Lentiviral gene therapy for the treatment of ABCA3-mediated surfactant deficiency

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy



Aimee Rebecca Ruffle

Gene Medicine Research Group

Radcliffe Department of Medicine

and

Wolfson College

University of Oxford

Supervisors: Prof Stephen C. Hyde, Prof Deborah R. Gill and Dr Thomas C. Roberts

Trinity Term 2024

Lentiviral gene therapy for the treatment of ABCA3-mediated surfactant deficiency

Aimee Rebecca Ruffle

Radcliffe Department of Medicine and Wolfson College, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2024

Abstract

Gene therapies are being developed for many rare genetic diseases and provide hope for patients and their families who currently have few treatment options available. Lung surfactant deficiencies comprise a group of rare genetic lung diseases, where babies present at birth with severe respiratory distress requiring mechanical ventilation for survival. Severe surfactant deficiency can result from mutations in the *ABCA3* gene which encodes the ABCA3 lipid transporter. The ABCA3 transporter is present in alveolar type II (ATII) cells, specifically, the limiting membrane of intracellular lamellar bodies (LBs) that assemble and store surfactant prior to secretion onto the lung surface. A recombinant lentiviral (rLV) vector based on Simian Immunodeficiency Virus (SIV) and pseudotyped with the F and HN glycoproteins from Sendai virus (rSIV.F/HN) was selected to accommodate the large ABCA3 cDNA. Candidate rLV genome plasmids were constructed for ABCA3 expression under control of the synthetic CG dinucleotide-free hCEF promoter and confirmed to express ABCA3 protein in a human lung cell line by western blotting. Alternative promoter sequences based on the endogenous promoter of the human *SFTPB* gene were also investigated, using rLVs expressing Green Fluorescent Protein as a reporter transgene. To evaluate the functional activity of ABCA3 expressed from the rLV vector, *in vitro* models of ATII cells were explored under both submerged and 'air-lifted' air-liquid interface culture conditions and gene editing was performed to ablate the *ABCA3* gene. Using transmission electron microscopy on cells cultured

under submerged conditions, LBs could be observed in the human A549 cell line but were not present in human H441 cells; although air-lifting H441 cells resulted in the emergence of some LB structures. An *in vitro* cell detoxification assay was established which showed reduced viability of cells expressing pathogenic variants of ABCA3 compared with wild-type ABCA3, after exposure to the cytotoxic drug doxorubicin. To explore potential to treat ABCA3-surfactant deficiency, highly purified and concentrated rSIV.F/HN.hCEF.ABCA3 vector was produced at large (5l) scale, resulting in vector titres suitable for testing *in vivo*. Intranasal instillation was used to deliver rSIV.F/HN expressing ABCA3 to the lungs of wild-type (BALB/c) mice, confirming expression of vector-derived *ABCA3* and no evidence of acute toxicity. A conditional *Abca3* knockout murine model was also developed where tamoxifen administration resulted in knockout of the *Abca3* gene to induce a respiratory disease phenotype. Overall, rLV vectors expressing ABCA3 are a promising strategy for the treatment of ABCA3-mediated surfactant dysfunction.

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Deborah Gill, Prof. Steve Hyde and Dr. Tom Roberts for their support and commitment to getting me through this DPhil in one piece. I have developed both personally and professionally over the years and am immensely grateful for all your help in making me into the person I am today.

To members of the Gene Medicine Research Group family past and present, I thank you for your friendship, wisdom and continual supply of motivational and anecdotal stories: Dr Altar Munis, Dr Kamran Miah, Dr Arlene Glasgow, Dr Jakob Hansen Haldrup, Dr Hamid Dolatshad, Dr Yue Du, Dr Toby Gamlen, Dr Rosie Munday, Dr Helena Meyer-Berg, Omar Habib, Eoin Mac Réamoinn, Gavin Turnbull, Marina Perez Cerezuela, Galina Boskh, Emily Castells, Shaz Tariq, Mariana Viegas, Rebecca Dean, Catriona Conway, and last but not least Tina Garland and Dr Stephanie Jones. Particular mention must go to, soon-to-be Dr, Wian Satyapertiwi who has been with me every step along the way, navigating the same bumpy and seemingly never-ending road to becoming a Doctor. It would be remise of me not to mention some of the good times we have shared – be that the annual barbeque with famous homemade limoncello, spontaneous cake meetings in the tearoom or the almost weekly debate on what is the most superior doughnut type.

Outside of the lab, my home in Oxford was found in the Wolfson College Boat Club. From my early days as a new recruit, to becoming a captain, and then president, through times of flooding and covid you have always been there to convince me that going out on the water on a dark, wet and cold morning would be an enjoyable experience. Competing in Bumps races both as a rower and a cox has been a highlight of my time at Oxford which I will look back on with fond memories.

From WCBC I found Sophie and Kirsty who have become some of my closest friends. To my wonderful friend Sophie, I thank you for regularly feeding me, making me laugh and listening to me complain about anything and everything. To Kirsty, since becoming housemates, and then cat parents, I have grown used to our evening catchups and slightly chaotic way of life. Tiggy and

Winnie, you never fail to put a smile back on my face at the end of a long day and I would like to think you will miss me when I am no longer there for you to annoy.

A thank you should also go to ‘The United Nations of Fox Hill’ – friends for life found during our undergraduate degrees. We may no longer live together but that has not stopped us from staying in close contact and supporting each other.

The final thank you belongs to my family. Your unwavering belief in me means the world and I would not be where I am today without you. To Grandma, the little letters and money ‘for the bottle bank’ are always a lovely surprise. To Mum, you have convinced me not to give up more times than I can count and I have no words to explain how grateful I am for everything you do for me.

Declaration

This thesis has been composed by myself and has not been used in any previous application for a degree. The results presented were obtained by myself and contributions made by others have been acknowledged within the text. All sources of published or unpublished information have been specifically acknowledged by means of a reference of personal communication notation.

List of Abbreviations

A549	Human lung carcinoma 549
ABC	ATP-binding cassette
ABCA3	ABC subfamily A member 3
Abca3-cKO	Conditional *Abca3* knockout
Abca3-iKO	Inducible *Abca3* knockout
ABCA4	ABC subfamily A member 4
ABCC1	ABC subfamily C member 1
ADP	Adenosine diphosphate
AEX	Anion exchange chromatography
ALI	Air-liquid interface
ATCC	American Type Culture Collection
ATI	Alveolar type I
ATII	Alveolar type II
ATMP	Advanced therapeutic medicinal product
ATP	Adenosine triphosphate
AUC	Area under the curve
BCA	Bicinchoninic acid
BLI	Bioluminescence imaging
BSA	Bovine serum albumin
CAR-T	Chimeric antigen receptor T-cell
CBh	Constitutive hybrid
cDNA	Complementary DNA
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
chILD	Childhood ILD
CMV	Cytomegalovirus-immediate early
CMV-	CMV with introns removed
CNV	Copy number variation
CPAP	Continuous positive airway pressure
CpG	Cytosine-guanine dinucleotide
cPPT	Central polypurine tract
CRISPR	Clustered regularly interspaced short palindromic repeats
Ct	Cycle threshold
DAPI	4’,6-Diamidino-2-phenylindole
ddPCR	Droplet digital PCR
DMEM	Dulbecco’s Modified Eagle’s Medium
DMSO	Dimethyl-sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
dPBS	Dulbecco's PBS
dPLA	Digital proximity ligation assay
DPPC	Dipalmitoylphosphatidylcholine
E.coli	*Escherichia coli*
e.g.	*Exempli gratia* (for example)
ECD	Extracellular domain
Ef1α	Elongation factor-1 alpha
EMA	European Medicines Agency
ER	Endoplasmic reticulum

ERQC	ER quality control system
ESC	Embryonic stem cell
EYFP	Enhanced YFP
F-12K	Kaighn's Modification of Ham's F-12 Medium
F/HN	Fusion and haemagglutinin/neuraminidase glycoproteins from Sendai virus
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
FDA	US Food and Drug Administration
FIV	Feline immunodeficiency virus
fLux	Firefly luciferase
gDNA	Genomic DNA
GFP	Green fluorescent protein
gRNA	Guide RNA
H&E	Harris haematoxylin and Eosin Y
H441	Human papillary adenocarcinoma 441
HBSS	Hank's balanced salt solution
hCEF	CMV enhancer with Ef1α promoter
HEK293	Human embryonic kidney 293
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HF	High-fidelity
HI	Heat inactivated
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase
ICC	Immunocytochemistry
ICE	Inference of CRISPR Edits
IHC	Immunohistochemistry
ILD	Interstitial lung disease
ILV	Intralumenal vesicle
Indel	Insertion-deletion
IP	Intraperitoneal
iPSC	Induced pluripotent stem cell
ITR	Inverted terminal repeat
ITS	Insulin-transferrin-selenium
KGF	Keratinocyte growth factor
KI	Knock-in
KO	Knockout
LAMP3	Lysosomal associated membrane protein 3
LB	Lamellar body
LMP	Ultra-low melting point
LoxP	Locus of X(cross)-over in P1
LS	Large-scale
LTR	Long terminal repeat
$MgCl_2$	Magnesium chloride
MHRA	Medicine and Healthcare products Regulatory Agency
MOI	Multiplicity of infection
MOPS	3-Morpholinopropane-1-sulfonic acid
mRNA	Messenger RNA
mSPB	Minimal SPB
MTT	3-(4, 5-Dimethylthiazol-2-yl)-2, 5-diphenyltatrazolium bromide
MVB	Multivesicular body
NaCl	Sodium chloride

NBD	Nucleotide-binding domain
NFkB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NFW	Nuclease-free water
NGS	Normal goat serum
NHEJ	Non-homologous end joining
NKX2.1	NK2 homeobox 1
NMD	Nonsense-mediated mRNA decay
NTC	No template control
NZY+	Casein enzymatic hydrolysate yeast extract
OCT	Optimal cutting temperature
OD	Optical density
p	Plasmid
PBS	Phosphate-buffered saline
PCLS	Precision-cut lung slices
PCR	Polymerase chain reaction
pDNA	Plasmid DNA
PFA	Paraformaldehyde
PIPES	1,4-Piperazinediethanesulfonic acid
qRT-PCR	Quantitative reverse transcription PCR
rAAV	Recombinant adeno-associated virus
rAdV	Recombinant adenovirus
RD	Regulatory domain
RDS	Respiratory distress syndrome
rLV	Recombinant lentivirus
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
ROCK	Rho kinase
ROI	Region of interest
RPM	Rotations per minute
RPMI-1640	Roswell Park Memorial Institute 1640 Medium
RT	Room temperature
rtTA	Reverse tetracycline-controlled transactivator protein
SALI	Surfactant air-liquid interface
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
sgRNA	Synthetic gRNA
SIN	Self-inactivating
SIV	Simian immunodeficiency virus
SIV_{AGM}	SIV from African green monkey
SIV_{CPZ}	SIV from chimpanzee
SIV_{SM}	SIV from sooty mangabey
SOC	Super Optimal broth with Catabolite repression
SP-A	Surfactant protein A
SP-B	Surfactant protein B
SP-C	Surfactant protein C
SP-D	Surfactant protein D
SpCas9	*Streptococcus pyogenes* Cas9
SS	Small-scale
TAE	Tris-acetate-ethylenediaminetetraacetic acid
TBS	Tris-buffered saline

TEER	Transepithelial electrical resistance
TEM	Transmission electron microscopy
TEPD	Transepithelial potential difference
TFF	Tangential flow filtration
T_m	Melting temperature
TMD	Transmembrane domain
TOPO	Topoisomerase-based
TRAC	T cell receptor α constant
TRE	Tetracycline-responsive element
TTF	Thyroid transcription factor
TTU	TaqMan transducing units
UK	United Kingdom
UPR	Unfolded protein response
USA	United States of America
UV	Ultraviolet
v/v	Volume: volume ratio
VCN	Vector copy number
VSV-G	Vesicular stomatitis virus
w/v	Weight: volume ratio
WPRE	Woodchuck hepatitis virus post-transcriptional regulatory element
WT	Wild-type
YFP	Yellow fluorescent protein

Table of Contents

Abstract ..... **I**
Acknowledgements ..... **III**
Declaration ..... **V**
List of Abbreviations ..... **VI**
Table of Contents ..... **X**
List of Figures ..... **XVI**
List of Tables ..... **XIX**
Chapter 1 Introduction **1**
1.1 Structure of the lung **1**
1.2 Role of pulmonary surfactant **3**
1.2.1 Surfactant composition 3
1.2.1.1 Surfactant proteins *3*
1.2.1.2 Surfactant lipids *4*
1.2.2 Surfactant homeostasis 8
1.2.3 ABCA3 11
1.2.3.1 Synthesis of ABCA3 ..... *11*
1.2.3.2 Structure of ABCA3 ..... *11*
1.3 Surfactant dysfunction **13**
1.3.1 Clinical presentation and current treatment 13
1.3.1.1 Respiratory distress syndrome *13*
1.3.1.2 Childhood interstitial lung disease *15*
1.3.2 Genetic disorders of surfactant dysfunction 15
1.3.2.1 SP-B deficiency *16*
1.3.2.2 SP-C-associated disease *18*
1.3.2.3 ABCA3-mediated surfactant dysfunction *18*
1.3.2.4 TTF-1-related lung disease *19*
1.4 Models of surfactant dysfunction **20**
1.4.1 Cellular models of the human lung parenchyma 20
1.4.1.1 Cell selection *20*
1.4.1.2 Air-liquid interface culture *22*
1.4.1.3 Alveolospheres *23*
1.4.1.4 Lung bud organoids *24*
1.4.2 Mouse models of *ABCA3*-mediated surfactant dysfunction 24
1.4.2.1 Knockout models *26*
1.4.2.1.1 Constitutive Abca3 knockout *26*

1.4.2.1.2 Conditional Abca3 knockout 27
1.4.2.2 Expression of ABCA3 mutations 31
1.5 Gene therapy for surfactant dysfunction 32
1.5.1 Fundamentals of gene therapy 33
1.5.1.1 Strategies for transgene delivery 34
1.5.2 Current gene therapy strategies for surfactant dysfunctions 36
1.5.2.1 SP-B deficiency 36
1.5.2.2 SP-C-associated disease 37
1.5.2.3 ABCA3-mediated surfactant dysfunction 38
1.5.3 Lentiviral vector-based gene therapy 38
1.5.3.1 Lentivirus biology 38
1.5.3.2 Recombinant lentiviral vectors 42
1.5.3.3 Current recombinant lentiviral vectors for ABC proteins 45
1.6 Aims 46
Chapter 2 Materials and methods 47
2.1 Common reagents and solutions 55
2.2 Molecular biology 55
2.2.1 Restriction enzyme digest 55
2.2.2 Polymerase chain reaction 56
2.2.3 Annealing guide RNA oligos 56
2.2.4 Gel electrophoresis 56
2.2.5 Gel extraction and PCR purification 57
2.2.6 DNA Ligation 57
2.2.7 Topoisomerase-based cloning 58
2.2.8 Site-directed mutagenesis 58
2.2.9 Preparation of selective Luria-Bertani solutions 58
2.2.10 Bacterial transformation 59
2.2.11 Colony PCR 59
2.2.12 Bacterial inoculation of growth media 60
2.2.13 Plasmid purification 60
2.2.14 Quantification of DNA and RNA 60
2.2.15 Sanger sequencing 60
2.2.16 Inference of CRISPR edits analysis 61
2.2.17 Cre recombination assay 61
2.3 Mammalian cell culture 61
2.3.1 Cell lines and culture conditions 61
2.3.1.1 Immortalised cells 62
2.3.1.2 Primary cells 63

2.3.2 Air-liquid interface cultures 63
2.3.2.1 Establishment of air-liquid interface cultures 63
2.3.2.2 Harvesting cells from air-liquid interface cultures 63
2.3.3 Transfection 64
2.3.4 Puromycin selection 64
2.3.5 Doxorubicin assay 65
2.4 Recombinant lentiviral vector production 65
2.4.1 Recombinant LV production in LV-MAX cells 66
2.4.2 Recombinant LV production in suspension-adapted HEK293T-17 cells 66
2.4.2.1 Small-scale (25ml) 66
2.4.2.2 Large-scale (5l) 66
2.4.3 Purification and concentration 67
2.4.3.1 Centrifugation 67
2.4.3.2 Anion exchange chromatography 67
2.4.3.3 Tangential flow filtration 67
2.4.4 Determination of viral titre 68
2.5 Gene expression analysis 68
2.5.1 DNA extraction 68
2.5.2 RNA extraction 68
2.5.3 Protein extraction 69
2.5.4 Quantitative reverse transcription PCR 69
2.5.5 Droplet digital PCR 69
2.5.6 Western blot 70
2.6 Histology 71
2.6.1 Cryosectioning 71
2.6.2 Haematoxylin and Eosin staining 71
2.6.3 Native fluorescence imaging 72
2.6.4 Immunostaining 72
2.6.4.1 Immunocytochemistry 72
2.6.4.2 Immunohistochemistry 73
2.6.5 Microscopy 73
2.6.5.1 Brightfield and fluorescence microscopy 73
2.6.5.2 Transmission electron microscopy 74
2.7 Animal studies 75
2.7.1 Maintenance and housing 75
2.7.2 Mouse strains 75
2.7.3 Inducible *Abca3* knockout model 76
2.7.3.1 Rederivation and breeding 76

2.7.3.2 Genotyping *77*
2.7.3.3 Tamoxifen administration *77*
2.7.3.3.1 Intraperitoneal injection *77*
2.7.3.3.2 Diet *78*
2.7.3.4 Animal monitoring *78*
2.7.4 Mouse anaesthesia 79
2.7.5 Intranasal administration of substances 79
2.7.6 *In vivo* bioluminescence imaging 79
2.7.7 Lung tissue harvesting 79
2.8 Precision cut lung slices 80
2.8.1 Generation of PCLS 80
2.8.2 PCLS culture 81
2.8.3 Viability assay 81
2.9 Statistical analysis 81
Chapter 3 Generation of recombinant lentivirus expressing ABCA3 82
3.1 Introduction 82
3.1.1 Considerations for production of recombinant lentivirus 82
3.1.2 Aims 83
3.2 Results 85
3.2.1 Design and construction of *ABCA3* rLV genome plasmids 85
3.2.1.1 Transgene design *85*
3.2.1.2 Promoter selection for rSIV.ABCA3 genome plasmid *85*
3.2.1.3 rSIV.ABCA3 genome plasmid construction *88*
3.2.2 Transient transfection of human immortalised cell lines with *ABCA3* rLV genome plasmids 88
3.2.2.1 Development of a western blot protocol for ABCA3 ..... *88*
3.2.2.2 Expression of ABCA3 from rLV genome plasmids *93*
3.2.3 Production and characterisation of rSIV.F/HN.hCEF.ABCA3 93
3.2.3.1 Comparison of rLV production platforms using rSIV.F/HN.hCEF.GFP *93*
3.2.3.2 Manufacturing yield of rSIV.F/HN with different transgenes *96*
3.2.3.3 Expression of ABCA3 from rSIV.F/HN.hCEF.ABCA3 ..... *99*
3.2.4 Assessment of alveolar specific promoters 101
3.2.4.1 Assessment of alveolar specificity by in vivo BLI *103*
3.2.4.2 Long-term expression profile of newly developed promoters *108*
3.3 Discussion 108
3.3.1 Improving lentiviral vector yields 108
3.3.2 Lentiviral vector design for clinical translation 111
3.3.3 Conclusion 113

Chapter 4 Development of cell models for ABCA3-mediated surfactant dysfunction 114

4.1 Introduction 114

4.2 Results 115

4.2.1 Primary alveolar pneumocytes as a model of alveolar type II cells 115

4.2.2 Evaluation of immortalised lung cell lines 119

4.2.3 Development of an *ABCA3* knockout cell line 127

4.2.3.1 gRNA selection and cloning 127

4.2.3.2 Approach 1: Large fragment deletion strategy 129

4.2.3.3 Approach 2: Disruption of exon 5 131

4.2.4 Expression of clinically relevant ABCA3 variants in H441 cells 134

4.2.4.1 Construction of rLV genome plasmids expressing pathogenic ABCA3 variants 134

4.2.4.2 Development and characterisation of ABCA3 variant cell lines 136

4.2.4.3 Impact of mutations on ABCA3 expression and function 139

4.2.5 Expression of functional ABCA3 from rSIV.F/HN.hCEF.ABCA3 in cell culture 140

4.3 Discussion 143

4.3.1 Maintaining an ATII cell phenotype in primary culture 143

4.3.2 Selection of H441 cells as an ATII cell model 145

4.3.3 Gene editing in H441 cells 146

4.3.4 Protein processing of pathogenic ABCA3 variants 148

4.3.5 Measuring ABCA3 transporter function 149

4.3.6 Conclusion 150

Chapter 5 *In vivo* lentiviral gene transfer for ABCA3-mediated surfactant dysfunction 151

5.1 Introduction 151

5.2 Results 152

5.2.1 Evaluation of rSIV.F/HN.hCEF.ABCA3 in BALB/c mice 152

5.2.2 Establishment of a conditional *Abca3* knockout mouse model 157

5.2.2.1 Breeding mice containing Sftpc-CreER^{T2}, floxed Abca3 and EYFP cassettes to generate a colony of Abca3-iKO mice 157

5.2.2.2 Characterisation of Abca3-iKO transgenic cassettes 160

5.2.3 Natural history of uninduced *Abca3*-iKO mice 162

5.2.4 Induction of a disease phenotype in *Abca3*-iKO mice 166

5.2.4.1 Approach 1: Intraperitoneal injection of tamoxifen 169

5.2.4.2 Approach 2: Delivery of tamoxifen in diet 179

5.3 Discussion 188

5.3.1 Molecular analysis of murine lung tissue 188

5.3.2 Tamoxifen route of administration 189
5.3.3 Disease phenotype in *Abca3*-iKO mice 191
5.3.4 Conclusion 192
Chapter 6 Discussion 193
6.1 Utility of a conditional *Abca3* knockout mouse for evaluating new therapeutics 193
6.2 The ideal murine model of ABCA3-mediated surfactant dysfunction 197
6.3 Improved animal models for respiratory diseases 199
6.4 Alternatives to animal models 201
6.5 Assays to assess the expression and function of ABCA3 204
References 208

List of Figures

Figure 1.1 Schematic of the human lung 2

Figure 1.2 Composition of pulmonary surfactant 5

Figure 1.3 Lamellar body formation 6

Figure 1.4 Surfactant homeostasis 9

Figure 1.5 Diagram of *ABCA3* processing and structure 12

Figure 1.6 Electron microscopy of lamellar bodies in human ATII cells 17

Figure 1.7 Schematic of the murine lung 25

Figure 1.8 Mechanism of tissue-specific, conditional *Abca3* knockout mice 28

Figure 1.9 Life cycle of HIV-1 40

Figure 1.10 Structure of the HIV-1 proviral genome 41

Figure 1.11 Schematic representation of the three generations of recombinant lentiviral vectors 43

Figure 3.1 Diagram of *ABCA3* processing and transgene design 84

Figure 3.2 Generation of recombinant lentiviral vector genomes encoding ABCA3 86

Figure 3.3 Structure of recombinant lentiviral vector genome plasmids encoding ABCA3 89

Figure 3.4 Development of assays to detect ABCA3 expression 91

Figure 3.5 Western blot analysis to confirm expression of ABCA3 from lentiviral vector genomes 94

Figure 3.6 Immunocytochemistry to confirm expression of ABCA3 from pCMV-.ABCA3 lentiviral genome plasmid 95

Figure 3.7 Comparison of production methods for producing recombinant lentivirus expressing GFP 97

Figure 3.8 Production of high yield lentivirus expressing ABCA3 98

Figure 3.9 ABCA3 expression following transduction of H441 cells with rSIV.F/HN 100

Figure 3.10 Design of novel promoters based on endogenous human *SFTPB* promoter (Altar Munis) 102

Figure 3.11 *In vivo* bioluminescence imaging of BALB/c mice dosed with recombinant lentivirus expressing *firefly* luciferase 105

Figure 3.12 Alveolar specificity of novel promoters 107

Figure 3.13 Expression profile of novel promoters in murine lung 109

Figure 4.1 Electron microscopy of human primary alveolar pneumocytes showing lamellar body structures....116

Figure 4.2 Immunofluorescence microscopy analysis of an ATII cell marker in alveolar pneumocytes 117

Figure 4.3 RNA expression of surfactant protein genes *ABCA3* and *SFTPB* in immortalised lung cell lines 120

Figure 4.4 Endogenous ABCA3 protein expression in immortalised lung cell lines....122

Figure 4.5 Electron microscopy of immortalised lung cell lines showing lamellar body structures 124

Figure 4.6 Electron microscopy of H441 SALI cultures over time....126

Figure 4.7 Genotyping potential H441 *ABCA3* knockout clones by endpoint PCR 130

Figure 4.8 Transfection of H441 cells with sgRNA-spCas9 RNP complexes targeting *ABCA3* exon 5 produces indels....132

Figure 4.9 Evaluation of *ABCA3* KO in a H441 cell clone using Sanger sequencing (Hamid Dolatshad) 135

Figure 4.10 Generation of recombinant lentiviral vector genomes encoding pathogenic ABCA3 variants....137

Figure 4.11 PCR amplification to confirm insertion of *ABCA3* variants into the H441 genome 138

Figure 4.12 Immunoblotting of H441 cells transduced with rHIV virus expressing ABCA3 variants 141

Figure 4.13 Cell viability of H441 ABCA3 variant cell lines in response to doxorubicin....142

Figure 4.14 Cell viability of H441 ABCA3 variant cell lines exposed to cytotoxic doxorubicin can be rescued by transduction with rSIV.F/HN.hCEF.ABCA3 144

Figure 5.1 H&E staining of WT BALB/c murine lungs transduced with recombinant lentivirus rSIV.F/HN.hCEF.ABCA3 expressing human ABCA3 153

Figure 5.2 Expression of murine *Abca3* and human *ABCA3* mRNA in mouse lungs transduced with rSIV.F/HN.hCEF.ABCA3 155

Figure 5.3 Expression of ABCA3 protein in BALB/c mice dosed with rSIV.F/HN.hCEF.ABCA3 156

Figure 5.4 Breeding strategy to generate a conditional *Abca3* knockout mouse model....158

Figure 5.5 Sanger sequencing traces of *Abca3*-iKO transgenic cassettes....161

Figure 5.6 Recombination of floxed *Abca3* pSP6.Abca3in3-7(*Abca3*-iKO) *in vitro* 163

Figure 5.7 Kaplan-Meier survival curve for uninduced *Abca3*-iKO mice 165

Figure 5.8 *Sftpc-CreER*^{T2} expression is leaky in uninduced *Abca3*-iKO mice....167

Figure 5.9 Weight and temperature of *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection 170

Figure 5.10 Evidence of pathology in lungs harvested from *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection 171

Figure 5.11 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection 173

Figure 5.12 Weight and temperature of *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection 176

Figure 5.13 No evidence of lung pathology in *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection 177

Figure 5.14 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection 178

Figure 5.15 Weight of *Abca3*-iKO mice administered a tamoxifen diet (0.25g/kg) for 8 or 14 days 180

Figure 5.16 Evidence of pathology in lungs harvested from male *Abca3*-iKO mice administered a tamoxifen diet for 8 or 14 days 181

Figure 5.17 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered a continuous tamoxifen diet 184

Figure 5.18 Weight of *Abca3*-iKO mice administered a pulsed tamoxifen diet (0.25g/kg) for a total of 15 days 186

Figure 5.19 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered a pulsed tamoxifen diet for a total of 15 days 187

Figure 6.1 *Abca3*-iKO male mice are more susceptible to disease induction with tamoxifen delivered by oral gavage compared with females (Kamran Miah) 196

List of Tables

Table 1.1 Features of diseases associated with pathogenic mutations in surfactant-related genes. 16

Table 1.2 Key characteristics of viral vector delivery platforms 36

Table 2.1 List of primers used in this study 47

Table 2.2 List of plasmids used in this study. 51

Table 2.3 List of recombinant lentiviral vectors used. 54

Table 2.4 List of cell lines and standard culture conditions used 62

Table 2.5 Preparation of Bis-Tris polyacrylamide gels 71

Table 2.6 List of antibodies used 71

Table 2.7 Transgenic alleles in *Abca3*-iKO mice 76

Table 3.1 Sequence of promoters used in rLV genome plasmids. 87

Table 4.1 Guide RNA sequences targeting *ABCA3* 128

Table 4.2 Selection criteria for SpCas9 gRNAs targeting *ABCA3* 128

Table 4.3 Editing efficiency of gRNAs targeting *ABCA3* in H441 cells. 129

Table 4.4 Copy number analysis of H441 ABCA3 variant cell lines. 139

Chapter 1 Introduction

1.1 Structure of the lung

The lungs function to supply oxygen and remove waste gases from the body, in addition to having auxiliary roles in defence against pathogens and particulate matter. Upon inhalation, oxygen travels down through the respiratory tree (Figure 1.1) until it reaches alveoli where gas exchange between inspired air and blood occurs. Collectively, the functional portion of the lung involved in gas exchange is known as the lung parenchyma. The lung parenchyma accounts for 80-90% of total lung volume [1] and comprises a large network of alveoli to maximise the surface area for gas exchange. On average, adult lungs contain an estimated 480 million alveoli [2] which cover a surface area of about $70m^2$ [3]. The number of alveoli increase as the lungs develop during alveolarization which begins at about 36 weeks gestation [4]. Alveolarization is considered largely complete at 2-3 years old in humans although there is evidence to suggest it continues at a slower rate into adolescence [4, 5].

Alveoli contain two specialised epithelial cell types – alveolar type I (ATI) and type II (ATII) pneumocytes that were first identified in 1953 [6]. Together they form a single layer epithelium which is covered in a thin layer of aqueous liquid – supporting both hydration of the epithelial layer and movement of immune cells such as alveolar macrophages. The ATI cells mediate gas exchange with a large, squamous morphology covering 95-98% of the alveolar surface to facilitate passive diffusion of gases into surrounding capillaries [7]. By contrast, ATII cells are small and cuboidal, accounting for 16% of total alveolar cells but only 5% of the alveolar surface area in the healthy adult lung [8]. The ATII cells perform a vital function in preventing alveolar collapse (Section 1.2.2) and act as progenitors for ATI cells, transdifferentiating after injury to regenerate the ATI cell population [9, 10].

