [Li 2001]. However, the finding that L-arginine supplementation does not improve collagen deposition in iNOS-/- mice to the same extent as in wild-type controls implicates NO directly as a factor in the promotion of wound healing by L-arginine [Shi 2000].

1.3.2 Overproduction of NO in IPF
As described in the previous section, under the appropriate conditions, NO is produced by several cells within the respiratory tract, including airway epithelial cells, inflammatory cells, fibroblasts and endothelial cells, and as such, there is accumulating evidence that IPF is a disease linked to excessive production of NO via up-regulation of iNOS. Evidence from the bleomycin mouse model of pulmonary fibrosis, has demonstrated that increased production of NO is linked to increased expression of iNOS in the alveolar epithelium and increased level of NO metabolites in bronchoalveolar lavage fluids (BALF) [Jang 2004, Pullamsetti 2011]; and genetic or pharmacological knockdown of iNOS attenuates lung collagen deposition [Genovese
In humans, a study of 48 IPF lungs compared to controls, demonstrated an increase in the expression of iNOS and the production of nitrotyrosine in the alveolar epithelium and inflammatory cells in the early to intermediate stages of the disease [Saleh 1997].

NO can react with superoxide (O2−) to produce peroxynitrite (ONOO−), a potent oxidant and the key source of NO-mediated injury [Beckman 1996]. Reactive oxygen species (ROS) and reactive nitrogen species (RNS) have been long been associated with IPF and therapy with the small molecule antioxidant N-acetylcysteine (NAC) has been used for many years. Studies in IPF lung have demonstrated signatures of both oxidative and nitrosative stress; BALF isolated from IPF patients demonstrates elevated levels of proteins damaged by oxidation [Maier 1991]; lung tissue from IPF patients demonstrates higher levels of nitrotyrosine, a by product of protein nitration by peroxynitrite [Saleh 1997] and lung epithelial cells from patients with IPF demonstrate signatures of chronic oxidative stress [Kuwano 2003]. There is also emerging evidence of ROS-mediated EMT [Radisky 2005, Canito 2010]. RNS and ROS are important modulators of TGF-β; leading to increased activation and increased expression of TGF-β in airway epithelial cells [Goroweic 2012], and thus NO may be a driver of TGF-β mediated EMT via generation of ROS/RNS.

More recently, evidence has emerged from in vivo human studies of increased NO in exhaled breath, exhaled breath condensate and serum from IPF patients. Measurement of the fraction of exhaled nitric oxide (FeNO) is performed in several lung diseases and the anatomical origin of NO production can be separated into alveolar and airway compartments by measuring the flow-independent NO parameters - alveolar concentration of exhaled NO (CalvNO) and maximum conducting airway wall flux (J’awNO) respectively [Tsoukias 1998]. A study of 30 patients with IIP (22 IPF and 8 idiopathic NSIP), demonstrated that IIP patients had significantly higher FeNO and CalvNO levels compared to healthy controls (p < 0.0001), indicating that the increased NO production was of alveolar origin [Cameli 2014].

The rapid reaction of NO and superoxide produces peroxynitrite which reacts with proteins to form the stable product 3-nitrotyrosine (3-NT). Exhaled breath condensate (EBC) is a novel, non-invasive technique to collect the water vapour and aerosolised particles generated from the respiratory tract to allow the assessment of cytokines, chemokines and inflammatory mediators as potential biomarkers in lung diseases. In a recent study of 20 patients with IPF, EBC levels of 3-NT, measured via enzyme
immune assay, were significantly higher compared to controls (2.5 versus 0.3 ng ml\(^{-1}\), \(p = 0.02\)) [Chow 2012]. EBC measurements of total nitrogen oxides (NOx) were also raised but did not reach significance. FeNO measurements were also elevated, confirming the results of earlier studies (although the contribution of alveolar NO to the FeNO was not assessed). Plasma levels of 3-NT have also recently been measured by mass spectrometry in a mixed cohort of 42 patients with ILD (32 IPF and 10 connective tissue disease associated ILD). Circulating 3-NT was significantly elevated in ILD patients compared to healthy controls (3.08-fold increase, \(p<0.0001\)) [Pennathur 2016].

Although the main intention of these recent studies has been to explore potential biomarkers in IPF, they provide valuable current evidence for the involvement of NO in the disease, specifically the overproduction of NO.

1.3.3 NO-ADMA-DDAH Pathway and IPF

The role of the NO-ADMA-DDAH pathway in IPF has yet to be fully understood. As discussed in the previous section, although there is accumulating evidence that there is overproduction of NO in IPF, which is likely to result from the up-regulation of iNOS, the contributions of ADMA and the DDAH isoforms to this process are only beginning to be explored in experimental models of pulmonary fibrosis.

A series of experiments by Pullamsetti et al. have supported a role for DDAH in pulmonary fibrosis. The only published human data thus far has been the comparison of DDAH isoform and iNOS expression in human lungs from IPF patients undergoing lung transplantation compared to donor lungs (n=6 lungs per group) [Pullamsetti 2011]. Expression of DDAH2 and iNOS mRNA was significantly increased in IPF lungs compared to donor lungs; whereas DDAH1 mRNA expression increased in the fibrotic lungs but did not reach significance. Protein levels of DDAH2 and iNOS were three times higher in IPF lungs than in donor lungs. Interestingly, DDAH1 protein levels appeared higher than DDAH2 in donor lungs, and remained at a similar level in the IPF lung, suggesting that there is a higher baseline level of DDAH1 than DDAH2 in the normal lung. Immunohistochemical staining showed that DDAH1 and DDAH2 were localised in the alveolar and bronchiolar epithelium, fibroblasts, endothelium, inflammatory cells in donor lungs. In IPF lungs, DDAH2 and iNOS were observed in close proximity to fibrotic scars, thickened septa, and fibroblast foci. Furthermore, co-staining of DDAH2 and iNOS with pro-SP-C, a specific marker of type 2 AECs, revealed the presence of DDAH2 and iNOS in type 2 AECs lining the thickened septa.
Bleomycin-challenged mice displayed strikingly similar findings to those demonstrated in the human IPF lung. Murine lung tissue sampled at 4, 7, 14, and 21 days after the bleomycin challenge demonstrated a significant and steady induction of DDAH2 mRNA and protein expression in response to bleomycin compared to control mice from days 7 to 21. DDAH1 mRNA expression appeared to increase in response to bleomycin but did not reach significance, and there was a higher level of DDAH1 than DDAH2 protein at baseline. In the control mice DDAH1 and DDAH2 immunohistochemical staining displayed the same distribution of localisation as the human donor lungs and an increase of both DDAH isoforms and iNOS was observed in the alveolar epithelium of bleomycin-injured lungs. Immunostaining with pro–SP-C demonstrated that DDAH2 and iNOS were present in type 2 AECs and increased in the lungs of bleomycin-challenged mice. Additionally, the increase in DDAH2 mRNA and protein expression on days 7 to 21 was mirrored by an increase in total DDAH enzyme activity (indirectly measured with a colorimetric assay to observe the formation of L-citrulline from ADMA), which peaked on days 7 to 21. The ADMA content in both lung tissue and BAL fluid initially increased after the bleomycin challenge on days 4 to 7, but decreased thereafter. Similarly, DDAH2 mRNA expression was increased, and ADMA levels were decreased, in type 2 AECs isolated from bleomycin-challenged lungs.

To investigate the effect of DDAH inhibition on bleomycin-induced fibrosis, bleomycin-challenged mice were treated with L-291, a selective DDAH1 inhibitor (10 mg/kg every third day for 21 days), ADMA (30 mg/kg every third day for 21 days) and saline (control). L-291 treated mice maintained lung compliance, and demonstrated reduced collagen deposition, tissue destruction and extent of fibrosis on histological analysis compared to saline and ADMA treatments. The collagen content in lung lysates, quantified with the Sircol assay, was also significantly decreased with L-291 treatment.

Importantly, none of these affects were demonstrated with ADMA - highlighting the fact that DDAH inhibition and ADMA treatment do not have one and the same effect on the development of lung fibrosis. The failure of exogenous ADMA treatment to attenuate bleomycin-induced pulmonary fibrosis corroborates previous evidence whereby direct infusion of ADMA resulted in elevated collagen deposition in mouse lungs and enhanced arginase activity [Wells 2009]. L-ornithine (the precursor for L-proline and collagen) is produced by activation of arginase pathway. The arginases and NOS share L-arginine as a common substrate. As ADMA competes with L-arginine for the
active site of NOS, elevated levels of ADMA may divert L-arginine to the arginase pathway, resulting in increased collagen deposition [Ricciardolo 2005].

To study the effect of DDAH overexpression in vivo, transgenic mice overexpressing DDAH1 (DDAH1-TG) were challenged with bleomycin and compared to wild-type mice. DDAH1-TG mice displayed significantly reduced lung compliance and significantly increased fibrotic scores and collagen deposition after 21 days indicating that overexpression of DDAH1, and possibly DDAH1 over-activity (not measured), aggravates bleomycin-induced fibrosis.

To determine whether the antifibrotic effects of DDAH inhibition were mediated via iNOS, bleomycin challenged iNOS wild type (iNOS-WT) and iNOS-knockout (iNOS-KO) mice were treated with either saline or L-291. As expected from prior studies, bleomycin challenged iNOS-KO mice displayed less fibrosis and lung collagen deposition compared to wild-type mice but more interestingly, these levels were not as low as treatment of bleomycin challenged iNOS-WT mice with L-291. These results indicate that the protective effect of DDAH1 inhibition on bleomycin-induced fibrosis is greater than that of iNOS-KO. Indeed, bleomycin challenged iNOS-KO mice treated with L-291 displayed enhanced protection against fibrosis than with iNOS-KO alone, back to the similar levels demonstrated with L-291 alone. Selective inhibition of DDAH1 appears to be protective against pulmonary fibrosis in the bleomycin model.

In summary, these important findings suggest that both DDAH1 and DDAH2 play an important role in the pathogenesis of pulmonary fibrosis. The human data demonstrates that DDAH1 appears to be the more abundant isoform expressed in human lungs at baseline and levels are similar in the IPF lung; whereas DDAH2 expression is induced several-fold and co-localises with iNOS in the IPF lung, specifically in type 2 AECs in the alveolar epithelium. The murine data comprehensively demonstrates that selective DDAH1 inhibition attenuates bleomycin-induced fibrosis and conversely, that DDAH1 overexpression aggravates it. Although several aspects need further exploration before a potential therapeutic role for DDAH inhibition can be considered, including the specific contributions of each DDAH isoform, taken together, this evidence indicates that DDAHs may play a key role in the development and progression of IPF.
1.3.4 NO-ADMA-DDAH Pathway and Epithelial to Mesenchymal Transition

At the present time, there is no published literature specifically on the role of DDAH in EMT. However, studies by Pullamsetti et al. have provided evidence that DDAH may be an important regulator of type 2 AEC biology and therefore involved in the pathogenesis IPF, a disease characterised by alveolar epithelial cell dysfunction, which includes the process of EMT.

Both DDAH isoforms are expressed in the type 2 AECs of healthy lungs, and in particular, the expression of DDAH2 is markedly increased in the type 2 AECs of IPF patients and mice with bleomycin-induced fibrosis [Pullamsetti 2011]. Furthermore, treatment of murine type 2 AECs with TGF-β1 significantly induced DDAH2 mRNA expression - suggesting that transcriptional modulation of DDAH2 is mediated by the key profibrotic cytokine that drives fibrogenesis and EMT. Importantly, pharmacological inhibition of DDAH1 with L-291 and DDAH2-selective siRNA inhibition of DDAH2 both attenuated serum-induced murine type 2 AEC proliferation and induced apoptosis in vitro, indicating a role for both DDAH isoforms in these processes.

As discussed previously, the type 2 AEC plays a key role in the pathogenesis of IPF - undergoing aberrant activation, hyperplasia, proliferation, epithelial-to-mesenchymal transition and increased apoptosis. Induction of DDAH2 by profibrotic mediators in vitro and overexpression of DDAH2 in hyperplastic type 2 AECs in vivo suggest that DDAH may be regulate these type 2 AEC processes, including EMT in IPF.

1.4 Human DDAH1 and DDAH2 Genes

1.4.1 DDAH Genes

*DDAH1* maps to chromosome 1p22 and the *DDAH2* maps to the major histocompatibility complex III region of chromosome 6p21.3 [Tran 2000]. The coding region of both *DDAH1* and *DDAH2* is divided into six exons and the intron/exon boundaries are highly conserved, having drifted by no more than 6 nucleotides. This striking level of conservation and the chromosomal localisation of the DDAHs supports gene duplication, which by phylogenetic analysis of DDAH sequences from diverse species, suggests that DDAH gene duplication occurred prior to the emergence of bony fish some 400 million years ago [Tran 2000, Hughes 1998].
The key observations that the two DDAH genes are highly conserved through evolution [Tran 2000] and that the protein sequences of both DDAH isoforms are highly conserved across mammalian species, with 98% similarity between murine and human DDAH amino acid sequences [Leiper 1999], suggests that DDAH1 and DDAH2 are of significant biological importance and functional disruption by single nucleotide polymorphisms (SNPs) may have a significant effect on the regulation of the methylarginines in vivo.

1.4.2 Human DDAH Genetic Studies
In view of the association between plasma ADMA levels and various disease states, genetic variation in the DDAH genes may potentially modulate endogenous ADMA concentrations and be of functional significance in the pathobiology of diseases. Consequently, there has been a dramatic increase in the number of genetic association studies exploring the relationship between DDAH SNPs with various human diseases and clinical phenotypes and/or ADMA levels. Of note, these studies are associative and are not designed to test causality. DDAH1 is the primary isoform responsible for metabolising ADMA [Leiper 2007], and accordingly, it has emerged upon review of the literature that DDAH1 SNPs associate with circulating ADMA levels whereas there is no evidence that DDAH2 SNPs do associate (Tables 5 and 6). A genetic association study of the DDAH genes in the context of pulmonary fibrosis has not been published thus far.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Minor allele frequency %</th>
<th>Associated tagged SNPs</th>
<th>Significant association with ADMA levels in a normal cohort</th>
<th>Significant association with disease state</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs530006</td>
<td>0.45</td>
<td>rs1146381</td>
<td>EARSII</td>
<td>T2DM¹; CVD in T2DM²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs7521189</td>
<td>PIVUS</td>
<td></td>
</tr>
<tr>
<td>rs6576765</td>
<td>0.44</td>
<td>rs233112</td>
<td>EARSII</td>
<td>CVD in T2DM²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs11161606</td>
<td>PIVUS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs233099</td>
<td>PIVUS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rs3768227</td>
<td>PIVUS</td>
<td></td>
</tr>
<tr>
<td>rs233109</td>
<td>0.43</td>
<td>None</td>
<td>PIVUS</td>
<td>Not known</td>
</tr>
<tr>
<td>rs669173</td>
<td>0.38</td>
<td>None</td>
<td>PIVUS</td>
<td>T2DM¹</td>
</tr>
<tr>
<td>rs877041</td>
<td>0.34</td>
<td>rs582145</td>
<td>PIVUS</td>
<td>Not known</td>
</tr>
<tr>
<td>rs527762</td>
<td>0.30</td>
<td>None</td>
<td>PIVUS</td>
<td>Not known</td>
</tr>
<tr>
<td>rs7555486</td>
<td>0.28</td>
<td>rs13373844</td>
<td>PIVUS</td>
<td>T2DM¹</td>
</tr>
<tr>
<td>SNP</td>
<td>Minor allele frequency %</td>
<td>Associated tagged SNPs</td>
<td>Association with ADMA levels in a normal cohort</td>
<td>Significant association with disease state</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>rs805305</td>
<td>0.32</td>
<td>rs805304, rs707916</td>
<td>EARSII (ns), PIVUS (ns)</td>
<td>Hypertension&lt;br&gt;Stroke&lt;br&gt;Septic shock&lt;br&gt;T2DM with CRF&lt;br&gt;Myocardial infarction</td>
</tr>
<tr>
<td>rs2272592</td>
<td>0.29</td>
<td>None</td>
<td>PIVUS (ns)</td>
<td>T2DM</td>
</tr>
<tr>
<td>rs9267551</td>
<td>0.22</td>
<td>None</td>
<td>PIVUS (ns)</td>
<td>T2DM</td>
</tr>
<tr>
<td>rs3131383</td>
<td>0.10</td>
<td>None</td>
<td>PIVUS (ns)</td>
<td>T2DM</td>
</tr>
</tbody>
</table>

Table 5. DDAH1 SNPs: associations with ADMA levels in normal cohorts and disease states

Minor allele frequency referenced from the CEU cohort, Utah residents with Northern and Western European ancestry in the International HapMap Project [Altshuler 2010]; ¹[Abhary 2010]; ²[Anderssohn 2014]; ³[Caplin 2010].

Table 6. DDAH2 SNPs: associations with ADMA levels in normal cohorts and disease states

Minor allele frequency referenced from the CEU cohort, Utah residents with Northern and Western European ancestry in the International HapMap Project [Altshuler 2010]; ¹[Maas 2009]; ²[Bai 2009]; ³[O’Dwyer 2006]; ⁴[Marra 2012]; ⁵[Perez-Hernandez 2014]; ⁶[Seo 2012]; ⁷[Andreozzi 2012]; ⁸[Abhary 2010]. Type 2 diabetes mellitus (T2DM), cardiovascular disease (CVD), chronic renal failure (CRF).

European Atherosclerosis Relatives Study II (EARSII): a case-controlled study of 407 white male students aged 18–28 with no documented cardiovascular disease whose fathers had suffered from a proven myocardial infarction with an equal number of white age-matched controls, all recruited from European universities in 1993 [Tiret 2000]. No heterogeneity was detected in ADMA levels between cases and controls, therefore the association between genotype and plasma ADMA concentrations was examined in the pooled population [Caplin 2010].

Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS): a population-based study of 1016 70-year-olds (50% women) in Sweden. Measurement of endothelium-dependent vasodilation using the invasive forearm technique with measurement of flow-mediated dilatation. Plasma L-arginine and ADMA levels were measured by high-performance liquid chromatography and 55 single nucleotide polymorphisms (SNPs) in the DDAH1 and DDAH2 genes were genotyped [Lind 2013].
1.4.3 Rationale for a Targeted Genetic Association Study

A targeted genetic association study typically adopts a candidate gene approach to assess the relationship between genetic variation in a particular gene and a disease or a disease-related phenotype [Soler 2012]. Confounding is a major concern in observational studies but as genotypes are assigned by an essentially random process, underlying patient factors are not likely to confound genetic associations. There are several advantages to this approach in the context of assessing the NO-ADMA-DDAH pathway in IPF.

In particular, measurement of ADMA levels in an IPF patient population has not been yet been performed but may be prone to several confounding factors. Measurement of serum ADMA levels is susceptible to systemic variables, not all of which can be adjusted for; underlying patient demographic and clinical factors, dietary intake of arginine, alternative sites of DDAH expression, and renal clearance will play a role in determining ADMA levels. Additionally, although bronchoalveolar lavage of the IPF lung is an attractive compartment in which to assess ADMA levels, BALF constituents will vary with sampling location (due to the heterogeneity of disease distribution and extent in the lung), and operator dependent factors such as varying dilution effects, indwelling times and patient tolerance of the procedure. Furthermore, it is increasingly recognised that the main limitation of all animal models of fibrosis is that they attempt to reproduce years of human injury to an unknown insult in only a brief period after a single insult. The widely used bleomycin-induced murine model of pulmonary fibrosis, a “one hit” and predominantly inflammatory model of lung injury, does not fully recapitulate the multiple pathways and complex pathobiology of IPF [Spagnolo 2015 (b)].

An alternative genetic approach is via genome wide association studies (GWAS) which adopt a hypothesis-free approach to studying association across the genome - although less prone to the reporting and publication biases which may occur with candidate gene studies, they require careful design and stringent statistical thresholds of significance to avoid false positive findings [Soler 2012]. Two major GWAS studies have been conducted in IPF and have identified genes involved in biological pathways that were not previously connected to the disease [Noth 2013, Fingerlin 2013]. Adopting a candidate gene approach with an a priori hypothesis based on a mechanism of biological relevance is therefore a valuable tool to investigate the NO-ADMA-DDAH pathway in IPF.
It can be hypothesised that \textit{DDAH} SNPs that are associated with ADMA levels may also affect disease phenotype; and accordingly, SNPs that modulate ADMA levels to a greater extent, may also have a greater effect on the disease phenotype. In the context of IPF, the directionality of ADMA levels observed in association with genetic variation in \textit{DDAH} is of particular importance. IPF appears to be a disease of NO overproduction and may thus potentially be associated with lower ADMA levels and DDAH over-activity. It has already been noted that both pharmacological and genetic knock-out of \textit{DDAH1} attenuates bleomycin-induced fibrosis whereas genetic overexpression \textit{DDAH1} aggravates it [Pullamsetti 2011]. Genetic variants of \textit{DDAH1} that are associated with lower ADMA levels may therefore be indicative of DDAH1 over-activity.
CHAPTER 2: HYPOTHESIS AND STUDY AIMS

2.1 Hypothesis

This project aims to investigate the nature of the relationship between the NO-ADMA-DDAH pathway and the progression of lung fibrosis; more specifically that variation in DDAH activity will impact on disease progression in idiopathic pulmonary fibrosis.

The hypothesis tested is that higher DDAH activity and lower ADMA levels are predictive of accelerated disease progression in idiopathic pulmonary fibrosis.

This hypothesis has been addressed using both in vitro and human studies.

2.2 In Vitro Study

The in vitro study aims to identify the mechanistic pathways linking DDAH activity to lung fibrosis in an experimental model of EMT in a human alveolar epithelial cell line.

The hypothesis tested is that increased DDAH activity mediates loss of endogenous NOS inhibition by ADMA and promotes the fibrotic phenotype in TGF-β mediated epithelial to mesenchymal transition.

2.3 Human Study

The human study aims to identify the impact of DDAH single nucleotide polymorphisms with a functional effect on ADMA levels on longitudinal lung function decline and survival in a cohort of well-phenotyped patients with idiopathic pulmonary fibrosis.

The hypothesis tested is that genetic variants associated with higher DDAH1 activity and lower ADMA levels and will be associated with accelerated lung function decline and mortality in idiopathic pulmonary fibrosis.
CHAPTER 3: MATERIALS AND METHODS

3.1 In Vitro Methods

3.1.1 Cell Line

Human type 2 alveolar epithelial cells from the A549 cell line (ATCC) were used in all in vitro experiments in this study. A549 cells are adenocarcinomic human type 2 alveolar epithelial cells that were originally removed and cultured from an explanted adenocarcinoma of the lung taken from a 58 year old Caucasian male in 1972. They are a commonly used cell line and when cultured in vitro, they grow as confluent monolayer cells adherent to the culture flask with characteristic type 2 alveolar epithelial cell morphology [Foster 1998].

3.1.2 Cell Line Culture

A549 cells were maintained in Ham’s F-12K (Kaihn’s) medium (Thermo Fisher Scientific) containing L-glutamine with 10% fetal bovine serum, and 100 U/ml penicillin and 100 ug/ml streptomycin at 37°C in a humidified 5% CO₂ atmosphere. Cells were cultured in T75 flasks and medium was changed every second day. The cells were passaged with trypsin/EDTA when reaching 75% confluence to a 1 in 4 split. In preparation for cell line experiments, passaged cells were cultured in six-well plates until 80% confluence. They were washed twice with phosphate buffered saline (PBS) and serum starved for 24 hours in serum-free medium containing 0.1% bovine serum albumin in order to achieve cell quiescence immediately prior to experimental treatments.

3.1.3 Cell Line Treatment

TGF-β induces A549 cells to undergo EMT in a time- and concentration-dependent manner [Kasai 2005, Kim 2007]. The A549 cells were serum starved for 24 hours before stimulation with TGF-β1 at 2.5 to 5ng/ml (R&D Systems) for 48 hours. A profile indicating transformation from an epithelial to mesenchymal phenotype was assessed. For further details, refer to experimental protocols outlined in the relevant subsections in Chapters 4 and 5.

3.1.4 Protein Harvesting from Cell Culture

Following experimental treatments, 1.5mL of media was aspirated from each well, centrifuged (14,000 rpm for 15 minutes at 4°C) and the supernatant stored at -80°C.
Adherent cells were washed twice with PBS prior to removal by scraping directly into 100uL RIPA lysis buffer (Thermo Fisher Scientific) containing 10% protease inhibitors (Roche Applied Sciences). The cell suspension was centrifuged again (14,000 rpm for 15 minutes at 4°C), the cell pellet was discarded and the cell lysate was aspirated, nanodropped to quantify the micrograms of protein per microliter, and stored at -80°C.

3.1.5 RNA Extraction from Cell Culture
Following experimental treatments, 1.5mL of media was aspirated from each well, centrifuged (14,000 rpm for 15 minutes at 4°C) and the supernatant stored at -80°C. Adherent cells were washed twice with PBS prior to removal by scraping directly into RLT buffer (Qiagen Ltd). RNA extraction was performed on the cell lysate using a column purification kit (RNeasy Mini Kit for RNA, Qiagen) according to the manufacturer’s protocols. The freshly made mRNA samples were immediately quantified and reverse transcribed to create cDNA from the RNA template, as described in later sections.

3.1.6 DNA Isolation from Human Whole Blood
Biobank samples of patient DNA were kindly donated by Professor Athol Wells at the Interstitial Lung Disease Unit at the Royal Brompton Hospital. DNA was isolated from human whole blood using a column purification kit (DNeasy kit for DNA, Qiagen Ltd) according to the manufacturer’s protocols and stored at -80°C at the Royal Brompton Hospital site.

3.1.7 Protein, RNA and DNA Quantification
Total protein, RNA and DNA concentrations were determined using a NanoDrop™ spectrophotometer (Thermo Scientific). This machine allows quantification of an unknown lysate using spectrometry at specific wavelengths to determine absorbance. Sample volumes of 2uL were loaded for analysis, measured in duplicate and then averaged.

3.2 Molecular Biology Methods

3.2.1 Reverse Transcription Quantitative PCR (RT-qPCR)
Relative mRNA expression was determined using a two-step PCR approach; first involving reverse transcription to create cDNA from the RNA template (RT-qPCR) followed by quantitative (qPCR) real-time PCR. The iScript cDNA Synthesis Kit
(BioRad) was used according to the manufacturer’s protocols with cycling conditions as set out below.

**Reaction Components**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>iScript reaction mix</td>
<td>4 μL</td>
</tr>
<tr>
<td>iScript reverse transcriptase</td>
<td>1 μL</td>
</tr>
<tr>
<td>Nuclease free water</td>
<td>x μL</td>
</tr>
<tr>
<td>Template RNA (1000ng)</td>
<td>x μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>20 μL (produces 50ng/μL cDNA)</td>
</tr>
</tbody>
</table>

**Reaction Protocol**

- **Incubation**: 5 minutes at 25°C
- **Thermal cycling x 1:**
  - Reverse transcription: 30 minutes at 42°C
  - RTase inactivation: 5 minutes at 85°C
- **End of reaction**: Hold at 4°C

Table 7. Reverse-transcription PCR protocol

### 3.2.2 Quantitative Polymerase Chain Reaction (qPCR)

The 7900HT Fast System (Applied Biosystems) was used for real-time quantitative PCR (qPCR). This type of PCR is capable of quantifying the amount of product at the end of each PCR cycle. The fluorophore SYBR green is a cyanine fluorescent dye, and once it has bound to double stranded DNA, it emits light in the green spectrum that is relative to DNA concentration. The iTaq Fast SYBR Green-based PCR Supermix with ROX reference dye (BioRad) was used according to the manufacturer’s protocols. Primers (R&D Systems) were added at 200nM along with 10ng cDNA to the Supermix. Samples were plated on to 96 well PCR plates and were run with the cycling conditions set out below.

Amplification curves were generated through the detection of SYBR green fluorescence and a cycle-threshold value (C<sub>T</sub>) read automatically at the base of the exponential phase of the curve. The C<sub>T</sub> value indicated the number of cycles required for the fluorescent signal to cross the threshold level above background fluorescence (set arbitrarily at 0.2). The C<sub>T</sub> value thus indicated that the quantity of the target gene in each sample had been determined. Amplification efficiency for each primer mix and a housekeeper gene was assessed by running serial dilutions of cDNA on the same PCR plate as the experimental sample. Standard curve C<sub>T</sub> values were plotted on a logarithmic scale and the slope of linear regression used to calculate relative gene
expression. Amplification curves were examined in duplicate for each pair of technical replicate samples. Technical replicate C_{T} values were averaged for each sample and normalised to a housekeeper gene to correct for the amount and quality of cDNA put into the reaction. Several “housekeeper” genes were optimised and 18S was selected as the most stably expressed gene against which to compare the relative fold expression of DNA.

### Reaction Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sybr Green and ROX mix</td>
<td>10 μL</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>0.4 μL</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>0.4 μL</td>
</tr>
<tr>
<td>Nuclease free water</td>
<td>7.2 μL</td>
</tr>
<tr>
<td>cDNA (1:10 dilution to 5ng/uL)</td>
<td>2 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>20 uL (10ng of cDNA per reaction)</td>
</tr>
</tbody>
</table>

### Reaction Protocol

<table>
<thead>
<tr>
<th>Step</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial step</td>
<td>2 minutes at 95°C</td>
</tr>
<tr>
<td>Thermal cycling x 40:</td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>3 seconds at 95°C</td>
</tr>
<tr>
<td>Annealing</td>
<td>30 seconds at 60°C</td>
</tr>
<tr>
<td>End of reaction</td>
<td>Disassociation step</td>
</tr>
</tbody>
</table>

Table 8. Quantitative PCR protocol

### 3.2.3 Western Blotting

Western blotting was used to detect and identify relative amounts of specific proteins in cell lysates. Proteins were first separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). SDS PAGE gels were made according to the molecular weight of the protein of interest as the speed and level of separation is determined by the percentage of acrylamide; a range of 7.5% to 12% was used for the resolving gel and 4% was used for the stacking gels. Protein lysates were prepared by adding Laemmli buffer and heating to 95°C to denature the tertiary protein structures. Equal amounts of protein (100ug) were loaded into each lane prior to electrophoresis. The protocol used is detailed below.

The relative amounts of a specific protein were detected by an infra-red technique using the Licor Odyssey detection and quantification system. The band densities for each protein were quantified and expressed as a ratio of band density to that of the housekeeper protein. Tubulin was chosen as the most stable of three other housekeepers that were tested. Antibodies for tubulin, α-SMA, E-cadherin, type 1
collagen, iNOS, eNOS, and nNOS were used (Abcam) and DDAH1 and DDAH2 (purified by our lab). Active full length protein standards were purchased for use as positive controls for human α-SMA, E-cadherin, type 1 collagen, iNOS, eNOS, and nNOS (Abcam).

<table>
<thead>
<tr>
<th>Protocol Step</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denature proteins</td>
<td>Add Laemlli buffer to protein lysate (3:1 ratio) and heat to 95°C for 4 minutes.</td>
</tr>
<tr>
<td>Separate proteins</td>
<td>Load 2uL Bio-Rad All Blue protein marker into first lane and load equal amounts of protein into others. Run SDS-PAGE gel in running buffer at 150V for up to 90mins.</td>
</tr>
<tr>
<td>Membrane transfer</td>
<td>Soak Immobilon-P membrane in 100% methanol. Transfer to membrane at 100V in transfer buffer for 1 hour. Surround tank with ice.</td>
</tr>
<tr>
<td>Block membrane</td>
<td>Block in 5% non-fat dried milk + PBS-Tween at room temperature for an 1 hour</td>
</tr>
<tr>
<td>Primary antibody incubation</td>
<td>Primary antibody incubation (in 5% non-fat dried milk + PBS-Tween) at 4°C overnight with agitation</td>
</tr>
<tr>
<td>First washes</td>
<td>PBS-Tween washes (3 x quick swills, 2 x 5 minutes washes with agitation)</td>
</tr>
<tr>
<td>Secondary antibody incubation</td>
<td>Secondary antibody incubation (in 5% non-fat dried milk + PBS-Tween) at room temperature for 1 hour with agitation</td>
</tr>
<tr>
<td>Second washes</td>
<td>PBS-Tween washes (3 x quick swills, 2 x 5 minutes washes with agitation)</td>
</tr>
<tr>
<td>Protein detection</td>
<td>Licor Odyssey detection and quantification system</td>
</tr>
</tbody>
</table>

**Table 9. Western blotting protocol**

<table>
<thead>
<tr>
<th>Human Primary Antibody</th>
<th>Dilution Used</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-tubulin (loading control)</td>
<td>1:10000</td>
<td>Abcam (ab7291)</td>
</tr>
<tr>
<td>α-SMA</td>
<td>10:10000</td>
<td>Abcam (ab5694)</td>
</tr>
<tr>
<td>E-cadherin</td>
<td>10:10000</td>
<td>Abcam (ab1416)</td>
</tr>
<tr>
<td>type 1 collagen</td>
<td>10:10000</td>
<td>Abcam (ab34710)</td>
</tr>
<tr>
<td>iNOS</td>
<td>10:10000</td>
<td>Abcam (ab3523)</td>
</tr>
<tr>
<td>eNOS</td>
<td>10:10000</td>
<td>Abcam (ab5589)</td>
</tr>
<tr>
<td>nNOS</td>
<td>10:10000</td>
<td>Abcam (ab3511)</td>
</tr>
<tr>
<td>DDAH1</td>
<td>15:10000</td>
<td>Polyclonal goat antibody to both isoforms</td>
</tr>
<tr>
<td>DDAH2</td>
<td>15:10000</td>
<td>generated by James Leiper and group</td>
</tr>
</tbody>
</table>

**Table 10. Primary antibodies and concentrations used**
3.2.4 KASP Genotyping Assay

DDAH SNPs were genotyped in patient DNA samples using a high-throughput commercial assay to achieve the highest accuracy and uniformity across the whole cohort in a single genotyping run. The Kompetitive Allele Specific PCR (KASP™) genotyping system is a fluorescence resonant energy transfer (FRET) based assay in which competitive allele-specific PCR enables bi-allelic discrimination of known SNPs at specific loci.

Aliquots of patient DNA were prepared for KASP genotyping in the Interstitial Lung Disease Unit laboratory at the Royal Brompton Hospital. The total DNA concentrations in each patient sample was quantified using a NanoDrop™ spectrophotometer (Thermo Scientific) and diluted to aliquots of 10ng/uL with nuclease free water. Sample volumes of 10uL were loaded onto 96 well plates which included two no template controls (NTCs) per plate. The plates were sealed and stored at -25°C until the time of delivery on dry ice to the service laboratory for processing.

The KASP genotyping system is comprised of two components:

1) the SNP specific KASP Assay mix - contains three assay-specific non-labelled oligonucleotides: two competitive, allele-specific forward primers and one common reverse primer. The allele-specific forward primers each have a unique tail sequence that corresponds with a universal FRET (fluorescence resonant energy transfer) cassette present in the KASP Master mix; one labelled with FAM™ dye and the other with HEX™ dye.

2) the KASP Master mix - contains the two universal FRET cassettes (FAM and HEX), ROX™ passive reference dye, Taq polymerase, free nucleotides and MgCl₂ in an optimised buffer solution.

The SNP-specific KASP Assay mix and the universal KASP Master mix are added to DNA samples, a thermal cycling reaction is then performed, followed by an end-point fluorescent read. During thermal cycling, the relevant allele-specific primer binds to the template and elongates, thus attaching the tail sequence to the newly synthesised strand. The complement of the allele specific tail sequence is then generated during subsequent rounds of PCR, enabling the FRET cassette to bind to the DNA. The FRET cassette is no longer quenched and emits fluorescence. Bi-allelic discrimination is achieved through the competitive binding of the two allele-specific forward primers. If
the genotype at a given SNP is homozygous, only one of the two possible fluorescent signals will be generated. If the genotype is heterozygous, a mixed fluorescent signal will be generated.

Figure 11. Mechanism of KASP genotyping

Once the fluorescence signals have been measured, the FAM fluorescence value and the HEX fluorescence value from each DNA sample are represented as an independent data point on a cluster plot. All samples that have the same genotype will
have generated similar levels of fluorescence and will therefore cluster together on the plot. Based on the relative position of these clusters, it is possible to determine the genotype of all of data points. A typical genotyping cluster plot is shown below.

![Genotyping Cluster Plot](image)

**Figure 12. Example of a typical genotyping cluster plot.**

### 3.3 Biochemical Methods

#### 3.3.1 Methylarginine Quantification

Liquid chromatography-mass spectrometry (LC-MS/MS) was used for the detection and quantification of methylarginines in cell lysates. The combination of high-pressure liquid-chromatography (to allow components of a sample to be separated based upon their binding affinity to a silica based column) and sequential mass spectrometry (to allow separation of ions according to their mass-to-charge ratio) is an extremely sensitive technique with a high sensitivity.

The Agilent 6400 Series Triple Quadrupole LC-MS/MS System was used for this study. Samples were prepared by methanol protein precipitation with an internal standard, 7-deuterated ADMA (Cambridge Isotope Laboratories) added to each sample to allow correction for methylarginine extraction efficiency. Following a protein precipitation in 1:5 dilution with methanol, the resulting pellet was removed by centrifugation and the supernatant was evaporated to dry in a heat block. The sample was then resuspended in mobile phase (0.1% formic acid) to run on the LC-MS/MS.

In the first stage, the sample was eluted off a hypercarb (Thermo Scientific) chromatography column using a mobile phase consisting of 0.1% formic acid, 1% acetonitrile (increasing to 50% between 5 and 10 minutes) over a total run time of 15
minutes. In the second stage, after elution, samples passed into the Agilent 6400 series triple quadrupole LC-MS/MS where they were vaporised and ionised. The MS parameters for detection were as follows: ADMA, mass-to-charge ratio (m/z): 203.3 to 46.0, collision energy (CE): 12; SDMA, m/z: 203.3 to 70.2, CE: 24; monomethylarginine, m/z: 189.3 to 70.2, CE: 24; arginine, m/z: 175.2 to 60.1, CE: 8; d7-ADMA, m/z: 210.0 to 46.0, CE: 24. The data was collected and analysed using Agilent’s Masshunter Qualitative analysis programme. Chromatograms were acquired and the amount of each methylarginine in the sample was determined by the total ion count within the relevant peak. Concentrations of the methylarginines were determined by running a standard curve within the sample run.

3.3.2 Nitric Oxide Quantification

NO is a highly reactive species and difficult to measure directly; it quickly oxidises to nitrite (NO₂⁻) and nitrate (NO₃⁻). These stable end products of NO metabolism are widely used as surrogates for NO formation and can be detected in several assays.

A chemiluminescence analyser, the Sievers NOA 280i (GE Analytical Instruments) was used to measure the total NOx content (both nitrite and nitrate combined) in small volumes of cell lysates and culture media. NO was re-derived from nitrites and nitrates by reduction in heated vanadium chloride before continuing into the system to react with ozone. The resulting gas-phase chemi-luminescent reaction emits in the red/infrared spectrum and is detected and quantified by a photo-multiplier system. Duplicate samples were prepared by methanol protein precipitation in a 1:5 dilution, the resulting pellet was removed by centrifugation and the supernatant was collected for analysis. The concentration of NOx (uM) was calculated by comparison to a standard curve generated from control solutions containing known amounts of sodium nitrate (0-200uM).