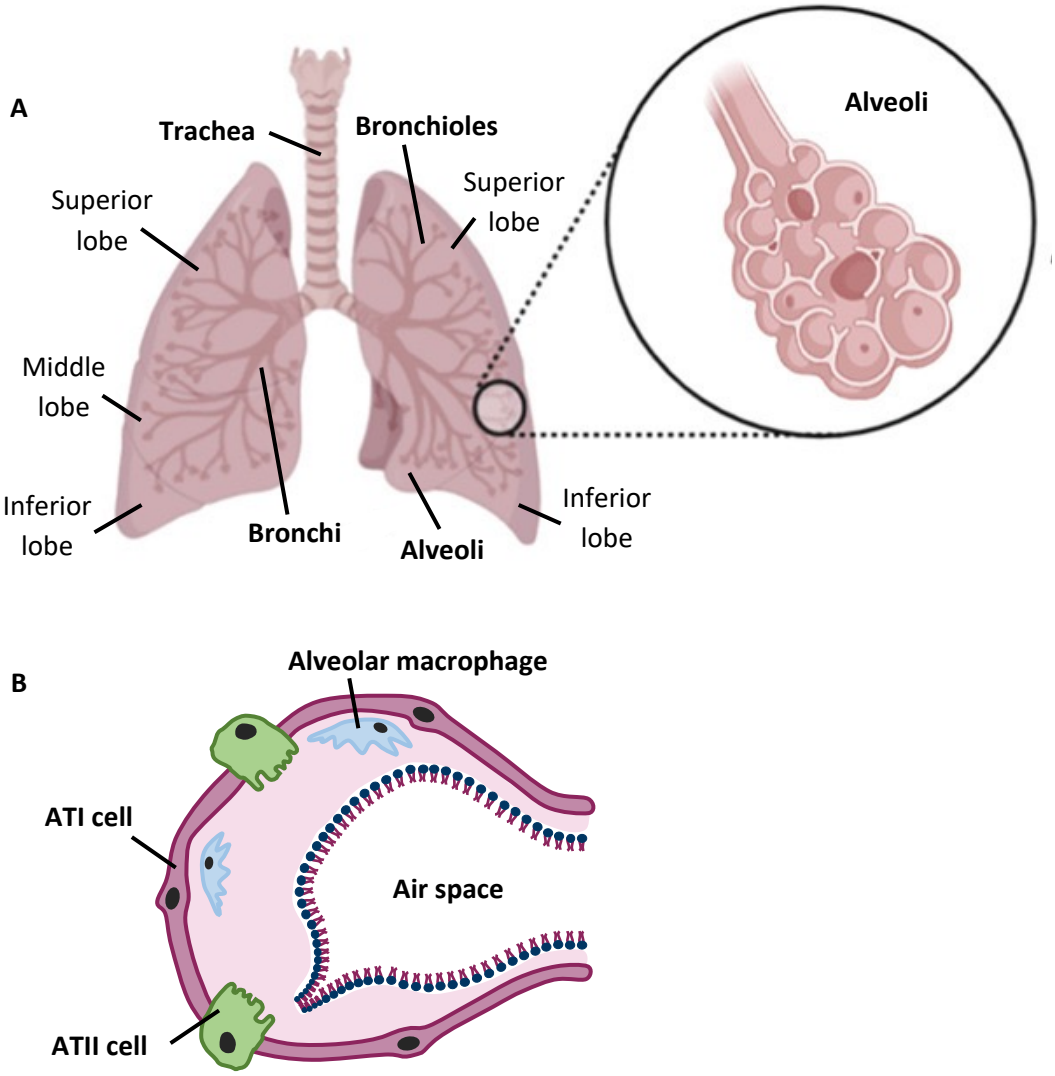


Figure 1.1 Schematic of the human lung

(A) The human lung is organised into five lung lobes; the right lung consists of three lobes (superior, middle, inferior) whilst the left lung consists of two lobes (superior, inferior). Upon inhalation, oxygen travels down the trachea and into the bronchi of each lung before progressing through a series of decreasing diameter bronchioles until they reach alveoli at the end of the respiratory tree. **(B)** Alveoli are the structural units of the lung parenchyma. They contain squamous alveolar type I (ATI) cells, which mediate gas-exchange with capillaries, and cuboidal alveolar type II (ATII) cells which produce surfactant to line the alveolar air surface.

The aqueous liquid covering the alveolar epithelium creates surface tension at the air-liquid interface (ALI), which is a particular physiological challenge at end-expiration when the alveolar surface area is at its smallest. To prevent alveolar collapse (atelectasis) surface tension must remain below 2mN/m [11]. Under physiological conditions, ATII cells produce pulmonary surfactant [12] which reduces surface tension to below 1mN/m at end-expiration [11], thus preventing alveolar collapse.

1.2 Role of pulmonary surfactant

1.2.1 Surfactant composition

Pulmonary surfactant is a surface-active, lipid/protein complex which forms a film on the air-facing surface of the alveolar fluid [13, 14]. Its role in reducing surface tension at end-expiration makes it critical for the maintenance of lung volumes during the respiratory cycle. An absence of functional surfactant results in respiratory distress syndrome (RDS), atelectasis and ultimately respiratory failure [15, 16].

1.2.1.1 Surfactant proteins

Surfactant is primarily composed of phospholipids and four key proteins – surfactant protein (SP)-A, SP-B, SP-C and SP-D (Figure 1.2) [17]. By molecular weight, SPs account for only 10% of surfactant but have important roles in surfactant structure, function and metabolism [18]. They are synthesised primarily in ATII cells but are also expressed at lower levels by non-ciliated, secretory club cells of the bronchiolar epithelium– the exception is SP-C which is considered ATII cell specific [19, 20]. Surfactant proteins -A and -D are hydrophilic [21], oligomeric proteins predominantly involved in modulating immune responses to pathogens and innate immune system activation [22]. They belong to the collectin family and contain a C-terminal lectin domain that binds carbohydrates found on bacterial and viral surfaces in a calcium-dependent manner [22, 23]. Unlike other surfactant components, SP-A and SP-D are freely secreted from ATII cells into the alveolar space and their presence has also been detected outside the lung at barrier surfaces of other organs [24, 25].

Surfactant proteins -B and -C are small, highly hydrophobic proteins [21], which associate with surfactant lipids in specialised lysosome-related organelles termed lamellar bodies (LBs) before secretion from ATII cells into the alveolar space [26]. In LBs, SP-B organises surfactant lipids into tightly packed concentric rings of membrane sheets (lamellae) reminiscent of an onion or tree cross-section when viewed by transmission electron microscopy (TEM; Figure 1.3) [26]. First synthesised as proproteins in the endoplasmic reticulum (ER) of ATII cells, SP-B and SP-C transition through the Golgi apparatus and into multivesicular bodies (MVBs) where they undergo a final cleavage event to generate mature proteins [26]. Within MVBs, SP-B and SP-C are localised to small intralumenal vesicles (ILVs) generated from the inward vesiculation of the MVB limiting membrane. The ILVs containing SP-B and SP-C become packaged in LBs when MVB and LB membranes fuse. The SP-B is essential for LB organisation and SP-C aids in rapid surfactant adsorption to the ALI; in conjunction with surfactant lipids, they contribute to efficient surface tension reduction [27, 28].

1.2.1.2 Surfactant lipids

Phospholipids are the most abundant component of surfactant (Figure 1.2) and a large contributor to its tension lowering properties [29]. 70% of surfactant lipids are composed of phosphatidylcholine [21, 30], of which 50% exists as dipalmitoylphosphatidylcholine (DPPC) [30]. The DPPC variant of phosphatidylcholine is the main contributor to surface tension reduction [29] as the saturated acyl chains can be tightly packed laterally in the surfactant film when the alveolar surface area is decreased during exhalation. Phospholipids are synthesised in the ER and transported into LBs by adenosine triphosphate (ATP)-binding cassette (ABC) subfamily A member 3 (ABCA3) glycoproteins which reside on the LB limiting membrane (Figure 1.3) [31].

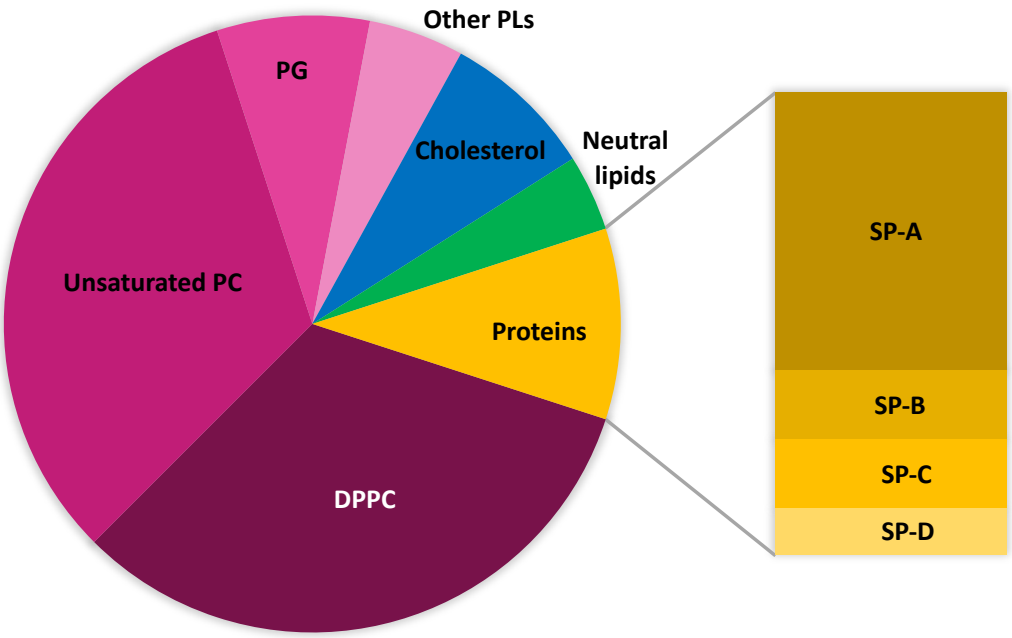


Figure 1.2 Composition of pulmonary surfactant

Pulmonary surfactant is comprised of 10% surfactant proteins (SPs; yellow) and 90% lipids, of which phospholipids (PLs; pink) account for approximately 75%. Unsaturated phosphatidylcholine (PC), the saturated PC form dipalmitoylphosphatidylcholine (DPPC) and phosphatidylglycerol (PG) are the three major PLs.

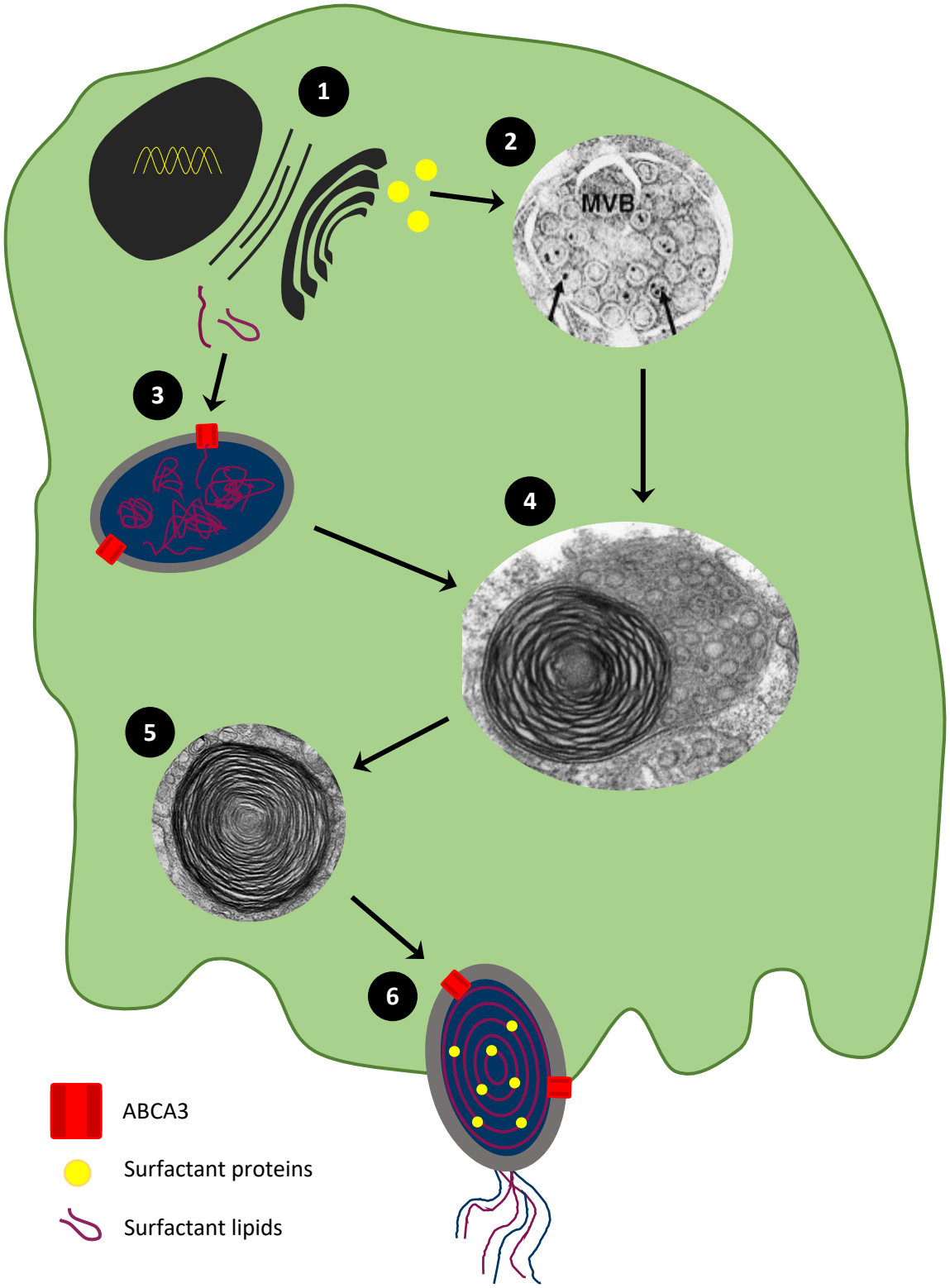


Figure 1.3 Lamellar body formation

Figure 1.3 Lamellar body formation

Surfactant proteins and phospholipids are **(1)** synthesised in the endoplasmic reticulum of ATII cells. Proproteins of SP-B and SP-C are packaged into intraluminal vesicles (ILVs) inside multivesicular bodies (MVBs) where they are cleaved to form the mature proteins. **(2)** Cryosection of foetal mouse lung stained for SP-C. SP-C localises to ILVs inside the MVB (adapted from [32]). **(3)** Phospholipids are transported into immature lamellar bodies (LBs) by ABCA3 which subsequently fuse with MVBs, allowing SP-B- and SP-C-containing ILVs to be incorporated. **(4)** Electron microscopy of murine lung showing fusion of an LB with a MVB (adapted from [33]). Inside LBs, surfactant lipids are organised into concentric rings of lamellae by SP-B. **(5)** Electron microscopy of a human LB obtained from healthy infants [34]. **(6)** Fusion of LBs with the cell membrane leads to release of surfactant into the alveolar space.

1.2.2 Surfactant homeostasis

Surfactant homeostasis at the ALI is maintained by its biosynthesis and secretion from ATII cells, in addition to its reuptake into ATII cells for recycling or alveolar macrophages for degradation (Figure 1.4) [35]. Surfactant release from LBs into the alveolar space occurs via exocytosis [36] when an elevation in cytoplasmic calcium triggers fusion of an LB with the ATII cell membrane. An increase in calcium is triggered by cell stretch during alveoli inflation [37]. Once in the alveolar space, surfactant components are rearranged into a lattice-like structure (tubular myelin) before undergoing a final rearrangement at the ALI to form the functional monolayer film of surfactant (Figure 1.4) [38]. Inactive small surfactant aggregates from the monolayer film are either catabolised in alveolar macrophages or endocytosed back into ATII cells and repackaged in MVBs [35].

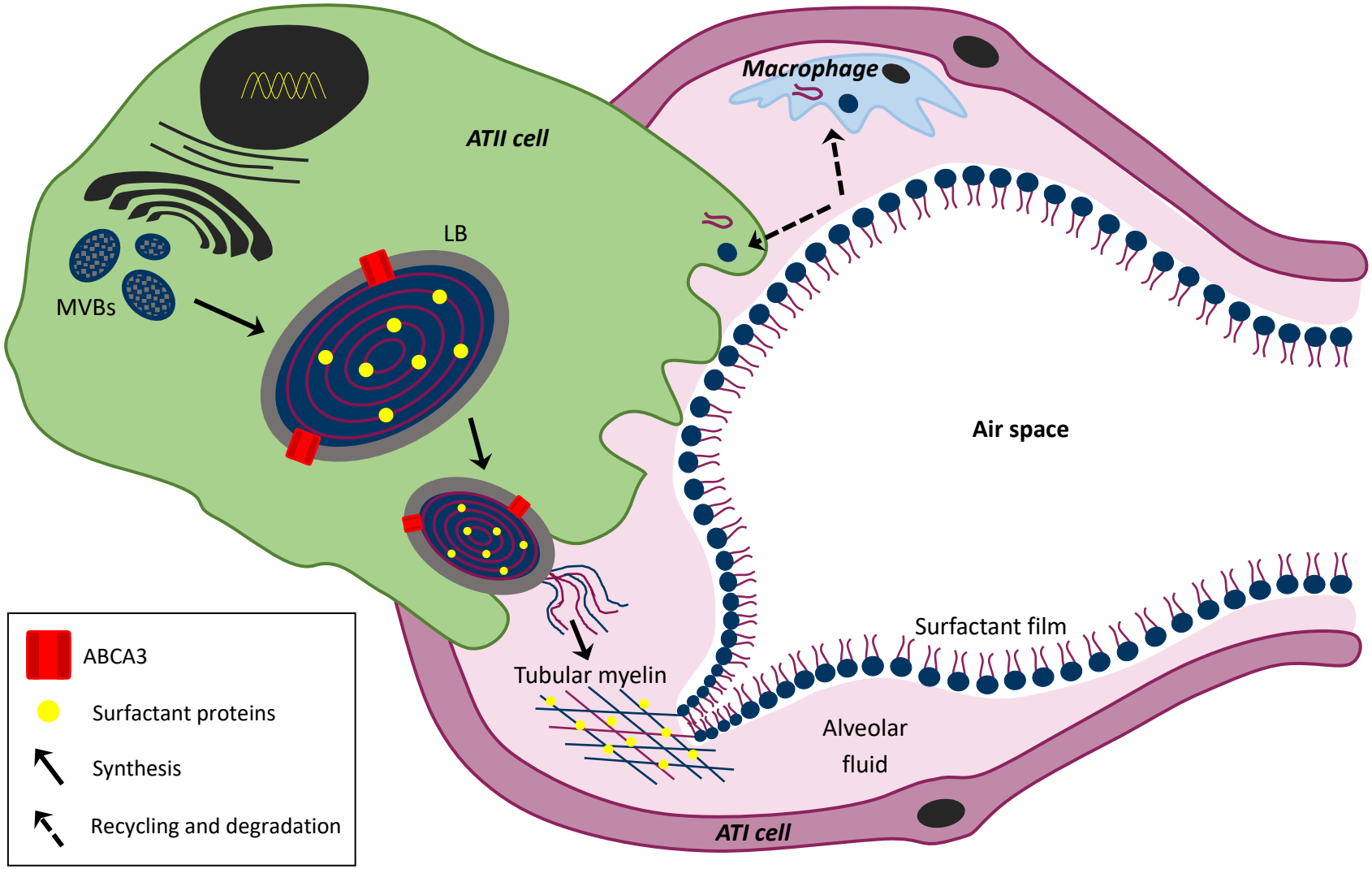


Figure 1.4 Surfactant homeostasis

Figure 1.4 Surfactant homeostasis

The components of surfactant are packaged into lamellar bodies (LBs) inside ATII cells prior to secretion into the alveolar space (Figure 1.3). After exocytosis of LBs surfactant undergoes structural rearrangement to form a film which adsorbs to the air-liquid interface. Surfactant released from LBs can be directly adsorbed to the surfactant film (not shown), in a process facilitated by SP-B oligomers, or unpacked in the alveolar fluid to form a tubular myelin intermediate prior to rearrangement into a surfactant film monolayer. The surfactant film is recycled when it becomes inactive in the air space. The majority of surfactant is recycled back into ATII cells where it can be repackaged into new LBs. Alternatively, the components of surfactant can be catabolised by alveolar macrophages.

1.2.3 ABCA3

The product of the *ABCA3* gene is an intracellular lipid transporter (Section 1.2.1.2) which is highly expressed on the limiting membrane of LBs in ATII cells [31] and at much lower levels throughout the body [39]. Outside of its role in surfactant production the function of ABCA3 is less well understood; patients with *ABCA3* mutations (Section 1.3.2.3) and *Abca3* knockout (KO) mice (Section 1.4.2.1) only present with a respiratory phenotype.

1.2.3.1 Synthesis of ABCA3

The ABCA3 protein is encoded by a large 80kb gene on chromosome 16p13.3 [40], comprising 30 coding and three non-coding exons (Figure 1.5), which direct the synthesis of a 1704 amino acid (180kDa) protein [41]. After synthesis and folding of the primary translation product in the ER, ABCA3 is routed via the Golgi apparatus where it is glycosylated (to form 190kDa protein) before trafficking to MVBs and then the outer membrane of LBs [29]. In post-Golgi compartments, the N-terminus is cleaved at K174 by the cysteine protease cathepsin L [42], generating a secondary protein product foreshortened by about 20kDa when visualised by western blot (190kDa and 170kDa apparent molecular weights; Figure 1.5). The purpose of this cleavage is poorly understood but is hypothesised to be a method of regulating ABCA3 protein activity because removal of the first luminal loop (Figure 1.5) likely alters protein function. When LBs fuse with the ATII cell membrane (Section 1.2.2), ABCA3 remains in the plasma membrane before subsequent recycling or degradation in lysosomes [43].

1.2.3.2 Structure of ABCA3

Structure prediction algorithms suggest ABCA3 is typical of most ABC transporters, a superfamily of integral membrane proteins responsible for ATP-powered translocation of substances across membranes, being formed from two homologous halves (Figure 1.5) [44]. Each half contains a transmembrane domain (TMD; containing 6 TM helices) and a cytosolic ABC nucleotide-binding domain (NBD), present in all ABC transporters [45], in addition to a regulatory domain (RD) and an

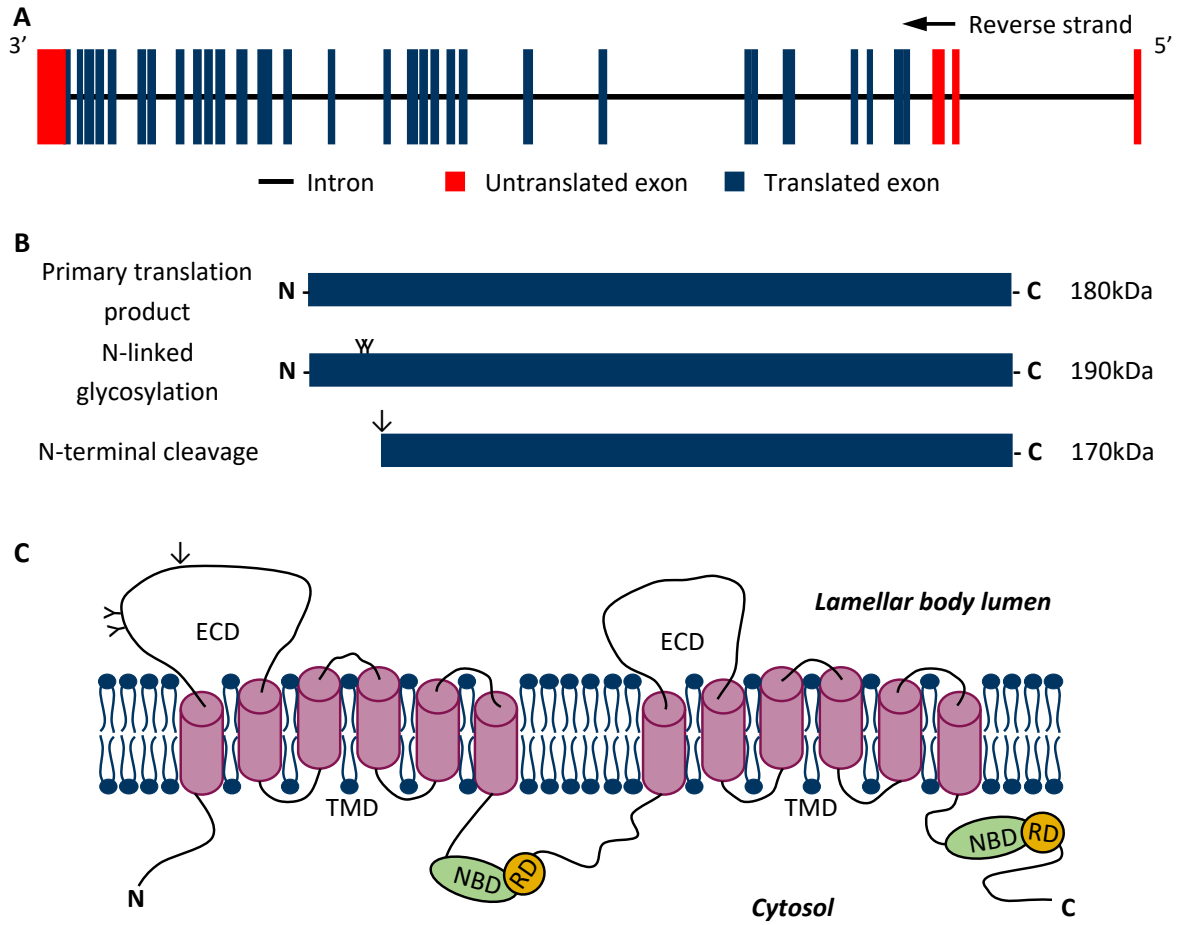


Figure 1.5 Diagram of *ABCA3* processing and structure

(A) *ABCA3* is an 80kb gene that contains 33 exons of which exons 1, 2 and 3 are fully non-coding and exon 33 is partially translated to form **(B)** a 190kDa protein glycosylated (Y) at asparagine residues 124 and 140. It is cleaved at lysine residue 174 (↓), generating a secondary, shorter ABCA3 protein product. **(C)** ABCA3 contains two transmembrane domains (TMDs), nucleotide-binding domains (NBDs), extracellular domains (ECDs) and regulatory domains (RDs).

extracellular domain (ECD) specific to the A subfamily (Figure 1.5) [46]. The NBD includes two conserved peptide motifs known as Walker A (binds phosphate in ATP) and Walker B (performs hydrolysis of ATP to adenosine diphosphate (ADP)), which were first identified as consensus sequences corresponding to ATP-binding sites in 1982 by Walker *et al.* [47]. In addition, ABCA transporters contain a unique Walker C signature motif which contributes to ATP-binding induced NBD dimerization and functions as a transactivation element during ATP hydrolysis [48]. With the energy from ATP hydrolysis, the TMDs create a channel within the LB membrane to allow inward translocation of lipids into LBs [29, 45].

1.3 Surfactant dysfunction

The importance of pulmonary surfactant to normal lung function is underpinned by the development of disease when surfactant composition or homeostasis is altered. Depending on the root cause of the problem, people may present with RDS at birth (Section 1.3.1.1) or develop chronic interstitial lung disease (ILD; Section 1.3.1.2). Symptoms of RDS include rapid, shallow breathing and blue-tinged extremities as the lungs cannot provide enough oxygen to the body. In contrast, a diagnosis of ILD is typically made in childhood or adolescence after patients experience repeated respiratory insufficiency, chronic cough and failure to thrive.

1.3.1 Clinical presentation and current treatment

1.3.1.1 Respiratory distress syndrome

Being born prematurely is the most common cause of RDS as the lungs are underdeveloped and unable to produce sufficient pulmonary surfactant to sustain life [49]. Furthermore, if the alveoli remain fluid-filled after birth any surfactant produced will be inactivated. Surfactant production usually begins between 24-28 weeks gestation and most babies produce enough surfactant to breathe normally by 34 weeks [50, 51]. Thus, babies born before 34 weeks typically present with RDS within hours of birth as their lungs either fail to inflate or struggle to remain inflated at end-exhalation when

the alveolar pressure increases [52]. Areas of atelectasis can be resolved, and further lung collapse prevented, using nasal continuous positive airway pressure (CPAP) [53]. Additionally, if babies are intubated, exogenous surfactant can be delivered directly to the lungs via the intubation tube to reduce surface tension until the lungs have matured and produce sufficient surfactant endogenously to support life [15]. Commercially available surfactants are animal-derived and have a somewhat different composition to human lung surfactant. For example, *Curosurf* (poractant alfa) [54], a widely used exogenous surfactant, contains only polar lipids and SPs due to its manufacturing method [55]. Despite the altered composition, when RDS arises from prematurity, animal-derived surfactant appears effective at lowering surface tension and preventing atelectasis.

More rarely, RDS can arise in term babies when there are mutations in genes necessary for surfactant production, homeostasis or metabolism [56]. This can result in changes to surfactant composition, LB morphology and secondary damage to ATII cells. Mutations in *SFTPB* (Section 1.3.2.1), *ABCA3* (Section 1.3.2.3) and *NK2 homeobox 1* (*NKX2.1*; Section 1.3.2.4) are all commonly associated with an RDS phenotype. Though the disease presentation is indistinguishable from that of premature infants, RDS with a genetic origin is refractory to exogenous surfactant therapy [57] and the outlook for patients is poor with most unable to survive the first few months. Although not well understood, the ineffectiveness of exogenous surfactant therapy may be due to the accumulation of endogenous dysfunctional surfactant, proteins and cellular infiltrates in airspaces (Section 1.3.2) which in turn inactivate exogenous surfactant [58, 59]. The only curative treatment is lung transplantation, however the instability of the affected infants and the requirement for long-term mechanical ventilation while waiting for relatively scarce neonatal donor organs to become available means this is rarely attempted outside of the United States of America (USA) [60]. In the United Kingdom (UK), the youngest recipient underwent a double lung transplant in 2016 at 5 months old [61]. Regrettably, lung transplantation in infants is associated with about 50% 5-year mortality [62] meaning many parents opt for palliative care.

1.3.1.2 Childhood interstitial lung disease

Childhood ILD (chILD) comprises a group of rare lung conditions found in infants, children and adolescents [63]. There are over 200 types of chILD, including surfactant dysfunction, but all cause problems in the airways, alveoli or lung interstitium. In less severe cases of surfactant dysfunction babies may appear healthy at birth but later develop chILD. Like RDS, which arises in term babies, if surfactant dysfunction is implicated in the development of chILD the cause is usually genetic. It may be that the gene implicated has a less essential role in surfactant homeostasis (*SFTPC*; Section 1.3.2.2) or that the mutated protein retains some function (ABCA3-mediated surfactant dysfunction; Section 1.3.2.3). If the cause of chILD is genetic the only curative treatment is a lung transplant. Although this is more feasible in older children due to the increased size of the chest cavity it would only be attempted if the condition cannot be managed conservatively. Supplemental oxygen, corticosteroids [64] and hydroxychloroquine [65, 66] have all been beneficial in treating milder cases of surfactant dysfunction. The mechanism of action for hydroxychloroquine is not fully understood but it has been suggested that in addition to anti-inflammatory properties it can also reduce the accumulation of mutant proteins [67].

1.3.2 Genetic disorders of surfactant dysfunction

Mutations in SPs and other proteins essential for surfactant homeostasis are rare but have all been implicated in surfactant dysfunction disorders [56]. The disease severity and clinical presentation depend on the gene and specific mutation involved (Table 1.1), but all result in altered surfactant composition and surface tension-lowering properties. In severe cases, surfactant dysfunction causes RDS at birth (Section 1.3.1.1) [68, 69], whereas in milder cases a diagnosis may be made in childhood or adolescence after presentation with chILD (Section 1.3.1.2) [70, 71]. Some common features of surfactant dysfunction include the accumulation of lipoproteinaceous material in alveolar airspaces (alveolar proteinosis), due to errors in the processing, packaging and clearance of surfactant components, and the emergence of foamy, lipid-laden alveolar macrophages, after impaired

catabolism of secreted surfactant components. Additionally, ATII cell hyperplasia and associated interstitial thickening is often observed – this likely arises from cells trying to repair damage induced by perturbed surfactant homeostasis, mechanical ventilation or oxygen therapy [72].