3.4 Human Study Methods

3.4.1 Patient Recruitment and Diagnostic Criteria

DNA samples for the IPF cohort were recruited from consecutive, newly presenting patients with IPF attending the interstitial lung disease (ILD) clinic at the Royal Brompton Hospital, London. The diagnoses were made from the IPF criteria as per the ATS/ERS/JRS/ALAT 2011 guidelines [Raghu 2011]. Caucasian individuals of Northern European descent only were included. The majority of patients with IPF received
combination immunosuppressive treatment and low-dose prednisolone (as was the current best treatment at the time). Patients on novel agents (nintedanib and pirfenidone) or in clinical trials were excluded. Whole human blood was collected; DNA was extracted and stored on the day of sample collection and KASP genotyping was performed as described in the methods above.

An additional cohort of patients with systemic sclerosis (SSc), specifically those with systemic sclerosis associated interstitial lung disease (SSc-ILD), was also genotyped for the exploratory hypothesis outlined in Chapter 9. DNA samples were collected from patients attending clinic at the Royal Brompton Hospital, London. The diagnosis of SSc was made from well-defined criteria [Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee, 1980]. The diagnosis of SSc-ILD was made by the presence of pulmonary fibrosis (PF) on HRCT imaging that had been documented on a radiology report or in the patients’ clinical notes. Caucasian individuals of Northern European descent only were included. Patients with SSc-ILD were treated with intended ‘best management’ with introduction/changes in immunosuppression dictated by progression of disease and side effects.

3.4.2 Clinical Assessment
All patients were assessed as part of routine clinical management upon presentation to the clinic. Pulmonary function tests (PFTs) and HRCT chests were performed at the Royal Brompton Hospital (RBH), a specialist tertiary referral centre in the UK with extensive expertise in performing, reporting and interpreting these tests.

PFTs (expressed as percent predicted) were performed according to standard criteria at the RBH lung function laboratory as per routine clinical practice. These were tests of pulmonary mechanics; forced vital capacity (FVC) and FEV1 (forced expiratory volume in 1 second) and a test of pulmonary diffusing capacity; transfer factor of the lung for carbon monoxide (TLCO).

Additionally the composite physiological index (CPI), a functional index of lung disease severity in patients with IPF and SSc-ILD was calculated; CPI=91.0−(0.65×TLCO% predicted)−(0.53×FVC% predicted)+(0.34×FEV1% predicted) [Wells 2003].
3.4.3 Data Collection and Clinical End-Points

PFT and HRCT data was collected for all patients up to the study cut off of September 2015. Any PFTs or HRCTs performed after this date were not used in the analysis. Baseline demographics (age, gender, smoking history and lung function tests on presentation to the clinic) were also collected. Longitudinal lung function and mortality data was collected for all patients according to the clinical endpoint criteria as discussed below.

1) Time to first decline of FVC

FVC is also used as a surrogate marker of mortality in clinical trials. Therapies for IPF are based on FVC and the current NICE recommendation is that all patients with FVC between 50–80% predicted should be considered for therapy. Of note, those with higher than 80% FVC at baseline therefore experience considerable FVC decline before being eligible for treatment using this definition. Significant functional deterioration was quantified using serial PFT indices commencing from first presentation to the Royal Brompton Hospital.

Time to first decline of FVC (TtD of FVC) was defined as the number of months taken for the relative percentage change of ≥10% in FVC on one subsequent time point from baseline. A relative decline of 10% in FVC establishes a high likelihood of true decline and thus denotes true disease progression, as opposed to confounding by any possible measurement variation [Wells 2013]

In IPF which is a relentlessly progressive disease, one subsequent time point from baseline was the minimum requirement for follow-up, and time to first observed decline was used as described. In SSc-ILD, to allow for possible response to treatment or spontaneous disease fluctuations, at least two follow-up time points from baseline were required and thus for SSc-ILD, time to irreversible decline was used [Stock 2013]. This was defined as the number of months taken for the relative percentage change of ≥10% in FVC, observed on at least two consecutive occasions. When follow-up ended with functional decline on a single occasion, this was accepted as a significant decline provided there was documented clinical or radiological evidence of worsening.

2) Length of survival from presentation

Mortality is the most clinically relevant and important end-point in IPF and this was available for the entire historical cohort of patients due to the poor median survival of IPF. In most clinical trials this is often not universally available.
Length of survival from presentation was defined as the number of months survived by the patient following presentation to the RBH ILD clinic. Survival data was not universally available for the SSc cohort, partly due to the burden of morbidity rather than mortality carried by this disease.

3) Time to first event
Time to the first event (also known as length of progression free survival) is traditionally defined as the length of time during and after the treatment of a disease that the patient lives with the disease but does not deteriorate. In this context, a composite end-point which combined time to first decline and survival from presentation was used. This enabled patients with severe disease who survived or patients with severe disease that prevented further PFT measurement, to be included.

3.4.4 Statistical Methods
Genotype frequencies were tested for deviation from Hardy–Weinberg equilibrium (HWE), by direct counting and the χ2 statistic. The parametric one-way ANOVA was used in Graph Pad Prism to test for differences in median lung function measures. Survival analyses and stepwise multivariate Cox proportional hazards analyses were carried out for each clinical end point as follows:

1) Survival Analyses
Survival curves (Kaplan-Meier) were created for each clinical end point using Graph Pad Prism. Censored data (i.e. those who were lost to follow-up or those where the clinical end point had not been reached) were included. This was followed by log rank testing to determine whether there was a significant difference between curves and measurement of hazard ratios (with 95% confidence intervals).

2) Cox Proportional Hazards Model
Next, a stepwise multivariate Cox proportional hazards analysis (Cox regression) was carried out for each clinical end point using Stata as follows:

In the first step, a univariate cox analysis was performed on each baseline variable. This was followed by a proportional hazards test to detect whether there was violation of the proportional hazards assumption. The proportional hazards assumption assumes that the effects of different variables on survival are constant over time, high p value = no evidence of departure from the assumption of proportional hazards, thereby the hazard was consistent over time.
In the second step, any potential confounding variables/disease severity at baseline were adjusted for. Multivariate Cox analysis was performed with only the significant univariate variables with p<0.05 remaining in the multivariate model. The multivariate analysis was then followed by a further proportional hazards test to determine if the assumption of proportional hazards had been met.

The pros and cons of performing a Bonferroni correction (BC) are widely known, it is known to be conservative and a potential problem of type 2 errors may arise [Bland 1995, Perneger 1998]. Furthermore, it may not be correct to perform a BC if SNPs are in linkage disequilibrium (LD) as they are not truly independent samples [Johnson 2010]. In view of the exploratory nature of this study and the a priori hypothesis, a BC was not performed.
CHAPTER 4: ESTABLISHMENT AND CHARACTERISATION OF AN IN VITRO MODEL OF EMT

4.1 Development of an Experimental Model of EMT

4.1.1 Introduction
EMT plays a key role in the normal homeostatic response to tissue injury and is activated during conditions of repair and remodelling. IPF is a disease that arises from a runaway repair and remodeling response triggered by alveolar epithelial cell injury. TGF-β is considered to be the prototypical cytokine for both the induction of EMT and driving the fibrotic process in IPF [Willis 2007]. As discussed, NO plays a key role in modulating key phases of the normal wound healing response in the context of alveolar epithelial injury and there is evidence of its overproduction in IPF. EMT is therefore an attractive biological pathway in which to interrogate the aberrant wound healing process of IPF, and in particular, the role of the NO-ADMA-DDAH pathway.

Currently, there are no appropriate experimental models for IPF i) due to the unknown specific aetiology of the disease, and ii) the traditional models are predominantly inflammatory in nature [Selman 2002]. The widely used bleomycin-induced model of experimental lung fibrosis is an inflammatory reaction to a sudden, acute lung injury. This is in contrast to IPF where the insidious, repetitive nature of alveolar epithelial cell injury and the absence of any ongoing detectable inflammation is characteristic [Thannickal 2004]. EMT is a conserved, fibrotic pathway that can be triggered by TGF-β, a key cytokine in IPF, and thus development of an in vitro model of EMT represents a useful experimental tool.

4.1.2 Protocols
Human type 2 alveolar epithelial cells (A549 cell line derived from human adenocarcinoma of the lung) were serum starved for 24 hours before stimulation with TGF-β1 at 5ng/ml (R&D Systems) for 48 hours. A profile indicating transformation from an epithelial to mesenchymal phenotype (E-cadherin, α-SMA, type 1 collagen, MMP-1,-2 and -7 and LOXL2), was assessed by qPCR, Western blotting and phase contrast light microscopy. Pilot experiments determined the optimal concentration and duration of TGF-β treatments necessary to produce the EMT profile of changes (pilot data not shown).
To investigate whether TGF-β treatment was inducing the Smad3 pathway, the proportion of phosphorylated Smad3 (pSmad3) of total Smad3 was determined by Western blotting. In addition, cells were pre-treated with 10uM SIS3, a novel specific inhibitor of TGF-β1/ALK-5 phosphorylation of Smad3 (Sigma Aldrich) for an hour after the serum starving step and prior to 5ng/ml TGF-β treatment, based on established methods in A549 cells [Zhang 2011(b)].

![EMT Protocol Diagram]

**Figure 13. Outline of EMT protocol used for in vitro experiments with A549 cell line**

Two technical replicates were cultured on each 6 well plate for each treatment. PCRs were performed on each technical replicate and were amplified in duplicate; the means of the duplicates, and then the subsequent mean value of the technical replicates, constituted one n number. Western blotting was performed on each technical replicate and the mean value used constituted one n number. Statistical analysis was performed using one-way analysis of variance (ANOVA) tests to analyse whether there were significant differences between the means of treatment groups compared to the control group.

### 4.1.3 Results

![E-cadherin Gene Expression Graph]

**Figure 14. E-cadherin gene and protein expression in TGF-β mediated EMT**

Significant decrease in E-cadherin gene expression ($p<0.0001$, $n=6$). On Western blotting with band intensities corrected for tubulin (50kDa), E-cadherin protein levels significantly decreased ($p<0.0001$, $n=6$).
Figure 15. α-SMA gene and protein expression in TGF-β mediated EMT

Increase in α-SMA gene expression (p=ns, n=6). On Western blotting with band intensities corrected for tubulin (50kDa), α-SMA protein levels significantly increased (p<0.018, n=5).

Figure 16. Type 1 Collagen expression in TGF-β mediated EMT

Significant increase in type 1 collagen (COL1-A1) gene expression (p<0.019, n=6). On Western blotting, no bands were demonstrated at 48 hours at concentrations of 0ng, 5ng and 10ng/ml TGF-β (p=ns, n=5). No collagen bands visible at 3, 4, 5, or 6 days (data not shown). Tubulin loading control at 50kDa.
Figure 17. MMP-1, MMP-2, MMP-7 and LOXL2 gene expression in TGF-β mediated EMT

Significant increase in expression of MMP-1 and MMP-2 (p<0.002 and p<0.006, n=3), no significant change in MMP-7 levels (p=ns, n=3). LOXL2 expression was significantly elevated with TGF-β treatment (p<0.0001, n=3).

Figure 18. TGF-β mediated Smad3 phosphorylation

Significant difference between percentage of pSmad3 of total Smad3 for TGF-β and TGF-β+SIS3 treatment groups compared to controls and between treatments (p<0.0001, n=6).
4.1.4 Discussion

Summary of Key Findings

i. A549 cells demonstrate plasticity in response to TGF-β treatment

ii. E-cadherin levels are reduced by TGF-β treatment

iii. α-SMA and type 1 collagen levels are increased by TGF-β treatment

iv. MMP-1, MMP-2 and LOXL2 gene expression is increased by TGF-β treatment

v. TGF-β induces Smad3 phosphorylation in the A549 cells

vi. TGF-β induces transformation from an epithelial to a mesenchymal phenotype

The results demonstrate the overall plasticity of A549 cells in response to TGF-β; the significant loss of E-cadherin and the gain of α-SMA and type 1 collagen in a concentration dependent manner was highly suggestive that the process of EMT was occurring. These findings corroborated previous findings from recent publications where TGF-β induced A549 cells to undergo EMT in a time- and concentration-dependent manner [Kasai 2005, Kim 2007]. Furthermore, the TGF-β induced expression of phosphorylated Smad3 was attenuated with SIS3 treatment - the significant increase in the proportion of phosphorylated Smad3 expression compared to total Smad3 expression indicated the TGF-β/Smad3 pathway (a pathway known to be key to both EMT) was being activated in the transition to a mesenchymal phenotype, resulting in induction of TGF-β responsive target genes such as α-SMA and type 1 collagen.

The gradual loss of E-cadherin is a universal feature of EMT, regardless of the initiating stimulus [Hay 1995]. The results demonstrated a significant decrease in E-cadherin expression at both transcriptional and translation levels. Gene expression of α-SMA demonstrated an increase, but importantly, the protein which was not present at baseline, was induced with TGF-β treatment. Expression of α-SMA is characteristic for activated myofibroblasts, the key effectors in laying down excessive ECM in IPF [Bartis 2013].

In view of the lack of consensus on marker specificity to indicate transition to a mesenchymal phenotype, MMP-1,-2 and -7 and LOXL2 were also selected. The elevation of MMP-1 gene expression is particularly interesting as current literature suggests an important biological role in IPF. MMP-1 mRNA is significantly upregulated in whole lung tissue from IPF patients compared with normal control lung [Konishi 2009] and expression of MMP-1 is mainly in the epithelium, not in the interstitial
compartment [Pardo 2008]. Similarly, MMP-2 (also known as gelatinase A) is widely expressed in fibrotic lungs, especially in areas of hyperplastic epithelial cells and by mesenchymal cells in the fibroblast foci and is implicated in basement membrane disruption [Selman 2000]. The elevation of MMP-1 and MMP-2 support the transition towards a fibrotic/mesenchymal phenotype.

Contrastingly, levels of MMP-7 (also known as matrilysin) were not elevated. In addition to MMP-1, MMP-7 is one of the most consistently elevated genes in fibrotic lungs and its levels relate to the severity of lung function impairment in IPF [Dancer 2011]. In IPF lungs, MMP-7 is mainly expressed by the abnormal alveolar epithelium [Zuo 2002]. The EMT model is not fully representative of the diverse pathological processes in IPF, and as such, the pleotropic biological interactions and roles of MMP-7 [Pardo 2008] will not have been replicated in a single cell line model.

EMT is thought to be a gradual process and an epithelial cell undergoing EMT will express a differing set of marker genes, levels of protein and intermediate cellular phenotypes at various points during the transition. On phase contrast light microscopy, the morphology of the cells was visibly changing towards a more fibroblastic appearance during cell culture treatments (images not shown). Although a several fold increase of type 1 collagen gene expression was observed, this was not replicated on Western blotting despite further time-concentration experiments up to 6 days in duration (data not shown). There are many different species of collagen, but type I and type 3 predominate within both healthy and fibrotic lungs. Collagen is initially secreted as soluble fibrillar precursors that self-associate to form a triple helix fibril – this renders the molecule insoluble and resistant to proteolytic attack by most enzymes except the metalloproteinases [Dancer 2009]. The nature of this insoluble molecule may have contributed to the Western blotting results. As previously discussed, although NO is critical for wound collagen deposition, clear-cut enhancement of collagen synthesis or gene expression has not been found and may be due to post-translational collagen regulation by NO [Witte 2002]. In the context of the aims of this experiment, the E-cadherin and α-SMA results provided evidence of the transition to a mesenchymal phenotype and this was further supported by elevation in LOXL2 gene transcription. The LOXL proteins facilitate the cross-linking of type 1 collagen molecules, and LOXL2, which is synthesised and secreted by fibroblasts, has been shown to be particularly important in driving collagen accumulation and deposition of the ECM [Barry-Hamilton 2010], and its induction suggests the process of mesenchymal transition had been activated.
An important point to note is that although the A549 cell line is a very commonly used human alveolar epithelial cell line, it is derived from a human adenocarcinoma of the lung and is clearly not entirely representative of the in vivo phenotype; the results must be interpreted accordingly with caution. However, over the course of these experiments, the overall indication was that the A549 cell line model was a consistent and reproducible in vitro model of EMT, and that TGF-β induced a loss of the epithelial marker E-cadherin and a gain of several mesenchymal markers in a concentration-dependent manner.

4.2 Characterisation of the NO-ADMA-DDAH Pathway in EMT

4.2.1 Introduction

After establishing the in vitro experimental model of EMT in the A549 cell line, the NO-ADMA-DDAH pathway was characterised. The role of the NO-ADMA-DDAH axis in TGF-β mediated EMT is unknown.

4.2.2 Protocols

From initial pilot experiments, the optimum conditions of TGF-β treatment to induce EMT (2.5-5ng TGF-β per ml of medium for 48 hours treatment duration) were selected to investigate levels of DDAH and NOS isoform expression, levels of intra- and extracellular NOX and the methylarginines in A549 cells.

Following TGF-β treatment, the transition from an epithelial to mesenchymal phenotype was checked by assessing E-cadherin and α-SMA expression. In view of the low basal gene expression of the NOS enzymes in A549 cells; a positive iNOS control was generated from the A549 cell line (confluent cells were cultured for 20 hours in serum-free medium with recombinant human TNF-α, IL-1B, and IFN-γ, each at 5 ng/ml concentration); mouse brain was used as a positive control for generating standard curves during qPCR for eNOS and nNOS. Active, full length human proteins for NOS enzymes were used as known proteins/controls in Western blotting. Statistical analysis of data was performed using one-way ANOVA tests to analyse whether there were any significant differences between the means of treatment groups compared to the control group.
4.2.3 Results

**Figure 19. DDAH1 expression in TGF-β mediated EMT**

No significant change in DDAH1 gene expression (p=ns, n=6). On Western blotting with band intensities corrected for tubulin (50kDa); there was no significant change in DDAH1 protein (p=ns, n=6).

**Figure 20. DDAH2 expression in TGF-β mediated EMT**

DDAH2 gene expression significantly increased (p=0.006, n=6). On Western blotting with band intensities corrected for tubulin (50kDa), DDAH2 protein levels demonstrated a multiple fold increase from baseline (p<0.004, n=6).

**Figure 21. Methylarginine levels in TGF-β mediated EMT**

Significant reduction in intracellular levels of both ADMA (n=4, p=0.016) and L-NMMA (n=5, p=0.016) but not SDMA (p=ns, n=5). No significant difference in extracellular levels of methylarginines (data not shown).
Figure 22. NOS gene expression in TGF-β mediated EMT

iNOS gene expression significantly increased several fold (p<0.002, n=5), significant decreases in nNOS (p<0.002, n=3) and eNOS (p<0.002, n=3).

Figure 23. iNOS protein expression in TGF-β mediated EMT

On Western blotting with band intensities corrected for tubulin (50kDa); a doublet band was consistently seen at the molecular mass of iNOS with a significant fold increase in protein levels (p<0.02, n=5). No induction of protein was seen with eNOS and nNOS (not shown).