Table 1.1 Features of diseases associated with pathogenic mutations in surfactant-related genes

Reduced (<).

Mutated gene	*SFTPB*	*SFTPC*	*ABCA3*	*NKX2.1*
Inheritance	Autosomal recessive	Autosomal dominant, *de novo*	Autosomal recessive	Autosomal dominant, *de novo*
Disease presentation	RDS	chILD	RDS, chILD	Brain-lung-thyroid syndrome
Lamellar body morphology	Large, disorganised MVBs	Typically unaffected	Small, electron dense core	Variable
Surfactant protein expression	< Mature SP-B, Misprocessed SP-C	Misprocessed SP-C	< ABCA3, < Mature SP-B, < Mature SP-C	Variable

1.3.2.1 SP-B deficiency

SP-B deficiency is a very rare autosomal recessive disease affecting ~1 in 1 million live births in the US [73]. The most common mutation (121ins2) causes a frameshift in exon 4 and premature stop codon [74], resulting in the complete absence of SP-B as the messenger ribonucleic acid (mRNA) undergoes nonsense-mediated mRNA decay (NMD) [75]. Invariably, disease-associated *SFTPB* mutations result in the complete loss of functional SP-B, neonatal RDS in full-term infants (Section 1.3.1.1) and death within 3 months (Table 1.1) [59]. An SP-B deficiency is most frequently associated with a histopathological diagnosis of congenital alveolar proteinosis [76]. ATII cells from SP-B deficient patients lack mature, well-organised LBs; instead they contain large, disorganised MVBs, as the phospholipid membrane fails to fuse correctly during packaging of surfactant lipids (Figure 1.6B) [77]; misprocessed SP-C accumulates in the alveolar space as SP-B cannot support proteolytic processing of

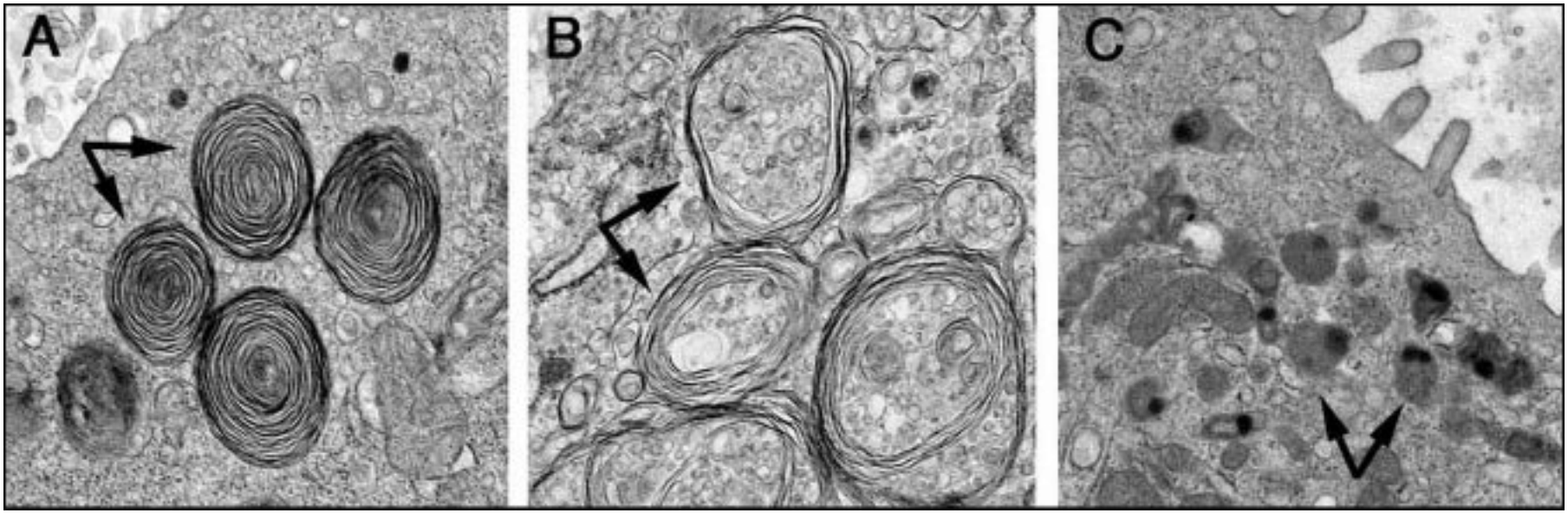


Figure 1.6 Electron microscopy of lamellar bodies in human ATII cells

(A) Mature lamellar bodies (LBs) from a healthy lung. **(B)** Large, disorganised multi-vesicular bodies (MVBs) present in a child with SP-B deficiency. **(C)** Small, abnormal LBS with an electron dense core observed in a patient with ABCA3-mediated surfactant dysfunction. Images obtained from: [78]

SP-C [79]; and, surfactant produced contains an abnormal phospholipid-to-protein ratio which is less effective at lowering surface tension [76].

1.3.2.2 SP-C-associated disease

SP-C-associated disease has a highly variable age of onset and severity but is primarily associated with chILD in infants and older children (Table 1.1; Section 1.3.1.2). Mutations in *SFTPC* often occur *de novo* and were first linked to lung disease in 2001 [71]. Unlike in SP-B deficiency, mutations in *SFTPC* are inherited in an autosomal dominant fashion with disease resulting from gain-of-function of the mutated SP-C protein [80]. Although mutations have been identified throughout the *SFTPC* gene, many of them are located within the BRICHOS domain which is important for the trafficking and processing of the SP-C proprotein (proSP-C) [81]. Depending on the mutation, mutated SP-C can cause protein misfolding, ER stress and activation of the unfolded protein response (UPR) pathway [82] or, alternatively, may be abnormally trafficked, blocking autophagy and leading to ATII cell apoptosis [83, 84]. The accumulation of misprocessed SP-C is associated with diffuse alveolar damage and a regenerating alveolar epithelium lined by hyperplastic ATII cells [72]. In adults with *SFTPC* mutations, pulmonary fibrosis is common [85] due to the inability of ATII cells to repair ongoing alveolar damage.

1.3.2.3 ABCA3-mediated surfactant dysfunction

ABCA3-mediated surfactant dysfunction is an autosomal recessive disorder (Table 1.1) which can present clinically as RDS at birth (Section 1.3.1.1) [69] or as chILD later in development (Section 1.3.1.2) [70]. Recent data on disease incidence is lacking but is estimated to be as high as 1 in 3500 in individuals of European descent making it the most common genetic cause of surfactant dysfunction [56, 86]. Mutations in *ABCA3* were first recognised as disease-causing in 2004 [69], and since then over 300 distinct pathogenic mutations have been identified [87]. Mutations are distributed throughout the length of the *ABCA3* gene and consist of missense, nonsense, frameshift and splice-site mutations, as well as insertions and deletions (indels) [88]. Mutations can be classified according to their expected functional impact – type I mutations result in mistrafficking of the mutant protein,

retention in the ER and no protein expression; type II mutations result in impaired function and ATPase-mediated phospholipid transport [89]. Due to the complete absence of ABCA3, type I mutations (e.g. L101P [69] and Y1515X [90]) are nearly always associated with RDS at birth and the most severe outcomes – in one study at least 17/19 babies identified as homozygous for the Y1515X allele died less than 3 months after birth [88]. In contrast, the phenotype observed with type II mutations (e.g. E292V [70], N568D [69], T1114M [70]) is more variable depending on the level of residual protein activity [88]. For example, a male compound heterozygous for the E292V and T1114M *ABCA3* variants was symptomatic at birth and required a lung transplant aged 12 years, whereas another compound heterozygous individual (E292V, E690K) become symptomatic aged 5 years and was still alive at the point of data collection 18 years later [70, 88].

The *ABCA3* loss-of-function mutations only appear to affect surfactant homeostasis in ATII cells, despite ubiquitous (though largely low-level) expression of the ABCA3 transporter in the body; in homozygous *Abca3* KO mice all organs, excluding the lungs, appeared normal [91, 92]. Small, abnormal LBs with densely packed phospholipid membranes and an electron dense core ('fried egg' appearance) [93] are a defining feature of ABCA3-mediated surfactant dysfunction (Figure 1.6C) [77]. Further features include a deficiency of phosphatidylcholine and phosphatidylglycerol in bronchoalveolar lavage fluid from infants born with RDS [94], presumably due to the defects in packaging surfactant lipids, and, in some patients, decreased expression of mature SP-B and SP-C [95] suggesting a symbiotic relationship between different surfactant-associated proteins.

1.3.2.4 TTF-1-related lung disease

Thyroid transcription factor (TTF)-1 is involved in lung development and regulates the expression of genes involved in surfactant synthesis [96]. Haploinsufficiency of *NKX2.1*, the gene encoding TTF-1, causes a rare form of progressive respiratory failure called brain-lung-thyroid syndrome [96, 97], associated with disruptions to surfactant synthesis and composition. Mutations frequently occur *de novo* and, similar to *SFTPC* mutations, are autosomal dominant. The lung disease phenotype arising

from *NKX2.1* mutations is highly variable in onset and severity and likely reflects how different SPs are impacted by the mutation – infants that present with neonatal RDS likely arise from delayed expression of ABCA3 and SP-B, whilst recurrent infections may reflect altered expression of SP-A or SP-D [98, 99].

1.4 Models of surfactant dysfunction

Much of our knowledge to date on surfactant homeostasis and dysfunction has arisen from early observations in patients with known genetic mutations. To assess potential targeted therapeutics, it is important to use *in vivo* and *in vitro* model systems that recapitulate key features of the diseased alveoli.

1.4.1 Cellular models of the human lung parenchyma

Knockout of endogenous protein expression, or overexpression of protein variants associated with surfactant dysfunction, is routinely used to study the downstream effects of altered protein expression and to assess correction achieved with candidate therapeutics. However, the relevance of such models to the human disease state can be variable and often depends on the genetic background and culture system used. As pulmonary surfactant is produced in ATII cells, a cellular model should aim to mimic the features of ATII cells in the body, or the anatomical complexities of the alveoli more widely.

1.4.1.1 Cell selection

Primary human ATII cells have been of limited use in pulmonary research because historically they do not proliferate for more than two passages in culture and also rapidly transdifferentiate into cells with an ATI-like phenotype [100]. Whilst this can now be minimised by maintaining cells in 3D culture (Section 1.4.1.3), the generation of disease models using clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9-based gene manipulation requires reliable cellular expansion in 2D culture – something that has not yet been achieved for primary human ATII cells. Apart from these

challenges in cell culture, human primary ATII cells remain difficult to isolate and are not widely available. Consequently, the field has favoured immortalised cell lines which are readily accessible, easy to culture and affordable. Many immortalised human cell lines have been used to model aspects of surfactant dysfunction disorders including human embryonic kidney 293 (HEK293) [101], human lung carcinoma 549 (A549) [102] and human papillary adenocarcinoma 441 (H441) [103] cells. Expression of *SFTPC* transgenes carrying patient-relevant mutations in HEK293 and A549 cells revealed that mutations in the BRICHOS domain activate the UPR pathway and induce ER stress resulting in the induction of proinflammatory cytokine signalling [101]. Likewise, expression of the type I ABCA3 variant L101P in A549 cells grown in submerged conditions confirmed that the mutated protein was completely retained in the ER, elevating ER stress and promoting apoptosis [102]. Despite the ease with which immortalised cells can be manipulated to express mutant protein variants or KO pre-existing protein expression, when grown in routine submerged culture they do not recapitulate the *in vivo* environment of ATII cells, as the apical cell surface is not exposed to air, and accordingly do not possess the necessary ATII cell machinery for surfactant production.

Since induced pluripotent stem cells (iPSCs) were first described in 2007 [104], the increased interest in stem cell technologies has provided a viable alternative to immortalised cells for modelling ATII cells and the parenchyma. Early work with embryonic stem cells (ESCs) demonstrated that ESCs could be induced to adopt features of alveolar epithelial cells; immunostaining by Van Haute *et al.* [105] showed that ESCs grown with an ALI differentiation protocol for 20 days expressed the ATII cell-specific marker SP-C at a level similar to 20-week foetal lung, whilst Wang *et al*. [106] used TEM to confirm the presence of LBs. However, the ethical concerns surrounding the isolation of human ESCs has limited their use in research. The reprogramming of terminally differentiated cells back to that of a progenitor cell (iPSCs) using the 'Yamanaka factors' [104] and subsequent redifferentiation into alternative cell lineages, removed the ethical barrier to stem cell models. Typically, iPSCs can be differentiated towards an alveolar epithelial phenotype by sequential exposure to various mediators, including the bone morphogenetic protein -4 inhibitor dorsomorphin and retinoic acid, to recapitulate

developmental pathways until a population of *NKX2.1*$^{+}$ lung progenitor cells are isolated [107]. From here, distal lung progenitor cells (*SOX9*$^{+}$) and iATII cells (*SFTPC*$^{+}$) can be isolated and cultured long-term as 3D alveolospheres (Section 1.4.1.3) or organoids (Section 1.4.1.4). Unlike immortalised cells, iATII cells possess functional LBs, can synthesise and secrete surfactant and activate immune signalling pathways in response to stimuli as similar to that observed in primary ATII cells [108]. However, the differentiation process is challenging, a >30-day protocol is required to recapitulate the milestones associated with *in vivo* development of distal lung [107] and can result in cells which transcriptionally more closely resemble foetal alveolar cells [108]. Importantly for drug discovery research, disease models derived from iATII cells can be obtained either by genetic manipulation of iPSCs prior to differentiation [109] or from patient-derived iPSCs when available [108, 109].

1.4.1.2 Air-liquid interface culture

Growing cells at the ALI, rather than as submerged cultures, promotes cell differentiation and more closely reflects the *in vivo* environment of ATII cells in the lung. To obtain an ALI model, cells are seeded on a semi-permeable support and medium removed from the apical surface after cell adhesion. The increased exposure to oxygen shifts cellular respiration into a more aerobic state which can induce cellular differentiation and promote the formation of tight junctions [110]. Furthermore, additional cells, such as mesenchymal cells, can be seeded on the bottom of the culture vessel, allowing for paracrine signalling between these cells and those on the semi-permeable support. For iATII cells cultured at the ALI, the ATII cell phenotype observed in 3D culture (Section 1.4.1.3) is maintained but the cells mature to better reflect adult, rather than foetal, primary cells [111].

Work by Ohlinger *et al*. [112] demonstrated that when A549 cells were grown in an ALI culture, cell surface morphology was altered, mRNA expression of ATII cell markers was increased along with membrane lipid staining. Furthermore, biological responses to stimuli can be different when cells are grown at the ALI [113] which has important implications for using cell models to screen therapeutic drug candidates. A human surfactant air-liquid interface (SALI) model based on human pulmonary

epithelial H441 cells has been used to model SP-B deficiency [103]. When these cells were grown under submerged conditions, SP-B was not correctly processed to form mature monomeric proteins, however, the proprotein, intermediates and monomeric forms of SP-B were all detected in SALI cultures. Furthermore, culturing H441 cells at the ALI resulted in polarised SALI cultures with robust epithelial barrier functions due to the formation of tight junctions between the cells – all characteristics of primary ATII cells. When SP-B deficiency was modelled by CRISPR-Cas9 mediated KO of *SFTPB*, transepithelial electrical resistance (TEER) in SALI cultures was reduced indicating a compromised epithelial barrier in this disease model.

1.4.1.3 Alveolospheres

Alveolospheres are a type of lung model formed when ATII cells are maintained in 3D culture. Generated from primary [114] or induced [108] ATII cells, these can be maintained long-term as alveolospheres since they continue to proliferate without loss of ATII markers such as SP-C. This contrasts with traditional 2D culture of primary ATII cells which do not proliferate for more than two passages and rapidly transdifferentiate into cells with an ATI phenotype [100]. Interestingly, the lack of any ATI cells emerging in human-derived alveolospheres can also be seen as a disadvantage as it indicates that the models are not adequately recapitulating the human alveolar stem cell niche [114]. Alveolospheres generated from primary ATII cells require the inclusion of support cells from epithelial or mesenchymal lineages for proliferation [114], however, iPSC-derived alveolospheres can be generated with [115] or without mesenchymal cell support [108]. In the context of surfactant dysfunction disorders, Jacob *et al*. [108] used iPSCs derived from patients carrying a homozygous *SFTPB* mutation (121ins2) to demonstrate that SP-B deficient alveolospheres were deficient in LB formation and SP-C processing. More recently, Sun *et al.* [109] published work on patient-derived and CRISPR-Cas9-generated iPSC lines carrying various *ABCA3* mutations in which the iATII cells could be maintained in alveolospheres or transitioned to 2D culture.

1.4.1.4 Lung bud organoids

Lung organoids are a multicellular 3D model grown from PSCs, which are phenotypically more representative of a developing lung, but do express ATII cell markers including SP-B, SP-C, ABCA3 and LBs [116]. Lung bud organoids derived from human ESCs [116, 117] are considered one of the most mature lung organoid models available and developmentally represent the lungs during the second trimester of pregnancy. In contrast to alveolospheres, lung bud organoids contain both airway (ciliated, club and goblet cells) and parenchymal (ATI and ATII) lung cell types, with ATII cell markers predominating, and have been used to create a model of SP-B deficiency for screening of gene therapy vectors [118].

1.4.2 Mouse models of *ABCA3*-mediated surfactant dysfunction

The structural anatomy of the mouse lung is different to that of the human lung. Like humans, there are five lobes in total in the mouse lung but contrastingly the mouse only has a single left lung lobe (Figure 1.7). The cellular organisation also varies, for example the mouse lung contains more club cells than the human lung [119]. Despite this, mice have been widely employed in lung research and in modelling surfactant dysfunction more specifically due to their transgenic capacity and short reproductive cycle. Models of SP-B deficiency [120], SP-C-associated disease [121] and ABCA3-mediated surfactant dysfunction (Section 1.4.2.1) have all been created by KO out of the endogenous mouse protein. As in humans (Section 1.3.2.1), *Sftpb* KO is lethal and associated with accumulation of misprocessed Sp-C in alveolar airspaces and altered surfactant composition [120], demonstrating that surfactant mouse models can successfully recapitulate key features of the human disease. Similarly, KO of *Sftpc* in mice has little impact on lung morphology, surfactant homeostasis, or survival [121], supporting evidence that gain-of-function mutant SP-C expression is responsible for many features of SP-C-associated disease (Section 1.3.2.2). One *Sftpb* KO model has also been used to show proof-of-principle for SP-B gene therapy; survival was prolonged by about 200 days after administration of a recombinant adeno-associated (rAAV) vector expressing human SP-B [122].

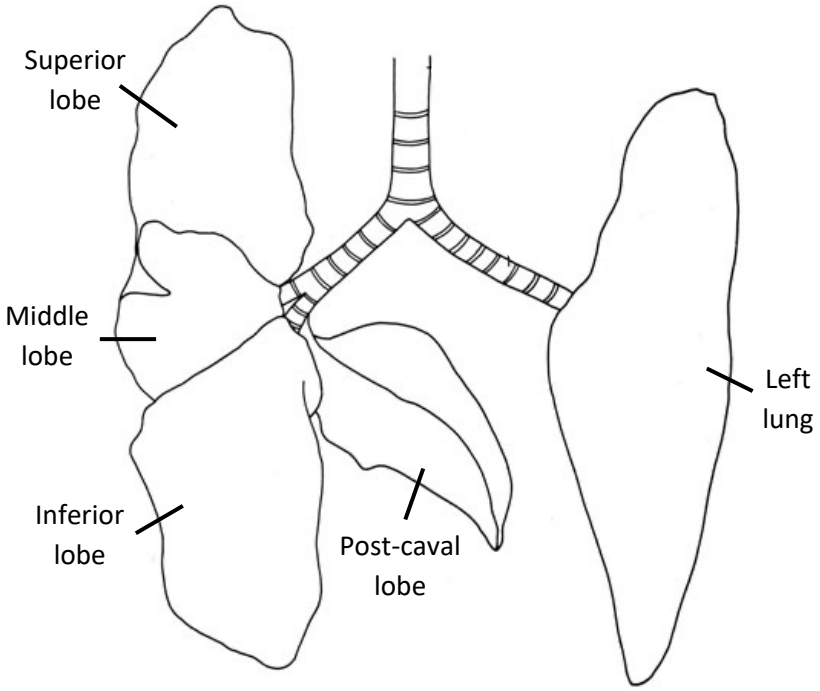


Figure 1.7 Schematic of the murine lung

The mouse lung is organised into five lung lobes. The right lung consists of four lobes comprising the superior lobe, middle lobe, inferior lobe and post-caval lobe. The left lung contains one large lobe. Image adapted from [123].

1.4.2.1 Knockout models

1.4.2.1.1 Constitutive Abca3 knockout

Constitutive *Abca3* KO mice (*Abca3*$^{-/-}$) were reported in 2007 by four groups [91, 92, 124, 125] – 3 years after *ABCA3* mutations were first confirmed to be disease-causing [69]. All reported that homozygous KO of *Abca3* was neonatal lethal with pups typically dying within 1 hour of birth due to a failure of the lungs to inflate. Similar to patients who present with RDS at birth, in *Abca3*$^{-/-}$ mice the surfactant composition was altered with a deficiency in phospholipid species and mature LBs were absent, instead replaced with 'fried egg' LBs (Section 1.3.2.3). Morphologically, the lungs from *Abca3*$^{-/-}$ mice were dark in colour and also had thick interalveolar septa, small air spaces and squamous ATI cells were absent from peripheral alveoli, indicative of pulmonary immaturity and suggesting an additional role for ABCA3 in lung development [91, 124]. Cheong *et al*. [91] also noted that *Abca3*$^{-/-}$ mice were not born at the expected Mendelian frequencies, suggesting a role for ABCA3 in embryonic development, however this was not noted by the other three groups. Interestingly, no abnormalities were seen outside of the lungs which suggests ABCA3 is redundant elsewhere in the body and correction of ABCA3 in the lungs should be sufficient to reverse disease. Whilst constitutive *Abca3*$^{-/-}$ mice have proven useful in better understanding the role of ABCA3 in the body and are a good reflection of clinical cases in which RDS is experienced at birth, the neonatal lethality makes them inappropriate models for assessing therapeutics unless *in utero* strategies are pursued.

Heterozygous counterparts to the above KO mice (*Abca3*$^{+/-}$) were developed and grew normally, showing no outward signs of respiratory distress. Expression of Abca3 was reduced to approximately 50% of levels in *Abca3*$^{+/+}$ mice [92], however the ultrastructure of LBs was normal and indistinguishable from control mice at birth suggesting that a greater than 50% reduction in functional ABCA3 is required to develop disease. Subtle differences between *Abca3*$^{+/-}$ and *Abca3*$^{+/+}$ mice were apparent in 6-month-old mice; LBs were smaller and phospholipid incorporation was reduced in heterozygous *Abca3* KO mice [91] but this had no clear impact on mouse phenotype. The lack of a

disease phenotype in *Abca3*$^{+/-}$ mice reflects the autosomal recessive nature of ABCA3-mediated surfactant dysfunction but also makes them a challenging disease model for evaluating therapeutics. Nevertheless, they have been used to investigate ABCA3 haploinsufficiency as a risk factor for lung injury [126].

1.4.2.1.2 Conditional Abca3 knockout

To investigate the impact of altered ABCA3 expression beyond the postnatal period, two conditional KO models have been described; a doxycycline-inducible model (*Abca3*$^{\Delta/\Delta}$) [127] and a tamoxifen-inducible model (*Abca3*-conditional KO *(Abca3*-cKO)) [128]. In these models, deletion of *Abca3* is spatially and temporally restricted by Cre recombinase activity. When active, Cre recombinase catalyses the recombination between locus of X(cross)-over in P1 (loxP) sites. The loxP sites are 34bp recognition sequences composed of two 13bp palindromic repeats separated by an 8bp asymmetric spacer sequence. When inserted either side of a sequence in the same orientation, loxP sites direct deletion of the 'floxed' deoxyribonucleic acid (DNA). In *Abca3*$^{\Delta/\Delta}$ and *Abca3*-cKO mice, exon 4 – 7 of *Abca3* are floxed, resulting in loss of *Abca3* expression after Cre recombination. In the doxycycline-inducible mice, reverse tetracycline-controlled transactivator protein (rtTA) is expressed under the *SFTPC* promoter and, in the presence of doxycycline, binds to the tetracycline-responsive element (TRE) to induce transcription of Cre recombinase and subsequent excision of floxed *Abca3* (Figure 1.8A). In the tamoxifen-inducible model, a Cre recombinase fusion protein (CreER^{T2}) is expressed under the *SFTPC* promoter. In the absence of tamoxifen, Cre recombinase is expressed but held in the cytoplasm due to its fusion with a mutated oestrogen receptor ligand binding domain (ER^{T2}). Tamoxifen is an oestrogen receptor antagonist and, as such, binds to ER^{T2} inducing the translocation of Cre recombinase to the nucleus where it catalyses recombination between loxP sites (Figure 1.8B).

When *Abca3*$^{\Delta/\Delta}$ dams were treated with doxycycline from embryonic day 6.5-14.5, to KO *Abca3* in ATII cells, the majority (~70%) pups died at birth with the same phenotype as *Abca3*$^{-/-}$ mice, however, approximately 30% of *Abca3*$^{\Delta/\Delta}$ pups survived for up to 9 months with survival being better in females

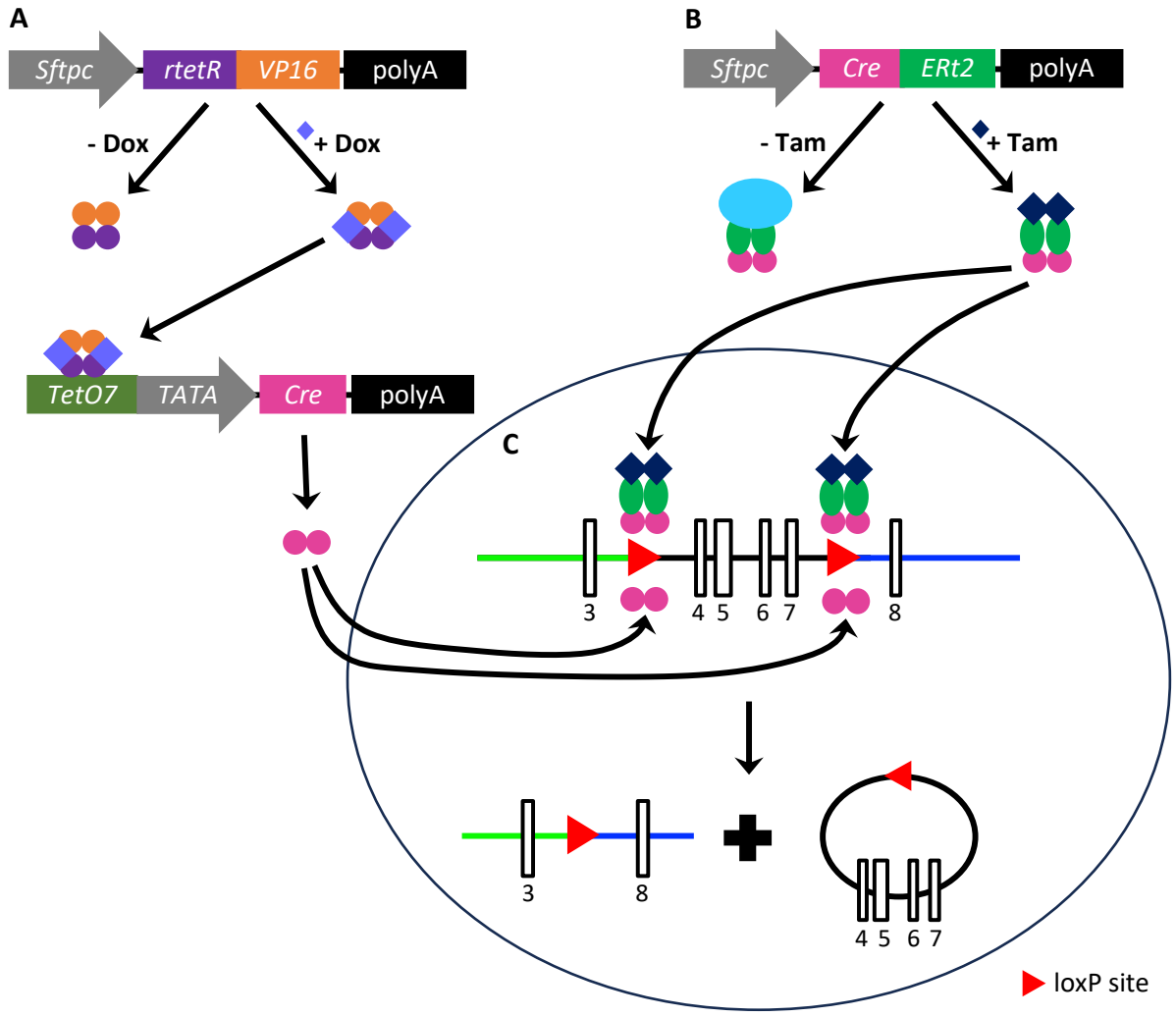


Figure 1.8 Mechanism of tissue-specific, conditional *Abca3* knockout mice

(A) Doxycycline (Dox)-inducible and **(B)** Tamoxifen (Tam)-inducible knockout of *Abca3* in ATII cells is mediated by Cre recombinase (Cre; pink). **(A)** Reverse tetracycline-controlled transactivator protein (rtTA) is a chimeric construct of a mutated version of the *E.coli tetR* gene (*rtetR*) and the *VP16* transactivation domain which is expressed under the control of the ATII cell-specific promoter *Sftpc*. In the absence of Dox, rtTA is expressed but unable to bind the tet responsive element (*TRE*) which is formed of seven tandemly repeated 19bp tetO sequences (*TetO7*) and a minimal *TATA* promoter. In the presence of Dox, rtTA dimers undergo a conformational change to bind *TetO7* driving the expression of Cre. **(B)** A fusion protein composed of Cre recombinase and a triple mutant form of the human oestrogen receptor ligand binding domain (ER^{T2}) is expressed under the control of the *Sftpc*

promoter. The mutant ER^{T2} is unable to bind the natural ligand (oestrogen) but will bind the synthetic oestrogen receptor ligand Tam. In the absence of Tam, $CreER^{T2}$ is held in the cytoplasm after ER^{T2} binds heat shock proteins which usually associate with the oestrogen receptor. In the presence of Tam, $CreER^{T2}$ binds to the active Tam metabolite 4-hydoxytamoxifen and can translocate to the nucleus. **(C)** In the nucleus, dimers of Cre (Dox-inducible mice) or $CreER^{T2}$ (Tam-inducible mice) catalyse the recombination between loxP recognition sequences flanking exons 4 – 7 of *Abca3*. This results in excision of exons 4 – 7 and knockout of *Abca3*.

than males [127]. In surviving mice, both *Abca3* mRNA levels and ATII cells stained positively for Abca3 were reduced by half compared with control mice adding to the evidence from *Abca3*$^{+/-}$ mice that expressing ABCA3 at 50% of control levels is enough to avoid early disease development. However, unlike *Abca3*$^{+/-}$ mice, *Abca3*$^{\Delta/\Delta}$ mice that survived beyond the postnatal period developed emphysema with enlarged air spaces apparent on lung sections. Furthermore, ATII cells isolated from surviving *Abca3*$^{\Delta/\Delta}$ mice contained either normal and giant LBs or, alternatively, no LBs were present. Assuming that both normal and giant LBs were present in ATII cells in which *Abca3* deletion was unsuccessful, this may suggest that cells with functional Abca3 can compensate for the lack of Abca3 in neighbouring cells. Interestingly, and in contrast to many chILD conditions which surviving *Abca3*$^{\Delta/\Delta}$ mice could be considered to represent, there were no signs of inflammatory cytokines or cell infiltrates in the air spaces of 9-month-old *Abca3*$^{\Delta/\Delta}$ mice. Finally, unlike in *Abca3*$^{-/-}$, *Abca3*$^{+/-}$ and *Abca3*$^{\Delta/\Delta}$ mice which died at birth, surviving *Abca3*$^{\Delta/\Delta}$ mice showed decreased *Sftpb* levels and increased *Sftpa* levels increased. For evaluation of therapeutics, prenatal KO of *Abca3* is probably not appropriate as the *Abca3*$^{\Delta/\Delta}$ mice which survived the natal period were phenotypically similar to control mice and the biochemical changes identified after tissue harvest are not reflectively of RDS or chILD patients with a deficiency in ABCA3. Furthermore, the high percentage of animals which die soon after birth would not be helpful in the evaluation of therapeutics and therefore goes against the principle of the 3Rs to replace, reduce and refine the use of animals in research [129].