Figure 24. NOX levels in TGF-β mediated EMT

Concentration of stable end products of NO metabolism; no significant difference was detectable in extra- or intracellular levels of nitrites/nitrates following EMT (p=ns, n=5).
4.2.4 Discussion

**Summary of Key Findings**

i. DDAH1 levels are unchanged during TGF-β mediated EMT
ii. DDAH2 is induced during TGF-β mediated EMT
iii. Asymmetric MAs are reduced during TGF-β mediated EMT
iv. iNOS is induced during TGF-β mediated EMT
v. NOX levels are unchanged during TGF-β mediated EMT

Investigation of the NO-ADMA-DDAH axis in the experimental model of EMT has produced several observations. A difference between the DDAH isoforms is evident. At baseline, DDAH1 appears to be the more abundant isoform in the type 2 AECs. Following TGF-β treatment, DDAH1 appears to increase slightly from its baseline level in the cell line. In contrast, levels of DDAH2 are minimal at baseline in the type 2 AECs but both DDAH2 gene and protein expression are increased several fold in a concentration-dependent manner with TGF-β treatment. These findings are consistent with the human data which demonstrates that DDAH1 is the more abundant DDAH isoform expressed in both normal human lungs and IPF lungs at baseline; whereas DDAH2 expression is induced several-fold and co-localises with iNOS in the IPF lung, specifically in type 2 AECs in the alveolar epithelium [Pullamsetti 2011]. These findings support that DDAH2 appears broadly involved in type 2 AEC growth and that its induction is mediated by the key profibrotic cytokine that drives EMT. Furthermore, it appeared that DDAH2 was a TGF-β responsive gene; expression of DDAH2 was attenuated with TGF-β + SIS3 treatment compared to TGF-β alone indicating that inhibiting phosphorylation of Smad3 also attenuated DDAH2 expression (data not shown, p<0.0001, n=6). It is interesting that both DDAH isoforms are expressed in type 2 AECs however, and previous work has shown that pharmacological inhibition of DDAH1 with L-291 and DDAH2-selective siRNA inhibition of DDAH2 both attenuated serum-induced murine type 2 AEC proliferation and induced apoptosis in vitro [Pullamsetti 2011], indicating a role for both DDAH isoforms in these processes. Protein-protein interactions between the DDAH isoforms and is unclear at present and together they both may play an important role in type 2 AECs.

A significant reduction was observed in the intracellular levels of both ADMA and L-NMMA but not SDMA. SDMA is not hydrolysed by DDAHs and is mainly cleared by renal excretion and deamination by AGXT2. DDAH acts to prevent the accumulation
of ADMA. DDAH1 appears to be the dominant isoform responsible for metabolising ADMA (and L-NMMA) based on in vivo findings and genetic association studies where DDAH1 SNP variants associate with ADMA levels [Leiper 2007, Caplin 2010, Lind 2013], however, a direct comparison of the contribution of each isoform to overall activity is unknown. The reduction in the level of the asymmetric MAs may be due several reasons; the slight increase in DDAH1 level from baseline; the induction of DDAH2; or the combined effects of both on hydrolysis of ADMA.

An interesting observation is that both DDAH2 and iNOS were induced together. The presence of a reactive cysteine residue (Cys-249) in the active site of DDAH is directly regulated by S-nitrosylation of this residue by NO. In cultured endothelial cells, human DDAH2 was reversibly S-nitrosylated following cytokine induced expression of iNOS [Lieper 2002]. The implication of these findings is that an increase in endogenous NO will increase S-nitrosylation and diminish DDAH activity, and this could lead to accumulation of ADMA and inhibition of NOS. However, ADMA did not appear to accumulate with DDAH2 induction in this model. Although regulation of DDAH2 activity by endogenous NO is known, it is not known if DDAH2 induction is also regulated by endogenous NO.

The observation that iNOS was induced in the experimental EMT model is interesting for several reasons. The induction of iNOS but not the other NOS isoforms is supportive of the literature where upregulation of iNOS is linked to pulmonary fibrosis. More specifically, iNOS is upregulated within the type 2 AECS both in experimental models of IPF and in the IPF lung [Saleh 1997, Pullamsetti 2011]. The complex interactions between iNOS, NO and endogenous TGF-β have been discussed i) alveolar epithelial injury results in upregulation of iNOS and excessive NO production which in turn activates latent TGF-β which also upregulates iNOS (figure 10), and ii) NO upregulates iNOS via NF-κB, taken together, both of these feedback loops may have contributed to the observed upregulation of iNOS and the transition towards the fibrotic phenotype. The upregulation of iNOS may also be considered in the context of the wound healing model, iNOS is upregulated early in response to cytokines and growth factors released in response to injury and expression may peak as early as 48 hours [Schwentker 2003].

However, although NO is known to activate latent TGF-β and promote iNOS in early wound injury, TGF-β has been demonstrated to reduce iNOS expression at the later stages of wound healing in fibroblasts during the switch from proliferative to synthetic...
function [Romanska 2002]. Exogenous TGF-β has in fact been reported to repress iNOS expression through a mechanism that is poorly understood, possibly by destabilising iNOS mRNA [Vodovotz 1997]. Although this mechanism was not reported in alveolar epithelial cells, if exogenous TGF-β was indeed acting to suppress iNOS transcription, the observed elevation of iNOS suggest this effect may have been overcome by other more dominant mechanisms of iNOS upregulation as discussed. Interestingly, the upregulation of iNOS was less at the higher concentration of TGF-β treatment of 5ng/ml as significantly higher levels of iNOS gene and protein expression were only apparent at 2.5ng/ml.

In view of the observed reduction of ADMA and L-NMMA levels, an increase in NO production would also be expected due to the loss of endogenous NOS inhibition by the asymmetric MAs. Similarly, in context of the upregulation of iNOS, NO levels would be expected to be greatly increased. However, this was not observed on measurement of NOX metabolites. This may be due a limitations of measurement, specifically the ability to detect the size of effect that may have been occurring in the cell line. Another possibility is that of a time point sampling problem, upregulation of iNOS in human lung fibroblasts in response to inflammatory cytokine stimulation has been demonstrated within 3hrs [Romanska 2002], and accordingly, detectable NO release may have also occurred earlier. In terms of the wound healing model of IPF, it stands to reason that if NO is one of the key promoters of lung fibrosis, then activation of iNOS gene should also occur very early in response to initial inflammatory injury in IPF.

The NO-ADMA-DDAH axis has not been investigated before in the context of EMT. DDAH1 is abundant at baseline and DDAH2 is induced – it is not clear why or how. Overall these results suggest that the DDAH isoforms may play an important regulatory role in type 2 AEC biology and therefore possibly in alveolar epithelial cell dysfunction, a key trigger for the fibrotic phenotype that becomes IPF.
CHAPTER 5: FUNCTIONAL SIGNIFICANCE OF THE NO-ADMA-DDAH PATHWAY IN EMT

5.1 Investigation of the effects of exogenous ADMA on EMT

5.1.1 Introduction
The results of Chapter 4 indicated that alterations in the NO-ADMA-DDAH axis were observed in the experimental model of EMT; DDAH1 was the more abundant isoform but DDAH2 was induced several fold, asymmetric MA levels were reduced and iNOS was induced. The observed induction of DDAH2 and the reduction in ADMA/L-NMMA levels is interesting; it is unknown whether DDAH2 induction was occurring in response to TGF-β during transition to a mesenchymal phenotype, or whether it was to prevent the accumulation of ADMA in order to restore the balance of the NO-ADMA-DDAH axis. Taken together with the induction in iNOS, overall, a pro-NO environment was observed and appeared to be occurring with EMT. It is unknown whether this environment was a feature or whether it was driving EMT. In order to determine the effects of pan-NOS inhibition on the TGF-β mediated transition from an epithelial to mesenchymal phenotype, the following experiments were conducted with exogenous ADMA.

5.1.2 Protocols
The in vitro experimental model of EMT was used for the following experiments. The human type 2 AECs were serum starved for 24hrs before stimulation with the following treatments for to 48hrs: Control (C); 5ng/ml TGF-β treatment alone (T); co-treatment with TGF-β 5ng/ml and 100uM/ml exogenous ADMA (T+A); and 100uM/ml exogenous ADMA treatment alone (A). The IC₅₀ of ADMA is dependent on the prevailing L-arginine concentration [Vallance 2004]. The L-arginine:ADMA ratio in vivo, is approximately in the region of 50:1. The L-arginine concentration in the culture medium was 2mM. Therefore the addition of 100uM ADMA resulted in a L-arginine:ADMA ratio of 20:1, close to physiological levels.

A profile indicating transformation to mesenchymal phenotype (E-Cadherin, α-SMA), and expression and protein levels of DDAH isoforms and iNOS was assessed. All band intensities for the Western blotting results below are corrected with tubulin (50kDa) as the housekeeper protein. Statistical analysis of data was performed using one-way analysis of variance (ANOVA) tests to analyse whether there were significant
differences between the means of treatment groups compared to the control group and additionally between the means of the TGF-β compared to TGF-β+ADMA treatment groups.

5.1.3 Results

Figure 25. Effect of ADMA treatment on E-cadherin protein expression in EMT model
Loss of E-cadherin with both TGF-β treatments and no difference between TGF-β alone and TGF-β+ADMA (p=ns). Significant increase with ADMA compared to control (p<0.0005, n=3). Tubulin loading control at 50kDa.

Figure 26. Effect of ADMA treatment on α-SMA protein expression in EMT model
Increase in α-SMA with TGF-β compared to control (p<0.003, n=3). No significant difference between TGF-β alone and TGF-β+ADMA (p=ns). Tubulin loading control at 50kDa.
Figure 27. Effect of ADMA treatment DDAH1 protein expression in EMT model
No difference in DDAH1 protein expression between treatment groups (p=ns, n=3). Tubulin loading control at 50kDa.

Figure 28. Effect of ADMA treatment on DDAH2 protein expression in EMT model
Increase in DDAH2 with TGF-β and TGF-β+ADMA compared to control (p<0.0024, n=6). No significant difference between TGF-β alone and TGF-β+ADMA (p=ns, n=6). Tubulin loading control at 50kDa.

Figure 29. Effect of ADMA treatment on iNOS protein expression in EMT model
Increase in iNOS with both TGF-β treatments compared to control, significant increase with TGF-β+ADMA treatment (p<0.0083, n=3). No significant difference between TGF-β alone and TGF-β+ADMA (p=ns, n=3). Tubulin loading control at 50kDa.

5.1.4 Discussion

Summary of Key Findings

i. E-cadherin levels are reduced by TGF-β and ADMA co-treatment
ii. E-cadherin levels are increased by exogenous ADMA treatment
iii. α-SMA levels are increased by TGF-β and ADMA co-treatment
iv. DDAH1 levels are unchanged by TGF-β and ADMA co-treatment
v. DDAH2 levels are increased by TGF-β and ADMA co-treatment
vi. iNOS levels are increased by TGF-β and ADMA co-treatment
vii. Exogenous ADMA does attenuate the induction of DDAH2
viii. Exogenous ADMA does not attenuate the TGF-β induced transformation from an epithelial to a mesenchymal phenotype

The significant loss of E-cadherin and the gain of α-SMA with TGF-β were consistent with previous results and indicated that a transition from an epithelial to mesenchymal phenotype was occurring in the TGF-β mediated EMT model. In the presence of TGF-β, the multiple fold induction of DDAH2 and iNOS but no change in the levels of DDAH1 was reproduced. Very similar results were demonstrated with TGF-β and exogenous ADMA co-treatment. This indicated that exogenous ADMA treatment did not attenuate the TGF-β induced transformation from an epithelial to a mesenchymal phenotype.

Another interesting observation is that expression of E-cadherin was increased by exogenous ADMA alone. This suggests that NO promotes loss of E-cadherin and that the effect was overcome by TGF-β treatment. It is unknown whether E-cadherin is a NO responsive gene and it appears from this data that removal of endogenous NO appears to promote its expression.

IPF appears to be a disease of NO overproduction and it was reasoned that altering the pro-NO environment with exogenous ADMA to achieve pan-NOS inhibition may be protective against a transition to the fibrotic phenotype but the findings did not support this. This may be because the effects of ADMA are not as straightforward as achieving pan-NOS inhibition alone. There are several reasons which may account for the results
observed with exogenous ADMA which broadly fall into two aspects; disruption of the L-arginine/ADMA balance and iNOS function.

L-arginine is a common substrate shared by NOS and arginase. Induction of arginase 1 and -2 has been demonstrated in the bleomycin model [Endo 2003]. L-ornithine, a precursor of L-proline (and thus collagen) is produced by the arginase pathway, and the supply of proline can be a crucial factor in the process of lung fibrosis. A recent study demonstrated that direct infusion of ADMA resulted in elevated collagen deposition in mouse lungs and enhanced arginase activity [Wells 2009]. This suggests that the effect of elevated levels of ADMA is to divert L-arginine to the arginase pathway and lead to increased collagen deposition. Although the cell line model is not fully representative of an in vivo mechanism, it is of particular relevance that TGF-β has been shown to potently upregulate arginase 1 and -2 expression in both primary and cell line murine fibroblasts [Kitowska 2007]. A similar event may have occurred in the EMT model but TGF-β induced arginase expression was not examined. It remains possible that exogenous ADMA treatment may have altered the L-arginine/ADMA balance and diverted L-arginine to this pathway.

In addition to producing NO, NOS can catalyse the formation of O2− in the presence of low L-arginine levels [Pou 1992]. When this occurs, the oxidation of L-arginine to NO is not complete and the activity of NOS is “uncoupled” [Vasquez-Vivar 1998]. Thus the generation of both O2− and NO by NOS can lead to ONOO− formation in L-arginine depleted cells. As ADMA competitively inhibits and displaces L-arginine from NOS, this has led to the hypothesis that ADMA may uncouple NOS and switch NO enzyme activity from the production of NO to O2− and subsequent ONOO−. Exogenous ADMA treatment (in the range of 10μM to 250μM) was demonstrated to show a dose-dependent increase the generation of O2− and ONOO− in a cytokine stimulated mouse epithelial cell line (LA-4) [Wells 2007]. RNS and ROS are important modulators of TGF-β in airway epithelial cells and fibroblasts in IPF. A recent study in the A549 cell line demonstrated that externally applied oxidative stress using hydrogen peroxide (H2O2) drove EMT via an increase in intracellular ROS generation and an increase TGF-β1 expression and activity [Gorowiec 2012]. Although not directly measured, it is possible that ADMA generated RNS and ROS, produced via NOS uncoupling, may be contributing to the EMT process.

ADMA derived oxidative stress is also of particular relevance in the context of the elevated DDAH2. The expression and function of DDAH2 is inhibited by peroxynitrite.
This potentially creates a feed-forward loop where oxidative stress generated from ADMA-induced uncoupling of NOS, inhibits DDAH2 mediated degradation of ADMA and thus exaggerates any increase in ADMA, further contributing to the EMT process. This could also be an explanation for lack of difference between the levels of induction of DDAH2 with TGF-β compared to TGF-β and ADMA co-treatment and the observation that EMT occurred with both. Results from chapter 4.2 demonstrated that ADMA levels decreased in association with DDAH2 induction suggesting that DDAH2 may have been partially active. If DDAH2 was indeed being induced in EMT, its activity in the presence of exogenous ADMA would have been negatively affected by oxidative stress. If this was the case then EMT occurred regardless of DDAH2 induction/activity and suggests that although induced by TGF-β, DDAH2 was not having a significant effect on driving EMT transition.

It is interesting that iNOS expression with TGF-β and ADMA co-treatment was enhanced compared to TGF-β treatment alone. It has been demonstrated that ADMA increases iNOS expression via increased NF-κB binding activity in the LA-4 lung epithelial cell line (where NF-κB binding activity was assessed by a transcription factor binding assay) [Klein 2010]. In this proposed mechanism for ADMA-enhanced iNOS expression, elevated ADMA contributes to the formation of superoxide and peroxynitrite via NOS uncoupling which in turn increases TNF-α-mediated NF-κB activation and the subsequent upregulation of iNOS.

There are additional limitations to the previously discussed problems of using a cell line model. A physiological level of ADMA was used and perhaps a supranormal level may have given different results. It is possible that intracellular levels of ADMA may not have been altered by exogenous ADMA treatment. ADMA has been shown to be effectively transported into cells through the inducible L-arginine transporter CAT-2B, suggesting that transport of both L-arginine and ADMA is mediated through the same transporter [Closs 1997]. Mass spectrometry was not performed, this would have given indication of how much ADMA was entering the cell. Interestingly, there are two reports suggesting possible impermeability of the respiratory epithelium to ADMA [Wells 2007, Dowling 1998].

Overall it appears that exogenous ADMA treatment has revealed some interesting observations but has also exposed/allowed discussion of the other confounding effects of ADMA. Treatment with exogenous ADMA in the cell line model was fraught with multiple treatment effects that could have affected the initial goal of the experiments –
to determine the effects of pan-NOS inhibition on the NO-ADMA-DDAH axis. The sledge-hammer approach of pan-NOS inhibition with ADMA needs refining and narrowing down in future experiments.

5.2 Investigation of the effects of selective iNOS inhibition in EMT

5.2.1 Introduction
The elevation of iNOS was a consistent observation from previous experiments in the EMT model. The relevance of its induction is not clear, the EMT model appears to support a pro-NO environment. In view of the confounding problems discussed with exogenous ADMA treatment, namely disruption of the L-arginine/ADMA balance and the possible impact of ADMA generated ROS/RNS on TGF-β, iNOS or DDAH2 transcription, a more selective approach was taken to determine the significance of the elevated iNOS expression in the system. It is not known what the effect of endogenous NO removal will be on the EMT process.

N-(3-(Aminomethyl)benzyl)acetamidine (1400W) is a slow, tight binding and highly selective inhibitor of human iNOS [Garvey 1997]. Furthermore, 1400W inhibits iNOS mediated NO formation without the production of superoxide [Wells 2007]. 1400W (W4262, Sigma Aldrich) was therefore used to determine the effect of selective iNOS inhibition on the TGF-β mediated transition from an epithelial to mesenchymal phenotype.

5.2.2 Protocols
The in vitro experimental model of EMT was used for the following experiments. The human type 2 AECs were serum starved for 24hrs before stimulation with the following treatments for to 48hrs: Control (C); 5ng/ml TGF-β treatment alone (T); co-treatment with TGF-β 5ng/ml and 100uM/ml 1400W (T+14); and 100uM/ml 1400W treatment alone (14).

Of note, 1400W has a binding constant of 2.0 uM and L-Arginine is a competitive inhibitor of 1400W binding with a binding constant of 3.0 uM [Garvey 1997]. The L-arginine concentration in the culture media for experiments was in the region of 2.4 uM and treatments with 1400W were at a concentration of 100uM/ml.
A profile indicating transformation to mesenchymal phenotype (E-Cadherin, α-SMA), and expression and protein levels of DDAH isoforms and iNOS was assessed. All band intensities for the Western blotting results below are corrected with tubulin (50kDa) as the housekeeper protein. Statistical analysis of data was performed using one-way ANOVA tests to analyse whether there were significant differences between the means of treatment groups compared to the control group and additionally between the means of the TGF-β compared to TGF-β+1400W treatment groups.

5.2.3 Results

**Figure 30. Effect of selective iNOS inhibition on E-cadherin protein expression in EMT**

Loss of E-cadherin with all treatments (p<0.0002, n=3). No difference between TGF-β alone and TGF-β+1400W (p=ns). Tubulin loading control at 50kDa.

**Figure 31. Effect of selective iNOS inhibition on α-SMA protein expression in EMT**

Increase in α-SMA with TGF-β compared to control (p<0.009, n=4). No significant difference between TGF-β alone and TGF-β+1400W (p=ns). Tubulin loading control at 50kDa.
Figure 32. Effect of selective iNOS inhibition on DDAH1 protein expression in EMT

No difference in DDAH1 protein expression between treatment groups (p=ns, n=3). Tubulin loading control at 50kDa.

Figure 33. Effect of selective iNOS inhibition on DDAH2 protein expression in EMT

Increase in DDAH2 with TGF-β compared to control (p<0.01, n=5). No significant difference between TGF-β alone and TGF-β+1400W (p=ns). Tubulin loading control at 50kDa.

Figure 34. Effect of selective iNOS inhibition on iNOS protein expression in EMT
Increase in iNOS with both TGF-β treatments compared to control, significant increase with TGF-β+1400W treatment (p<0.01, n=5). No significant difference between TGF-β alone and TGF-β+1400W (p=ns). Tubulin loading control at 50kDa.

5.2.4 Discussion

Summary of Key Findings

i. E-cadherin levels are reduced by TGF-β and 1400W co-treatment
ii. α-SMA levels are increased by TGF-β and 1400W co-treatment
iii. DDAH1 levels are unchanged by TGF-β and 1400W co-treatment
iv. DDAH2 levels are decreased by TGF-β and 1400W co-treatment
v. iNOS levels are increased by TGF-β and 1400W co-treatment
vi. Selective iNOS inhibition does not attenuate the TGF-β induced transformation from an epithelial to a mesenchymal phenotype

The significant loss of E-cadherin and the gain of α-SMA with TGF-β were consistent with previous results and indicated that a transition from an epithelial to mesenchymal phenotype was occurring in the TGF-β mediated EMT model. In the presence of TGF-β, the previously observed multiple fold induction of DDAH2 and iNOS but no change in the levels of DDAH1 was reproduced. These findings taken together appear to broadly indicate that selective iNOS inhibition did not attenuate the TGF-β induced transformation from an epithelial to a mesenchymal phenotype.

A previous observation was that expression of E-cadherin was increased by exogenous ADMA treatment. It is unknown whether E-cadherin is a NO responsive gene and it appeared from this data that removal of endogenous NO (although not directly measured in this experiment) may promote its expression. However, this effect was not demonstrated by 1400W. This is suggestive that ADMA may promote E-Cadherin via an iNOS independent mechanism. The mechanism for this is unknown and requires further investigation.

Although the differences in α-SMA and DDHA2 levels between TGF-β and TGF-β+1400W co-treatments did not reach statistical significance, levels of both did appear to be lower with selective iNOS inhibition. These are both TGF-β responsive genes, it is possible that removal of endogenous NO may have impacted the activation of latent TGF-β and affected transcription of these genes (see figure 35).
The finding that DDAH2 levels may be responsive to the levels of endogenous NO is interesting. As discussed in Chapter 4, iNOS derived NO can lead to S-nitrosylation of the active site and inhibit DDAH2 activity in cultured endothelial cells [Leiper 2002]. However, it is not known whether endogenous NO can affect induction of DDAH2 as well as activity. Less DDAH2 induction with the same dose of TGF-β in a low NO environment suggested that NO is needed for maximal TGF-β mediated DDAH2 induction. It is also possible that DDAH2 activity in the presence of 1400W (removal of endogenous NO) would have not have been as impacted as it might otherwise have been in a high NO environment. If this was the case then EMT occurred regardless of DDAH2 induction or activity and suggests that although DDAH2 was being induced, it was not having a functional effect or driving the EMT transition. Although a similar result was not seen with exogenous ADMA, this may have been due to the NOS uncoupling problem as discussed which may have affected the active site. An alternative explanation for this finding is that results have broadly revealed that DDAH2 is a TGF-β responsive gene in this model and it is possible that removal of endogenous NO resulted in less latent TGF-β activation and thus less expression of DDAH2.

As discussed in the context of the normal wound healing process, the relationship between TGF-β, NO and iNOS is complicated. NO can activate latent TGF-β. TGF-β is both reported to inhibit and promote expression of iNOS [Vodovotz 1997, Romanska 2012]. It has been reported that NF-κB, the transcription factor responsible for transcription of iNOS, accumulates in the nucleus in a cell undergoing EMT [Lee 2006]. A summary of these interactions in the type 2 AEC is illustrated in the figure below. In the EMT model, it is proposed that iNOS expression was occurring due to TGF-β mediated EMT and subsequent NF-κB accumulation. Endogenous NO may have provided further drive to activate latent TGF-β thereby forming a feed-forward loop for further EMT/iNOS expression. Selective iNOS inhibition may have decreased the transcription of TGF-β responsive genes in EMT, but it did not affect the expression of iNOS. NO has biphasic role on iNOS induction via NF-κB, it is possible that by inhibiting iNOS activity, the resulting fall in endogenous NO may have further promoted iNOS gene transcription and account for this result.
Although the indication was that selective iNOS inhibition did not attenuate the TGF-β induced transformation from an epithelial to a mesenchymal phenotype, there may however be a downstream significance of elevated iNOS in the transitioned phenotype. Induction of iNOS plays a key role in the wound healing model, and is known to be an early event in fibroblasts and is upregulated in experimental models of IPF. Interestingly, the induction of both DDAH2 and iNOS in response to TGF-β appeared to be a consistent finding, although the functional significance of this is unknown, given the chromosomal position of the DDAH2 gene (proximity to the major histocompatibility complex III locus), DDAH2 may have an important role in regulating iNOS mediated responses in the type 2 AECs.
CHAPTER 6: SELECTING DDAH POLYMORPHISMS FOR THE HUMAN STUDY

6.1 Introduction

A step-wise methodical approach was taken to select SNPs in $DDAH1$ and $DDAH2$ to analyse in the human study. This was to ensure coverage of the whole gene in an organised way and to avoid a selection bias of SNPs.

![Diagram showing the step-wise approach taken to select DDAH SNPs for analysis]

Figure 36. Summary of the step-wise approach taken to select DDAH SNPs for analysis

6.2 Selection of SNPs for genotyping

6.2.1 The International HapMap Project

SNPs in the $DDAH$ genes were searched for on the International HapMap Project Web site [Thorisson 2005, www.hapmap.ncbi.nlm.nih.gov]. The goal of the International HapMap Project was to develop a haplotype map of the human genome, HapMap, which would describe the common patterns of human DNA sequence variation and make this information freely available in the public domain [International HapMap Consortium 2003]. The genotypes of millions of sequence variants, their frequencies and the degree of association between them, were determined in DNA samples from populations with ancestry from parts of Africa, Asia and Europe. Through this huge, international, multi-centre collaboration, millions of SNPs were discovered and many
GWAS studies used this dataset in research for disease association. SNPs were genotyped in three phases from 2005 to 2009 in a variety of populations. The CEU (Utah residents with Northern and Western European ancestry) dataset was used for the purposes of this study as this most population most closely matches the UK’s IPF population.

6.2.2 DDAH1 and DDAH2 SNPs
All SNPs in DDAH1 and DDAH2 were searched for on the HapMap website. In order to capture regulatory regions in the flanking sequences, the intergenic areas both upstream and downstream to the next flanking genes were included; 250 genotyped SNPs were identified in DDAH1 and 9 in DDAH2.

Figure 37. DDAH1 gene and SNPs
Location: 1p22.3 Blue triangles denote db SNPs, red are genotyped SNPs
[Adapted from HapMap]

Figure 38. DDAH2 gene and SNPs
Location: 6p21.3 Blue triangles denote db SNPs, red are genotyped SNPs
[Adapted from HapMap]

6.2.3 Linkage disequilibrium plots for DDAH1 and DDAH2 SNPs
The DDAH SNP data was imported from HapMap into Haploview (software version 4.2) to create graphical representation of the linkage disequilibrium (LD) relationships between SNPs. LD for a pair of SNP alleles is a measure of deviation from random association (i.e., no recombination) [HapMap glossary].
Figure 39. DDAH1 LD Plot

Each SNP is labelled by rs number and Haploview identifier (number in bold). The relative position on the gene is indicated by vertical bars. Each square has $r^2$ value for the pair of SNPs.
Each SNP is labelled by rs number and Haploview identifier (number in bold). The relative position on the gene is indicated by vertical bars. Each square has $r^2$ value for the pair of SNPs.

SNPs with MAF = 0% cannot be analysed by Haploview which accounts for 3 that are not present in this LD plot.

LD can be measured by the $r^2$ score. Each square represents the pairwise LD ($r^2$ LD) between two SNPs. The $r^2$ score is shown on a grey scale, with higher $r^2$ LD represented by a darker blocks. The white blocks show $r^2$=0 (random assortment) and the black blocks show $r^2$=1 (total LD). From this score, tag SNPs (tSNPs) can be indentified. A tSNP is the minimum SNP needed to identify a haplotype, for example, $r^2$= 1 indicates two SNPs perfectly “tags” each other.

6.2.4 Identification of tSNPs in DDAH1 and DDAH2

The tagger programme in Haploview 4.2 was used to tag SNPs across the DDAH1 and DDAH2 genes on the basis of linkage disequilibrium patterns observed in the CEU samples genotyped as part of the International HapMap Project.

I) For DDAH1, SNPs with MAF > 20% were considered; 16 tag SNPs captured all 118 possible alleles with an $r^2$ score of at least 0.8 (mean $r^2$ = 0.938).

II) For DDAH2, all SNPs were considered; 4 tag SNPs captured all 6 possible alleles with an $r^2$ score of at least 0.8 (mean $r^2$ = 1).
<table>
<thead>
<tr>
<th>SNP</th>
<th>Sequence</th>
<th>Gene</th>
<th>Chr</th>
<th>Position</th>
<th>Functional Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs530006</td>
<td>TAACTCTTGACAGAACACTGGAAG[G/T]ATTCTCTCTGCAAAACTACTCCTC</td>
<td>DDAH1</td>
<td>1</td>
<td>85447825</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs11161618</td>
<td>AACCCTGAAAAACTACATGAGCACAAC[G/C]GGAGATGCATGTGTTGGTGGGAGA</td>
<td>DDAH1</td>
<td>1</td>
<td>85449719</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs669173</td>
<td>ACAAAGCTCAACGCAGCTGACACAA[G/C]ATCTAGTTATTTATTTGTAGCC</td>
<td>DDAH1</td>
<td>1</td>
<td>85433745</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs6576765</td>
<td>GGAATATGCACTAGCTGAGGAAAG[A/T]GATGTTTTTTCTGCTATTTA</td>
<td>DDAH1</td>
<td>1</td>
<td>85405329</td>
<td>intron variant, upstream variant 2KB</td>
</tr>
<tr>
<td>rs506733</td>
<td>GGGCTCAATGTTTCCCACACTG[CC/T]AGTGAAGAATCTAGAAGGGGTTCC</td>
<td>DDAH1</td>
<td>1</td>
<td>8546978</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs877041</td>
<td>TGAGCTCTTTTCTGACTCTTTTCTAATTTGCTGAGCTATTTTCT</td>
<td>DDAH1</td>
<td>1</td>
<td>85458237</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs755486</td>
<td>TTAACCTGCTTTTCTGCAAACACACTA[G/C]CAGAGCTCTGATCTACCAGTGTG</td>
<td>DDAH1</td>
<td>1</td>
<td>85453287</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs1241321</td>
<td>TCCAGACATATGCTTTTTTTTCT[CC/T]GGGGTACCCTTCTCTTCTCTCT</td>
<td>DDAH1</td>
<td>1</td>
<td>85462196</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs1523993</td>
<td>CCGGTGCTAGTATTGAGTCTGT[CC/G]CTGTTTTTTGATTTGACTAGG</td>
<td>DDAH1</td>
<td>1</td>
<td>85433445</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs480414</td>
<td>TACATAATGTTGGGAGAAGAGAATCTAGAAGGGTGTG</td>
<td>DDAH1</td>
<td>1</td>
<td>85461780</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs527762</td>
<td>AAAACAAAAAGATACAGATCTCT[CC/T]CTAGGCTTTGAGCTCTTCCACA</td>
<td>DDAH1</td>
<td>1</td>
<td>85463531</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs12130380</td>
<td>GAAGAAATGCGAGGGAGCTGTCAGAGCAG[CC/T]TTGTGAAATTTTAAAAAAATGAGG</td>
<td>DDAH1</td>
<td>1</td>
<td>85435626</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs233051</td>
<td>TGGTGTTGAGTAAATCTAAGCTAATG[A/G]ATAACCTCAACACAATCTAGTGTG</td>
<td>DDAH1</td>
<td>1</td>
<td>85372730</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs2300639</td>
<td>CCTCTGGCATTAAAGCAGCTAAT[CC/G]CAGAGCTCTGATCTACCAGTGTG</td>
<td>DDAH1</td>
<td>1</td>
<td>85357993</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs986639</td>
<td>GGACGGGATTTGAGTTGAGATGTGAGCAG[CC/T]GGGGTGAAAGGAAGAGGGAAGGGCAGG</td>
<td>DDAH1</td>
<td>1</td>
<td>85370709</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs539714</td>
<td>GTTACTACCAAGCTGCTCTAAAAC[CC/T]TGCAATTTACAGGCAAGAATTAAATG</td>
<td>DDAH1</td>
<td>1</td>
<td>85438655</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs28366162</td>
<td>GTGGAGAATAATGACAGCAAGACAC[CC/T]AGAGCTCTGATAGTGCAGACACTA</td>
<td>DDAH1</td>
<td>1</td>
<td>85372813</td>
<td>intron variant</td>
</tr>
<tr>
<td>rs805305</td>
<td>TTCTGTCCCCATAACCTTTGCGGGTC[CC/A]GGGGCTGAGCTCAGTCTCTCTCGGG</td>
<td>DDAH2</td>
<td>6</td>
<td>31729610</td>
<td>intron variant, upstream variant 2KB</td>
</tr>
<tr>
<td>rs9267551</td>
<td>GGATCGGAGCTGCTGCTTTTTATGGC[CC/T]GGGGTCTCTGTGAGGCACTCATA</td>
<td>DDAH2</td>
<td>CLIC1</td>
<td>6</td>
<td>31730180</td>
</tr>
<tr>
<td>rs805304</td>
<td>GGGGGGCTGCTGCTGCTTTTTATGGC[CC/A]GGGGTGCTCTGTGAGGCACTCATA</td>
<td>DDAH2</td>
<td>CLIC1</td>
<td>6</td>
<td>31730111</td>
</tr>
<tr>
<td>rs2272592</td>
<td>AATCAAGTGCTTTTTACCTAAGGGG[CC/A]TTTAAAGTTGCTGGGCACTGAAATG</td>
<td>DDAH2</td>
<td>CLIC1</td>
<td>6</td>
<td>31730575</td>
</tr>
</tbody>
</table>

Table 11. tSNPs genotyped in DDAH1 and DDAH2 genes

For DDAH2, in view of the slightly differing functional consequences of rs805304 and rs805305, both were selected for genotyping and SNP rs28366162 was not selected in view of a very low MAF of <0.05% in the CEU cohort.
6.3 Genotyping results for tSNPs in IPF Cohort

6.3.1 Genotype distribution and allele frequency data for DDAH1 tSNPs

![Graph showing DDAH1 SNP genotype distribution](image)

**Figure 41. Distribution of DDAH1 SNP genotypes amongst in IPF Cohort**

For each SNP, each bar in turn represents the numbers of patients with the major homozygote, heterozygote and minor homozygote.

<table>
<thead>
<tr>
<th>DDAH1</th>
<th>Genotype Frequency</th>
<th>Allele Frequency</th>
<th>ADMA Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs530006</td>
<td>0.26 0.47 0.24 0.03</td>
<td>0.51 0.49</td>
<td>EARSII, PIVUS</td>
</tr>
<tr>
<td>rs11161618</td>
<td>0.24 0.54 0.20 0.01</td>
<td>0.52 0.48</td>
<td>Not known</td>
</tr>
<tr>
<td>rs669173</td>
<td>0.27 0.53 0.19 0.01</td>
<td>0.54 0.46</td>
<td>PIVUS</td>
</tr>
<tr>
<td>rs6576765</td>
<td>0.34 0.49 0.16 0.01</td>
<td>0.59 0.41</td>
<td>EARSII, PIVUS</td>
</tr>
<tr>
<td>rs506733</td>
<td>0.31 0.53 0.10 0.06</td>
<td>0.61 0.39</td>
<td>Not known</td>
</tr>
<tr>
<td>rs877041</td>
<td>0.39 0.49 0.10 0.03</td>
<td>0.65 0.35</td>
<td>PIVUS</td>
</tr>
<tr>
<td>rs7555486</td>
<td>0.43 0.49 0.07 0.01</td>
<td>0.68 0.32</td>
<td>PIVUS</td>
</tr>
<tr>
<td>rs1241321</td>
<td>0.43 0.49 0.06 0.03</td>
<td>0.69 0.31</td>
<td>Not known</td>
</tr>
<tr>
<td>rs1523993</td>
<td>0.44 0.49 0.06 0.01</td>
<td>0.70 0.30</td>
<td>Not known</td>
</tr>
<tr>
<td>rs480414</td>
<td>0.43 0.49 0.04 0.04</td>
<td>0.70 0.30</td>
<td>Not known</td>
</tr>
<tr>
<td>rs527762</td>
<td>0.47 0.44 0.06 0.03</td>
<td>0.71 0.29</td>
<td>PIVUS</td>
</tr>
<tr>
<td>rs12130380</td>
<td>0.51 0.47 0.01 0.00</td>
<td>0.75 0.25</td>
<td>EARSII, PIVUS</td>
</tr>
<tr>
<td>rs233051</td>
<td>0.54 0.37 0.04 0.04</td>
<td>0.76 0.24</td>
<td>EARSII</td>
</tr>
<tr>
<td>rs2300639</td>
<td>0.54 0.44 0.01 0.00</td>
<td>0.76 0.24</td>
<td>PIVUS</td>
</tr>
<tr>
<td>rs986639</td>
<td>0.63 0.36 0.01 0.00</td>
<td>0.81 0.19</td>
<td>Not known</td>
</tr>
<tr>
<td>rs539714</td>
<td>0.71 0.21 0.04 0.03</td>
<td>0.85 0.15</td>
<td>EARSII</td>
</tr>
</tbody>
</table>

**Table 12. DDAH1 SNP genotype and allele frequency in IPF cohort**

Unknown refers to the “no call” rate (i.e. the genotype could not be confidently called)
6.3.2 Genotype distribution and allele frequency data for DDAH2 tSNPs

![DDAH2 SNP Genotype Distribution](image)

Figure 42. Distribution of DDAH2 genotypes in IPF cohort

For each SNP, each bar in turn represents the numbers of patients with the major homozygote, heterozygote and minor homozygote.

<table>
<thead>
<tr>
<th>DDAH2</th>
<th>Genotype Frequency</th>
<th>Allele Frequency</th>
<th>ADMA Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Major</td>
<td>Hetero</td>
<td>Minor</td>
</tr>
<tr>
<td>rs805304</td>
<td>0.44</td>
<td>0.44</td>
<td>0.11</td>
</tr>
<tr>
<td>rs805305</td>
<td>0.41</td>
<td>0.43</td>
<td>0.10</td>
</tr>
<tr>
<td>rs2272592</td>
<td>0.60</td>
<td>0.33</td>
<td>0.03</td>
</tr>
<tr>
<td>rs9267551</td>
<td>0.69</td>
<td>0.29</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 13. DDAH2 SNP genotype and allele frequency in IPF cohort

Unknown refers to the “no call” rate (i.e. the genotype could not be confidently called)

6.3.3 Testing for deviation from Hardy-Weinberg Equilibrium

The Hardy–Weinberg equilibrium (HWE) states that allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences. Testing for deviation from the HWE was performed using a chi-squared test \( \chi^2 \) with a significant p-value cut-off of 0.05 to indicate whether there was a significant difference between the observed genotype frequencies obtained from the IPF cohort and the expected genotype frequencies for a population in HWE. All genotype frequencies for all DDAH1 and DDAH2 SNPs were in HWE.
6.4 Final selection of tSNPs for the human study

6.4.1. Rationale for selection of DDAH1 SNPs

The relatively small sample size of the IPF cohort (70 patients) was a key factor in setting the MAF cut off at which to analyse SNPs. The false-positive rate is influenced by MAF and analysing the more common SNPs may result in significantly fewer false positives. The power to detect a given genetic effect with a given study size depends to a great extent on the MAF of the risk allele tested, specifically, loci with a low MAF (<10%) have significantly lower power to detect weak genotypic risk ratios than loci with a high MAF (>40%) [Tabangin 2009]. Taking into account the MAFs in the IPF cohort (and the absolute numbers of patients with the minor homozygote), a cut off of MAF>40% was selected.

As discussed, DDAH1 SNPs that significantly associate with ADMA levels in a normal cohort were selected as they are more likely to do so in a disease cohort and thus allow for any clinical differences in phenotype to be detected. Measurement of ADMA levels in the EARSII study was previously conducted by our group. The European Atherosclerosis Relatives Study II (EARSII) was a case-controlled study of 407 Caucasian male students aged 18–28 with no documented cardiovascular disease whose fathers had suffered from a proven myocardial infarction with an equal number of Caucasian age-matched controls, all recruited from European universities in 1993 [Tiret 2000]. No heterogeneity was detected in ADMA levels between cases and controls, therefore the association between genotype and plasma ADMA concentrations was examined in the pooled population in our recent publication [Caplin 2010].