Postnatal KO of *Abca3* has not been described for *Abca3*$^{\Delta/\Delta}$ mice, however, it was used in *Abca3*-cKO mice [128]. Knockout of *Abca3* in adult 6-8-week-old *Abca3*-cKO mice resulted in a delayed, but similar, phenotype to that seen in *Abca3*$^{-/-}$ mice at birth. After receiving a tamoxifen diet for 12-15 days, which reduced *Abca3* mRNA levels in purified ATII cells by 90%, mice developed fatal respiratory failure with lethargy, dyspnoea and a hunched appearance all noted prior to death. Alveolar wall thickening and signs of inflammation were observed on lung sections, in addition to an increase in neutrophils and a proinflammatory transcriptome in bronchoalveolar lavage fluid. Similar to the KO models described above, and in patients with a deficiency in functional ABCA3 [94], surfactant

composition was altered in induced *Abca3*-cKO mice; there was a decrease in phosphatidylcholine and phosphatidylglycerol phospholipid species. The phenotypic and biochemical changes described in adult *Abca3*-cKO mice administered a tamoxifen diet for at least 6 days reflect some of the key pathogenic features of ABCA3-mediated surfactant dysfunction. Furthermore, inducing *Abca3* deletion postnatally coupled with the delayed decline in animal wellbeing means there is a window of opportunity to dose and evaluate the efficiency of potential disease therapeutics. Interestingly, it is important to note that when deletion of *Abca3* in *Abca3*-cKO mice [128] was sublethal (tamoxifen exposure limited to 4 days) *Abca3* mRNA levels were initially significantly decreased but were restored to control levels 12 weeks after tamoxifen administration. This suggests that unedited ATII cells which retain Abca3 expression have a selective growth advantage and can restore the Abca3-expressing ATII cell population. Additionally, heterogenous alveolar wall thickening and inflammation present 3 weeks after tamoxifen administration were absent 9 weeks later, showing that the lung could repair the alveolar damage induced by *Abca3* KO without fibrosis developing. For gene therapy approaches to treating ABCA3-mediated surfactant dysfunction, this indicates that it may only be necessary to target a small proportion of ATII cells to initiate disease recovery. However, this also means that caution is required when evaluating gene therapeutics in a conditional *Abca3* KO model to ensure the disease phenotype is not self-resolving.

1.4.2.2 Expression of ABCA3 mutations

Insertion of mutations or transgenic elements into the mouse genome is an alternative method of generating murine disease models. One such 'knock-in' (KI) model based on the common E292V missense mutation (Section 1.3.2.3) has been engineered to mimic the chILD phenotype observed in some patients with ABCA3-mediated surfactant dysfunction [130]. Beers *et al*. [130] inserted the E292V mutation into the endogenous *Abca3* locus in addition to an intronic *pgk-Neomycin (Neo)* selection cassette. The resulting homozygous mice ($Abca3^{E292V}$-$rNeo^{+/+}$) survived into adulthood without overt signs of respiratory distress, however, the total phospholipid content in bronchoalveolar

lavage fluid was reduced and LBs were smaller but more abundant compared with control mice. In contrast to *Abca3* KO models, the change in phospholipid content was not associated with a decrease in Sp-B expression and intriguingly *Abca3* mRNA levels were increased in *Abca3*^{E292V}-*rNeo*$^{+/+}$ mice compared with controls. Similar to patients with *ABCA3* mutations, lung morphology was altered in *Abca3*^{E292V}-*rNeo*$^{+/+}$ mice; large foamy macrophages were detected in addition to patchy fibrotic lung remodelling and ATII cell hyperplasia. Furthermore, as observed in 9-month-old induced *Abca3*$^{\Delta/\Delta}$ mice, enlarged air spaces characteristic of emphysema were apparent in 32-week-old *Abca3*^{E292V}-*rNeo*$^{+/+}$ mice. Importantly, when the *pgk-Neo* cassette was removed (E292V$^{+/+}$) the changes in phospholipid content and lung histology previously observed were not present. This suggests that whilst *Abca3*^{E292V}-*rNeo*$^{+/+}$ mice recreated some chILD features associated with *ABCA3* mutations in patients this was due to the *pgk-Neo* cassette and not the presence of the *Abca3* E292V mutation.

In E292V$^{+/+}$ mice [131] there were no overt signs of respiratory distress, nor abnormal histological features, changes in total phospholipid content or *Abca3* mRNA levels. However, the small LBs with electron-dense inclusions detected in *Abca3*^{E292V}-*rNeo*$^{+/+}$ mice were still present in E292V$^{+/+}$ mice and signs of mild alveolar inflammation were detected in older, 32-week-old mice. Tomer *et al.* [131] propose that the E292V mutation in *ABCA3* disrupts ATII cell homeostasis rather than surfactant homeostasis, thus explaining the delayed development of lung abnormalities which are a consequence of defective autophagy and deterioration of ATII cell function. The 32-week timeframe for any biochemical changes to emerge coupled with no outward disease phenotype makes *Abca3*^{E292V}-*rNeo*$^{+/+}$ and E292V$^{+/+}$ mice unsuitable for evaluating therapeutic products for ABCA3-mediated surfactant dysfunction.

1.5 Gene therapy for surfactant dysfunction

Gene therapy is the treatment of disease by transfer of genetic material into cells. It includes the replacement or inactivation of disease-causing genes, and the introduction of modified genetic material to help treat disease. Gene therapy has the potential to permanently treat ABCA3-mediated

surfactant dysfunction. As a monogenic, loss-of-function disease it is hypothesised that introducing a non-mutated copy of *ABCA3* into ATII cells could restore protein function, resulting in correct packaging of phospholipids into LBs, and restoration of surfactant and ATII cell homeostasis. Depending on the duration of effect, this could be curative for patients or, as a minimum, extend the transplantation window for those presenting with intractable RDS at birth.

1.5.1 Fundamentals of gene therapy

The first successful gene therapy clinical trial began in 1990 when *ex vivo*-corrected T cells were administered to a 4-year-old girl born with severe combined immunodeficiency [132]. Although efficacy and/or safety issues precluded the advancement of early gene therapy products, significant advances have since resulted in many gene therapies being approved by regulatory agencies for a range of diseases. For example, Zolgensma, an *in vivo* delivered rAAV vector, was approved in 2019 by the US Food and Drug Administration (FDA) [133] and in 2020 by the European Medicines Agency (EMA) [134] for spinal muscular atrophy in infants under 2 years; it has been shown to improve motor function [135]. Likewise, chimeric antigen receptor T-cell (CAR-T) therapies for cancers, in which isolated patient T-cells undergo *ex vivo* gene transfer to express cancer-targeting antibodies, have been very successful in the clinic [136] and multiple licensed forms are marketed in North America and throughout Europe.

Gene therapy products encompass two key components – the genetic cargo and the delivery modality. The genetic cargo can include genetic information encoding antibodies, gene editing machinery or more classically a non-mutated copy of a gene. Loss-of-function diseases, including SP-B deficiency (Section 1.3.2.1) or ABCA3-mediated surfactant dysfunction (Section 1.3.2.3), should be amenable to gene addition therapy in which the mutated gene is left *in situ* and a non-mutated copy delivered. However, for gain-of-function diseases, such as SP-C-associated disease (Section 1.3.2.2), the mutant allele must be silenced to have the most benefit. In these gain-of-function diseases delivery of CRISPR-

Cas9 gene editing machinery, or molecules achieving mRNA/protein knockdown via, for example, RNA interference or antisense nucleotide-based approaches, may be more appropriate.

1.5.1.1 Strategies for transgene delivery

Efficient delivery of transgenes to the target organ and cell type is vital for successful *in vivo* gene therapies. Several vector systems have demonstrated effective gene transfer to the respiratory epithelium including non-viral cationic lipids [137] and recombinant adenoviral (rAdV) [138], AAV [122] and lentiviral (rLV) vectors [139]. Non-viral vectors are most commonly used to deliver nucleic acid-based therapeutics, by complexing the nucleic acid with lipid/polymeric material to protect it, improve transfection efficiency and increase expression of the therapeutic component. Although non-viral delivery mechanisms have proven safe in clinical trials [137], repeat administration is generally required for sustained transgene expression. Furthermore, the *in vivo* expression level is often lower than can be achieved with viral vectors and may be sub-therapeutic [139, 140].

Recombinant AdV vectors have been shown to transduce ATII cells in a range of models and species including non-human primates [141], mice [138] and human lung explants [138]. The rAdV vectors have a large packaging capacity (approximately 36kb) which could accommodate all the surfactant-associated genes; however, they are non-integrating and trigger strong host immune responses making them less suited for long-term gene expression (Table 1.2) [142]. In contrast, rAAV vectors, which are also non-integrating, have a better safety profile for prolonged gene expression and consequently rAAV gene therapy products targeting surfactant dysfunctions are actively being pursued (Section 1.5.2.1) [122]. Recombinant AAV vectors contain a single-stranded viral DNA genome which is converted to double-stranded DNA and organised into circular episomes once in the nucleus [142]. Apart from the episomal nature of AAV which limits gene expression in dividing cells, one of the main challenges of standard rAAV vectors is the small packaging capacity – approximately 4.7kb (Table 1.2). Whilst this would package *SFTPB* or *SFTPC* transgenes (1.1kb, 591bp), *ABCA3* would be too large (5.1kb) and some gene editing machinery would also be close to the upper limit. Several strategies to

overcome the size limitations of AAV have been explored, including 'oversized' rAAV vectors and dual/split vector systems. Simply inserting large cargo into standard rAAV vectors to produce oversized rAAV vectors has resulted in significantly lower than anticipated vector yields and truncation of packaged genomes [143]. Dual/split vector systems in which the transgene is divided between multiple rAAV vectors and co-transduced have shown more promise [144]. Depending on the split vector system used, reassembly of the full-length transgene *in vivo* can be dependent on concatamerisation and subsequent splicing of intervening inverted terminal repeat (ITR) elements during mRNA processing [145], homologous recombination between overlapping DNA fragments [146], or protein trans-splicing using split inteins [147]. Whilst split rAAV vector systems do work in murine models there remains a question over the safety and efficiency of these platforms for human translation.

Unlike rAdV and rAAV vectors, rLV vectors are retroviruses that integrate into the host cell genome leading to sustained transgene expression – a potential advantage in cells with a fast turnover rate compared with non-integrating, episomal vectors [142]. Although homeostatic turnover of ATII cells is relatively slow, the ATI cell progenitor function, which is activated following acute lung injury to restore the lung epithelium (Section 1.1), means that periods of rapid proliferation are possible. Importantly, unlike the retroviral gene therapy predecessor (gamma-retroviral vectors), use of which caused oncogenic genotoxicity, rLV vectors appear to have no integration preference for oncogenic loci [139]. The packaging capacity for rLV is also large enough to accommodate all surfactant-related genes. This vector is discussed in more depth in Section 1.5.3 and key characteristics highlighted in Table 1.2.

Table 1.2 Key characteristics of viral vector delivery platforms

Characteristic	rAdV	rAAV	rLV
Genome	Double-stranded DNA	Single-stranded DNA	Single-stranded RNA
Packaging capacity	36kb	4.7kb	9kb
Transduction	Dividing and non-dividing cells	Dividing and non-dividing cells	Dividing and non-dividing cells
Expression	Transient	Transient	Stable
Existence in host cell	Episomal	Episomal	Integrated
Immunogenicity	High	Low	Low

1.5.2 Current gene therapy strategies for surfactant dysfunctions

1.5.2.1 SP-B deficiency

There are currently no approved gene therapies for surfactant dysfunction disorders, however preliminary work in murine models has demonstrated the potential benefits of gene therapeutics. Most research has focused on SP-B deficiency (Section 1.3.2.1) [148]; its status as a loss-of-function disease, the small size of the *SFTPB* gene and the severity of the disease phenotype make it an appealing gene therapy target. The gene therapies below were evaluated in a conditional Sp-B KO mouse model [120] in which mice typically survive for about 5 days after Sp-B KO is induced [122]. The first attempts at correcting Sp-B deficiency using rAdV vectors were unsuccessful in expressing SP-B protein *in vivo* [122]. Since then, delivery of chemically modified *SFTPB* mRNA [149] and electroporation of complementary DNA (cDNA) [150], in addition to genome editing of the transgenic *Sftpb* loci [151], have resulted in modest improvements in survival of Sp-B-deficient mice (approximately 30 days).

The most promising data has emerged from the use of viral vectors. Intratracheal delivery of rAAV expressing SP-B increased the median survival of neonatal Sp-B-deficient mice to more than 200 days [122] and was associated with restoration of LB formation and improved lung structure. The eventual

decline in SP-B expression and overall gene therapy efficacy was presumed to be due to dilution of the episomal vector following ATII cell turnover. Unfortunately, whilst the immunogenicity of rAAV is considered to be low, repeat administration of rAAV vectors is not usually possible without addition of immunosuppressants because of the emergence of neutralising antibodies [152]. Thus, if found to be efficacious in humans, rAAV vectors expressing SP-B would likely be limited to expanding the window for lung transplantation. More recently, successful murine survival in this model showed delivery of rLV vector encoding *SFTPB* leading to 61% survival 1 year after induction of Sp-B KO [153]. Given the high rate of protection from the otherwise lethal disease phenotype in this model, it was not possible to calculate median survival over the 1-year in-life phase of this study; however, life-long protection after a single rLV vector dose might be achievable, based on the previously reported duration of reporter gene expression from the same vector [139]. This vector has also previously demonstrated efficacy in an SP-B KO human cell model (Section 1.4.1.2) [103]. Importantly, and in contrast to rAAV, repeat administration of rLV to the murine lung is possible without the need for adjuvants or immune suppression [154, 155], suggesting that repeat dosing in patients might be possible if needed.

1.5.2.2 SP-C-associated disease

As mentioned above (Sections 1.3.2.2 and 1.5.1) designing gene therapies for SP-C-associated disease is more challenging than SP-B deficiency because it is a gain-of-function disease. To date, the only gene therapy strategy proven effective in a murine disease model is *in utero* rAdV-mediated gene editing [156]. Mice heterozygous for the p.I73T *Sftpc* mutation have an aberrant injury-repair response resulting in fibrotic remodelling, and under certain conditions are born with severe diffuse parenchymal lung damage and die soon after birth. Delivery of CRISPR-Cas9 machinery to inactivate the $Sftpc^{I73T}$ allele in embryonic day 16 mice improved neonatal survival, ATII cell morphology and lung maturation.

1.5.2.3 ABCA3-mediated surfactant dysfunction

As with SP-C-associated disease, little progress has been made in developing gene therapeutics for ABCA3-mediated surfactant dysfunction. Instead, there are reports of repurposing small molecule therapies developed for cystic fibrosis (CF); a disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR)* gene, encoding the ABC transporter ABCC7. Ivacaftor and genistein lead to an increase in CFTR transport activity in some CF patients with functional mutations (akin to type II *ABCA3* mutations) [157]. Whilst these drugs did improve ABCA3 function for some type II ABCA3 variants, the effectiveness was mutation-dependent and would not be a viable treatment for type I mutations in which protein expression is altered (Section 1.3.2.3). For development of gene therapeutics, the main challenge is the large size of *ABCA3*. Domm [158] used a split rAAV vector approach to deliver *ABCA3* to BALB/c mice, using the AAV6.2FF serotype with enhanced tropism for the lung [159] that is also being used in rAAV approaches for SP-B deficiency (Section 1.5.2.1). The split-vector constructs were designed to have a 500bp region of homology to undergo homologous recombination *in vivo* to express full length ABCA3 in the lung parenchyma. Although the full-length protein was detected by western blot, it is unknown how much ABCA3 was synthesised and whether this was at therapeutically beneficial levels as the split-vector system was not evaluated in a disease model. The size constraints associated with rAAV vectors could be resolved by exploring *ABCA3* gene addition approaches using rLV vectors (Section 1.5.3.2).

1.5.3 Lentiviral vector-based gene therapy

1.5.3.1 Lentivirus biology

Lentiviruses are enveloped, single-stranded RNA viruses that belong to the family *Retroviridae* [142]. They can be distinguished from other retroviruses by their ability to transduce both dividing and non-dividing cells and (in their native forms) are defined by their long incubation period prior to disease presentation. Lentiviruses are present in primates, ungulates and felines [160], with primates being the natural host for several closely related simian and human immunodeficiency viruses (SIV and HIV,

respectively) which cause acquired immunodeficiency syndrome. In monkeys, SIVs are divided into lineages depending on their host species, including SIV_{AGM} from African green monkey, SIV_{CPZ} from chimpanzee and SIV_{SM} from sooty mangabey. In the natural host, SIV infection is non-pathogenic however transmission to different species results in disease [161]. Transmission of SIV to humans resulted in the emergence of HIV-1 and HIV-2; two genetically distinct viruses closely related to SIV_{CPZ} [162] and SIV_{SM} [163] respectively.

Lentiviral infection [164, 165] occurs when mature virions enter the host cell after engagement of envelope glycoproteins with cognate receptors on the cell surface (Figure 1.9). Following endocytosis, the virus is uncoated and viral RNA reverse transcribed into proviral double-stranded DNA. During uncoating some viral proteins dissociate from the viral core and complex with the newly formed proviral DNA to facilitate nuclear import and integration into the host cell genome. Transcription and translation of the provirus is then mediated by host cell machinery before new infectious virions can be assembled at the host cell membrane and released into the extracellular space in a process called 'budding'. During budding host cell membrane proteins can be incorporated into the envelope of the virion potentially altering subsequent cellular interactions and infection.

The LV genome consists of two identical, single-stranded RNA molecules that are encapsulated within the core of the virus particle. Encapsulation of RNA into virions is directed by secondary RNA structures (collectively known as the RNA packaging signal, Psi, ψ) located upstream of the main viral genome [166]. The proviral DNA genome (Figure 1.10) is then generated following reverse transcription of the viral RNA genome, degradation of the RNA, and integration of the DNA into the host genome (Figure 1.9). The basic structure of the integrated LV provirus genome is very similar between HIV/SIV subtypes. *Gag*, *Pol* and E*nv* genes are essential genes present in all retroviruses encoding structural proteins, enzymes for reverse transcription and integration into the host cell genome, and viral envelope glycoproteins, respectively [167]. In addition to these genes, LV genomes also encode regulatory (*Rev*, *Tat*) genes necessary for initiation of viral replication and accessory (*Vif, Vpr, Vpu/Vpx,*

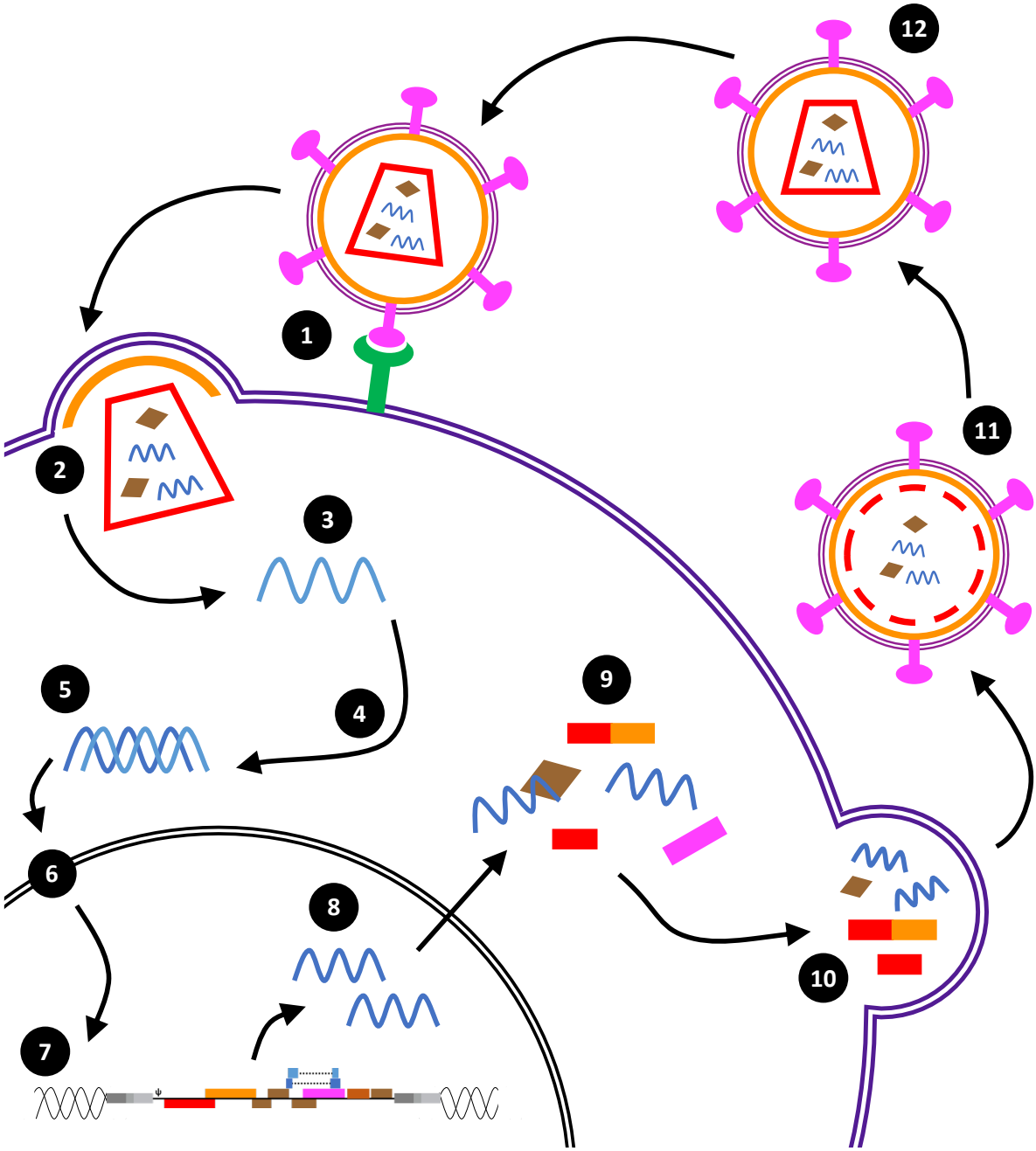


Figure 1.9 Life cycle of HIV-1

(1) Envelope glycoproteins on mature virions bind receptors on the cell surface leading to **(2)** endocytosis and **(3)** viral uncoating. **(4)** Reverse transcription of the RNA genome produces **(5)** a double-stranded DNA proviral genome which **(6)** enters the nucleus and **(7)** integrates into the host genome. After **(8)** replication, the viral components are **(9)** exported from the nucleus and **(10)** assembled into new virions at the cell membrane. **(11)** Budding and **(12)** maturation of immature virions result in virions which can go on to infect more cells.

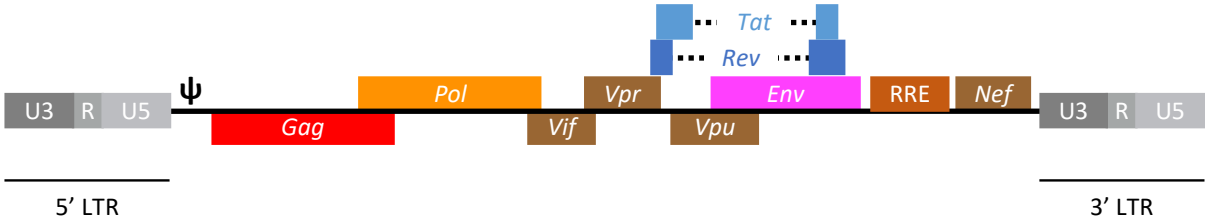


Figure 1.10 Structure of the HIV-1 proviral genome

The HIV-1 proviral genome encodes three essential structural genes (*Gag, Pol, Env*) in addition to regulatory (*Tat, Rev*; blue) and accessory (*Vif, Vpr, Vpu, Nef*; brown) genes. Gag precursor contains the viral core proteins and the GagPol precursor also includes the protease, reverse transcriptase and integrase proteins. The genome is flanked by long terminal repeats (LTRs) comprising U3, R and U5 regions. The *Rev* response element (RRE) is a *cis*-acting RNA element which binds to *Rev*.

Nef) genes which influence pathogenicity, viral replication and budding [165]. Finally, the genome is flanked by long terminal repeats (LTRs), formed of U3, R and U5 domains, which contain a promoter, enhancer and polyadenylation motifs, that drive transcription and are important for reverse transcription [168].

1.5.3.2 Recombinant lentiviral vectors

Wild-type (WT) LV can be converted into a transgene delivery tool by separating the viral components into multiple plasmid constructs and making the resultant rLV replication-incompetent [166]. At least three generations of rLV have been reported to-date (Figure 1.11); each platform was originally designed for production of HIV-1-derived rLV vectors although the concepts are applicable to other LV lineages. First-generation rLVs [169] retained a significant proportion of the HIV-1 genome and were produced by transient transfection of three plasmid constructs – genome plasmid, envelope plasmid, packaging plasmid (Figure 1.11). The transgene was inserted between identical HIV-1 LTR sequences in the genome plasmid which also contained all *cis*-acting viral elements to favour high genome expression during vector production and efficient transduction and integration of vectors. The packaging plasmid contained the structural, enzymatic, regulatory and accessory viral genes, and the envelope plasmid encoded viral glycoproteins from vesicular stomatitis virus (VSV-G) which permits transduction of cells expressing the low-density lipoprotein receptor [170], rather than HIV-1 *Env* which targets CD-4^{+} T-cells. Of note, the open reading frames for HIV-1 *Env* and *Vpu* were blocked in the packaging plasmid. Whilst stable transduction of non-dividing cells could be demonstrated with first-generation rLV vectors [169], the retention of many non-essential accessory genes made them unsafe for clinical translation.

In second-generation rLV vectors many of the non-essential accessory genes from the packaging plasmid were removed – only *Gag*, *Pol*, *Rev* and *Tat* were retained (Figure 1.11) [171]. *Tat* is required to transactivate the U3 promoter within the 5' LTR of the vector genome plasmid and *Rev* is essential for post-transcriptional nuclear export of viral mRNA. Additional safety features were then

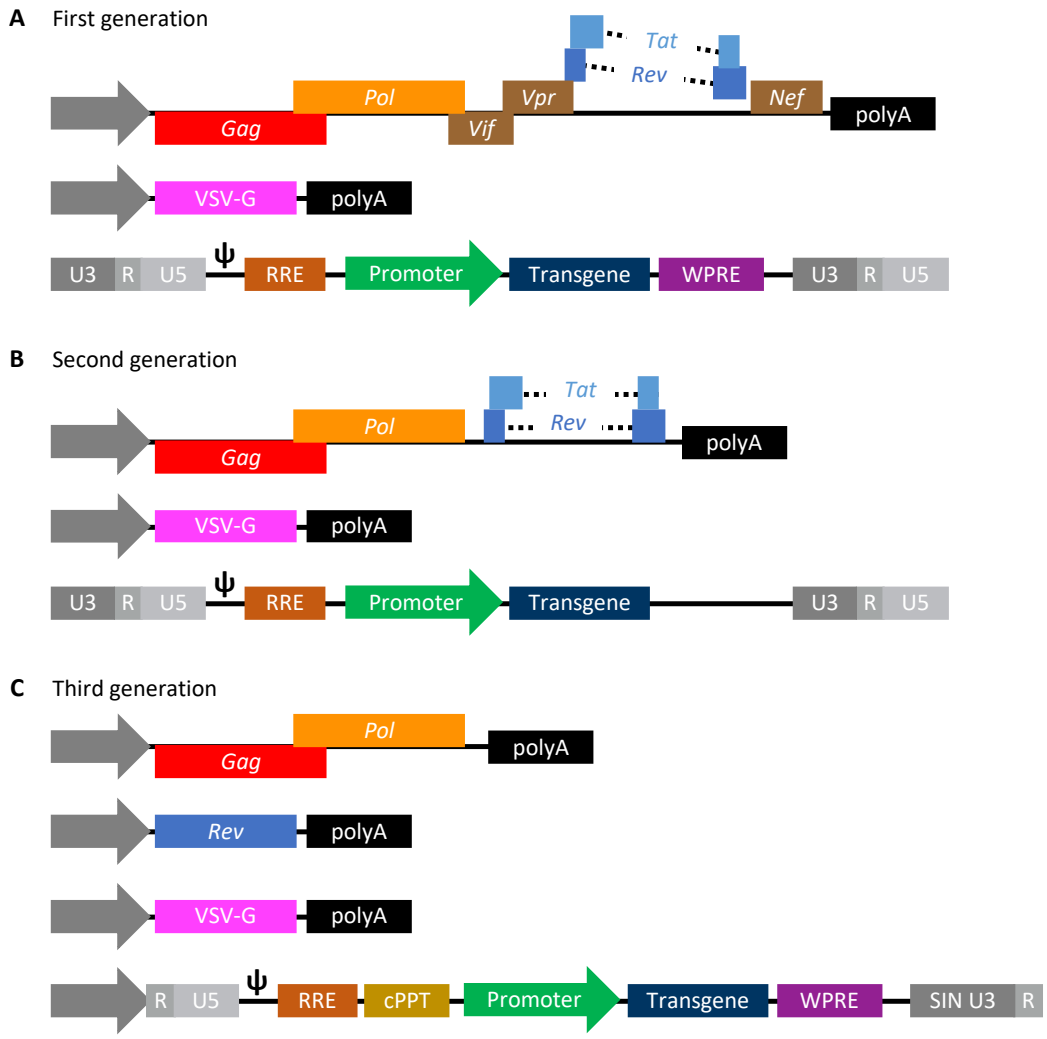


Figure 1.11 Schematic representation of the three generations of recombinant lentiviral vectors

(A) First-generation recombinant lentiviral vectors are produced from three plasmids. The packaging plasmid containing *Gag* and *Pol* retains many HIV-1 accessory genes (brown). The viral pseudotype glycoproteins (VSV-G shown) and transgene are encoded on separate envelope and genome plasmids respectively. **(B)** In the second-generation system non-essential accessory genes were removed. **(C)** To improve safety, third-generation recombinant lentiviral vectors are produced from four plasmids; *Rev* (blue) is removed from the packaging plasmid and expressed under a heterologous promoter. Additional features were introduced into the genome plasmid including removal of the 5' U3 region, modification of the 3' U3 region to make it self-inactivating (SIN), and inclusion of central polypurine tract (cPPT) and Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) sequences. RRE: *Rev* response element.

implemented in a third-generation rLV vector system to further reduce the risk of generating replication-competent rLV: transcription factor binding sites were deleted from the U3 region (ΔU3) of the 3' LTR to make it self-inactivating (SIN) [172]; the system was made *Tat*-independent by replacing the U3 domain of the 5' LTR with a heterologous promoter [173]; Rev expression was driven by a heterologous promoter in a fourth plasmid (Figure 1.11) [173]. Furthermore, to increase viral titres and transgene expression levels, *cis*-acting modifications were made to the genome plasmid including addition of a central polypurine tract (cPPT) to improve nuclear import prior to proviral DNA integration [174] and a Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) to improve nuclear export of the transgene [175]. All rLV vectors discussed in this thesis refer to third-generation vectors unless stated otherwise.