The DDAH1 SNPs that satisfied the criteria of i) MAF>40% in IPF cohort and ii) having a significant effect on ADMA levels in the EARSII cohort were rs530006 and rs6576765. In terms of directionality of effect on ADMA levels, the minor homozygote was significantly associated with lower ADMA levels for rs530006, and the major homozygote was significantly associated with lower ADMA levels for rs6576765.

Of note, the DDAH1 gene variant with increased renal DDAH1 mRNA transcription that was associated with lower plasma ADMA levels and a steeper rate of renal function decline - rs17384213 [Caplin 2010] did not emerge as a common variant in the IPF cohort (tSNP rs539714, MAF 15%).
Figure 43. DDAH1 SNP rs530006: ADMA concentrations by genotype
Adjusted plasma ADMA concentrations (geometric mean and 95% CI) p<0.005* by ANOVA between genotypes in EARSII participants [Data from Caplin 2010].
Genotype frequencies: AA 27.9; AG 48.9; 23.2.

Figure 44. DDAH1 SNP rs6576765: ADMA concentrations by genotype
Adjusted plasma ADMA concentrations (mean and 95% CI) p<0.005* by ANOVA between genotypes in EARSII participants [Data from Caplin 2010].
Genotype frequencies: TT 41.0; CT 44.3; CC 14.6.

6.4.2 Rationale for selection of DDAH2 SNPs
As discussed, DDAH2 SNPs are not known to associate with ADMA levels and thus were analysed on an exploratory basis in the IPF cohort and not based on a specific hypothesis or a prediction based on directionality of ADMA levels. In view of the differing functional consequences of rs805304 and rs805305, both were selected for analysis.
CHAPTER 7: DDAH POLYMORPHISMS IN IDIOPATHIC PULMONARY FIBROSIS

7.1 Introduction
IPF appears to be a disease of NO overproduction and modulation of NO bioavailability via DDAH1 activity may be a key pathogenetic mechanism as supported by in vivo evidence; pharmacological and genetic knock-out of DDAH1 attenuates bleomycin-induced fibrosis whereas genetic overexpression of DDAH1 aggravates it [Pullamsetti 2011]. It has been hypothesised that genetic variants of DDAH1 that are associated with lower ADMA levels, and therefore indicative of DDAH over-activity, may be associated with adverse disease phenotype. Furthermore, SNPs that modulate ADMA levels to a greater extent, may also have a greater effect on the disease phenotype. The aims of this study were to identify the impact of DDAH1 SNPs with a functional effect on ADMA levels (as identified in Chapter 6) on longitudinal lung function decline and survival in a cohort of IPF patients. The same cohort of IPF patients was used and the same analyses were performed on the DDAH2 SNPs as per the DDAH1 SNPs. However as there is no known directionality of effect on ADMA by DDAH2 homozygotes, genotype analyses were performed only.

7.2 Baseline clinical data

7.2.1 Patient demographics

![Histogram of Total IPF Cohort: Age at Presentation](image)

**Figure 45. Baseline clinical data: age at presentation**

Mean 61.99 years (SD 8.03) Range 43 to 86 years
Figure 46. Baseline clinical data: gender

Figure 47. Baseline clinical data: smoking history

Figure 48. Baseline clinical data: summary of baseline pulmonary function
### 7.2.2 Clinical end points

**Figure 49. Clinical end point: time to first decline in FVC**
Defined as the number of months taken for the relative change of ≥10% in FVC on one subsequent time point from baseline (total cohort for survival curve analysis n=62)

**Figure 50. Clinical end point: length of survival from presentation**
Defined as the number of months survived by the patient following presentation to the RBH ILD clinic (total cohort for survival curve analysis n=67)

**Figure 51. Clinical end point: time to first event**
Defined as the first of either event to occur – time to first decline (the relative change ≥10% in FVC on one subsequent time point from baseline) or mortality from all causes (total cohort for survival curve analysis n=67)
7.3 Genotyping results for rs530006

7.3.1 Genotype distribution

Figure 52. rs530006 genotype distribution

7.3.2 Baseline PFTs by genotype

Figure 53. Baseline PFTs by rs530006 genotype
No significant difference in PFTs between rs530006 genotypes on one-way ANOVA (p=ns)
7.3.3 Time to decline of FVC

Figure 54. Time to decline of FVC by carriage of rs530006 genotype
Significant difference between TtD of FVC by carriage of rs530006 genotype
Log-rank (Mantel-Cox) test: p<0.0039*
Median (months): TT 13.73; GT 38.63; GG 34.5

Figure 55. Time to decline of FVC by carriage of rs530006 minor homozygote
Significant difference between TtD of FVC with carriage of the TT homozygote
Log-rank (Mantel-Cox) test: p<0.0009*
Median (months): TT 13.73; GG/GT 36.37
Hazard Ratio (Mantel-Haenszel): 4.675 (95% CI 1.876 to 11.65)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Univariate Models</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>_t</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>0.990</td>
<td>0.018</td>
</tr>
<tr>
<td>1.sex</td>
<td>1.073</td>
<td>0.371</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.960</td>
<td>0.325</td>
</tr>
<tr>
<td>fev1</td>
<td>0.995</td>
<td>0.007</td>
</tr>
<tr>
<td>fvc</td>
<td>0.988</td>
<td>0.007</td>
</tr>
<tr>
<td>tlco</td>
<td>0.994</td>
<td>0.008</td>
</tr>
<tr>
<td>cpi</td>
<td>1.014</td>
<td>0.009</td>
</tr>
<tr>
<td>ttrs530006</td>
<td>3.283</td>
<td>1.244</td>
</tr>
<tr>
<td>gtrs530006</td>
<td>0.735</td>
<td>0.233</td>
</tr>
<tr>
<td>ggrs530006</td>
<td>0.632</td>
<td>0.225</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Model correcting for Time Varying Co-variates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Haz. Ratio</td>
</tr>
<tr>
<td>_t</td>
<td></td>
</tr>
<tr>
<td>main variable</td>
<td></td>
</tr>
<tr>
<td>ttrs530006</td>
<td>3.594</td>
</tr>
<tr>
<td>fvc</td>
<td>1.001</td>
</tr>
<tr>
<td>cpi</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 14. rs530006 TtD of FVC: multivariate Cox proportional hazards analysis
Following Cox univariate analysis, significant and potentially confounding time varying co-variates (FEV1, FVC, CPI) that violated the assumption of PH were adjusted for in a stepwise multivariate analysis. ttrs530006 remained significant following multivariate analysis and correction for time varying co-variates (p<0.001)
7.3.4 Length of survival

**Figure 56. Length of survival by rs530006 genotype**

Significant difference between survival from presentation by carriage of rs530006 genotype

Log-rank (Mantel-Cox) test: p<0.0418*

Median (months): TT 26.0; GT 57.2; GG 42.4

**Figure 57. Length of survival by carriage of rs530006 minor homozygote**

Significant difference between survival from presentation with carriage of the TT homozygote

Log-rank (Mantel-Cox) test: p<0.0118*

Median (months): TT 26.0; GG/GT 45.1

Hazard Ratio (Mantel-Haenszel): 2.431 (95% CI 1.218 to 4.855)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Univariate Models</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>age</td>
<td>1.010</td>
<td>0.017</td>
</tr>
<tr>
<td>1.sex</td>
<td>1.490</td>
<td>0.439</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.670</td>
<td>0.195</td>
</tr>
<tr>
<td>fev1</td>
<td>0.990</td>
<td>0.006</td>
</tr>
<tr>
<td>fvc</td>
<td>0.987</td>
<td>0.006</td>
</tr>
<tr>
<td>tlco</td>
<td>0.976</td>
<td>0.008</td>
</tr>
<tr>
<td>cpi</td>
<td>1.027</td>
<td>0.009</td>
</tr>
<tr>
<td>ttrs530006</td>
<td>2.060</td>
<td>0.604</td>
</tr>
<tr>
<td>gtrs530006</td>
<td>0.715</td>
<td>0.184</td>
</tr>
<tr>
<td>ggrs530006</td>
<td>0.839</td>
<td>0.250</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Multivariate Model</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>fvc</td>
<td>1.012</td>
<td>0.013</td>
</tr>
<tr>
<td>tlco</td>
<td>1.001</td>
<td>0.026</td>
</tr>
<tr>
<td>cpi</td>
<td>1.038</td>
<td>0.039</td>
</tr>
<tr>
<td>ttrs530006</td>
<td>1.859</td>
<td>0.598</td>
</tr>
</tbody>
</table>

Table 15. rs530006 length of survival: multivariate Cox proportional hazards analysis
Following Cox univariate analysis, significant and potentially confounding time varying co-variates (FVC, TLCO and CPI) were adjusted for in a multivariate analysis. ttrs530006 remained borderline significant following multivariate analysis to correct for time varying co-variates (p<0.054)
7.3.5 Time to first event

**Figure 58. Time to first event by rs530006 genotype**

Significant difference between time to first event by carriage of rs530006 genotype

Log-rank (Mantel-Cox) test: p<0.0007*

Median (months): TT 12.7; GT 27.93; GG 33.43

**Figure 59. Time to first event by carriage of rs530005 minor homozygote**

Significant difference between time to first event with carriage of the TT homozygote

Log-rank (Mantel-Cox) test: p<0.0001*

Median (months): TT 12.7; GG/GT 27.93

Hazard Ratio (Mantel-Haenszel): 4.765 (95% CI 2.139 to 10.61)
Variable | Cox Regression Univariate Models | PH Test
--- | --- | ---
| _t | Haz. Ratio | Std. Err. | z | P>z | [95% Conf. Interval] | prob>chi2 |
age | 0.992 | 0.016 | -0.520 | 0.600 | 0.962 | 1.023 | 0.376 |
1.sex | 1.558 | 0.471 | 1.470 | 0.142 | 0.862 | 2.817 | 0.136 |
1.smoking | 0.912 | 0.255 | -0.330 | 0.742 | 0.527 | 1.579 | 0.924 |
fev1 | 0.992 | 0.006 | -1.310 | 0.191 | 0.980 | 1.004 | 0.297 |
fvc | 0.989 | 0.006 | -1.910 | 0.056 | 0.977 | 1.000 | 0.086 |
tlco | 0.986 | 0.007 | -1.860 | 0.062 | 0.972 | 1.001 | 0.748 |
cpi | 1.017 | 0.008 | 2.100 | 0.036 | 1.001 | 1.033 | 0.301 |
ttrs530006 | 3.318 | 1.100 | 3.620 | 0.000 | 1.732 | 6.356 | 0.555 |
gtrs530006 | 0.628 | 0.165 | -1.770 | 0.077 | 0.375 | 1.052 | 0.439 |
ggrs530006 | 0.808 | 0.236 | -0.730 | 0.465 | 0.456 | 1.432 | 0.151 |

| Variable | Cox Regression Multivariate Model | PH Test |
--- | --- | ---
| _t | Haz. Ratio | Std. Err. | z | P>z | [95% Conf. Interval] | individual | global |
cpi | 1.015 | 0.008 | 1.740 | 0.081 | 0.998 | 1.031 | 0.332 | 0.523 |
ttrs530006 | 3.039 | 1.016 | 3.330 | 0.001 | 1.579 | 5.852 | 0.485 |

| Variable | Cox Regression Multivariate Model (inc borderline PFTS) | PH Test |
--- | --- | ---
| _t | Haz. Ratio | Std. Err. | z | P>z | [95% Conf. Interval] | individual | global |
fvc | 1.004 | 0.012 | 0.350 | 0.724 | 0.981 | 1.028 | 0.823 | 0.175 |
tlco | 1.003 | 0.028 | 0.110 | 0.914 | 0.950 | 1.059 | 0.65 | 0.065 |
cpi | 1.022 | 0.039 | 0.590 | 0.556 | 0.950 | 1.101 | 0.120 |
ttrs530006 | 3.116 | 1.073 | 3.300 | 0.001 | 1.587 | 6.118 | 0.282 |

Table 16. rs530006 time to first event: multivariate Cox proportional hazards analysis
Following Cox univariate analysis, significant and potentially confounding time varying co-
variates (FVC, TLCO and CPI) were adjusted for in a multivariate analysis. ttrs530006 remained
significant following multivariate analysis to correct for time varying co-
variates (p<0.001)
7.4 Genotyping results for rs6576765

7.4.1 Genotype distribution

![rs6576765 genotype distribution](image)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>24</td>
</tr>
<tr>
<td>TA</td>
<td>34</td>
</tr>
<tr>
<td>AA</td>
<td>11</td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 60. rs6576765 genotype distribution

7.4.2 Baseline PFTs by genotype

![rs6576765 FEV1 by genotype](image)

![rs6576765 FVC by genotype](image)

![rs6576765 TLCO by genotype](image)

![rs6576765 CPI by genotype](image)

Figure 61. Baseline PFTs by rs6576765 genotype

No significant difference in PFTs between rs6576765 genotypes on one way ANOVA (p=ns)
7.4.3 Time to decline of FVC

Significant difference between TtD of FVC by carriage of rs6576765 genotype
Log-rank (Mantel-Cox) test: p<0.0249*
Median (months): TT 20.67; TA 38.63; AA 33.43

Figure 62. Time to decline of FVC by carriage of rs657675 genotype

Significant difference between TtD of FVC with carriage of the TT homozygote
Log-rank (Mantel-Cox) test: p<0.0075*
Median (months): TT 20.67; TA/AA 34.5
Hazard Ratio (Mantel-Haenszel): 2.576 (95% CI 1.288 to 5.153)
### Variable Cox Regression Univariate Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Haz. Ratio</th>
<th>Std. Err.</th>
<th>z</th>
<th>P&gt;z</th>
<th>prob&gt;chi2</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>0.990</td>
<td>0.018</td>
<td>-0.530</td>
<td>0.595</td>
<td>0.955</td>
<td>1.027</td>
</tr>
<tr>
<td>age</td>
<td>1.073</td>
<td>0.371</td>
<td>0.200</td>
<td>0.839</td>
<td>0.545</td>
<td>2.113</td>
</tr>
<tr>
<td>1.sex</td>
<td>0.960</td>
<td>0.325</td>
<td>-0.120</td>
<td>0.905</td>
<td>0.495</td>
<td>1.863</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.995</td>
<td>0.007</td>
<td>-0.610</td>
<td>0.539</td>
<td>0.981</td>
<td>1.010</td>
</tr>
<tr>
<td>fev1</td>
<td>0.988</td>
<td>0.007</td>
<td>-1.780</td>
<td>0.075</td>
<td>0.975</td>
<td>1.001</td>
</tr>
<tr>
<td>fvc</td>
<td>0.994</td>
<td>0.008</td>
<td>-0.690</td>
<td>0.490</td>
<td>0.978</td>
<td>1.011</td>
</tr>
<tr>
<td>tico</td>
<td>1.014</td>
<td>0.009</td>
<td>1.470</td>
<td>1.440</td>
<td>0.995</td>
<td>1.032</td>
</tr>
<tr>
<td>cpi</td>
<td>2.303</td>
<td>0.737</td>
<td>2.610</td>
<td>0.009</td>
<td>1.230</td>
<td>4.311</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>0.664</td>
<td>0.208</td>
<td>-1.310</td>
<td>0.191</td>
<td>0.359</td>
<td>1.227</td>
</tr>
<tr>
<td>aars6576765</td>
<td>0.572</td>
<td>0.258</td>
<td>-1.240</td>
<td>0.215</td>
<td>0.236</td>
<td>1.383</td>
</tr>
<tr>
<td>fvc</td>
<td>1.000</td>
<td>0.001</td>
<td>1.750</td>
<td>0.081</td>
<td>1.000</td>
<td>1.003</td>
</tr>
<tr>
<td>cpi</td>
<td>1.000</td>
<td>0.001</td>
<td>-0.440</td>
<td>0.661</td>
<td>0.999</td>
<td>1.001</td>
</tr>
</tbody>
</table>

### PH Test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard Ratio</th>
<th>Std. Err.</th>
<th>z</th>
<th>P&gt;z</th>
<th>prob&gt;chi2</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>0.990</td>
<td>0.018</td>
<td>-0.530</td>
<td>0.595</td>
<td>0.955</td>
<td>1.027</td>
</tr>
<tr>
<td>age</td>
<td>1.073</td>
<td>0.371</td>
<td>0.200</td>
<td>0.839</td>
<td>0.545</td>
<td>2.113</td>
</tr>
<tr>
<td>1.sex</td>
<td>0.960</td>
<td>0.325</td>
<td>-0.120</td>
<td>0.905</td>
<td>0.495</td>
<td>1.863</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.995</td>
<td>0.007</td>
<td>-0.610</td>
<td>0.539</td>
<td>0.981</td>
<td>1.010</td>
</tr>
<tr>
<td>fev1</td>
<td>0.988</td>
<td>0.007</td>
<td>-1.780</td>
<td>0.075</td>
<td>0.975</td>
<td>1.001</td>
</tr>
<tr>
<td>fvc</td>
<td>0.994</td>
<td>0.008</td>
<td>-0.690</td>
<td>0.490</td>
<td>0.978</td>
<td>1.011</td>
</tr>
<tr>
<td>tico</td>
<td>1.014</td>
<td>0.009</td>
<td>1.470</td>
<td>1.440</td>
<td>0.995</td>
<td>1.032</td>
</tr>
<tr>
<td>cpi</td>
<td>2.303</td>
<td>0.737</td>
<td>2.610</td>
<td>0.009</td>
<td>1.230</td>
<td>4.311</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>0.664</td>
<td>0.208</td>
<td>-1.310</td>
<td>0.191</td>
<td>0.359</td>
<td>1.227</td>
</tr>
<tr>
<td>aars6576765</td>
<td>0.572</td>
<td>0.258</td>
<td>-1.240</td>
<td>0.215</td>
<td>0.236</td>
<td>1.383</td>
</tr>
<tr>
<td>fvc</td>
<td>1.000</td>
<td>0.001</td>
<td>1.750</td>
<td>0.081</td>
<td>1.000</td>
<td>1.003</td>
</tr>
<tr>
<td>cpi</td>
<td>1.000</td>
<td>0.001</td>
<td>-0.440</td>
<td>0.661</td>
<td>0.999</td>
<td>1.001</td>
</tr>
</tbody>
</table>

**Table 17. rs6576765 TtD of FVC: multivariate Cox proportional hazards analysis**

Following Cox univariate analysis, significant and potentially confounding time varying co-variates (FEV1, FVC, and CPI) were adjusted for in a multivariate analysis. ttrs6576765 remained significant following multivariate analysis to correct for time varying co-variates (p<0.014).
7.4.4 Length of survival

**Figure 64. Length of survival by rs6576765 genotype**

No significant difference in survival from presentation by carriage of rs6576765 genotype

Log-rank (Mantel-Cox) test: p=0.0839

Median (months): TT 32.965, TA 45.7; AA 37.73

**Figure 65. Length of survival by carriage of rs6576765 major homozygote**

No significant difference between survival from presentation with carriage of TT homozygote

Log-rank (Mantel-Cox) test: p=ns

Median (months): TT 32.965; GG/GT 43.95
### Table 18. rs6576765 length of survival: multivariate Cox proportional hazard analysis

Following Cox univariate analysis, significant and potentially confounding time varying co-variates (FVC, TLCO, and CPI) and were adjusted for in a multivariate analysis. atrs6576765 did not remain significant following multivariate analysis to correct for time varying co-variates (p<0.073).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Univariate Models</th>
<th></th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
<td>z</td>
</tr>
<tr>
<td>age</td>
<td>1.010</td>
<td>0.017</td>
<td>0.580</td>
</tr>
<tr>
<td>1.sex</td>
<td>1.490</td>
<td>0.439</td>
<td>1.350</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.670</td>
<td>0.195</td>
<td>-1.370</td>
</tr>
<tr>
<td>fev1</td>
<td>0.990</td>
<td>0.006</td>
<td>-1.690</td>
</tr>
<tr>
<td>fvc</td>
<td>0.987</td>
<td>0.006</td>
<td>-2.190</td>
</tr>
<tr>
<td>tlco</td>
<td>0.976</td>
<td>0.008</td>
<td>-2.870</td>
</tr>
<tr>
<td>cpi</td>
<td>1.027</td>
<td>0.009</td>
<td>2.930</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>1.535</td>
<td>0.417</td>
<td>1.580</td>
</tr>
<tr>
<td>atras6576765</td>
<td>0.552</td>
<td>0.149</td>
<td>-2.200</td>
</tr>
<tr>
<td>aars6576765</td>
<td>1.431</td>
<td>0.485</td>
<td>1.060</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Multivariate Model</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>fvc</td>
<td>1.009</td>
<td>0.012</td>
</tr>
<tr>
<td>tlco</td>
<td>1.000</td>
<td>0.026</td>
</tr>
<tr>
<td>cpi</td>
<td>1.036</td>
<td>0.037</td>
</tr>
<tr>
<td>atras6576765</td>
<td>0.602</td>
<td>0.170</td>
</tr>
</tbody>
</table>

Following Cox univariate analysis, significant and potentially confounding time varying co-variates (FVC, TLCO, and CPI) and were adjusted for in a multivariate analysis. atrs6576765 did not remain significant following multivariate analysis to correct for time varying co-variates (p<0.073).
7.4.5 Time to first event

**Figure 66. Time to first event by rs6576765 genotype**

Significant difference between time to first event by carriage of rs6576765 genotype

Log-rank (Mantel-Cox) test: p<0.0212*

Median (months): TT 15.8; TA 29.5; AA 27.07

**Figure 67. Time to first event by carriage of rs6576765 major homozygote**

Significant difference between time to first event with carriage of the TT homozygote

Log-rank (Mantel-Cox) test: p<0.0058*

Median (months): TT 15.8; TA/AA 28.45

Hazard Ratio (Mantel-Haenszel): 2.336 (95% CI 1.279 to 4.268)
### Table 19. rs6576765 time to first event: multivariate Cox proportional hazard analysis

Following Cox univariate analysis, the significant and potentially confounding time varying co-variates (CPI) was adjusted for in a multivariate analysis. ttrs530006 remained significant following multivariate analysis and correction for time varying co-variates ($p<0.006$).