Third-generation rLV vectors are considered to have low oncogenic potential despite integration into the host genome. To-date, out of 258 patients dosed with third-generation SIN rLV vectors (without strong LTR-derived promoter/enhancers) there have been no recorded incidence of oncogenesis (Steve Hyde, Personal communication). In comparison, 16 cases and 3 deaths have been reported in 118 patients administered gamma-retroviral vectors. In addition to the safety features (modified LTRs, SIN, described above) included in third-generation rLV vectors, the apparently random integration of rLV into open chromatin also limits the oncogenic potential of the vector – proportionally, introns account for a larger percentage of open chromatin than exons [176]. This contrasts with gamma-retroviral vectors which preferentially integrate near transcriptional start sites, thus increasing the potential for modifying expression of oncogenes [177].

To improve transduction of target cells, the tropism of rLV vectors can be altered in a process called pseudotyping [178]. Replacing the envelope glycoproteins with those from a heterologous virus influences the efficiency of vector internalisation as the cognate cell surface receptors required for endocytosis are different [179, 180]. The Indiana VSV-G pseudotype exploited in the initial development of rLV vector systems remains frequently used due to its broad tissue specificity and

high transduction capacity [180]. However, its broad cellular tropism can also be viewed as a disadvantage, particularly for *in vivo* gene therapy applications, if restricted expression of the therapeutic cargo is desired to limit off-target effects. Alternative pseudotypes for rLV have been evaluated, including addition of the fusion (F) and haemagglutinin/neuraminidase (HN) glycoproteins from Sendai virus, generating the F/HN pseudotype for pulmonary gene therapy [155] and the baculovirus-derived GP64 envelope protein for transduction of liver endothelial cells [181] and the lung [182]. In addition to influencing the transduction profile, changing the pseudotype can impact the viral titre and intracellular behaviour. For example, VSV-G-pseudotyped rLV vectors can be readily produced to a high titre in part due to their stability during ultracentrifugation purification [183] and rLV vectors pseudotyped with rabies G protein undergo retrograde transport along cell axons potentially allowing for peripheral administration of gene therapies targeting the central nervous system [184].

1.5.3.3 Current recombinant lentiviral vectors for ABC proteins

The efficacy of gene therapy rLV vectors have already been evaluated for some ABC protein-associated diseases [139, 185]. Due to their large size, ABC transporter transgenes are unsuitable for delivery via single rAAV vectors and whilst they can be packaged in rAdV vectors, the immunogenic profile associated with these vectors *in vivo* does not support long-term transgene expression (Section 1.5.1.1, Table 1.3). Stargardt disease is a type of macular dystrophy that arises from mutations in ABC subfamily A member 4 (*ABCA4*) [186]. Exclusively expressed in retina photoreceptors, ABCA4 is a flippase, exporting retinylidene-phosphatidylethanolamine into the cell cytoplasm. In *Abca4* KO mice, a brownish pigment called lipofuscin accumulates in retinal pigment epithelium, causing cell damage [187]. Kong *et al*. [185] designed a recombinant equine infectious anaemia viral vector pseudotyped with VSV-G encoding the 6.8kb coding sequence of *ABCA4*. One year after subretinal administration of the rLV expressing ABCA4 to *Abca4* KO mice the accumulation of lipofuscin in the eye was reduced

compared with mock-treated eyes and at levels comparable to that seen in WT control mice, thus demonstrating the potential utility of rLV-based gene therapy for ABC protein-associated diseases.

More progress to demonstrate the clinical translatability of delivering rLV vectors expressing ABC proteins *in vivo* has been made for CF. Cooney *et al*. [188] describes a feline immunodeficiency virus (FIV)-based vector pseudotyped with the GP64 envelope protein that encodes CFTR. In a newborn pig CF model, delivery of aerosolised rFIV vector pseudotyped with GP64 and expressing CFTR, resulted in partial correction of defective anion transport and altered host immune defences. Furthermore, a phase 1 clinical trial involving an rSIV vector expressing CFTR and pseudotyped with Sendai F/HN will be evaluated in people with CF lung disease (Lenticlair) [189]. The vector has already shown promise at resolving CF phenotypes in an *in vitro* iodide efflux assay, in *Cftr* KO mice and in human intestinal organoids derived from people with CF [139].

1.6 Aims

Surfactant dysfunction caused by mutations in *ABCA3* is the most common type of genetic surfactant dysfunction disorder and is fatal within the first year of life in many cases. Despite this, there are no curative treatments available apart from lung transplantation, which is challenging and not widely available. The development of a lentiviral gene addition therapy for delivery and expression of ABCA3 in ATII cells in the lung parenchyma offers a potentially curative solution for patients. To begin evaluating the feasibility of such a therapy, candidate lentiviral vectors expressing ABCA3 were generated (Chapter 3), before cellular (Chapter 4) and murine (Chapter 5) models of ABCA3-mediated surfactant dysfunction were explored to assess the efficacy of a candidate lentiviral vector.

Chapter 2 Materials and methods

A list of all primers (Table 2.1), plasmids (Table 2.2) and rLV vectors (Table 2.3) used is included below for reference.

Table 2.1 List of primers used in this study

Use/Target	Name	Direction	Sequence (5'-3')
Plasmid cloning (PCR amplification)			
ABCA3 transgene	Fw_20220510	Forward	GGGCAGGCTAGCCACCATGG
ABCA3 transgene	Rv_WT_20220510	Reverse	CGTGATCTAGATCGCCCCTCCTCTGCGGTGG
Y1515X *ABCA3* variant transgene	Rv_Y1515X_20220510	Reverse	CGTGATCTAGACGTCCTGACCAGCTTGTTGGC
Abca3 intron 3	mABCA3_ins3_F	Forward	CACTTTTCCCTCTCTGGCCTTG
Abca3 intron 7	mABCA3_in7_R6	Reverse	CATGAAGGGGTCACGCAGAG
Colony PCR/Plasmid sequencing			
pGM1323, 955	GM18-1162	Forward	TTGCTAACGCAGTCAGTGCT
pGM1324, 1325, 956, 957	GM16-1025	Forward	CCACAGATCCATATCCGCGGCAA
pGM1323-1325, 955-957	GM16-1032	Reverse	ATAGGGAGGGGGAAAGCGAAAGTC
pGM1131-1134	GM16-1075	Forward	TAAGGCAGGCAGGAGTGTTCC
pSpCas9 PX459 V2.0 plasmids	GM16-937	Reverse	TACCGTAAATACTCCACC
pGM1122-1127	GM20-1351	Forward	CGCAAATGGGCGGTAGGCGTG
pGM1122-1127	GM20-1361	Reverse	GTACTCGGTCATGGTAAGCT
pCR-Blunt II-TOPO plasmids	M13F	Forward	GTAAAACGACGGCCAG
pCR-Blunt II-TOPO plasmids	M13R	Reverse	CAGGAAACAGCTATGAC
Site-directed mutagenesis			
ABCA3 L101P	L101PF_hABCA3_211	Forward	CGCATGTTGATCACAGGTGCCCTGCGCACTG
ABCA3 L101P	L101PR_hABCA3_211	Reverse	CAGTGCGCAGGGCACCTGTGATCAACATGCG
ABCA3 E292V	E292VF_hABCA3_211	Forward	CATCATGCGCATGTACACCTTCAGCCTCCTTTC
ABCA3 E292V	E292VR_hABCA3_211	Reverse	GAAAGGAGGCTGAAGGTGTACATGCGCATGATG
ABCA3 N568D	N568DF_hABCA3_211	Forward	CCCGGCACCGTCGTGGCCCAGCA
ABCA3 N568D	N568DR_hABCA3_211	Reverse	TGCTGGGCCACGACGGTGCCGGG
ABCA3 T1114M	T1114MF_hABCA3_211	Forward	GCCAGGATGGAGAACATGCTGGCCAAGAATG
ABCA3 T1114M	T1114MR_hABCA3_211	Reverse	CATTCTTGGCCAGCATGTTCTCCATCCTGGC

ABCA3 Y1515X	Y1515XF_hABCA3_211	Forward	GTTACCACCACTCTACGTCCTGACCAGCTTGTTG
ABCA3 Y1515X	Y1515XF_hABCA3_211	Reverse	CAACAAGCTGGTCAGGACGTAGAGTGGTGGTAAC
ICE (PCR amplification)			
ABCA3 intron -1	UpEx1_BlastF_211213	Forward	GAAATGGAGAAACGGCCAAAG
ABCA3 intron -1	UpEx1_BlastR_211213	Reverse	CCTGCAACCCGCTACTG
ABCA3 exon 4	Ex4_BlastF_211213	Forward	CTGCACCTGGCCAAAAAGAC
ABCA3 exon 4	Ex4_BlastR_211213	Reverse	GGCAGGAAGAGTTCCAGGAC
ABCA3 exon 5	Ex5_BlastF_211213	Forward	GTCCTGGAACTCTTCCTGCC
ABCA3 exon 5	Ex5_BlastR_211213	Reverse	GGACCCAAAGGAGAGACTGC
ABCA3 exon 5	SynthKO_ABCA3_F2	Forward	CACCTCGCATGTTGATCACA
ABCA3 exon 5	SynthKO_ABCA3_R2	Reverse	CGCTCCTCCTCTGGAAGAAC
ABCA3 exon 8	Ex8_BlastF_211213	Forward	GAAGTCACTCTGTTGCCCCA
ABCA3 exon 8	Ex8_BlastR_211213	Reverse	CACCAAGCCTTTGGACATGG
TRAC locus	Synth+_TRAC_F	Forward	TCAGGTTTCCTTGAGTGGCAGG
TRAC locus	Synth+_TRAC_R	Reverse	TAAGGCCGAGACCACCAATCAG
ICE (sequencing)			
ABCA3 exon 4	Ex4_BlastF_211213	Forward	CTGCACCTGGCCAAAAAGAC
ABCA3 exon 5	SynthKO_ABCA3_Sq2	Forward	CTGGGAAGAACTACACCCTGCAG
ABCA3 exon 8	Ex8_BlastF_211213	Forward	GAAGTCACTCTGTTGCCCCA
TRAC locus	Synth+_TRAC_Sq	Forward	CTGGCCGTGAACGTTCACTGAAATCATGGC
Genotyping of gDNA samples			
ABCA3 intron -1	ABCA3KO_gt_In-1_F	Forward	CAACCATCCCCGAACCCCCC
ABCA3 intron 8	ABCA3KO_gt_In+8_R	Reverse	GGGAGGCTGAGATGGGAGAATTTC
ABCA3 exon 2	ABCA3KO_gt_Ex2_F	Forward	GATGTCCTGCGAGGTCACGGTG
ABCA3 exon 4	ABCA3KO_gt_Ex4_R	Reverse	GTGTAGTTCTTCCAGAGGAGGAGC
ABCA3 exon 3	ABCA3_KOgt_Ex3c_F	Forward	GCTGAGCAACTGAACCTGAAAC
ABCA3 exon 6	ABCA3_KOgt_Ex6c_R	Reverse	GTCCTCAAAGTCCTTCTCGGAG
CMV-	GM19-1241	Forward	GTACACCAATGGGCGTGGAT
ABCA3 exon 5	qRT-PCR_R_Ex5	Reverse	GGGCACATTTTCCGACTGAATC
mAbca3 intron 3	mAbca3-F	Forward	AGCACTTTTCCCTCTCTGGCCTTGAG
mAbca3	mAbca3-R	Reverse	TGCCCACCCAGCACCATGCT
Cre	CRE comm-F (Cre-A)	Forward	TGCCACGACCAAGTGACAGCAATG
Cre	CRE comm-R (Cre-B)	Reverse	AGAGACGGAAATCCATCGCTCG
EYFP	EYFP MUT-F	Forward	AGGGCGAGGAGCTGTTCA
EYFP	EYFP MUT-R	Reverse	TGAAGTCGATGCCCTTCAG
Abca3 intron 7	F1_mAbca3_in7sq	Forward	TACCACCTGCGCTTCAGTTACAC
Abca3 intron 7	F2_mAbca3_in7sq	Forward	GAGACTGAAGGCTGGCATACCAC
Abca3 intron 7	F3_mAbca3_in7sq	Forward	GTCACAACAGTTGCCGTGTCAAC
Abca3 intron 7	F4_mAbca3_in7sq	Forward	GAGAGACTGCTCCGAGGTGAGAG
Abca3 intron 7	F5_mAbca3_in7sq	Forward	ACTGCTGTTTCAAAGACCCGAGT
Abca3 intron 7	F6_mAbca3_in7sq	Forward	GGCTACACAGAGAAACCTTGCCA
Abca3 intron 7	F7_mAbca3_in7sq	Forward	CCAAGAAGCAAGTTGGGGAGGAA
Abca3 intron 7	F8_mAbca3_in7sq	Forward	TGCCTTACAGCTTGGACTCATGG
SpCas9 gRNA plasmid cloning			
ABCA3 intron -1	Int-1_Fw_hABCA3	Forward	CACCGCGAACCGCGTACTCGCAGGG
ABCA3 intron -1	Int-1_Rv_hABCA3	Reverse	AAACCCCTGCGAGTACGCGGTTCGC
ABCA3 exon 4	Ex4_Fw_hABCA3	Forward	CACCGCGTCCCTTTCAGCAAGACGA

ABCA3 exon 4	Ex4_Rv_hABCA3	Reverse	AAACTCGTCTTGCTGAAAGGGACGC
ABCA3 exon 5	Ex5_Fw_hABCA3	Forward	CACCGCCGTCACTGAGACAGTGCGC
ABCA3 exon 5	Ex5_Rv_hABCA3	Reverse	AAACGCGCACTGTCTCAGTGACGGC
ABCA3 exon 5	Ex5.1_Fw_hABCA3	Forward	CACCGGCCCGGGTAGATGGTGGCGT
ABCA3 exon 5	Ex5.1_Rv_hABCA3	Reverse	AAACACGCCACCATCTACCCGGGCC
ABCA3 exon 5	Ex5.2_Fw_hABCA3	Forward	CACCGTGTTTTCTGGGATCCTCATC
ABCA3 exon 5	Ex5.2_Rv_hABCA3	Reverse	AAACGATGAGGATCCCAGAAAACAC
ABCA3 exon 5	Ex5.3_Fw_hABCA3	Forward	CACCGCTTCCAGAAGCGGAAGGTCC
ABCA3 exon 5	Ex5.3_Rv_hABCA3	Reverse	AAACGGACCTTCCGCTTCTGGAAGC
ABCA3 exon 8	Ex8_Fw_hABCA3	Forward	CACCGAAGGGGTCTGCGATGAACGG
ABCA3 exon 8	Ex8_Rv_hABCA3	Reverse	AAACCCGTTCATCGCAGACCCCTTC
qRT-PCR			
CFTR	hCFTRInt1-9987F	Forward	CTTCCCCCATCTTGGTTGTTC
CFTR	hCFTRInt1-10100R	Reverse	TGACAGTTGACAATGAAGATAAAGATGA
CFTR	hCFTRInt1-10022T	Probe	VIC-TGTCCCCATTCCAGCCATTTGTATCCT-TAMRA
WPRE	DNAVEC_WPRE_203F	Forward	TGGCGTGGTGTGCACTGT
WPRE	DNAVEC_WPRE_281R	Reverse	CCCGGAAAGGAGCTGACA
WPRE	DNAVEC_WPRE_223T	Probe	FAM-TTGCTGACGCAACCCCCA
β-actin	Bact_qRT_F	Forward	AGAGCTACGAGCTGCCTGAC
β-actin	Bact_qRT_R	Reverse	AGCACTGTGTTGGCGTACAG
ABCA3 exon 4	ABCA3_qRT_F_Ex4	Forward	CTCCTCTGGAAGAACTACACCC
ABCA3 exon 5	ABCA3_qRT_R_Ex5	Reverse	GGGCACATTTTCCGACTGAATC
SFTPB	SPB_qRT_F	Forward	AATTCCCCATTCCTCTCCCCTAT
SFTPB	SPB_qRT_R	Reverse	GATGCCGCCCGCCAC
β-actin (mouse)	Bact_qPCR_F1_2102	Forward	GCGAGCACAGCTTCTTTG
β-actin (mouse)	Bact_qPCR_R1_2102	Reverse	ATGCCGGAGCCGTTGTC
Abca3 exon 6	mABCA3_qRT_Fex6_202011	Forward	CGAGGACTACATTCGCTATGAC
Abca3 exon 7	mABCA3_qRT_Rex7_202011	Reverse	GTCCGGGGCTTGGAAAAAGT
Abca3 exon 10	mABCA3_qRT_Fex10_202011	Forward	TCAGTAAAGCCAACATAGCAGC
Abca3 exon 11	mABCA3_qRT_Rex11_202011	Reverse	GAGGAGCCACGAAAAAGTAGG
ddPCR			
WPRE	WPRE_d_01_SH	Forward	CTATGTGGATACGCTGCTTTA
WPRE	WPRE_d_01_SH	Reverse	GAGACAGCAACCAGGATTTA
WPRE	WPRE_d_01_SH	Probe	HEX-CATGCTATTGCTTCCCGTATGGCT-BHQ-1
CFTR	hCFTRint1	Forward	CTTCCCCCATCTTGGTTGTTC
CFTR	hCFTRint1	Reverse	TGACAGTTGACAATGAAGATAAAGATGA

CFTR	hCFTRint1	Probe	FAM-TGTCCCCATTCCAGCCATTTGTATCCT-BHQ-1
Abca3 post recombination loxP	loxP3_F1_ddPCR	Forward	GTCAAGGCTGGAAGTTCTGT
Abca3 post recombination loxP	loxP3_R1_ddPCR	Reverse	TGACAAGGGGCCAATTTTGT
Abca3 post recombination loxP	loxP3_probe_ddPCR	Probe	HEX-ACGCCTCGTCAACACCGTACGCTTGT-BHQ-1
Abca3 intron 3 loxP	loxP1_F4_ddPCR	Forward	GCGGTGTACCATAACTTCGT
Abca3 intron 3 loxP	loxP1_R3_ddPCR	Reverse	GCGGTGTACCATAACTTCGT
Abca3 intron 3 loxP	loxP1_probe_ddPCR	Probe	FAM-TGGCGAGCCACAGAGCATGGTGCT-BHQ-1

Table 2.2 List of plasmids used in this study

The cytomegalovirus-immediate early (CMV) promoter with introns removed is referred to as CMV-.

Plasmid	Laboratory reference	Promoter	Transgene	Source
Cloning plasmids				
pABCA3_1	pGM1298	-	*ABCA3_1*	Twist Bioscience
pABCA3_2	pGM1299	-	*ABCA3_2*	Twist Bioscience
pCMV-.SFTPB	pGM751	CMV-	*SFTPB*	Gift from Altar Munis
phCEF.SFTPB	pGM749	hCEF	*SFTPB*	Gift from Altar Munis
pEf1α.SFTPB	pGM753	Ef1α	*SFTPB*	Gift from Altar Munis
pCMV-.ABCA3_1	pGM1323	CMV-	*ABCA3_1*	This study
phCEF.ABCA3_1	pGM1324	hCEF	*ABCA3_1*	This study
pEf1α.ABCA3_1	pGM1325	Ef1α	*ABCA3_1*	This study
pCMV-.coACE2-F2A-PuroR	pGM943	CMV-	*CoACE2-F2A-PuroR*	Gift from Kamran Miah
pCMV-.L101P	pGM992	CMV-	*ABCA3* containing L101P mutation	This study
pCMV-.E292V	pGM993	CMV-	*ABCA3* containing E292V mutation	This study
pCMV-.N568D	pGM994	CMV-	*ABCA3* containing N568D mutation	This study
pCMV-.T1114M	pGM995	CMV-	*ABCA3* containing T1114M mutation	This study
pCMV-.Y1515X	pGM996	CMV-	*ABCA3* containing Y1515X mutation	This study
pCMV-.*f*Lux	pGM350	CMV-	*f*Lux	Gift from Ian Pringle
pmSPB.GFP	pGM769	mSPB	GFP	Gift from Mariana Viegas
pAlv1.GFP	pGM1119	Alv1	GFP	Gift from Mariana Viegas
pAlv2.GFP	pGM1120	Alv2	GFP	Gift from Mariana Viegas
pAlv3.GFP	pGM1121	Alv3	GFP	Gift from Mariana Viegas

Genome plasmids for rSIV.F/HN production				
pCMV-.GFP	pGM311	CMV-	*GFP*	Gift from Ian Pringle
phCEF.GFP	pGM357	hCEF	*GFP*	Gift from Ian Pringle
phCEF.CFTR	pGM326	hCEF	*CFTR*	Gift from Ian Pringle
pCMV-.ABCA3	pGM955	CMV-	*ABCA3*	This study
phCEF.ABCA3	pGM956	hCEF	*ABCA3*	This study
pEf1α.ABCA3	pGM957	Ef1α	*ABCA3*	This study
pCMV-.fLux	pGM350	CMV-	fLux	Gift from Ian Pringle
phCEF. fLux	pGM356	hCEF	fLux	Gift from Ian Pringle
pmSPB.fLux	pGM1131	mSPB	fLux	This study
pAlv1.fLux	pGM1132	Alv1	fLux	This study
pAlv2.fLux	pGM1133	Alv2	fLux	This study
pAlv3.fLux	pGM1134	Alv3	fLux	This study
Accessory plasmids for rSIV.F/HN production				
pCBA.SIV.coGagPol	pGM691	CBA	Codon optimised SIV *Gag* and *pol*	Gift from Ian Pringle
pCMV.Rev	pGM299	CMV	SIV *Rev*	Gift from Ian Pringle
pCAG.SeV.Fct4	pGM301	CAG	Sendai virus F envelope protein	Gift from Mary Connolly
pCAG.SeV.EnvctHN	pGM303	CAG	Sendai virus HN envelope protein	Gift from Mary Connolly
Genome plasmids for rHIV.VSV-G production				
pCMV-.GFP	pGM378	CMV-	*GFP*	Gift from Ian Pringle
pCMV-.ABCA3-F2A-PuroR	pGM1122	CMV-	*ABCA3*, Puromycin resistance	This study
pCMV-.L101P-F2A-PuroR	pGM1123	CMV-	*ABCA3* containing L101P mutation, Puromycin resistance	This study
pCMV-.Y1515X-F2A-PuroR	pGM1124	CMV-	*ABCA3* containing Y1515X mutation, Puromycin resistance	This study
pCMV-.E292V-F2A-PuroR	pGM1125	CMV-	*ABCA3* containing E292V mutation, Puromycin resistance	This study
pCMV-.N568D-F2A-PuroR	pGM1126	CMV-	*ABCA3* containing N568D mutation, Puromycin resistance	This study
pCMV-.T1114M-F2A-PuroR	pGM1127	CMV-	*ABCA3* containing T1114M mutation, Puromycin resistance	This study

Accessory plasmids for rHIV.VSV-G production				
pCMV.HIV.GagPol	pGM281	CMV	HIV *Gag* and *Pol*	Gift from Ian Pringle
pRSV.Rev	-	RSV	HIV *Rev*	Gift from Ian Pringle
pMD2-G	-	CMV	VSV-G	Gift from Didier Trono (Addgene plasmid 12259)
PCR-Blunt II-TOPO plasmids				
pSP6.Abca3in3-7 (BALB/c)	pGM1240	SP6	*Abca3* intron 3 – intron 7 from BALB/c mice	This study
pSP6.Abca3in3-7(*Abca3*-iKO)	pGM1241	SP6	*Abca3* intron 3 – intron 7 from *Abca3*-iKO mice (containing LoxP sites)	This study
pSP6.Cre	pGM1326	SP6	Cre amplicon from *Abca3*-iKO mice	This study
pSP6.EYFP	pGM1327	SP6	*EYFP* amplicon from *Abca3*-iKO mice	This study
pSP6.Abca3in3	pGM1328	SP6	*Abca3* intron 3 amplicon from *Abca3*-iKO mice	This study
PX459 V2.0 (62988; Addgene) CRISPR/Cas9 plasmids				
Plasmid	**Laboratory reference**	**gRNA target**	**gRNA target sequence**	**Source**
pU6.gRNA(ABCA3in-1).CBh.spCas9-T2A-PuroR	pGM1042	*ABCA3* intron -1	CGAACCGCGTACTCGCAGGG	This study
pU6.gRNA(ABCA3ex4).CBh.spCas9-T2A-PuroR	pGM1043	*ABCA3* exon 4	CGTCCCTTTCAGCAAGACGA	This study
pU6.gRNA(ABCA3ex5).CBh.spCas9-T2A-PuroR	pGM1044	*ABCA3* exon 5	CCGTCACTGAGACAGTGCGC	This study
pU6.gRNA(ABCA3ex8).CBh.spCas9-T2A-PuroR	pGM1045	*ABCA3* exon 8	AAGGGGTCTGCGATGAACGG	This study
pU6.gRNA(ABCA3ex5.1).CBh.spCas9-T2A-PuroR	pGM1207	*ABCA3* exon 5	GCCCGGGTAGATGGTGGCGT	This study
pU6.gRNA(ABCA3ex5.2).CBh.spCas9-T2A-PuroR	pGM1208	*ABCA3* exon 5	TGTTTTCTGGGATCCTCATC	This study
pU6.gRNA(ABCA3ex5.3).CBh.spCas9-T2A-PuroR	pGM1209	*ABCA3* exon 5	CTTCCAGAAGCGGAAGGTCC	This study

Table 2.3 List of recombinant lentiviral vectors used

Brackets indicate when production data only was used.

Vector	Laboratory reference		Strain	Pseudotype	Promoter	Transgene	Genome plasmid	Source
	Vector ID	Production						
rSIV.F/HN.hCEF.CFTR	vGM195	(W269, W215, W214)	SIV	F/HN	hCEF	*CFTR*	pGM326	Gift from Steve Hyde
rSIV.F/HN.hCEF.GFP	vGM196	W304, (W365, W361, W353)	SIV	F/HN	hCEF	*GFP*	pGM357	Gift from Kamran Miah
rSIV.F/HN.hCEF.ABCA3	vGM336	W321	SIV	F/HN	hCEF	*ABCA3*	pGM956	This study
	vGM438	W372						
rSIV.F/HN.CMV-.fLux	vGM231	W341	SIV	F/HN	CMV-	fLux	pGM350	Gift from Kamran Miah
rSIV.F/HN.hCEF.fLux	vGM223	W340	SIV	F/HN	hCEF	fLux	pGM356	Gift from Mariana Viegas
rSIV.F/HN.mSPB.fLux	vGM396	W349	SIV	F/HN	mSPB	fLux	pGM1131	This study
rSIV.F/HN.Alv1.fLux	vGM397	W343	SIV	F/HN	Alv1	fLux	pGM1132	This study
rSIV.F/HN.Alv2.fLux	vGM398	W347	SIV	F/HN	Alv2	fLux	pGM1133	This study
rSIV.F/HN.Alv3.fLux	vGM399	W345, W362	SIV	F/HN	Alv3	fLux	pGM1134	This study
rHIV.VSV-G.CMV-.GFP	vGM234	-	HIV-1	VSV-G	CMV-	*GFP*	pGM378	Gift from Kamran Miah
rHIV.VSV-G.CMV-.ABCA3-F2A-PuroR	vGM449	-	HIV-1	VSV-G	CMV-	*ABCA3*	pGM1122	This study
rHIV.VSV-G.CMV-.L101P-F2A-PuroR	vGM450	-	HIV-1	VSV-G	CMV-	*ABCA3*-L101P	pGM1123	This study
rHIV.VSV-G.CMV-.E292V-F2A-PuroR	vGM451	-	HIV-1	VSV-G	CMV-	*ABCA3*-E292V	pGM1125	This study
rHIV.VSV-G.CMV-.N568D-F2A-PuroR	vGM452	-	HIV-1	VSV-G	CMV-	*ABCA3*-N568D	pGM1126	This study
rHIV.VSV-G.CMV-.T1114M-F2A-PuroR	vGM453	-	HIV-1	VSV-G	CMV-	*ABCA3*-T1114M	pGM1127	This study
rHIV.VSV-G.CMV-.Y1515X-F2A-PuroR	vGM454	-	HIV-1	VSV-G	CMV-	*ABCA3*-Y1515X	pGM1124	This study

2.1 Common reagents and solutions

Chemicals, reagents and solvents were of analytical grade and approved for tissue culture use when necessary. All chemicals were obtained from Sigma (Gillingham, UK) unless stated otherwise. When required, sterilisation of reagents and solutions was achieved by autoclaving for 15-30 minutes at 121°C, or by filtration through a 0.22µm surfactant-free cellulose acetate filter (Thermo Scientific, Paisley, UK). Solutions were stored at room temperature (RT; 20-25°C) unless otherwise stated. Where concentrations are indicated by percentages, they refer to weight:volume (w/v) ratios for solids at RT and volume:volume (v/v) ratios for liquids at RT.

Buffers were prepared using Milli-Q reverse osmosis ultrapure water (Merck, GIllingham, UK) and all other solutions in nuclease-free water (NFW; Promega, Southampton, UK), or absolute ethanol (VWR, Lutterworth, UK) if non-dissolvable in water, as directed by the manufacturer, unless otherwise stated. Phosphate-buffered saline (PBS) was prepared from tablets as directed or, for cell culture purposes, obtained from Gibco (Dulbecco's PBS, dPBS; Paisley, UK) as indicated. Methanol-free paraformaldehyde (PFA) was prepared to a final 4% solution by dissolving in PBS at 60°C and adjusting to pH7, before storage at -20°C.

2.2 Molecular biology

2.2.1 Restriction enzyme digest

Restriction enzymes were purchased from New England Biolabs (NEB, Hitchin, UK), and used according to manufacturer's instructions. For plasmid cloning, 1-5µg DNA was digested in a 20-50µl reaction for 1-2 hours at the recommended temperature. For analytical restriction enzyme reactions, ~100ng of DNA was used with total enzyme content not exceeding 10% of the final volume.