Several time-varying co-variates had borderline p values in the univariate analysis, a second multivariate model was run, this time correcting for FVC, TLCO, CPI and atrs6576765, ttrs6576765 remained significant with $p<0.047$.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Univariate Models</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>age</td>
<td>0.992</td>
<td>0.016</td>
</tr>
<tr>
<td>1.sex</td>
<td>1.558</td>
<td>0.471</td>
</tr>
<tr>
<td>1.smoking</td>
<td>0.912</td>
<td>0.255</td>
</tr>
<tr>
<td>fev1</td>
<td>0.992</td>
<td>0.006</td>
</tr>
<tr>
<td>fvc</td>
<td>0.989</td>
<td>0.006</td>
</tr>
<tr>
<td>tlco</td>
<td>0.986</td>
<td>0.007</td>
</tr>
<tr>
<td>cpi</td>
<td>1.017</td>
<td>0.008</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>2.112</td>
<td>0.584</td>
</tr>
<tr>
<td>atrs6576765</td>
<td>0.608</td>
<td>0.158</td>
</tr>
<tr>
<td>aars6576765</td>
<td>0.852</td>
<td>0.283</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Multivariate Model</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>cpi</td>
<td>1.019</td>
<td>0.009</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>2.169</td>
<td>0.606</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cox Regression Multivariate Model (inc borderline PFTS)</th>
<th>PH Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>_t</td>
<td>Haz. Ratio</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>fvc</td>
<td>0.998</td>
<td>0.012</td>
</tr>
<tr>
<td>tlco</td>
<td>1.005</td>
<td>0.029</td>
</tr>
<tr>
<td>cpi</td>
<td>1.021</td>
<td>0.039</td>
</tr>
<tr>
<td>ttrs6576765</td>
<td>2.169</td>
<td>0.847</td>
</tr>
<tr>
<td>atrs6576765</td>
<td>0.989</td>
<td>0.382</td>
</tr>
</tbody>
</table>
7.5 Genotyping results for rs805804

7.5.1 Genotype distribution

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>31</td>
</tr>
<tr>
<td>AC</td>
<td>31</td>
</tr>
<tr>
<td>CC</td>
<td>8</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 68. rs805804 genotype distribution

7.5.2 Baseline PFTs by genotype

Figure 69. Baseline PFTs by rs804305 genotype

No significant difference in all PFTs between rs804305 genotypes on one way ANOVA (p=ns) (data for FEV1, TLCO and CPI not shown)
7.5.3 Time to decline of FVC results

Figure 70. Time to decline of FVC by carriage of rs805304 genotype
No significant difference between TtD of FVC by carriage of rs805304 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): AA 24.7; AC 32.0; CC 20.67

7.5.4 Length of survival results

Figure 71. Length of survival by rs805304 genotype
No significant difference between survival from presentation by carriage of rs805304 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): AA 39.27; AC 38.13; CC 56.65
7.6 Genotyping results for rs805805

7.6.1 Genotype distribution

![Genotype distribution graph](image)

**Figure 72.** rs805305 genotype distribution

<table>
<thead>
<tr>
<th>Genotype</th>
<th>CC</th>
<th>GC</th>
<th>GG</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>29</td>
<td>30</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

7.6.2 Baseline PFTs by genotype

![PFTs by genotype](image)

**Figure 73.** Baseline PFTs by rs805305 genotype

No significant difference in all PFTs between rs805305 genotypes on one-way ANOVA (p=ns) (data for FEV1, TLCO and CPI not shown)
7.6.3 Time to decline of FVC results

Figure 74. Time to decline of FVC by carriage of rs805305 genotype
No significant difference between TtD of FVC by carriage of rs805305 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): CC 24.7; GC 27.57; GG 22.6

7.6.4 Length of survival results

Figure 75. Length of survival by rs805305 genotype
No significant difference between survival from presentation by carriage of rs805305 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): CC 39.27; CG 37.93; GG 57.2
7.7 Discussion

Summary of Key Findings

I. rs530006 (DDAH1) was associated with shorter time to lung function decline and shorter length of survival

II. rs6576765 (DDAH1) was associated with shorter time to lung function decline

III. DDAH2 SNPs were not associated with time to lung function decline or length of survival

IV. Baseline lung function did not vary by DDAH1 or DDAH2 SNPs

The baseline clinical data for the IPF cohort was typical of the UK IPF patient population. Following exclusion of 3 patients on novel agent therapies and any patients with no available data, patient numbers available for survival curve analyses were: TtD of FVC n=62, length of survival n=67, and time to first event n=67. The clinical end point data was checked for outliers; one IPF patient had an extremely long survival but was included in the analyses due to diagnostic certainty by multidisciplinary team consensus. The genotype distributions for the SNPs were in HWE, and although not strictly a control group, the DDAH minor allele frequencies were similar to the CEU population in Hap Map; rs530006 - 0.45 and 0.49; rs6576765 - 0.44 and 0.41; rs805305 - 0.32 and 0.33; and rs805304 - 0.31 and 0.34, in the HapMap CEU and IPF cohorts respectively. Baseline PFTs did not vary according to genotype for any SNP, this may have been as expected as the values were taken at baseline on presentation to the clinic, and in any case, were corrected for in the multivariate Cox analysis.

The genotyping results observed for rs530006 were consistent in their directionality of effect on the clinical end points. There was a significant difference between each of the rs530006 genotypes for all three clinical end points; TtD of FVC (p<0.0039); length of survival from presentation (p<0.0418); and time to first event (p<0.0007). Several significant observations were made with carriage of the minor TT homozygote that was hypothesised to be associated with lower ADMA levels and thus greater lung function decline and mortality. There was a significant difference in the time to first decline of FVC according to carriage of the minor TT homozygote on univariate analysis (HR 4.675, 95% CI 1.876 to 11.65, p<0.0009). This remained significant following adjustment for significant time varying co-variates FEV1, FVC, CPI on multivariate Cox analysis (p<0.001). There was also significant difference in length of survival from presentation according to carriage of the minor TT homozygote on univariate analysis
(HR 2.431, 95% CI 1.218 to 4.855, p<0.0118). This reached the limits of statistical significance following adjustment for significant time varying co-variates FVC, TLCO and CPI on multivariate Cox analysis (p<0.054). Median survival with the TT homozygote was 26 months compared to 45.1 months for GG/GT genotypes. Accordingly, there was a corresponding difference in time to first event with carriage of the minor TT homozygote (HR 4.765, 95% CI 2.139 to 10.61, p<0.0001). Overall the rs530006 TT homozygote demonstrated a clear separation in terms of survival curves from the GT and GG genotypes and was associated the worst clinical outcome in all 3 clinical end points which withstood correction in multivariate Cox analysis.

The genotyping results for rs6576765 demonstrated a significant difference between each of the genotypes for two of the clinical end points; TtD of FVC (p<0.0249); and time to first event (p<0.0212), results for length of survival from presentation did not reach significance (p<0.0839). Several observations were made with carriage of the major TT homozygote that was hypothesised to be associated with lower ADMA levels and thus greater lung function decline and mortality. There was a significant difference in the time to first decline of FVC according to carriage of the major TT homozygote on univariate analysis (HR 2.576, 95% CI 1.288 to 5.153, p<0.0075). This remained significant following adjustment for significant time varying co-variates FEV1, FVC, CPI on multivariate Cox analysis (p<0.014). There was no significant difference between survival from presentation with carriage of TT homozygote. There was a significant difference between time to first event with carriage of the TT homozygote (HR 2.336 (95% CI 1.279 to 4.268, p<0.0058). This remained significant following adjustment for significant time varying co-variates FVC, TLCO and CPI on multivariate Cox analysis (p<0.047). The major TT homozygote appeared to have a greater effect on time to lung function decline rather than on length of survival.

For both DDAH2 SNPs, there was no evidence of difference in time to first decline of FVC or length survival from presentation by carriage of rs805305 or rs805304 genotypes. The proposed hypothesis was that DDAH1 SNPs would be likely to have a greater functional effect on ADMA levels and thus clinical outcome measures. The DDAH2 SNPs were tested on an exploratory basis and the observations from the experimental EMT model and published literature by Pullamsetti et al., indicate a potential role of DDAH2 in the pathogenesis of IPF, although the mechanisms remain unclear.
The main limitation of the data is that a relatively small cohort size in which to detect a difference was used. Furthermore, there was a lack of a replication cohort in which to cross-check the findings. However, there are several advantages of this historic IPF cohort, namely the longitudinal lung function and mortality data, the lack of potential interference to clinical outcomes by any disease modifying therapies and the diagnostic certainty of a consensus multidisciplinary team diagnosis by a specialist tertiary referral centre in the UK. Survival is very robust end point and FVC is a surrogate for mortality, thus it could be considered that a further strength of the data set is that the results demonstrate a similar answer to the same question that was asked in two different ways, the directionality of the results for each clinical end point was the same.

The results for both \textit{DDAH1} SNPs appeared to be similar but rs530006 had the stronger results, particularly for length of survival. Upon selection, the \textit{DDAH1} SNPs were not in high LD with each as the $r^2$ value was set as 0.80. However, review of the LD plot data for \textit{DDAH1} demonstrates an $r^2$ value of 0.50, suggesting the pair of SNPs are in a degree of LD with each other. A genuine association signal will tend to reflect the local characteristics of the genome, and any nearby SNPs that are in LD, may also manifest the association signal. It is unclear if the effects that have been observed are independent for each SNP or a combined effect. However, the positions of the two SNPs are on differing ends of the gene and furthermore, the directionality of effect of the minor homozygote allele differs for each SNP, suggesting that the SNPs may be acting independently.

Interestingly, although there was separation of the survival curves, a dose effect of the effect homozygote was not observed from the survival curves for either \textit{DDAH1} SNP. This may have been due to the ‘heterozygote advantage’ that a heterozygote genotype confers in nature. The possible heterozygote advantage here may relate to improved vascular function associated with the allele for lower ADMA levels. In future work, it will be interesting to genotype rs530006 in a UK control cohort and compare genotype frequencies to the IPF population, in order to detect any possible association with the minor TT homozygote with IPF which could potentially designate it as a genetic risk allele for IPF.
CHAPTER 8: DDAH POLYMORPHISMS IN SYSTEMIC SCLEROSIS ILD

8.1 Introduction
Systemic sclerosis (SSc) is a multisystem, autoimmune connective tissue disease, characterised by excessive extracellular matrix deposition, with a highly variable pattern of organ involvement and prognosis. The leading cause of mortality is due to pulmonary involvement, from either pulmonary fibrosis or pulmonary hypertension [Steen 2007]. The pathogenesis of pulmonary fibrosis in SSc involves a complex combination of epithelial and endothelial cell injury with inflammatory and immune activation. Occurring in response to unknown initiating factors, the interaction between vascular, epithelial, and immune dysfunction leads to dysregulated fibroblast activation and increased extracellular matrix production [Renzoni 2007]. SSc-associated interstitial lung disease (SSc-ILD) is a major cause of morbidity and mortality in patients with SSc and although most patients have only limited involvement, in a significant minority, SSc-ILD is progressive and requires prompt treatment with immunosuppressive therapies. Early identification of patients at risk of progressive ILD in order to allow prompt treatment is a current clinical challenge. Being able to identify those at risk of ILD progression will also allow avoidance of potentially toxic treatments in those with limited ILD with an indolent course. Identification of SNPs associated with disease progression will answer an unmet clinical need.

The triad of vascular, epithelial, and immune dysfunction in SSc-ILD is of particular interest in the context of the DDAH isoforms. Of particular relevance, DDAH2 localises to 6p21.3, a region which contains many genes involved in the immune and inflammatory responses and has been linked with susceptibility to several autoimmune diseases including rheumatoid arthritis [Tran 2003]. It is widely acknowledged that IPF and SSc-ILD are distinct diseases and thus a study on an SSc cohort was conducted with an exploratory hypothesis to test whether the previous findings were specific to the IPF cohort or related to pulmonary fibrosis in general. The same SNPs analysed in the IPF cohort were analysed in the SSc cohort.

The hypothesis tested is that genetic variants associated with higher DDAH1 activity and lower ADMA levels and will be associated with accelerated lung function decline in SSc-ILD.
8.2 Baseline clinical data (SSc Cohort)

8.2.1 Patient demographics

**Figure 76. Baseline clinical data (SSc cohort): Age**

Mean 49.16 years (SD 12.48)

**Figure 77. Baseline clinical data (SSc cohort): Gender**

**Figure 78. Baseline clinical data (SSc cohort): Smoking history**
Figure 79. Baseline clinical data (SSc cohort): Pulmonary function tests

8.2.2 Clinical end point

Figure 80. Clinical end point: time to decline of FVC (SSc Cohort)
SSc-ILD was defined as: HRCT evidence of PF on CT report or result noted in clinical notes. Time to irreversible decline of FVC was defined as: relative change of ≥10% in FVC on two consecutive time points from baseline (total cohort for survival curve analysis n=255)
8.3 Genotyping results for rs530006 (SSc cohort)

8.3.1 Genotype distribution

![Genotype Distribution](image)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG</td>
<td>78</td>
</tr>
<tr>
<td>GT</td>
<td>133</td>
</tr>
<tr>
<td>TT</td>
<td>46</td>
</tr>
<tr>
<td>Unknown</td>
<td>7</td>
</tr>
</tbody>
</table>

Figure 81. rs530006 genotype distribution (SSc Cohort)

8.3.2 Baseline PFTs by genotype

![Baseline PFTs by Genotype](image)

No significant difference in FVC between rs530006 genotypes on one way ANOVA (p=ns)
8.3.3 Time to decline of FVC results

Figure 83. TtD of FVC by carriage of rs530006 genotype (SSc Cohort)
No significant difference between TtD of FVC by carriage of rs530006 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): TT 93.53; GT 69.17; GG 82.73

Figure 84. TtD of FVC by carriage of rs530006 minor homozygote (SSc Cohort)
No significant difference between TtD of FVC with carriage of the TT homozygote
Log-rank (Mantel-Cox) test: p=ns
Median (months): TT 93.53; GG/GT 77.83
8.4 Genotyping results for rs6576765 (SSc Cohort)

8.4.1 Genotype distribution

Figure 85. rs6576765 genotype distribution (SSc cohort)

8.4.2 Baseline PFTs by genotype

Figure 86. Baseline PFTs by rs6576765 genotype (SSc Cohort)

No significant difference in FVC between rs6576765 genotypes on one way ANOVA (p=ns)
8.4.3 Time to decline of FVC results

**Figure 87.** TtD of FVC by carriage of rs657675 genotype (SSc Cohort)

No significant difference between TtD of FVC by carriage of rs657675 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): TT 86.23; TA 79.03; AA 67.63

**Figure 88.** TtD of FVC by carriage of rs6576765 major homozygote (SSc Cohort)

No significant difference between TtD of FVC with carriage of the TT homozygote
Log-rank (Mantel-Cox) test: p=ns
Median (months): TT 86.23; TA/AA 78.5
8.5 Genotyping results for rs805304 (SSc Cohort)

8.5.1 Genotype distribution

![Genotype Distribution Chart]

**Figure 89.** rs805304 genotype distribution (SSc Cohort)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>85</td>
</tr>
<tr>
<td>AC</td>
<td>131</td>
</tr>
<tr>
<td>CC</td>
<td>40</td>
</tr>
<tr>
<td>Unknown</td>
<td>8</td>
</tr>
</tbody>
</table>

8.5.2 Baseline PFTs by genotype

![FVC by Genotype Chart]

**Figure 90.** Baseline PFTs by rs805304 genotype (SSc Cohort)

No significant difference in FVC between rs805304 genotypes on one way ANOVA (p=ns)
8.5.3 Time to decline of FVC results

![Graph showing time to decline of FVC by genotype](image1.png)

**Figure 91. Time to decline of FVC by carriage of rs805304 genotype (SSc Cohort)**

No significant difference between TtD of FVC by carriage of rs805304 genotype

Log-rank (Mantel-Cox) test: p=ns

Median (months): AA 73.33; AC 92.53; CC 79.03

8.6 Genotyping results for rs805305 (SSc Cohort)

8.6.1 Genotype distribution

![Bar chart showing rs805305 genotype distribution](image2.png)

**Figure 92. rs805305 genotype distribution (SSc Cohort)**
8.6.2 Baseline PFTs by genotype

Figure 93. Baseline PFTs by rs805305 genotype (SSc Cohort)
No significant difference in FVC between rs805305 genotypes on one way ANOVA (p=ns)

8.6.3 Time to decline of FVC results

Figure 94. Time to decline of FVC by carriage of rs805305 genotype (SSc Cohort)
No significant difference between TtD of FVC by carriage of rs805305 genotype
Log-rank (Mantel-Cox) test: p=ns
Median (months): CC 77.83; GC 90.73; GG 79.03
8.7 Discussion

Summary of Key Findings

I. *DDAH1* SNPs were not associated with shorter time to lung function decline
II. *DDAH2* SNPs were not associated with shorter time to lung function decline
III. Baseline lung function did not vary by *DDAH1* or *DDAH2* SNPs

Of the 300 SSc patients, 264 were defined as having SSc-ILD following review of documented clinical/HRCT findings. The baseline clinical data for the SSc-ILD cohort was typical of the UK SSc-ILD patient population. Following exclusion of 9 patients with no data, patient numbers available for time to irreversible decline analysis were n=255. Baseline PFTs did not vary according to genotype for any SNP.