2.2.2 Polymerase chain reaction

For use in plasmid cloning, DNA sequences were amplified via polymerase chain reaction (PCR) using 10ng of DNA template and the high-fidelity polymerase Phusion Hot Start II DNA polymerase (F549; Thermo Scientific). Reactions contained a final concentration of 200µM deoxynucleotide triphosphates (dNTPs; NEB), 0.5µM of forward and reverse primers (Table 2.1), and 0.02U/µl polymerase. For PCR amplification of genomic DNA (gDNA), typically 50-100ng gDNA template was amplified using GoTaq G2 Hot Start Green Master Mix (M7423; Promega) according to manufacturer's instructions. Genomic DNA obtained from H441 cells after transfection of *Streptococcus pyogenes* Cas9 (SpCas9)-ribonucleoprotein (RNP) complexes (Synthego, Redwood City, California, USA) was amplified using AmpliTaq Gold 360 Master Mix (4398881; Applied Biosystems, Paisley, UK), in-line with guidance from Synthego. Primer melting temperatures (T_m) were calculated using the T_m calculator online tool matching the polymerase (Thermo Fisher Scientific, Promega). Cycling conditions were designed according to the manufacturer's instructions, taking into consideration the calculated primer T_m and the size of the desired PCR product.

2.2.3 Annealing guide RNA oligos

For generation of CRISPR expression plasmids (Table 2.2), 10µM of forward and reverse oligos, containing the gRNA sequence, were annealed with T4 polynucleotide kinase (NEB) in associated ligation buffer. Thermocycling was performed at 37°C for 30 minutes, then 95°C for 5 minutes, before ramping down to 25°C at a rate of 5°C/min. To allow for downstream ligation of oligo duplexes into BbsI-digested plasmids (Section 2.2.1), the forward oligo contained a 5' CACCG overhang, and the reverse oligo contained 5' AAC and 3' C overhangs.

2.2.4 Gel electrophoresis

Fragments of DNA were resolved by agarose gel electrophoresis. 1-2% agarose gels were generated by heating agarose in a microwave with 1x Tris-acetate-ethylenediaminetetraacetic acid (TAE)

buffer (Thermo Fisher Scientific) until the agarose fully dissolved. After heating, SYBR Safe DNA gel stain (S33102; Invitrogen, Paisley, UK) was added to the agarose solution before being poured into a mould to set. Prior to electrophoresis, DNA samples were mixed with purple loading dye (NEB). Electrophoresis was typically completed at 10-15V/cm until sufficient resolution was achieved. A 1kb or 100bp DNA ladder (NEB) was included on each gel to estimate fragment size. Fragments were visualised under ultraviolet (UV) light at 302nm using an iBright FL1500 imaging system (Invitrogen).

2.2.5 Gel extraction and PCR purification

After restriction enzyme digestion or PCR amplification, DNA fragments were typically resolved by agarose gel electrophoresis (Section 2.2.4). When DNA fragments were purified directly following PCR (Section 2.2.2) or electrophoresis (Section 2.2.4), QIAquick (Qiagen, Hilden, Germany) PCR Purification or Gel Extraction Kits were used, respectively, according to the manufacturer's instructions. Purified DNA fragments were eluted in 10-30µl NFW, depending on concentration of DNA required, and stored at -20°C.

2.2.6 DNA Ligation

Purified DNA fragments were ligated using T4 DNA ligase and associated buffers (NEB) according to manufacturer's instructions. Typically, a 1:3 molar ratio of plasmid vector: insert DNA with 100ng of vector and 1µl of T4 DNA ligase (440U) in a total volume of 20µl was used. The mass of insert DNA required was calculated using the NEBioCalculator: Ligation (NEB). For ligation of gRNA oligos into CRISPR plasmids, annealed oligo duplexes (Section 2.2.3) were diluted 1:20 in NFW before 1µl was ligated with 50ng of a BbsI-digested CRISPR expression plasmid (Table 2.2). Reactions were incubated at 16°C degrees overnight or 10 minutes at RT before use in bacterial transformation (Section 2.2.10).

2.2.7 Topoisomerase-based cloning

The ZeroBlunt topoisomerase-based (TOPO) PCR Cloning Kit (Invitrogen) was used to insert blunt-end PCR products directly into a specialised plasmid vector (PCR-Blunt-II-TOPO; Invitrogen). PCR products were amplified using Phusion Hot Start II Polymerase (Section 2.2.2) and purified as described in Section 2.2.5. Subsequently, 6µl TOPO cloning reactions, containing 50ng PCR product and 1µl PCR-Blunt-II-TOPO, were incubated in a salt solution (200mM sodium chloride (NaCl), 10mM magnesium chloride ($MgCl_2$)) for 30 minutes (RT) and used to transform TOP10 chemically competent *Escherichia coli* (*E.coli*) (Section 2.2.10).

2.2.8 Site-directed mutagenesis

Site-directed mutagenesis (SDM) was completed on plasmid DNA (pDNA) using the QuikChange II XL SDM Kit from Agilent (200522; Stockport, UK), according to manufacturer's instructions with the adaption described below. 50µl PCR reactions were prepared containing 200ng pDNA, 125ng forward and reverse primers, 1µl dNTPs, 5µl Agilent QuikSolution and 1µl *PfuUltra* high-fidelity (HF) DNA polymerase. Reactions were cycled according to manufacturer's instructions with an adapted extension time of 2.5min/kb calculated based on plasmid length. Primers were designed using Agilent software. Following PCR, reactions were digested with *DpnI* at 37°C for 1 hour. *DpnI*-treated DNA was used to transform XL10-Gold Ultracompetent cells (Agilent), according to manufacturer's instructions, except Super Optimal broth with Catabolite repression (SOC) medium was used instead of Casein enzymatic hydrolysate yeast extract (NZY+) broth.

2.2.9 Preparation of selective Luria-Bertani solutions

Luria-Bertani broth was prepared using 1% Tryptone, 1% NaCl and 0.5% granulated yeast extract. Prior to use, Luria-Bertani broth was sterilised by autoclaving and filter sterilised kanamycin (50µg/ml) or ampicillin (100µg/ml) subsequently added. Luria-Bertani agar was prepared by addition of 1.5% Select Agar to Luria-Bertani broth before sterilisation. To prepare Luria-Bertani agar plates, Luria-Bertani agar was heated until molten and cooled to 50°C before addition of

antibiotics. The molten Luria-Bertani agar was then poured into 9cm petri dishes and allowed to set before storage at 4°C. Before use, plates were dried for 1 hour at 42°C.

2.2.10 Bacterial transformation

Plasmid DNA constructs were used to transform chemically competent *E.coli* according to manufacturer's instructions. Stbl3 cells were used for all transformations with the following exceptions: C2925 cells (NEB) were used to generate methylation negative plasmids; TOP10 cells (Invitrogen) were used after TOPO cloning (Section 2.2.7); XL10-Gold ultracompetent cells (Agilent) were used after SDM (Section 2.2.8). 100µl of transformed cells was spread onto Luria-Bertani agar plates (Section 2.2.9) containing the appropriate antibiotic and incubated in an inverted position overnight at 37°C. After confirming no colonies were present on negative control Luria-Bertani agar plates (mock transformation of *E.coli*), single colony transformants were isolated using a pipette tip for colony PCR (Section 2.2.11) and/or further amplification.

2.2.11 Colony PCR

Colony PCR was used to screen colonies when *E.coli* transformation efficiency was low (<5 colonies obtained). First, isolated transformed *E.coli* colonies (Section 2.2.10) were individually streaked onto new selective Luria-Bertani agar plates and incubated overnight at 37°C for downstream bacterial inoculation (Section 2.2.12). After streaking to single colonies, the *E.coli* colonies were added to PCR reactions containing GoTaq G2 Hot Start Green Master Mix (Promega) and PCR amplification performed (Section 2.2.2) with an initial denaturation time of 10 minutes. Primers were designed to amplify a region encompassing the desired insertion (Section 2.2.6). Products were visualised by gel electrophoresis (Section 2.2.4).

2.2.12 Bacterial inoculation of growth media

A single transformed *E.coli* colony, grown overnight on selective Luria-Bertani agar plates (Section 2.2.10), was inoculated into 5ml, 200ml or 400ml Luria-Bertani broth (Section 2.2.9) containing the appropriate antibiotic. Cultures were incubated overnight in an orbital shaking incubator (37°C, 225RPM) before pDNA was extracted (Section 2.2.13).

2.2.13 Plasmid purification

Plasmid DNA from 5ml cultures were purified using the Wizard Plus SV MiniPrep DNA Purification System (Promega), according to manufacturer's instructions, and typically eluted in 50-100µl NFW. For larger scale preparations of pDNA, plasmid purification kits from Qiagen (Maxi, Mega) were used as per manufacturer's guidance. pDNA was typically eluted in 100-500µl NFW. All purified DNA was stored at -20°C.

2.2.14 Quantification of DNA and RNA

The concentration of DNA and RNA was measured using a NanoDrop Spectrophotometer 2000 (Thermo Scientific). The purity of samples was assessed by the optical density $(OD)_{260}/OD_{280}$ ratio, where values >1.8 for DNA and >2.0 for RNA were considered pure. Each 1.5µl sample was read in duplicate or triplicate, and the average used to determine the final values.

2.2.15 Sanger sequencing

Gene expression cassettes in newly constructed plasmids, as well as purified PCR products, were sent to Source BioScience for Sanger sequencing. Plasmid templates were provided at 100ng/µl and PCR products at 10ng/µl, for 5µl per sequencing reaction. Primers were supplied at 3.2pmol/µl, 5µl per reaction. Deoxyguanosine triphosphate chemistry was used to improve the sequencing data quality.

2.2.16 Inference of CRISPR edits analysis

The PCR products amplified from non-edited control and CRISPR-edited gDNA samples (Section 2.2.2) were submitted for Sanger sequencing (Section 2.2.15) and the resultant sequencing data analysed by Inference of CRISPR Edits (ICE) software (Synthego).

2.2.17 Cre recombination assay

TOPO-cloned plasmids (Section 2.2.7) containing loxP sites were incubated with Cre recombinase (NEB) as per the manufacturer's instructions. In brief, 250ng pDNA was incubated with 0.8 units Cre recombinase at 37°C for 30 minutes then 70°C for a further 10 minutes. The number of DNA products were then visualised by agarose gel electrophoresis (Section 2.2.4) to assess recombination. A control plasmid, pLOX (NEB), was included as a positive control for successful recombination and pDNA without loxP sites used as a negative control.

2.3 Mammalian cell culture

2.3.1 Cell lines and culture conditions

All cell culture reagents were obtained from Gibco and cell lines obtained from the American Type Culture Collection (ATCC; Manassas, Virginia, USA) unless otherwise stated. Adherent and suspension cell lines were cultured in a static humidified incubator (37°C, 5% CO_2) and orbital shaking, humidified incubator (37°C, 8% CO_2, 135RPM) respectively. Cells were routinely screened for mycoplasma (G238; Applied Biological Materials Inc, Vancouver, Canada) and visually checked for signs of contamination. Live cell counts were determined either with trypan blue staining and manual counting with a Neubauer haemocytometer, or by automated cell counting using a NucleoCounter NC-200 (ChemoMetec, Allerod, Denmark) and Via1-Cassettes (ChemoMetec) according to manufacturer's instructions.

Cell lines were passaged biweekly at ca. 70-90% confluence between 1:3 and 1:10 as appropriate. For adherent cell lines, cells were washed with dPBS prior to incubation with TrypLE (trypsin

derivative) at 37°C until cells were visibly detached (approximately 5-20 minutes). For suspension cell lines, cells were typically centrifuged at 500*g* for 5 minutes before resuspension in fresh media. For cryopreservation, cells were washed and detached as described above, resuspended in culture medium supplemented with 10% dimethyl-sulfoxide (DMSO) and frozen in freezing containers with isopropanol at -80°C, before long-term storage in liquid nitrogen.

2.3.1.1 Immortalised cells

Adherent HEK293T cells (HEK293T/17, CRL-11268) were cultured in Dulbecco's Modified Eagle's Medium (DMEM). Human A549 cells (CCL-185) were cultured in Kaighn's Modification of Ham's F-12 Medium (F-12K). Human H441 cells (HTB-174) were cultured in Roswell Park Memorial Institute 1640 Medium (RPMI-1640) ATCC modification. All media were supplemented with 100U/ml penicillin, 100mg/ml streptomycin, 2mM L-glutamine and 10% heat inactivated (HI; 60°C, 1 hour) Foetal Bovine Serum (FBS; Sigma) except in H441 maintenance medium where L-glutamine was excluded and 10% non-heat inactivated FBS (Sigma) was used (Table 2.4).

Serum-free, suspension-adapted HEK293T cells (HEK293T/17 SF, ACS-4500; herein referred to as HEK293T-17) were cultured in FreeStyle 293 Expression Medium and LV-MAX viral production cells (A35347; Gibco) in LV-MAX production medium without serum or antibiotics.

Table 2.4 List of cell lines and standard culture conditions used

Cell line	Culture	Maintenance Medium	Use
HEK293T-17	Suspension	FreeStyle 293 expression medium	Recombinant LV production
LV-MAX	Suspension	LV-MAX production medium	Recombinant LV production
HEK293T	Adherent	DMEM, 2mM L-glutamine, 100 U/ml penicillin, 100mg/ml streptomycin, 10% HI FBS	Multi-purpose
A549	Adherent	F-12K, 100 U/ml penicillin, 2mM L-glutamine, 100mg/l streptomycin, 10% HI FBS	ATII cell model
H441	Adherent	RPMI-1640, 100 U/ml penicillin, 100mg/ml streptomycin, 10% FBS	ATII cell model

2.3.1.2 Primary cells

Human alveolar pneumocytes (EP71AL; Epithelix, Geneva, Switzerland) were cultured in Pneumocyte culture medium (Epithelix) according to supplier's instructions. In short, cells were passaged at 80% confluence using dPBS and TrypLE as described (Section 2.3.1) before centrifugation (500*g*, 10 minutes) in an excess of DMEM supplemented with 10% HI-FBS and resuspension in Pneumocyte culture medium.

2.3.2 Air-liquid interface cultures

2.3.2.1 Establishment of air-liquid interface cultures

Air-liquid interface cultures were established by culturing cells on 12-well or 24-well Transwell inserts (CLS3460, CLS3470; Corning, Flintshire, UK) as described by Munis *et al.* [103]. In brief, approximately $1x10^5$ or $5x10^4$ total cells/well in maintenance media were seeded onto 12-well or 24-well Transwells respectively. 48-72 hours after seeding (day 3/4), cells were air-lifted by removal of medium from the apical side of the Transwell insert, and replacement of medium on the basolateral side of the insert with 'polarisation' medium comprising: the relevant base media supplemented with 2mM L-glutamine, 100U/ml penicillin, 100mg/ml streptomycin, 1% insulin-transferrin-selenium (ITS), 4% FBS and 1µM dexamethasone. A 20µg/ml solution of dexamethasone (D8893) was prepared in absolute ethanol and sterile RPMI-1640 medium, as per manufacturer's instructions, and stored at -20°C. Typically, polarisation media was replaced three times per week for a minimum of 14 days before ALI cultures were used in experiments. All ALI cultures established from H441 cells are referred to as SALIs.

2.3.2.2 Harvesting cells from air-liquid interface cultures

Cells were harvested from ALI cultures for subsequent DNA (Section 2.5.1), RNA (Section 2.5.2) or protein extraction (Section 2.5.3). Basolateral medium was removed before the apical Transwell surface was washed with dPBS and incubated with TrypLE at 37°C for 45 minutes. Following

incubation with TrypLE, the apical Transwell surface was gently scraped with a pipette tip and cells resuspended in dPBS.

2.3.3 Transfection

Adherent cells were seeded ($5x10^5$ cells/well) in 6-well plates 24 hours before transient transfection of pDNA with Lipofectamine 3000 (Invitrogen). Cells were seeded to achieve 70-90% confluence at the point of transfection. Typically, for 6-well plates, 1µg of DNA complexed with 6.25µl of Lipofectamine 3000 was added per well. Complexes were prepared in OptiMEM as per manufacturer's instructions and reagents scaled as advised for alternative culture vessels. Transfection of H441 cells was completed according to the H441-specific optimised protocol published by Thermo Fisher Scientific. Cells were typically analysed 48-72 hours after transfection.

For transfection of SpCas9-RNP complexes (Synthego), cells were reverse transfected using Lipofectamine CRISPRMAX reagent (Invitrogen) according to the manufacturer's 'CRISPR Editing of Immortalized Cells with RNPs using Lipofection Protocol' (Synthego). Briefly, SpCas9 and synthetic gRNA (sgRNA) were complexed according to manufacturer's instructions before reverse transfection of cells ($6x10^4$cells/well) with 3.9pmol sgRNA and 3pmol SpCas9 per well in a 24-well plate format.

For transient transfection of pDNA for viral vector production see Section 2.4.

2.3.4 Puromycin selection

Puromycin selection was completed on H441 cells after transfection or transduction with vectors containing selection cassettes for puromycin resistance. 24 hours post-transfection, cells were temporarily cultured in media containing 2.5µg/ml puromycin for 3 days before being returned to standard complete media. From 48 hours after transduction, H441 cells were continuously maintained in media containing 2µg/ml puromycin.

2.3.5 Doxorubicin assay

H441 cells were seeded in triplicate in 96-well plates at a density of $1.4x10^4$ cells per well. 24 hours after seeding, cells were treated with different concentrations of doxorubicin hydrochloride (herein referred to as doxorubicin; 0-20µM) for 3 hours. Cells were subsequently washed with dPBS and incubated for a further 24 hours in maintenance medium under standard conditions. Cell viability was then assessed by reduction of 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltatrazolium bromide (MTT) by metabolically active cells (CellTiter 96 Non-radioactive Cell Proliferation Assay; Promega) according to manufacturer's instructions. Absorbance was measured at 570nm and 650nm using a SpectraMAX i3x microplate reader (Molecular Devices, San Jose, California, USA), and percentage viability calculated for each condition, using non-doxorubicin treated cells as the control, as described by Promega.

2.4 Recombinant lentiviral vector production

Third-generation rLV vectors based on HIV-1 (rHIV) or SIV_{AGM} (rSIV) were produced in LV-MAX cells (Section 2.4.1) or suspension-adapted HEK293T-17 cells (Section 2.4.2) at a scale of 25ml (small-scale, SS), 1.5l (LV-MAX large-scale, LS) or 5l (HEK293T-17 LS). Recombinant LV was pseudotyped with F/HN or VSV-G and were produced after transient transfection of producer cells with genome, Gag-Pol, Rev and pseudotyping plasmids (Table 2.2) at a ratio of 20:9:6:12 unless otherwise stated. Where multiple pseudotyping plasmids were used (F/HN) the ratio was divided equally. Cell supernatants were harvested 48 hours post-transfection. One hour before harvest of cell supernatant, 1mM $MgCl_2$ and Benzonase (5U/l; Merck) were added to reduce nucleic acid contaminants. Crude rLV was harvested using a GE Ultra 5µm filter (GE Healthcare) and an AcroPak 0.8/0.45µm filter (Pall, Portsmouth, UK). HIV-1 rLV preparations were further purified by centrifugation (Section 2.4.3.1), whilst LS productions underwent column anion exchange chromatography (AEX; Section 2.4.3.2) and tangential flow filtration (TFF; Section 2.4.3.3) for

purification and concentration. After production, vectors were aliquoted for single-use and stored at -80°C.

The viral vectors used in this study are summarised in Table 2.3. When manufactured at LS, rLV was produced (Section 2.4.1 and 2.4.2.2), purified (Section 2.4.3.2 and 2.4.3.3) and titrated (Section 2.4.4) by the Virus Production Team in the Gene Medicine Research Group (NDLCS, University of Oxford, UK).

2.4.1 Recombinant LV production in LV-MAX cells

Production using LV-MAX cells were completed according to manufacturer's instructions for the LV-MAX Lentiviral Production System (Gibco). Cells were seeded in polycarbonate Erlenmeyer flasks (Corning) at a density of $4x10^6$cells/ml before transfection 24 hours later of 0.63µg total pDNA/$1x10^6$ cells with LV-MAX transfection reagent. Five hours post-transfection LV-MAX enhancer was added.

2.4.2 Recombinant LV production in suspension-adapted HEK293T-17 cells

For production using suspension-adapted HEK293T-17 cells, cells were transfected using PEIpro (PolyPlus-transfection) with a total of 0.33µg pDNA/$1x10^6$ cells at a PEIpro:DNA volume:mass ratio of 2:1. Approximately 18 hours post-transfection 5mM sodium butyrate was added to cells to boost production yields.

2.4.2.1 Small-scale (25ml)

For SS rLV production, HEK293T-17 cells were seeded at $1x10^6$ cells/ml in polycarbonate Erlenmeyer flasks 24 hours prior to transfection.

2.4.2.2 Large-scale (5l)

For LS 5l rLV production, HEK293T-17 cells were seeded at 2-$4x10^6$ cells/ml in single-use WAVE Cellbag bioreactors (GE Healthcare, Amersham, UK) containing FreeStyle media supplemented with

2.3ml/l 10% Pluronic F-68 (Gibco). Mounted on a WAVE 20/50 rocking platform (20RPM, 8° angle; GE Healthcare), Cellbags were inflated with 5% CO_2 and incubated at 37°C for 48 hours prior to transfection. The pH of the growth media was automatically maintained at 7.2 by a WAVEPODII controller (GE Healthcare) using CO_2 and sodium bicarbonate.

2.4.3 Purification and concentration

2.4.3.1 Centrifugation

For production of rLV vectors based on HIV-1, after 0.45µm filtration, material was purified by overnight centrifugation at 4800*g* at 4°C. The resulting pellet was subsequently resuspended in OptiMEM by pipetting (approximately 60x) before aliquoting and storage at -80°C.

2.4.3.2 Anion exchange chromatography

Anion exchange chromatography was performed using an ÄKTA Pure 150L chromatography station (GE Healthcare) with 5ml Mustang Q XT5 quaternary ammonium AEX membranes (Pall). Before use, viral containing supernatants were diluted (1:1) in 50mM Tris (pH7.5, 4°C) to reduce the conductivity and ensure efficient viral particle capture by the AEX membranes. Viral supernatant was loaded onto the membranes at a flow rate of 25ml/min before washing with a low-salt buffer (150mM NaCl, 25mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), pH7.5) and elution in high-salt buffer (1M NaCl, 25mM HEPES, pH7.5). Ultraviolet absorbance peaks were used to identify vector-containing fractions.

2.4.3.3 Tangential flow filtration

After AEX (Section 2.4.3.2), rLV was concentrated and purified via TFF using a KR2i TFF system (Repligen, Waltham, Massachusetts, USA). In brief, vector was ultra-filtered to 10ml using a 500kDa pore, 20cm^2 membrane (Spectrum Lifesciences, Rancho Dominguez, California, USA) at 15ml/min before diafiltration against 100ml TSSM buffer (20mM tromethamine, 100mM NaCl, 10mg/ml

sucrose, 10mg/ml mannitol) for approximately 10 diavolumes and followed by a final ultrafiltration volume reduction step yielding approximately 5ml of final rLV.

2.4.4 Determination of viral titre

The titre of rLV vectors was determined by quantitative reverse transcription PCR (qRT-PCR) (Section 2.5.5) detection of WPRE in integrated rLV genomes. LV-MAX cells were transduced with dilutions of purified rLVs and gDNA extracted (Section 2.5.1) after 72 hours. The number of integrated WPRE sequences was quantified, in triplicate, against a WPRE sequence DNA standard curve and endogenous gDNA control sequence (*CFTR* intron 1) using primers listed in Table 2.1 (DNAVEC_WPRE, hCFTRInt1). 'TaqMan transducing units' (TTU) per ml of vector preparation were determined by generating a linear regression curve from the dilution series in Prism 8 software (GraphPad Software, San Diego, California, USA).

2.5 Gene expression analysis

2.5.1 DNA extraction

Extraction of DNA from cultured cells was completed using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. The AllPrep DNA/RNA Mini Kit (Qiagen) was used for simultaneous extraction of DNA and RNA from murine tissues. Tissues were homogenised in Buffer RLT Plus (Qiagen) using Lysing Matrix D tubes (MP Biomedicals, Cambridge, UK) and the FastPrep Instrument (MP Biomedicals) for 40 seconds at 6.0m/sec.

2.5.2 RNA extraction

Total RNA was extracted from cultured cells using the RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. On-column DNase I digestion was performed to prevent gDNA carry-over according to manufacturer's instructions. For extraction of RNA from tissue the AllPrep DNA/RNA Mini kit (Qiagen) was used for simultaneous purification of gDNA and total RNA according to manufacturer's instructions (Section 2.5.1).

2.5.3 Protein extraction

Cultured cells were suspended in a lysis buffer (25mM Tris-HCl [pH 7.5], 150mM NaCl, 1% Triton X-100) supplemented with cOmplete mini protease inhibitor cocktail (Roche, Switzerland) and incubated at 4°C for 30 minutes. Supernatant was harvested from cell suspensions after centrifugation (17,000*g*, 20 minutes, 4°C) and stored at -80°C.

Murine tissue was homogenised in passive lysis buffer (E194A, Promega) supplemented with protease inhibitor using Lysing Matrix D tubes and the FastPrep Instrument for 40 seconds at 6.0m/sec. Homogenates were subsequently transferred to a fresh Eppendorf tube and centrifuged (17,000*g*, 20 minutes, 4°C) before supernatant was stored at -80°C.

2.5.4 Quantitative reverse transcription PCR

Quantitative RT-PCR was typically completed on RNA samples using Power SYBR Green RNA-to-Ct kit (Applied Biosystems) in a QuantStudio 7 Flex Real Time PCR System (Thermo Scientific), apart from for the determination of viral titre (Section 2.4.4) when TaqMan FAST Universal PCR Master Mix (Applied Biosystems) was used. Reactions were prepared with three technical replicates using 100ng RNA/well and a final primer concentration of 150nM (Table 2.1), according to manufacturer's instructions (Applied Biosystems). The thermal dissociation protocol was run at the end of the PCR to assess the specificity of products present in the reaction. Target genes with a cycle threshold (Ct) ≥40 were considered undetectable. The relative expression level of target genes was determined using the $2^{-\Delta\Delta Ct}$ method, with β-actin as the reference gene. Expression was typically normalised to that of a naïve control.

2.5.5 Droplet digital PCR

All droplet digital PCR (ddPCR) reagents were obtained from Bio-Rad (Watford, UK) unless stated otherwise. Copy number variation (CNV) ddPCR was completed on gDNA samples using an automated droplet generator (Bio-Rad) and a QX200 droplet reader (Bio-Rad) for analysis.

Reactions were prepared using ddPCR Supermix for probes (no deoxyuridine triphosphate) with 60ng gDNA, 900nM primers and 250nm probes (see Table 2.1) according to manufacturer's instructions. In brief, droplets were generated before undergoing PCR amplification and subsequent CNV analysis. Reactions were completed with at least two technical replicates.

2.5.6 Western blot

Approximately 10-40µg of protein samples were analysed by immunoblotting after total protein concentrations were determined using the Pierce bicinchoninic acid (BCA) Protein Assay Kit (Thermo Scientific), according to manufacturer's instructions. Prior to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), samples were typically incubated under reducing conditions at 37°C for 15 minutes in Laemmli buffer (#1610747, Bio-Rad) supplemented with a final concentration of 2.5% 2-mercaptoethanol. Samples were separated on 8% Bis-Tris polyacrylamide gels, prepared according to Table 2.5 using reagents obtained from Bio-Rad, in 3-morpholinopropane-1-sulfonic acid (MOPS) buffer (Bio-Rad). The SDS-PAGE was completed at 100V for 15 minutes, then 170V until the desired level of separation was observed. Unless stated otherwise, gels were transferred onto 0.4µm nitrocellulose membranes (Bio-Rad) by wet transfer at 350mAmps. Alternatively, after equilibration in 20% ethanol for 10 minutes, gels were transferred onto 0.2µm nitrocellulose membranes (Bio-Rad) using the iBlot 2 Gel Transfer Device (Invitrogen), according to manufacturer's instructions. A dry transfer programme of 20V (2 minutes), 23V (4 minutes), 25V (4 minutes) was used. Membranes were blocked in 5% skimmed milk powder in tris-buffered saline (TBS) supplemented with 0.1% Tween-20 for 1 hour. Membranes were then incubated with primary antibodies at 4°C overnight, before incubation with appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies for 1 hour at RT. Bands were visualised using the Clarity Western ECL Substrate kit (Bio-Rad), and images captured using the iBright FL1500 imaging system. Primary and secondary antibodies used are listed in Table 2.6.

Table 2.5 Preparation of Bis-Tris polyacrylamide gels

Reagent	8% Resolving gel	4% Stacking gel
Acryl/Bis 30% (37.5:1)	6ml	1.33ml
Tris 1.5M pH 8.8	7.5ml	NA
Tris 0.5M pH6.8	NA	1.25ml
10% Sodium dodecyl sulphate	200µl	50µl
Distilled water	6ml	7.4ml
10% Ammonium persulphate	200µl	150µl
TEMED	20µl	15µl

Table 2.6 List of antibodies used

Use	Target	Antibody	Dilution	Supplier	Catalogue number
Western blot	ABCA3	Rabbit anti-ABCA3	1:1000	Abcam	ab99856
	ABCA3	Rabbit anti-ABCA3	1:500	Invitrogen	PA5-52478
	β-tubulin	Rabbit anti- β-tubulin	1:500	Invitrogen	Ab6046
	Secondary	Swine anti-rabbit immunoglobulins/HRP	1:3000	Dako	P0399
Human ICC	ABCA3	Rabbit anti-ABCA3	1:200	Invitrogen	PA5-52478
	proSP-C	Rabbit anti-proSP-C	1:200	Millipore	AB3786
	Secondary	Goat anti-rabbit AlexaFluor555	1:500	Invitrogen	A21428
Murine IHC	ABCA3 (mouse/human)	Rabbit anti-ABCA3	1:100	Seven Hills Bioreagents	WRAB-70565
	Secondary	Goat anti-rabbit AlexaFluor555	1:500	Invitrogen	A21428

2.6 Histology

2.6.1 Cryosectioning

Cryosections of 7µm were generated using a CryoStar NX50 (Thermo Scientific) with the specimen holder set to -30°C. Sections were collected after tissue was trimmed in 30µm increments until the region of interest was located. Consecutive cryosections were collected onto Superfrost Plus glass slides (VWR), air-dried for 30 minutes and stored at 4°C before staining.

2.6.2 Haematoxylin and Eosin staining

Murine lung cryosections (Section 2.6.1) were stained with Harris haematoxylin and Eosin Y (0.5% alcoholic; H&E) using a progressive protocol. Cryosections were first air-dried for 5 minutes and

dehydrated in 100%, then 70%, ethanol for 2 minutes each, before three 2-minute washes in Milli-Q water. Samples were subsequently immersed in haematoxylin for 15 seconds and rinsed in tap water (John Radcliffe Hospital, Oxford, UK), before multiple (5) wash steps in Milli-Q water as described above. Cryosections were finally immersed in Eosin for 45 seconds before rehydrating in increasing concentrations of ethanol (70-100%), clearing in Xylene for 4 minutes and mounting with DPX mounting medium.

2.6.3 Native fluorescence imaging

Cryosections (Section 2.6.1) and precision-cut lung slices (PCLS; Section 2.8) from inducible *Abca3* KO (*Abca3*-iKO) mice were imaged for native yellow fluorescent protein (YFP) expression after treatment with tamoxifen. The PCLS were imaged live *in situ*, but cryosections were first incubated in 300nM 4',6-diamidino-2-phenylindole (DAPI; Invitrogen) diluted in TBS for 5 minutes before mounting with ProLong Gold or Diamond Antifade Mounting Medium (Invitrogen). Mounted samples were cured for a minimum of 24 hours before images were captured (Section 2.6.5.1).

2.6.4 Immunostaining

Adherent cells and cryosections were stained in a humid environment using appropriate cell culture plasticware or a humid Staintray staining system (Simport, Quebec City, Canada), respectively.

2.6.4.1 Immunocytochemistry

For immunocytochemistry (ICC), cells were first fixed with 1% PFA for 20 minutes, permeabilised using 0.1% triton X-100 in PBS for 10 minutes and then blocked for 1 hour in 1% bovine serum albumin (BSA) dissolved in PBS. For all subsequent steps solutions were diluted in PBS containing 0.1% triton X-100 and 1% BSA. After blocking, cells were incubated with the appropriate primary and then secondary antibodies (see Table 2.6) for 1 hour each. Between all steps cells were washed twice in PBS. Before imaging, nuclei were stained with 300nM DAPI for 5 minutes.

2.6.4.2 Immunohistochemistry

For cryosections, slides were washed in TBS for 30 minutes before being blocked and permeabilised in 1% BSA, 5% normal goat serum (NGS), 0.1% Triton X-100 in TBS for 20 minutes. This solution was used as the diluent in subsequent steps. After washing in TBS for 5 minutes, slides were incubated with the appropriate primary and secondary antibodies (see Table 2.6) for 2 hours each, with an intervening 40-minute wash in TBS. Slides were then incubated with DAPI and mounted as described in Section 2.6.3.

2.6.5 Microscopy

2.6.5.1 Brightfield and fluorescence microscopy

All fluorescence and brightfield images were captured using an EVOS FL Auto 2 Imaging System (Invitrogen) with associated DAPI (AMEP4950), green fluorescent protein (GFP; AMEP4951), Texas Red (AMEP4955) and YFP (AMEP4954) light cubes. For comparison of samples, light intensity, exposure and gain settings were fixed. A minimum of three images were taken per sample with the field of view typically selected at random based on brightfield imagery or DAPI staining to identify location of cells. For capture of stitched images, the sample location was first determined using brightfield microscopy before the field was selected using the EVOS FL Auto 2 software (Invitrogen). The software was then used to automate image capture and stitching of the selected fields. Experiments generally included three technical replicates. Processing and analysis of images was carried out using Fiji software [190].

For analysing the percentage of pro-SPC positive cells, the overall cell number per image was first calculated using the nuclear counterstain. Then the threshold fluorescence for pro-SPC staining was determined from secondary antibody only controls and the percentage of pro-SPC positive cells above threshold calculated for each image. The analysis excluded particles formed from less than 100 pixels.

2.6.5.2 Transmission electron microscopy

After fixation, all sample preparation and TEM was completed by the Dunn School Electron Microscopy Facility (University of Oxford, Oxford, UK) using a JEOL JEM-1400Flash 120kV TEM equipped with a Gatan Rio camera. Cells were cultured on Transwell inserts in submerged conditions (Section 2.3.1) or in an ALI culture (Section 2.3.2). Samples were fixed in 4% formaldehyde and 2.5% glutaraldehyde in 0.1M 1,4-Piperazinediethanesulfonic acid (PIPES) buffer (pH7.2), applied to the apical and basolateral surface, for 1 hour at RT before storage at 4°C and processing. For time course studies, samples were immersed in a storage buffer (0.25% glutaraldehyde in 0.1M PIPES buffer, pH7.2) at 4°C, 24 hours after fixation, to ensure concurrent sample processing.

For processing (Dunn School), the samples were washed with buffer (0.1M PIPES buffer pH7.2) then incubated with buffer containing 50mM glycine. Samples were then washed again in buffer. During this step the Transwell membranes were excised from the surrounding plastic insert using a scalpel and cut into strips for further processing. Sample strips were stained with 1% (w/v) osmium tetroxide and 1.5% potassium ferrocyanide in buffer, washed extensively with Milli-Q water and further stained with 0.5% uranyl acetate. The samples were washed again with Milli-Q water and dehydrated through an ethanol series. Following dehydration samples were infiltrated with and embedded in low viscosity epoxy resin (TAAB) ahead of polymerisation at 60°C for 24 hours. Sections of 90nm were cut from the resin blocks using a Leica UC7 Ultramicrotome and collected onto 3mm copper grids. The sections were then post-stained with lead citrate and imaged.

2.7 Animal studies

2.7.1 Maintenance and housing

All animal experiments were carried out under United Kingdom Home Office approved project (PP9473919) and personal licences in accordance with the Animals (Scientific Procedures) Act 1986. Studies involving the induction of *Abca3* deficiency (Chapter 5) were authorised to involve significant physiological dysregulation – such sensitive studies were also subject to review by ACER - the central University of Oxford Animal welfare and ethics review board and included extensive (typically 2-4 times/day) animal monitoring by an extended team of personal licence holders (Section 2.7.3.4). All other studies were limited to induce a maximum of moderate physiological dysregulation. Mice were housed by the University of Oxford Biomedical Services (BMS; John Radcliffe Hospital, Oxford, UK) and fed on standard chow and water *ad libitum*, unless stated otherwise, under housing and observational conditions in accordance with the UK Home Office ethical and welfare guidelines.

All animals were identified by ear notching before study initiation, and prior to the commencement of procedures mice were allowed to acclimatise for at least 7 days. At the experimental end point, or when mice were below 85% of their weight at study initiation, animals were sacrificed by cervical dislocation.

2.7.2 Mouse strains

Wild-type, pathogen-free, female BALB/c mice purchased from Envigo (Loughborough, UK) were used for alveolar promoter development studies (Section 3.2.4) and vector transduction studies (Section 5.2.1). Mice were aged 6 weeks at study initiation.

An *Abca3*-iKO mouse model (Section 5.2.2) was rederived in-house (Section 2.7.3.1) and used in preliminary disease induction studies with tamoxifen. All *Abca3*-iKO mice were aged 8-25 weeks at the start of experimentation.

2.7.3 Inducible *Abca3* knockout model

In *Abca3*-iKO mice, LoxP sites flank exons 4 and 7 of the *Abca3* gene, containing the start codon, transmembrane domain 1 and extracellular domain 1. LoxP sites also flank a strong tpA stop sequence upstream of enhanced YFP (EYFP) in the *Rosa26* locus [191], preventing EYFP transcription in an uninduced state. Recombination between LoxP sites (model induction) is dependent on exposure to Cre recombinase. In addition to the two floxed transcriptional elements (Table 2.7), *Abca3*-iKO mice contain a Cre recombinase oestrogen receptor fusion protein ($CreER^{T2}$) under the transcriptional control of the endogenous murine *Sftpc* promoter – an ATII cell restricted promoter (Table 2.7).

Under basal conditions, the Cre recombinase fusion protein is sequestered outside the nucleus and unable to interact with LoxP sites in *Abca3* and *Rosa26* loci. In the presence of tamoxifen, Cre recombinase relocates to the nucleus resulting in excision of floxed elements, KO of *Abca3* and induction of EYFP expression.

Table 2.7 Transgenic alleles in *Abca3*-iKO mice

Common name	Allele	MGI allele ID	Reference
$Abca3^{flox}$	$Abca3^{flox}$	-	[127, 128] Jeffrey Whitsett
Sftpc-CreER^{T2}	$Sftpc^{tm1(Cre/ERT2,rtTA)Hap}$	MGI:5444645	[192, 193] Harold Chapman
EYFP	$Gt(ROSA)26Sor^{tm1(EYFP)Cos}$	MGI:2449041	[191] Frank Costantini

2.7.3.1 Rederivation and breeding

With assistance from BMS staff, homozygous *Abca3*-floxed ($Abca3^{flox/flox}$) mice and heterozygous *Abca3*-floxed ($Abca3^{flox/WT}$), Sftpc-Cre positive (*Sftpc-CreER*$^{T2+}$), *R26R-EYFP* positive ($EYFP^{+}$) mice (kind gifts from Jeffrey Whitsett; see Table 2.7) were used to generate mice with an $Abca3^{flox/flox}$, *Sftpc-CreER*$^{T2+/+}$, $EYFP^{+/+}$ genotype (*Abca3*-iKO). $Abca3^{flox/flox}$ mice were first bred with $Abca3^{flox/WT}$,

Sftpc-CreER^{T2+}, *EYFP^{+}* mice before *Abca3$^{flox/flox}$*, *Sftpc-CreER^{T2+}*, *EYFP^{+}* mice underwent live rederivation to remove any adventitious organisms (see Section 5.2.2.1 for more details; Figure 5.4)

After rederivation, mice were bred in pairs, for up to 12 months of age, until a fully homozygous genotype was observed. Pups were weaned and ear notches taken for identification and genotyping (Section 2.7.3.2) at 3-4 weeks. Unless stated otherwise, all mice used in experimental studies were predicted to be *Abca3$^{flox/flox}$*, *Sftpc-CreER$^{T2+/+}$*, *EYFP$^{+/+}$*.

2.7.3.2 Genotyping

Genomic DNA was extracted from ear punch biopsies using the QIAamp DNA Mini kit (Qiagen), with tissue lysis at 56°C for 3 hours, before being used in PCR reactions (Section 2.2.2) with primers targeting the LoxP site upstream of exon 4 in *Abca3*, *CreER^{T2}* and *EYFP* (see Table 2.1). Mice homozygous or heterozygous for the *Abca3^{flox}* allele were identified by the presence of one (400bp) or two (300bp, 400bp) PCR products respectively after agarose gel visualisation (Section 2.2.4). Mice were identified as positive (PCR product) or negative (no PCR product) for *Sftpc-CreER^{T2}* and *EYFP* alleles. After *Abca3$^{flox/flox}$*, *Sftpc-CreER^{T2+}*, *EYFP^{+}* colonies were established, ear punch biopsies were sent to Transnetyx (Cordova, Tennessee, USA) for genotyping (real-time PCR-based approach) to understand the zygosity at the *Sftpc* and *Rosa26* loci.

2.7.3.3 Tamoxifen administration

2.7.3.3.1 Intraperitoneal injection

A 20mg/ml solution of tamoxifen was prepared by dissolving in preheated corn oil (50°C), before agitation at 37°C overnight and storage at -80°C. Prior to use, 20mg/ml tamoxifen was thawed at 37°C (135RPM) before short-term storage (<1 week) at 4°C. Animals were dosed with 100mg/kg or 200mg/kg of 20mg/ml tamoxifen by intraperitoneal (IP) injection for a maximum of 5 days consecutively with or without anaesthesia (Section 2.7.4). On each injection day, a maximum

volume of 2ml of tamoxifen was injected into the peritoneal cavity after manual restraint using an insulin syringe with 29G needle (Braun, Sheffield, UK).

2.7.3.3.2 Diet

Animals were administered a tamoxifen diet containing 0.25g/kg tamoxifen and 5% sugar (E8220 Version 352; SAFE Complete Care Compliance; Aston Pharma, London, UK) *ad libitum* for 8 or 14 days consecutively. Alternatively, the tamoxifen diet was pulsed for five days before returning to standard chow for two days, for a total of 15 days. Following the described period on a tamoxifen diet, mice were maintained on standard chow until the study endpoint. Mice and food were weighed regularly to monitor intake of tamoxifen chow.

2.7.3.4 Animal monitoring

Administration of tamoxifen (Section 2.7.3.3) was anticipated to induce a severe respiratory disease phenotype in *Abca3*-iKO mice. To limit prolonged severe physiological impairment and establish a humane endpoint for respiratory distress, after tamoxifen administration body weight, temperature and respiratory effort were typically assessed 2-4 times a day at set timepoints. Surface body temperature was obtained from the perianal region using an infrared thermometer (193-8691; RS Components, Corby, UK). Core body temperatures were monitored using temperature programmable microchips (United Information Devices (UID), Lake Villa, Illinois, USA) that were implanted subcutaneously over the flank of *Abca3*-iKO mice prior to tamoxifen delivery. Mice were first anaesthetised (Section 2.7.4) and placed prone onto heated mats, before a microchip pistol grip injector (UID) was used to implant the internal temperature transponder. A non-invasive URH-1HP Reader (UID) was used to scan microchips and obtain temperature data. Respiratory distress was assessed using criteria established by Kang *et al.* [122]. In brief, mice were scored from M0 (no phenotype) to M3 (severe phenotype) based on the respiratory rate, any periods of inactive behaviour and cyanosis.

2.7.4 Mouse anaesthesia

For administration of substances (Section 2.7.5), excluding via diet, and *in vivo* bioluminescence imaging (BLI; Section 2.7.6), anaesthesia was induced by inhalation of 3-5% isoflurane (Abbot, Maidenhead, UK) and maintained with 3% isoflurane. Depth of anaesthesia was continuously monitored by observing breathing rate and testing pedal reflex responses.

2.7.5 Intranasal administration of substances

Nasal instillation was used to deliver rLV and D-luciferin (PerkinElmer, Waltham, Massachusetts, USA) to the mouse lung. Recombinant LV was formulated (Section 2.4) and diluted to an appropriate dose in TSSM before repeat administration for four consecutive days. Substances were delivered in a 100µl final volume and administered as described by Xenariou *et al.* [194]. In brief, anaesthetised mice (Section 2.7.4) were held vertically with the mouth closed, and a continuous droplet of substance applied topically to the nose.

2.7.6 *In vivo* bioluminescence imaging

In vivo BLI was completed in mice dosed with rLV expressing firefly Luciferase (fLux), under anaesthesia (Section 2.7.4), 10 minutes after intranasal administration of 15mg/ml D-luciferin (Section 2.7.5). An IVIS Lumina LT Imaging System (PerkinElmer) in conjunction with Living Image v4.5 (PerkinElmer) software was used to visualise and capture photon emissions. Luciferase activity in the nose and lung of each mouse was calculated by drawing a region of interest (ROI) and determining the average luminescence as photons/sec/cm^2. Dividing the average luminescence in the lung by the average luminescence in the nose was used as a proxy for alveolar specificity (see Section 3.2.4 for details).

2.7.7 Lung tissue harvesting

After cervical dislocation, mice were dissected to expose the lungs and trachea. A cannula attached to a blunt 21G needle (Becton Dickinson, Plymouth, UK) was inserted into a small incision made

underneath the larynx and secured to the trachea with sutures (Ethicon, Raritan, New Jersey, USA). The trachea and lungs were subsequently removed and rinsed in dPBS before the superior, middle, inferior and post-caval lobes were isolated and immersed in RNAlater solution (Invitrogen) for later extraction of DNA (Section 2.5.1), RNA (Section 2.5.2) and protein (Section 2.5.3). As per manufacturer's instructions, tissue immersed in RNAlater solution was stored at 4°C for 24-72 hours before long-term storage at -80°C. Following removal of tissue for gene expression analysis (Section 2.5), the left lung was slowly inflated with 4% PFA, by attaching a 2ml syringe to the inserted cannula, and then submerged in 4% PFA overnight. The next day, the PFA was removed, and the trachea and left lung reinflated and submerged in 30% sucrose diluted in dPBS overnight. On the final day, lungs were inflated with a 1:1 solution of 30% sucrose and optimal cutting temperature (OCT) embedding medium (Thermo Scientific), submerged in cryomolds (Sigma) pre-filled with OCT medium and stored at -80°C before cryosectioning (Section 2.6.1).

2.8 Precision cut lung slices

2.8.1 Generation of PCLS

Prior to harvest of fresh, unfixed lung tissue, 6% ultra-low melting point (LMP) agarose (Type IX-A; A2576; Sigma), supplemented with 100U/ml penicillin and 100mg/ml streptomycin, was prepared in Hank's Balanced Salt Solution (HBSS; Gibco) by heating and equilibrating to 42°C in a temperature-controlled bead bath. Mice were subsequently culled by cervical dislocation and the lungs inflated (Section 2.7.7) *in situ* with the 42°C LMP agarose, before the chest cavity was covered in ice until the agarose had set. Once set, agarose-filled lung lobes were mounted onto a Compresstome VF-300-OZ vibrating microtome (Precisionary, Ashland, Massachusetts, USA) for sectioning. Thick 500µm PCLS sections were cut using an oscillation and speed setting of five. PCLS were then transferred into Falcon tubes containing 7ml of PCLS culture medium (high glucose DMEM (D5796; Sigma) supplemented with 100U/ml penicillin, 100mg/ml streptomycin, 4mM L-glutamine, 20µg/ml gentamicin and 1.25µg/ml amphotericin B) for washing overnight at 37°C by

continuous inversion (18x/min) using a tube rotator (115 VAC; Cole-Parmer, St Neots, UK). The PCLS were subsequently transferred to static culture (Section 2.8.2).

2.8.2 PCLS culture

The PCLS were cultured in a static humidified incubator (37°C, 5% CO_2) in PCLS culture. One PCLS was cultured in 1ml medium per well of a 24-well plate (Corning). Medium was replaced three times/week.

2.8.3 Viability assay

Viability of PCLS was assessed 24-72 hours after generation (Section 2.8.1) using the Live/Dead Viability/Cytotoxicity kit for mammalian cells (L3224; Invitrogen). Untreated PCLS were selected at random and incubated for 45 minutes at 37°C in PCLS media containing 4µM calcein-AM and 4µM ethidium homodimer-1. PCLS were then imaged *in situ* using an EVOS FL Auto 2 Imaging System (Section 2.6.5.1). Calcein-AM hydrolysis, indicative of live cells, was detected using the GFP light cube. Ethidium homodimer-1 fluorescence, indicative of dead cells, was detected using the Texas Red light cube. A PCLS treated with 4% PFA for 20 minutes was stained and imaged in parallel as a control for what 'dead' tissue might look like using this kit.

2.9 Statistical analysis

All statistical analyses were performed using Prism 8 (GraphPad Software). Parametric tests were performed unless data were not normally distributed (Shapiro-Wilk test) or the sample size was less than 3. If the distribution significantly deviated from normal, a logarithmic transformation was first performed, and normality tests repeated, before parametric or equivalent non-parametric statistical analysis were completed. The statistical test used to analyse each data set is stated in the text. A calculated p value of <0.05 was deemed a significant difference and indicated on figures where appropriate (*=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001). Error bars in figures represent the standard error of the mean (SEM) unless stated otherwise.

Chapter 3 Generation of recombinant lentivirus expressing ABCA3

3.1 Introduction

3.1.1 Considerations for production of recombinant lentivirus

For production of third-generation lentiviral gene therapy vectors, via transient transfection, three key components are required - packaging plasmids, envelope plasmids and a genome plasmid (Figure 1.11; Section 1.5.3.2) The packaging plasmids encode for Gag, Pol and Rev which are necessary for the viral core structure and reverse transcription. Envelope plasmids express glycoproteins which are incorporated into the surface of lentiviral particles during assembly and dictate vector tropism. By pseudotyping rLV with envelope proteins from different viruses, transduction efficiency in specific target organs can be improved and off-target effects reduced. Recombinant SIV pseudotyped with F/HN proteins [195], known as rSIV.F/HN, has resulted in efficient lung tropism in mice [155, 196], including the efficient transduction of both airway and parenchymal lung cell types [139]. For ABCA3 gene addition therapy, ATII cells in the lung parenchyma are the target cell population (Section 1.2.3) making rSIV.F/HN an appropriate pseudotype.

The genome plasmid expresses the transgene of interest under the control of a mammalian promoter and is flanked by LTR sequences. During rLV production, the sequence between the LTRs is packaged into lentiviral particles. The efficiency of packaging is influenced by the cargo size (LTR-LTR) and for vectors based on HIV-1 the maximum packaging capacity is typically stated as 8-10kb. The human *ABCA3* gene is large (80kb) with a coding sequence of 6.6kb (ENST00000301732.10; Figure 3.1). Although this is within the estimated packaging capacity for rLV it is known that packaging efficiency decreases with increased cargo size, leading to potential problems with low vector yield during production [197, 198]. In the development of gene therapies for Duchenne muscular dystrophy, minigene versions of the very large dystrophin gene which lack non-essential

exons were successfully created to facilitate genome packaging [199]. Here (Section 3.2.1.1), truncation of *ABCA3* was explored for ABCA3 vector design to increase the likelihood of obtaining high yields of rLV. Achieving a high yield is important for both vector usage in preclinical *in vivo* models and for clinical translation where the dosing volume is often restricted.

3.1.2 Aims

Studies in this chapter describe the generation of rLV for expression of ABCA3. This includes the design and construction of rLV genome plasmids encoding the *ABCA3* cDNA, production of rSIV.F/HN expressing ABCA3 protein and evaluation of promoters with improved specificity for alveolar cell activity.

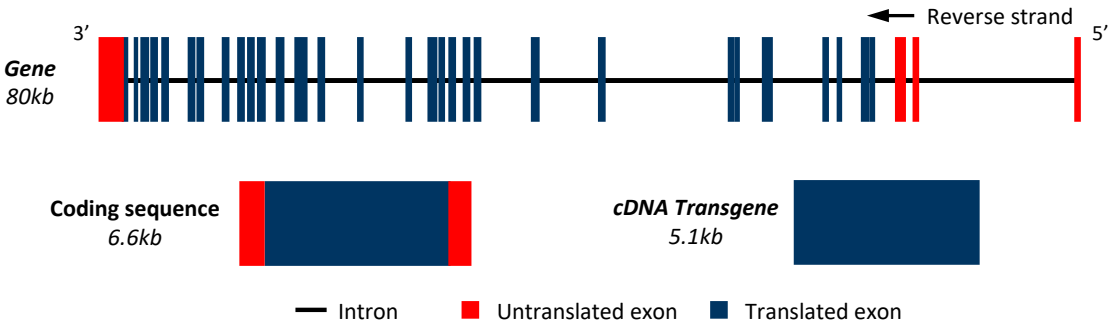


Figure 3.1 Diagram of *ABCA3* processing and transgene design

The 80kb *ABCA3* gene is located on the reverse strand of chromosome 16p13.3 in humans and contains 33 exons which are transcribed to form mRNA. Exons 1, 2 and 3 are fully non-coding and exon 33 is partially translated. Untranslated regions of *ABCA3* were excluded from the transgene design to form a complementary DNA (cDNA) sequence 5.1kb in length. Diagrams scaled to be representative of individual intron and exon sizes.

3.2 Results

3.2.1 Design and construction of *ABCA3* rLV genome plasmids

3.2.1.1 Transgene design

To generate rLV for expression of ABCA3, genome plasmids were designed and constructed using the WT, human *ABCA3* coding sequence as identified on the Ensembl (ENSG00000167972) and NCBI (NM_001089) databases. For more efficient packaging in rLV, the cDNA transgene was reduced in size from 6.6kb to 5.1kb by removing untranslated regions of the cDNA sequence (exons 1 – 3, exon 33). A Kozak consensus sequence (GCCACCATGG) was added 5' of the start codon to enhance translation.

The *ABCA3* cDNA was synthesised as two fragments (*ABCA3_1*, *ABCA3_2*), in plasmids pABCA3_1 (pGM1298) and pABCA3_2 (pGM1299; Table 2.2; Twist Bioscience, San Francisco, USA), and an endogenous AscI restriction site used to join the two fragments together within the genome plasmid in a stepwise process (Figure 3.2). For cloning compatibility with standard laboratory plasmids and protocols, a NheI restriction site was added to the 5' end of *ABCA3_1*, and PspOMI sites to the 3' ends of *ABCA3_1* and *ABCA3_2* (Figure 3.2). No endogenous NheI or PspOMI restriction sites were identified in the cDNA sequence.

3.2.1.2 Promoter selection for rSIV.ABCA3 genome plasmid

Persistent expression of ABCA3 is anticipated to be desirable in ATII cells of the lung for treating ABCA3 deficiency. Three promoters were selected for incorporation into rLV genome plasmids to drive expression of ABCA3 – CMV-, elongation factor-1 alpha (Ef1α) and CMV enhancer with EF1α promoter (hCEF; Table 3.1). The CMV- and Ef1α promoters can provide robust expression levels in multiple cell types and have been used in clinical trials and approved therapeutics [200, 201], but they contain unmethylated cytosine-guanine dinucleotide (CpG) motifs which are susceptible to CpG-methylation induced silencing [202, 203]. The hCEF promoter [202] incorporates

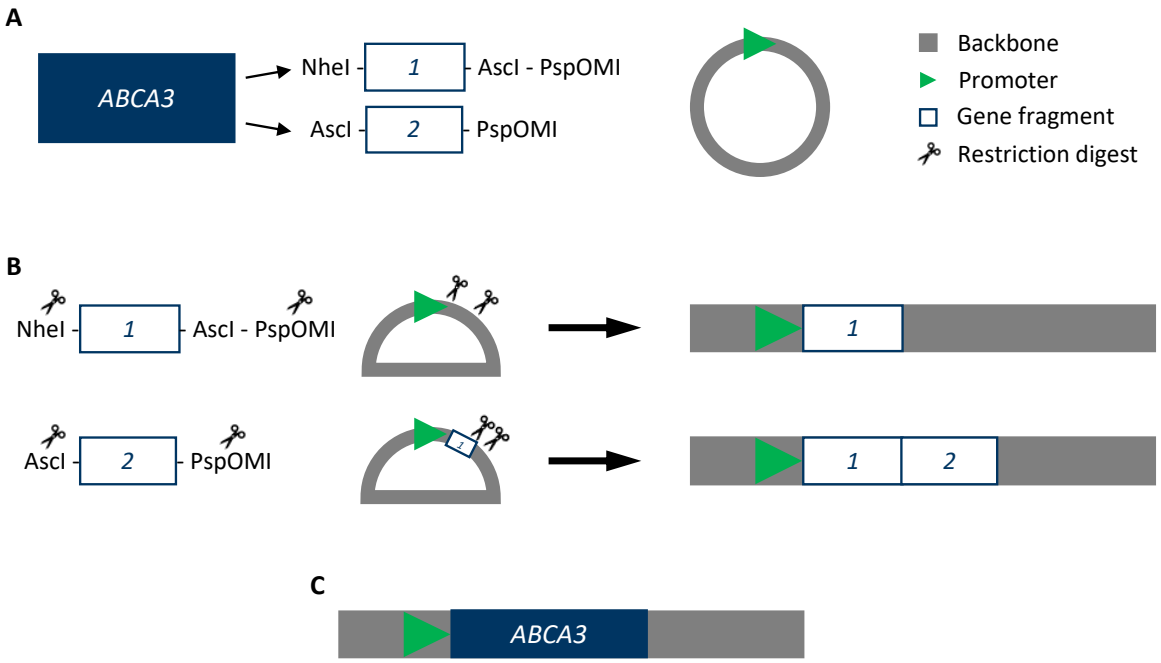


Figure 3.2 Generation of recombinant lentiviral vector genomes encoding ABCA3

(A) The *ABCA3* transgene was synthesised in two fragments (*ABCA3_1* (pGM1298), *ABCA3_2* (pGM1299)) with a 5' NheI and 3' PspOMI restriction site added to the end of *ABCA3_1* for compatibility with lentiviral backbone plasmids pGM749, pGM751 and pGM753 (Table 2.2). An AscI restriction site, present in the endogenous cDNA sequence, was included 3' of *ABCA3_1* and 5' of *ABCA3_2* to facilitate joining the two fragments. **(B)** First, pABCA3_1 (pGM1298) and the lentiviral backbone plasmids (pGM749, pGM751, pGM753) were ligated after double restriction digestion with NheI (5') and PspOMI (3'). The intermediate plasmids (pGM1323, pGM1324, pGM1325) and pABCA3_2 (pGM1299) were then digested with AscI (5') and PspOMI (3') before ligation to generate the **(C)** final recombinant lentiviral genome plasmids encoding ABCA3 (pGM955, pGM956, pGM957; Table 2.2).

CpG-depleted versions of the human CMV enhancer and Ef1α promoter, thus making it inherently resistant to silencing induced by CpG-methylation, and was selected after demonstration of high and sustained levels of expression in the murine lung [139]. After rLV delivery, hCEF-directed transgene expression is high in the airways, but also detectable in the parenchyma [139]. In the context of non-viral lung gene delivery, hCEF has also been successfully used in clinical trials in people with CF [137].

Table 3.1 Sequence of promoters used in rLV genome plasmids.

Enhancer sequences are written in italics.

Promoter	Sequence
CMV-	*tctcaatattggccattagccatattattcattggttatatagcataaatcaatattggctattggccattgcatacg ttgtatctatatcataatatgtacatttatattggctcatgtccaatatgaccgccatgttggcattgattattgacta gttattaatagtaatcaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggt aaatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaac gccaatagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagt gtatcatatgccaagtccgcccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatg accttacgggactttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggca gtacaccaatgggcgtggatagcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggag tttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaataaccccgccccgttgacgcaaatgggcgg taggcgtgtacggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcactagaagctttattgc ggtagtttatcacagttaaattgctaacgcagtcag*tgcttctgacacaacagtctcgaacttaagctgcagaagtt ggtcgtgaggcactgggcagg
Ef1α	tcccgtgaggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggt cggcaattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttt tcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttgccgccag aacacagg
hCEF	*gttacataacttatggtaaatggcctgcctggctgactgcccaatgacccctgcccaatgatgtcaataatgatgt atgttcccatgtaatgccaatagggactttccattgatgtcaatgggtggagtatttatggtaactgcccacttggc agtacatcaagtgtatcatatgccaagtatgccccctattgatgtcaatgatggtaaatggcctgcctggcattat gcccagtacatgaccttatgggactttcctacttggcagtacatctatgtattagtcattgctattaccatggg*gaatt cactagtggagaagagcatgcttgagggctgagtgcccctcagtgggcagagagcacatggcccacagtccctgag aagttgggggggaggggtgggcaattgaactggtgcctagagaaggtggggcttgggtaaactgggaaagtgatgtg gtgtactggctccacctttttccccagggtgggggagaaccatatataagtgcagtagtctctgtgaacattc

3.2.1.3 rSIV.ABCA3 genome plasmid construction

The *ABCA3* cDNA was cloned into SIV lentiviral backbone plasmids containing the CMV- (pCMV-.SFTPB; pGM751), hCEF (phCEF.SFTPB; pGM749) and Ef1α (pEf1α.SFTPB; pGM753) promoters (Figure 3.2). First, pABCA3_1 (pGM1298) and backbone plasmids (Table 2.2) were digested with NheI (5') and PspOMI (3') (Section 2.2.1) and, after agarose gel electrophoresis (Section 2.2.4), the desired fragments gel purified (Section 2.2.5) and ligated (Section 2.2.6) to form intermediate plasmids (pCMV-.ABCA3_1 (pGM1323), phCEF.ABCA3_1 (pGM1324) and pEf1α.ABCA3_1 (pGM1325; Table 2.2). Successful insertion was confirmed by Sanger sequencing (Section 2.2.15; GM16-1025, GM18-1162, GM16-1032; Table 2.1). Subsequently, the intermediate and pABCA3_2 (pGM1299) plasmids were digested with AscI (5') and PspOMI (3'), ligated and sequenced (GM16-1025, GM18-1162, GM16-1032; Table 2.1) to generate the final *ABCA3* rLV genome plasmids (Figure 3.3): pCMV-.ABCA3 (pGM955), phCEF.ABCA3 (pGM956), pEf1α.ABCA3 (pGM957; Table 2.2).