The genotype distributions for the SNPs in the SSc-ILD cohort were in HWE. The *DDAH1* minor allele frequencies were similar to the CEU population in Hap Map; rs530006 - 0.45 and 0.44; rs6576765 - 0.44 and 0.42 in the HapMap CEU and SSc-ILD cohorts respectively. The *DDAH2* minor allele frequencies were; rs805305 - 0.32 and 0.41; and rs805304 - 0.31 and 0.41, in the HapMap CEU and SSc-ILD cohorts respectively. Interestingly, both *DDAH2* SNP minor allele frequencies appeared to be higher in the SSc-ILD cohorts compared to the HapMap CEU cohorts. Genotyping the two *DDAH2* SNPs in a control cohort from the UK in order to detect any possible association of rs805304 and rs805304 with SSc-ILD will be interesting and has not been previously done.

Overall, the results demonstrated no evidence of a significant difference between time to irreversible decline of FVC by genotype for either *DDAH1* or *DDAH2* SNPs. Further studies would require a replication cohort in which to cross-check the findings. Although both IPF and SSc-ILD share similar phenotypic features of pulmonary fibrosis, they are two distinct diseases in terms of pathogenesis, genetics, clinical features, longitudinal disease behaviour and treatment. Ongoing vascular and immune dysfunction are key pathogenetic mechanisms in SSc compared to IPF and thus NO in the context of these two diseases plays a very different role. Whereas IPF may be associated with NO excess, high levels of NO may potentially confer an advantage in SSc due to the other main pulmonary component of the disease – systemic sclerosis associated pulmonary hypertension (SSc-PH).
Of particular relevance to these findings, SSc-ILD has emerged to be genetically distinct from IPF in the published literature. The recently reported MUC5B promoter SNP is strongly and robustly associated with IPF but is not associated with SSc-ILD [Peljto 2012, Stock 2013] and none of the SNPs identified in a large GWASs of idiopathic interstitial pneumonia were associated with SSc-ILD [Wu 2016]. The findings of this exploratory study are supportive of this. Furthermore, they add strength to the evidence of the potential role played by DDAH in IPF.

Overall, although this exploratory study has yielded negative results in relation to the proposed hypothesis, these are useful observations for the SSc-ILD field in general. This is the first time that DDAH SNPs have been known to be genotyped in SSc-ILD. A key strength of this data set is the large cohort size of well-phenotyped patients with longitudinal clinical data from a specialist tertiary referral centre in the UK.
CHAPTER 9: DISCUSSION AND CONCLUSION

9.1 Summary of key findings

The aim of this thesis was to investigate the nature of the relationship between the NO-ADMA-DDAH pathway and the progression of lung fibrosis; more specifically, to determine whether higher DDAH activity would promote disease progression in IPF. The evidence accumulated in the course of this work is summarised as follows:

I. DDAH2 and iNOS are induced during TGF-β mediated EMT
II. NOS inhibition does not attenuate TGF-β mediated EMT
III. DDAH1 SNP rs530006 is associated with accelerated lung function decline and mortality in IPF
IV. DDAH SNPs are not associated with lung function decline in SSc-ILD

9.2 Discussion

9.2.1 In Vitro study

The experimental model of EMT used in the in vitro study corroborated previously published findings whereby TGF-β induced A549 cells to undergo EMT in a time- and concentration-dependent manner [Kim 2007]. The pattern of observations was highly suggestive that the transition from an epithelial to a mesenchymal phenotype was occurring and included the following key findings; loss of E-cadherin, a universal feature of EMT [Hay 1995], induction of α-SMA, a characteristic feature of activated myofibroblasts [Bartis 2014] and the activation of the TGF-β/Smad3 pathway, a key pathway in EMT [Ferrarelli 2014]. Although type 1 collagen gene expression occurred but protein was not detected (for the reasons already discussed), the elevation of MMP-1, MMP-2 and LOXL2 gene expression also supported the transition towards a fibrotic/mesenchymal phenotype. A key limitation was use of the A549 cell line as representative of type 2 AECs, it is clearly not representative of in vivo mechanisms and the results must be interpreted cautiously. However, the pattern of results observed over the course of these experiments indicated that overall, the A549 cell line model was a reproducible in vitro model of EMT as TGF-β induced a loss of the epithelial marker E-cadherin and a gain of several mesenchymal markers in a concentration-dependent manner that was consistent throughout the in vitro study.
The experimental model produced several observations which support that DDAH is important to type 2 AEC biology. DDAH1 was the more abundant isoform in the type 2 AECs at baseline and levels increased only slightly following TGF-β treatment. In striking contrast, levels of DDAH2 were minimal at baseline and following TGF-β treatment, both DDAH2 gene and protein expression increased several fold in a concentration-dependent manner. Pullamsetti et al. have provided previous evidence that DDAH may be an important regulator of type 2 AEC biology; i) treatment of murine type 2 AECs with TGF-β1 significantly induced DDAH2 mRNA expression, ii) pharmacological inhibition of DDAH1 with L-291 and DDAH2-selective siRNA inhibition of DDAH2 both attenuated serum-induced murine type 2 AEC proliferation *in vitro*, and iii) both DDAH isoforms are expressed in healthy type 2 AECs, and in particular, DDAH2 expression is markedly increased in the type 2 AECs of IPF patients and mice with bleomycin-induced fibrosis [Pullamsetti 2011].

An interesting observation was that both DDAH2 and iNOS were induced together. DDAH2 expression is induced several-fold and co-localises with iNOS in the IPF lung, specifically in type 2 AECs in the alveolar epithelium [Pullamsetti 2011]. Induction of iNOS plays a key role in the wound healing model, and upregulation of iNOS within the type 2 AECS has been observed in experimental models of IPF and in the IPF lung [Saleh 1997]. Given the chromosomal position of the *DDAH2* gene in proximity to the major histocompatibility complex III locus, DDAH2 may have an important role in regulating iNOS mediated responses in the type 2 AECs.

Regulation of DDAH activity by endogenous NO is known - a reactive cysteine residue (Cys-249) in the active site of DDAH is directly regulated by S-nitrosylation by NO derived from iNOS [Leiper 2002]. Accordingly, an increase in endogenous NO will potentially increase S-nitrosylation and diminish DDAH activity, which in turn may lead to the accumulation of ADMA and inhibition of NOS to prevent further production of NO. However, ADMA did not appear to accumulate in this model. DDAH1 appears to be the dominant isoform responsible for metabolising and preventing the accumulation of ADMA thus it is interesting that a reduction in the level of the asymmetric MAs was observed with increasing TGF-β treatment in view of the only very slight increase in DDAH1 from baseline.

The reactive cysteine residue in DDAH also confers susceptibility to oxidative and nitrosative stress and thus RNS/ROS generated from ADMA-induced uncoupling of NOS, may also diminish DDAH activity. Thus in terms of the functional significance of
the observed induction in DDAH2 in the experimental model, DDAH activity in the presence of exogenous ADMA would potentially have been negatively affected by oxidative stress. Despite this, TGF-β mediated EMT occurred regardless, suggesting that diminishing DDAH function did not impact the process.

Although not reaching statistical significance, DDAH2 levels appeared to be lower with selective iNOS inhibition and TGF-β co-treatment compared to TGF-β treatment alone. This observation that DDAH2 induction may be responsive to the level of endogenous NO is interesting. Although endogenous NO can regulate DDAH2 activity via the reactive cysteine residue, it is not known whether endogenous NO can affect DDAH2 induction as well. Results have broadly revealed that DDAH2 is a TGF-β responsive gene, it is possible that removal of endogenous NO may have negatively impacted the activation of latent TGF-β and thus transcription of TGF-β target genes, including DDAH2. Attenuated levels of TGF-β mediated DDAH2 induction in a low NO environment suggest that endogenous NO derived from iNOS is required for maximal TGF-β mediated DDAH2 induction. Of note a similar result was not demonstrated with exogenous ADMA treatment. This may be due to other confounding factors such as the RNS/ROS generated from ADMA-induced uncoupling of NOS, as previously discussed.

Linking back to the in vitro hypothesis tested, DDAH activity did indeed appear to be associated with the fibrotic phenotype in TGF-β mediated epithelial to mesenchymal transition. However, the functional significance in terms of whether it was promoting the fibrotic phenotype is unclear. Taken together, if selective iNOS inhibition affected DDAH induction and if exogenous ADMA treatment affected DDAH activity, transition to the mesenchymal phenotype occurred regardless of DDAH2 induction or activity. Although transcriptional modulation of DDAH2 appears to be mediated by TGF-β, the key profibrotic cytokine that drives EMT, the role of DDAH overall in this process remains unclear. Of note a downstream effect that DDAH may have in the transitioned cell has not been explored.

In summary, Pullamsetti et al. had already established that DDAH1 and DDAH2 inhibition both attenuated serum-induced murine type 2 AEC proliferation in vitro [Pullamsetti 2011] but did not investigate transition to a mesenchymal phenotype. The role of DDAH specifically in context of EMT by simultaneously assessing markers to indicate transition has not yet been published and overall, these findings are interesting and the evidence indicates a role for both DDAH isoforms in the type 2 AEC and EMT.
9.2.2 Human Study

The genotyping study was conducted in a relatively small but robust IPF cohort with corresponding longitudinal lung function and mortality data. Importantly, in the era of novel agents for IPF, patients were not receiving treatment with any disease modifying therapies. Mortality is the paramount clinical end point that is often difficult to achieve in clinical trials and thus FVC decline is widely used as a surrogate for mortality; data for both was available for the IPF cohort. This enabled the key question, determining whether genetic variants associated with DDAH activity would impact disease progression in IPF, to be asked (and answered) in two separate ways.

The genotyping results for both DDAH1 SNPs were consistent in their directionality of effect on all three clinical end points. The results for rs530006 emerged as the stronger results both statistically and in terms of clinical importance, particularly the association with shorter length of survival in IPF. Carriage of the minor TT homozygote was hypothesised to be associated with lower ADMA levels and thus greater lung function decline and mortality. There was a significant difference in all three clinical end points according to carriage of the minor TT homozygote on univariate analysis which remained significant following adjustment for significant time varying co-variates on multivariate Cox analysis. The median length of survival with the TT homozygote was 26 months compared to 45.1 months for GG/GT genotypes, indicating a clinically meaningful difference in length of survival for patients with IPF and a particularly aggressive disease progression with carriage of the rs530006 TT homozygote. Overall the rs530006 TT homozygote demonstrated a clear separation in terms of survival curves from the GT and GG genotypes and was associated the worst clinical outcome in all 3 clinical end points which withstood correction in multivariate Cox analysis.

There was no evidence of a significant difference between lung function decline by genotype for either DDAH1 or DDAH2 SNPs in a large cohort of well-phenotyped SSc-ILD patients with longitudinal clinical data from a specialist tertiary referral centre in the UK. The results for the SSc-ILD cohort are important in the context of IPF, as they support the genetic distinction between the two diseases which share the phenotypic feature of pulmonary fibrosis. The MUC5B promoter SNP is strongly associated with IPF but is not associated with SSc-ILD [Peljto 2012] and none of the SNPs identified in a large GWAS of IIPs were associated with SSc-ILD [Wu 2016]. The findings of the SSc-ILD cohort are supportive of this.
The contribution of rs530006 to further understanding IPF pathogenesis, in addition to the finding that it is associated with accelerated disease progression, is important. The biological roles of many of the IPF associated common genetic variants identified by GWAS are unclear and it is possible that the relevant biological effect of individual SNPs including those in DDAH1 are subtle or present only in the context of complex gene–environment interactions. The lung is major source of ADMA and ADMA has been quantified in BALF [Bulau 2007], thus it could be possible that small changes in MA metabolism in the lung due to DDAH variants might well have a significant clinical effect. In many ways, the NO-ADMA-DDAH axis can be viewed as a self-regulating system, and in IPF, it may be the case that a DDAH variant may tip the balance of the system and contribute to an environment which leads to accelerated disease progression.

A high-throughput screen has recently identified proton pump inhibitors (PPIs) as DDAH inhibitors; PPIs bound reversibly to purified human DDAH1, inhibited DDAH activity, increased intracellular ADMA concentration and reduced intracellular NO [Ghebremariam 2013]. A recent study by the same authors postulated this as a plausible mechanism to partly explain the protective effect that PPIs may have in IPF [Ghebremariam 2015]. Abnormal gastroesophageal reflux (GER) has been reported in up to 90% of patients with IPF and is a risk factor for microaspiration-associated lung injury, a mechanism that has been implicated in IPF pathogenesis and disease progression [Tobin 1998, Raghu 2015]. Antacid treatment appears to improve outcome measures of lung function and survival in IPF but the evidence has not been conclusive [Lee 2011, Lee 2013] and thus it has been only conditionally recommended in the current IPF guidelines, acknowledging the potential increased risk for pneumonia with antacid therapy [Raghu 2015]. The findings by Ghebremariam et al. are supportive of the hypothesis that lowering DDAH activity (DDAH inhibition) may benefit IPF patients in the real world. PPIs are inexpensive and widely available drugs, the reported DDAH inhibition was a class effect and included the commonly used agents, omeprazole and lansoprazole. The findings for the DDAH1 variant in this study form a case for PPI use in patients with the rs530006 TT genotype (hence presumed DDAH over-activity) regardless of the results of GER investigations as PPI therapy may confer a clinical benefit in this subsection of patients with a more aggressive disease phenotype.

Overall, these are potentially important findings which need to be cross-checked in a replicate IPF cohort. It is unknown whether the rs530006 TT homozygote is a potential risk allele for IPF and in view of the findings of this study, it will be interesting to
determine this by comparing genotype frequencies in the IPF population with a control cohort. The potential issues with measuring ADMA levels have already been discussed and hence formed the rationale for the use of DDAH variants as a genetic tool. But in view of the hypothesis for the human study, measurement of the ADMA:SDMA ratio as an index of DDAH activity would be interesting to perform in BALF taken at baseline for the IPF cohort in order to assess whether DDAH activity varies by rs530006 genotype. Following corrections to allow for inter-operator/sampling differences, these results would give an indication of DDAH1 variant activity at the single time point at which the samples were taken and would thus however need to be interpreted accordingly.

The MUC5B promoter variant rs35705950 is present in ~50% of individuals with IPF and is recognised as the strongest known genetic risk factor for the development of familial and sporadic IPF but counterintuitively is associated with improved survival [Mathai 2015]. The finding that the rs530006 TT homozygote is linked with shorter survival despite the possibility that 50% of these patients may carry a protective MUC5B allele is important to consider. The MUC5B status of the carriers/non-carriers of the rs530006 TT homozygote is not known. It will be interesting to explore whether the accelerated disease progression and shorter survival is related to MUC5B status, specifically whether the rs530006 results are being confounded by MUC5B or still are significant despite it. This is an emerging conundrum in IPF that is not unique to this study; it is increasingly recognised that clinical trials should consider controlling for the presence of prognosis-modifying genetic variants in addition to controlling for confounding baseline clinical variables, as progression-free survival is an important primary end-point and may be affected by MUC5B status [Mathai 2015]. It is interesting to note that the MUC5B status of the patient populations in the landmark nintedanib and pirfenidone clinical trials is not known.

The genotyping results for rs530006 suggest that depending on allele carriage, the risk of disease progression and mortality is significantly different in IPF patient populations. Being able to identify those patients at risk of accelerated disease progression and mortality is particularly important in the current age of novel therapeutic options for IPF. In those with early IPF and worse genetic risk factors, therapeutic intervention could be taken earlier to slow FVC decline and maintain maximal lung function. In particular, prediction of mortality is challenging in IPF due to the heterogeneity of disease progression, and accurately predicting mortality will determine the urgency of lung transplantation, and be helpful to managing end-of-life decisions.
Additionally, if different genetic variants confer different prognoses and perhaps different molecular mechanisms of disease, it could potentially follow that therapeutic efficacy could vary as well [Mathai 2015]. Different genetic signatures between patients may reflect the underlying biology of the disease and allow a personalised medicine approach to therapy, for example, ‘molecular phenotyping’ in oncology has enabled novel therapies to be targeted at individuals most likely to respond. This is particularly attractive in the context of DDAH, as DDAH inhibition is a potential therapeutic target and the lung is an organ in which topical delivery is possible, thus avoiding systemic side effects of drug administration.

9.3 Further Work

Several observations from the work presented in this thesis deserve further study. The main observation from the in vitro study; the induction of both DDAH2 and iNOS in the type 2 AEC in response to TGF-β, needs further investigation in order to fully understand the mechanisms of induction, the biological roles, the interactions between each and the downstream effects in the transitioned cell/fibrotic phenotype. Both appear to play an important role in the wound healing response of the type 2 AEC, and given the chromosomal position of the DDAH2 gene in close proximity to the major histocompatibility complex III locus, DDAH2 may have an important role in regulating iNOS mediated responses in the type 2 AECs. Although DDAH2 appears to be a TGF-β responsive gene, it is currently unknown whether it contains a Smad3 binding site. Although it is known to affect DDAH2 activity via the reactive cysteine residue in the active site, it is also not known whether iNOS derived endogenous NO can affect induction of DDAH2.

The genotyping results for rs530006 are promising and require validation in a larger, replicate cohort of IPF patients. Taken alone, they support that DDAH may play an important biological role in IPF, and are thus a platform for further investigation of the NO-ADMA-DDAH pathway in IPF. It is unknown whether the rs530006 TT homozygote is a potential risk allele for IPF and it will be interesting to determine this by comparing genotype frequencies with a control cohort. A prospective study to measure BALF ADMA levels in IPF and perhaps more importantly the ADMA:SDMA ratio as an indicator of DDAH activity would be interesting to perform. If ADMA levels are associated with IPF this will reinforce the biological plausibility of this target and BALF
measurements may enable disease stratification. The potential discovery of ADMA as an accurate, non-invasive biomarker in IPF will fulfil several unmet needs; aid in diagnosis, reliably measure disease severity, refine prognostic evaluation, improve disease monitoring and assess therapeutic response.

9.4 Conclusion

The overall hypothesis tested in this thesis was: “higher DDAH activity and lower ADMA levels are predictive of accelerated disease progression in idiopathic pulmonary fibrosis”. The main findings of this thesis support this hypothesis i) DDAH may play a role in the transition towards the fibrotic phenotype in an in vitro study of type 2 AECs and ii) a DDAH1 variant is associated with accelerated lung function decline and mortality in a human study of patients with IPF.

Overall, the work in this thesis has provided some novel insights into the potential role of DDAH in IPF. The aetiology of IPF is complex and heterogeneous, more specifically, the complicated relationship between inherited genetic variants and environmental factors is emerging. DDAH biology may be one of the multiple pathways that converge to the fibrotic phenotype that we know as IPF.
REFERENCES


Wells AU. Any fool can make a rule and any fool will mind it. *BMC Med*. 2016 Feb 10;14:23.


ABSTRACTS AND PRESENTATIONS

National

The nitric oxide-asymmetric dimethylarginine-dimethylarginine dimethylaminohydrolase (NO-ADMA-DDAH) axis in TGF-β mediated epithelial-mesenchymal transition [Abstract]
Lota HK, Leiper JL
Poster discussion at the British Lung Foundation Alumni Conference, London 2016

International

Functional significance of the nitric oxide-asymmetric dimethylarginine-dimethylarginine dimethylaminohydrolase (NO-ADMA-DDAH) axis in TGF-β mediated epithelial-mesenchymal transition [Abstract]
Lota HK, Leiper JL
Poster discussion at the British Thoracic Society Winter Meeting, London 2015

Role of the nitric oxide-asymmetric dimethylarginine-dimethylarginine dimethylaminohydrolase (NO-ADMA-DDAH) axis in TGF-β mediated epithelial-mesenchymal transition [Abstract]
Lota HK, Leiper JL
Poster discussion at the European Respiratory Society Congress, Amsterdam 2015

Future Publications