3.2.2 Transient transfection of human immortalised cell lines with *ABCA3* rLV genome plasmids

3.2.2.1 Development of a western blot protocol for ABCA3

To demonstrate expression of ABCA3 protein from the rLV genome plasmids, a western blot protocol was required. Adherent HEK293T cells were transiently transfected (Section 2.3.3) with a plasmid expressing ABCA3 (pCMV-.ABCA3; pGM955) or a plasmid expressing GFP (pCMV-.GFP; pGM311) (Table 2.2) as an indicator of transfection efficiency. Lysates of RNA (Section 2.5.2) and protein (Section 2.5.4) were extracted from the transfected cells for qRT-PCR (Section 2.5.4) and western blot analysis of ABCA3, respectively. Analysis by qRT-PCR (ABCA3_qRT_F_Ex4, ABCA3_qRT_R_Ex5, Bact_qRT_F, Bact_qRT_R; Table 2.1) showed significantly increased expression of *ABCA3* in cells transfected with pCMV-.ABCA3 (pGM955) compared with GFP-transfected

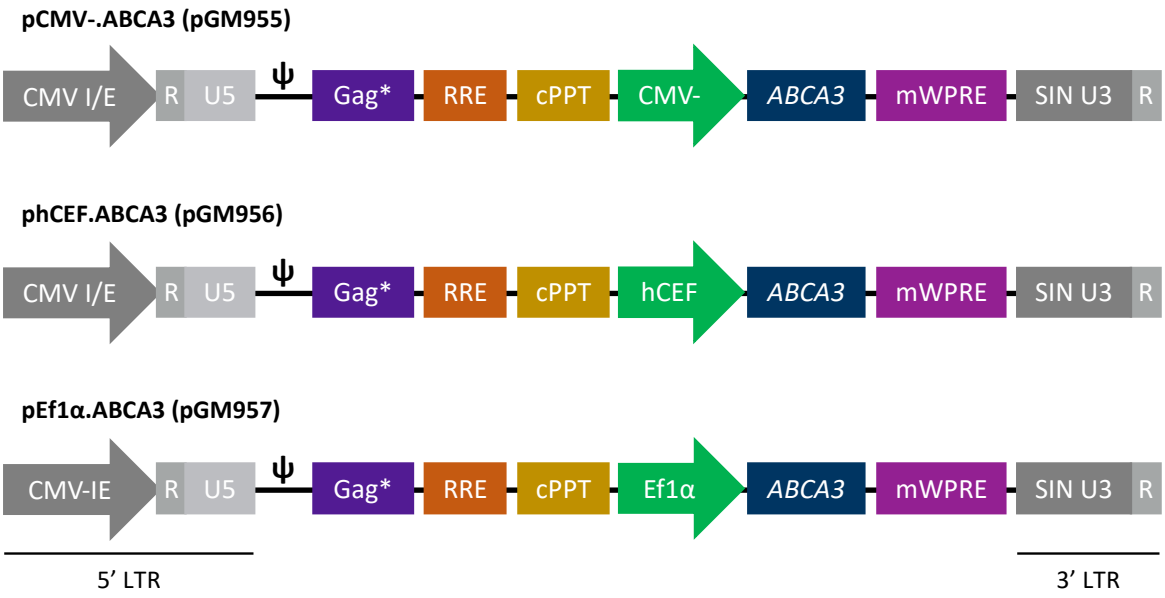


Figure 3.3 Structure of recombinant lentiviral vector genome plasmids encoding ABCA3

Lentiviral vector genome plasmids encoding ABCA3 (pCMV-.ABCA3, pGM955; phCEF.ABCA3, pGM956; pEf1α.ABCA3, pGM957) contained the following components: 5' long terminal repeat (LTR) comprising a heterologous CMV immediate-early (CMV-IE) promoter, R and U5; Psi (ψ); partial Gag (Gag*); *rev* response element (RRE); central polypurine tract (cPPT); transgene promoter (CMV-, hCEF or Ef1α); transgene (*ABCA3*); modified woodchuck hepatitis virus posttranscriptional regulatory element (mWPRE); 3' LTR comprised of a self-inactivating (SIN) U3 and R.

HEK293T cells using a Welch's t-test ([t (5)=7.40, p=0.0007]; Figure 3.4) indicating successful *ABCA3* transcription from the rLV genome plasmid.

The detection of transmembrane proteins by western blot can be affected by the formation of high molecular mass aggregates and poor conformational stability [204, 205]. To establish robust conditions for ABCA3 detection, different sample preparation conditions and primary anti-ABCA3 antibodies were evaluated. Before immunoblotting, proteins are typically denatured at 95°C [205], but detection of another ABC protein (ABCC7 also known as CFTR) by western blotting has been improved by lowering the denaturation temperature from 95°C [205] to below 60°C [206]. To evaluate the impact of denaturation temperature on ABCA3 detection, immunoblots were performed after sample preparation at 95°C (5 minutes), 70°C (15 minutes) and 37°C (15 minutes; Figure 3.4). A correlation was observed between decreasing incubation temperature and increasing detection of ABCA3 at the anticipated molecular weights (190kDa, 170kDa; Figure 3.4; Section 1.2.3). ABCA3 was most readily detected when protein lysates were denatured at 37°C and this temperature was used for all future experiments.

Two polyclonal antibodies against ABCA3, tested in parallel, detected both the 190kDa and 170kDa isoforms of ABCA3 (Figure 3.4). Blotting with antibody ab99856 appeared to result in fewer non-specific bands compared with antibody PA5-52478 suggesting ab99856 was more specific for ABCA3. However, at equivalent exposure times, the protein bands were also less intense for ab99856-stained samples indicating antibody ab99856 was less sensitive at detecting ABCA3. Due to its increased sensitivity at detecting ABCA3, PA5-52478 was used for future western blot experiments unless stated otherwise.

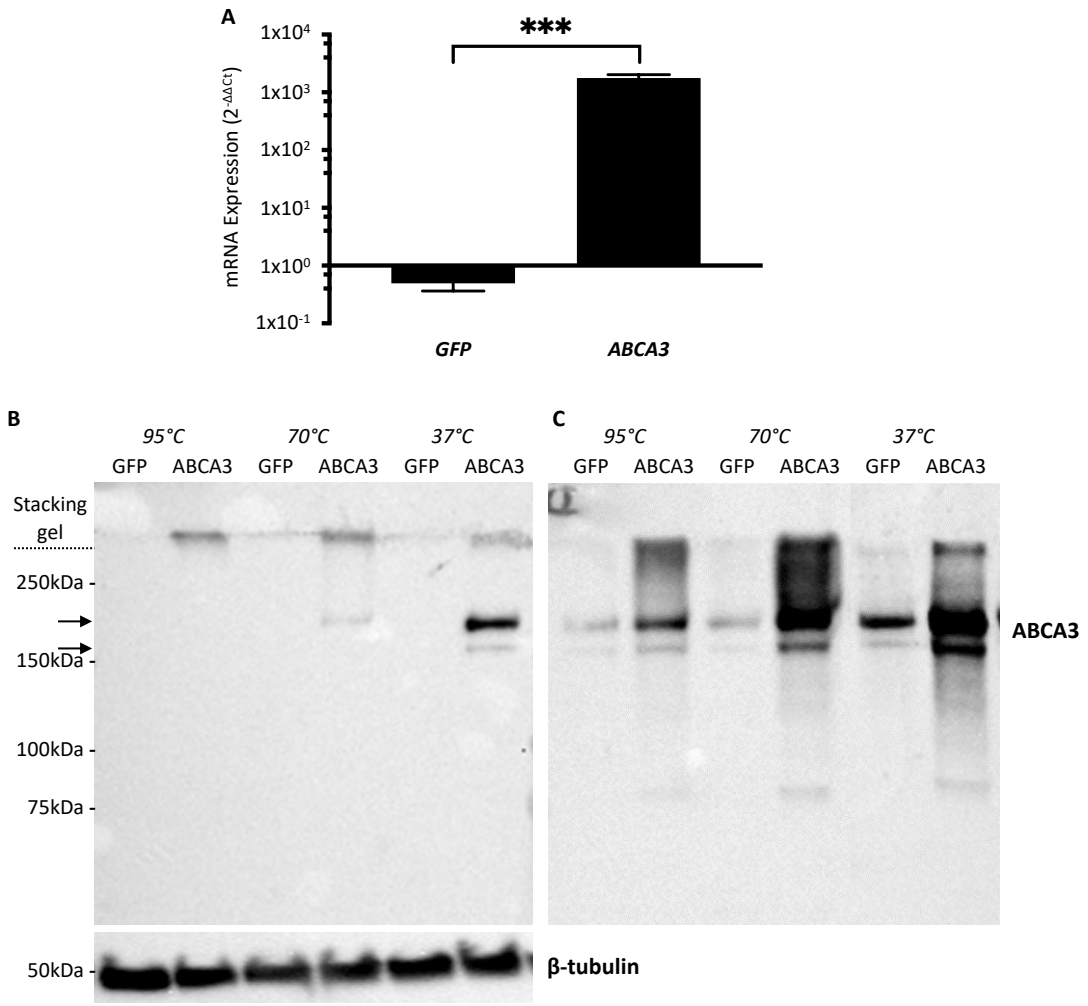


Figure 3.4 Development of assays to detect ABCA3 expression

Figure 3.4 Development of assays to detect ABCA3 expression

Adherent HEK293T cells were transfected (Section 2.3.3) with vector genome plasmid pCMV-.ABCA3 (pGM955; ABCA3) or a control plasmid pCMV-.GFP (pGM311; GFP; Table 2.2). 48-72 hours after transfection **(A)** RNA and **(B, C)** protein lysates were extracted. **(A)** *ABCA3* expression was determined via qRT-PCR and normalised to that of naïve HEK293T cells using $2^{-\Delta\Delta Ct}$ analysis (Section 2.5.4). Quantitative RT-PCR was performed with three technical replicates and data shown as mean ± SEM of n=2 independent experiments. A significant difference in *ABCA3* expression was determined by Welch's t-test [t (5)=7.40, p=0.0007, ***]. **(B, C)** Total protein lysates (40µg) were denatured under reducing conditions for western blot analysis using ABCA3 primary antibodies **(B)** ab99856 or **(C)** PA5-52478 (Table 2.5). Denaturation conditions of 95°C (5 mins), 70°C (15 mins) and 37°C (15 mins) were compared. Blotting for β-tubulin using primary antibody ab6046 was used as a protein loading control. The stacking gel and size of protein markers used are labelled with arrows at approximately 170kDa and 190kDa. Data shown are representative of two independent repeats.

3.2.2.2 Expression of ABCA3 from rLV genome plasmids

To confirm expression of ABCA3 from genome plasmids generated in Section 3.2.1 (pGM955, pGM956, pGM957), HEK293T cells were transiently transfected with each plasmid using Lipofectamine 3000 (Section 2.3.3) before protein lysates were extracted for western blot analysis (Section 2.5.3 and 2.5.6). Protein samples were immunoblotted for ABCA3 using the conditions optimised in Section 3.2.2.1 and anti-ABCA3 antibody ab99856 (Table 2.5). Antibody ab99856 was selected as ABCA3 expression levels induced by transfection were anticipated to be high. Figure 3.5 shows the expected 170kDa and 190kDa products (Section 1.2.3) in all HEK293T cells transfected with rLV *ABCA3* genome plasmids indicating successful transcription, translation, and processing of the *ABCA3* transgene. These results were further supported by ICC (Section 2.6.4.1) of A549 cells transiently transfected with pGM955 (pCMV-.ABCA3) for ABCA3 (Figure 3.6), where abundant ABCA3 staining was observed in transfected cells compared with naïve control cells.

3.2.3 Production and characterisation of rSIV.F/HN.hCEF.ABCA3

The ability to manufacture high yields of rLV at scale is crucial for successful *in vivo* lentiviral gene therapy applications. This was important to evaluate for vectors expressing ABCA3 since experience with manufacturing in-house, research grade rLV expressing the similar ABCC7 transmembrane protein (CFTR) indicates that a large, complex transgene cargo can result in reduced viral yields compared with equivalent production of vectors with smaller, simple transgenes, such as GFP (Steve Hyde, Personal communication).

3.2.3.1 Comparison of rLV production platforms using rSIV.F/HN.hCEF.GFP

To inform *ABCA3* vector manufacturing, two different production platforms were initially compared at the 25ml scale (SS) using a simple GFP transgene. This included a commercially available rLV production process (LV-MAX) which uses a derivative of the HEK293F cell line (herein referred to as LV-MAX cells; Section 2.4.1) as well as our routine, in-house process using

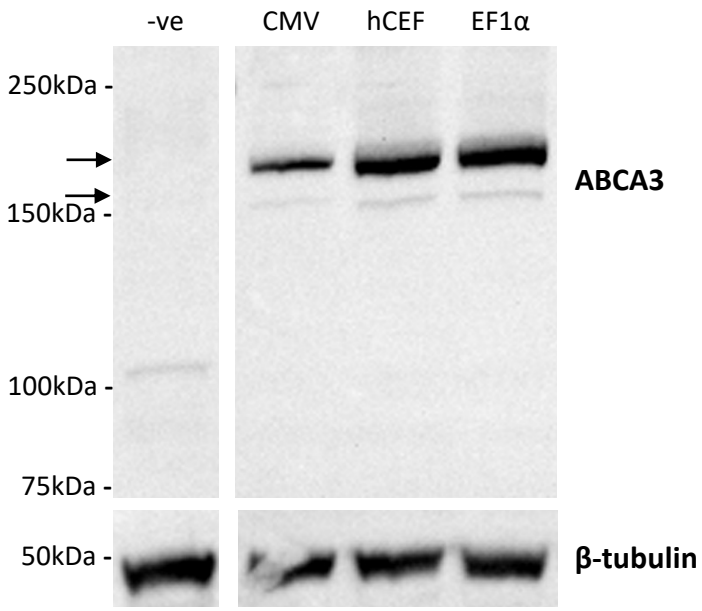


Figure 3.5 Western blot analysis to confirm expression of ABCA3 from lentiviral vector genomes

Adherent HEK293T cells were transfected (Section 2.3.3) with vector genome plasmids expressing ABCA3 under the control of CMV- (pGM955), hCEF (pGM956) or Ef1α (pGM957) promoters (Table 2.2). Total protein lysates were extracted 72 hours after transfection and 30µg denatured under reducing conditions (Section 2.5.6) for western blot analysis using ABCA3 primary antibody ab99856. Naïve HEK293T cells were included as a background control (-ve). Blotting for β-tubulin (ab6046, Table 2.5) was used as a protein loading control. The stacking gel and size of protein markers used are labelled with arrows at approximately 170kDa and 190kDa.

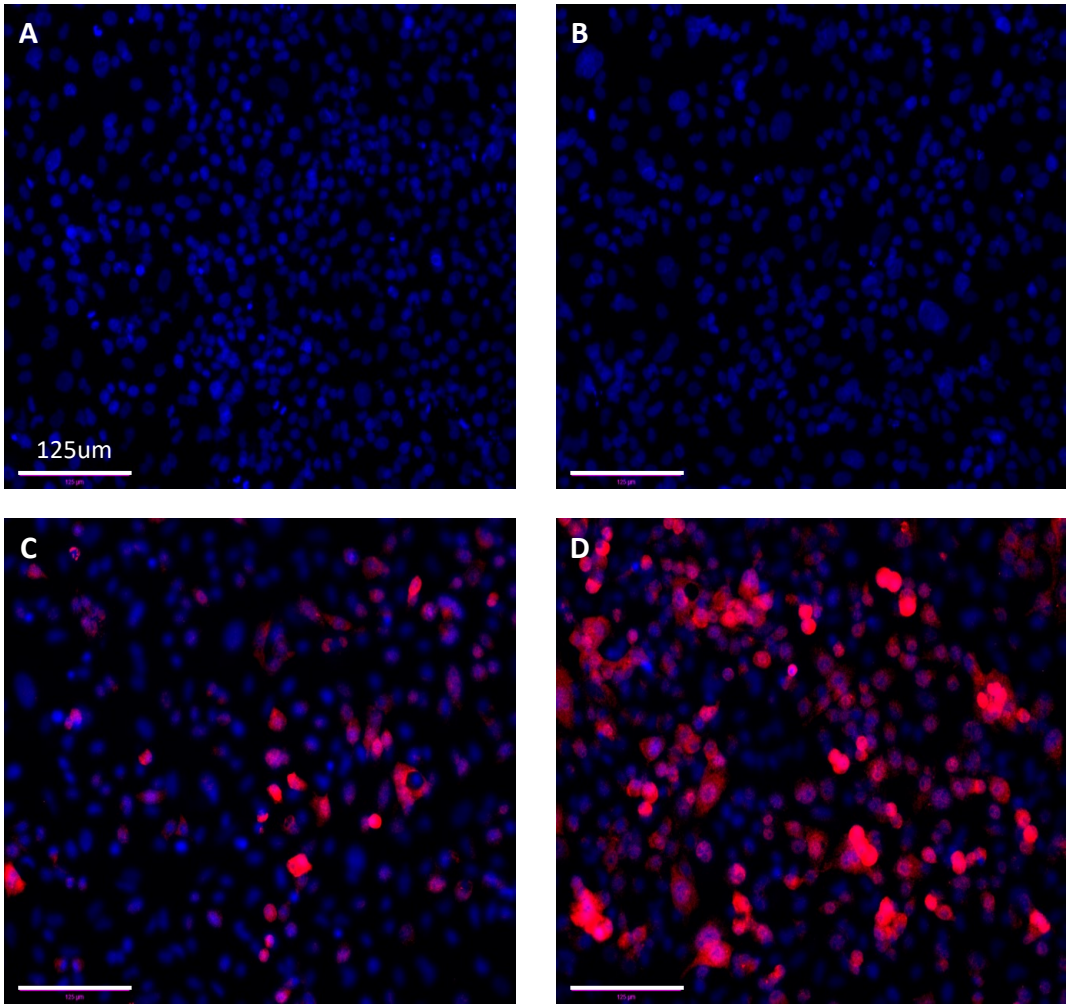


Figure 3.6 Immunocytochemistry to confirm expression of ABCA3 from pCMV-.ABCA3 lentiviral genome plasmid

Immunocytochemical analysis (Section 2.6.4.1) was completed in triplicate on **(A, C, D)** A549 cells transfected with a vector genome plasmid expressing ABCA3 (pCMV-.ABCA3, pGM955; Table 2.2) or **(B)** naïve A549 cells. 72 hours after transfection cells were fixed, permeabilised and immunostained for **(B-D)** ABCA3 expression (red) using the primary antibody PA5-52478 and AlexaFluor555-conjugated secondary antibody (A21428). Nuclei were counterstained with Hoechst (blue). **(A)** Secondary antibody only stained transfected cells. Representative images from three replicates (three images/replicate) are shown. All scale bars indicate 125µm.

suspension-adapted HEK293T-17 cells (Section 2.4.2). Briefly, genome plasmid pGM357 (phCEF.GFP), along with four other plasmids required for rLV production (pGM691, pGM299, pGM301, pGM303; Table 2.2), were transiently transfected into HEK293T-17 (Section 2.4.2.1) or LV-MAX cells (Section 2.4.1) at the recommended ratio (Genome:Gag-Pol:Rev:F:HN 20:9:6:6:6) before subsequent harvest (Section 2.4). Following production of vGM196 (rSIV.F/HN.hCEF.GFP; Table 2.3), viral titre was determined using a transducing unit assay where integrating DNA copy number was measured by qRT-PCR (Section 2.4.4). When the relative yields (TTU/manufacturing l) were calculated, no significant difference between production processes was identified (LV-MAX=$1.91x10^9$ TTU/l, HEK293T-17=$7.97x10^9$ TTU/l; [t (4)=2.04, p=0.1106]; Figure 3.7). This suggested that the commercially available process was unlikely to provide any improvements in relative yield for rLV expressing complex transgenes such as ABCA3. Thus, our in-house HEK293T-17 process was used to confirm that the rLV genome plasmids constructed in Section 3.2.1 were suitable for manufacturing of rLV expressing ABCA3 at high yield. In addition, this in-house production process also has validated LS capability (5l production) that would be useful for *in vivo* studies described in later Chapters.

3.2.3.2 Manufacturing yield of rSIV.F/HN with different transgenes

Using the production processes described above (Section 2.4), differences in yield were measured from rLV expressing ABCA3 (vGM336, rSIV.F/HN.hCEF.ABCA3) alongside similar rVL vector configurations expressing GFP (vGM196, rSIV.F/HN.hCEF.GFP) and CFTR (vGM195, rSIV.F/HN.hCEF.CFTR; Table 2.3). The yield from rSIV.F/HN.hCEF.CFTR (vGM195), expressing CFTR with similar properties to ABCA3 and which has already progressed to 200l scale GMP manufacture, was considered a good benchmark for deciding whether rSIV.F/HN.ABCA3 (vGM336) was a viable gene therapy product from a manufacturing perspective.

As seen in Figure 3.8, all three rLV configurations were produced at both SS and LS. As expected, yields increased when vectors were produced at LS (Steve Hyde, Personal communication).

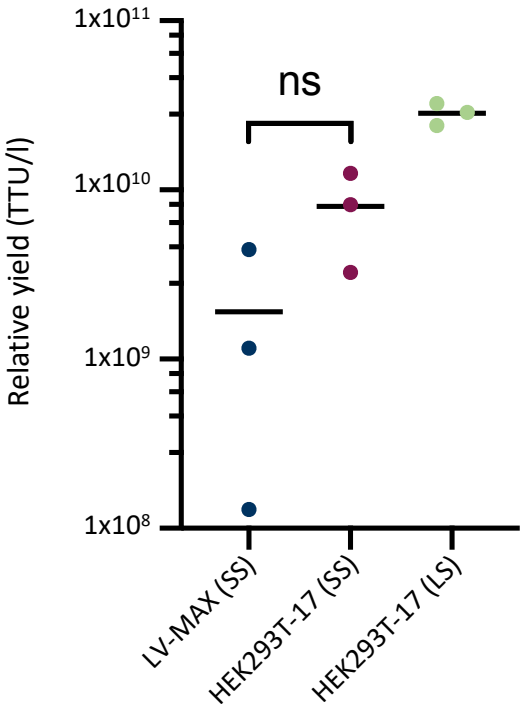


Figure 3.7 Comparison of production methods for producing recombinant lentivirus expressing GFP

LV-MAX cells were transduced with dilutions of purified rSIV.F/HN.hCEF.GFP (vGM196) in the presence of 8µg/ml of polybrene. After 72 hours genomic DNA was extracted and the integration of WPRE DNA quantified by qRT-PCR (Section 2.4.4). Relative yield (TTU/l) was calculated using linear regression and volume of purified lentivirus obtained during production. Data are shown as individual values and mean for three independent productions per method, with four technical replicates for qRT-PCR assessment. Significant differences between the values for 25ml SS productions were determined by Independent t-test [t (4)=2.04, p=0.1106]. A calculated p value of >0.05 was deemed non-significant (ns). SS, small scale (25ml production); LS, large scale (5l production); ns, non-significant; TTU, TaqMan transducing units.

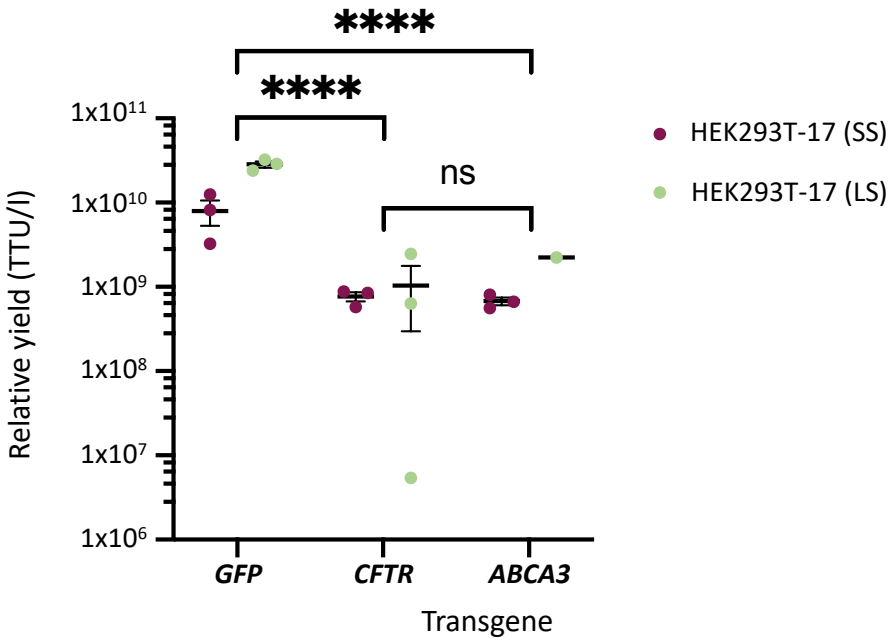


Figure 3.8 Production of high yield lentivirus expressing ABCA3

Suspension HEK293T-17 cells were transduced with dilutions of purified rSIV.F/HN.hCEF expressing GFP (vGM196), CFTR (vGM195) and ABCA3 (vGM336) transgenes in the presence of 8µg/ml of polybrene. After 72 hours genomic DNA was extracted and the integration of WPRE DNA quantified by qRT-PCR (Section 2.4.4). Relative yields (TTU/l) were calculated using linear regression and the volume of purified rLV obtained during production. Data are shown as individual values and mean ± SEM for three independent productions per method, with four technical replicates for qRT-PCR assessment. Significant differences between the values were determined by two-way ANOVA with Tukey's post hoc analysis for transgenes. A statistically significant interaction between the effects of transgene and production method on relative yield was identified, [F (2, 10)=21.48, p=0.0002]. A calculated p value of >0.05 was deemed non-significant (ns). Calculated p values of <0.0001 (****) were deemed significant. SS, small scale (25ml production); LS, large scale (5l production); TTU, TaqMan transducing units.

Furthermore, the transgene had an impact on the vector yield obtained – for SS HEK293T-17 productions, compared with production of rSIV.F/HN.hCEF.GFP, the yield was 10.4-fold lower when the transgene was CFTR and 11.8-fold lower with an ABCA3 transgene. A two-way ANOVA test confirmed a statistically significant interaction between the effects of transgene and production scale on relative yield [F (2, 10)=21.48, p=0.0002], suggesting that changing the transgene has a differing impact on overall yield depending on the virus production scale used. Tukey's post hoc analysis identified the decrease in yield between GFP and CFTR or ABCA3 transgenes to be significant (p<0.0001). Importantly, however, there was no statistically significant difference between CFTR (vGM195) and ABCA3 (vGM336) yields (p=0.9599; Figure 3.8). In conclusion, HEK293T-17 cells can be used to produce high yields of rLV expressing ABCA3.

3.2.3.3 Expression of ABCA3 from rSIV.F/HN.hCEF.ABCA3

To confirm ABCA3 expression from rSIV.F/HN.hCEF.ABCA3, vector produced at LS and purified by AEX (Section 2.4.3.2) and TFF (Section 2.4.3.3; vGM336, W321; Table 2.3) was serially diluted from multiplicity of infection (MOI) 40 to MOI 1 (based on TTU levels determined in LV-MAX cells; Section 2.4.4) and used to transduce H441 human lung cells. The H441 cell line was used as it is receptive to the rSIV.F/HN pseudotype and has been used as an ATII cell model (Section 1.4.1.2) [103]. The H441 cells transduced with rSIV.F/HN.hCEF.GFP (vGM195, W304) at MOI 40 were used as an indicator of transduction efficiency. Protein lysates were immunoblotted using PA5-52478 anti-ABCA3 primary antibody (Table 2.5) using the western blot protocol (Section 2.5.6) optimised in Section 3.2.2.1. As shown in Figure 3.9, rSIV.F/HN.hCEF.ABCA3 directed dose-dependent expression of both 170kDa and 190kDa ABCA3 isoforms compared with the endogenous control (β-tubulin) samples. This result supports the progression of rSIV.F/HN.hCEF.ABCA3 (vGM336) into disease models and animal studies.

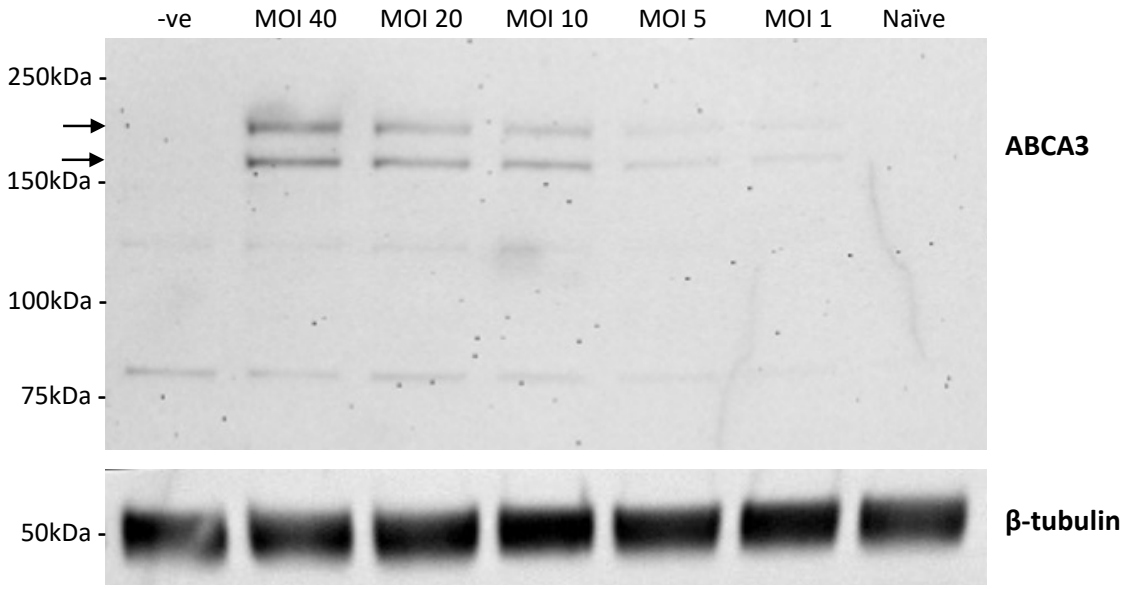


Figure 3.9 ABCA3 expression following transduction of H441 cells with rSIV.F/HN

H441 cells were transduced with serially diluted rSIV.F/HN.hCEF.ABCA3 (vGM336) up to MOI 40, or an equivalent negative control GFP vector (vGM196; Table 2.3) at MOI 40. After 72 hours, total protein lysates were extracted from transduced and naïve H441 cells and 30µg of protein blotted for ABCA3 under reducing conditions using primary antibody PA5-52478. Blotting for β-tubulin (ab6046, Table 2.5) was used as a protein loading control. The stacking gel and size of protein markers used are labelled with arrows at approximately 170kDa and 190kDa.

3.2.4 Assessment of alveolar specific promoters

As described earlier (Section 1.2.3), ABCA3 expression is highly enriched in ATII cells. Using a strong ubiquitous promoter (CMV-, Ef1α, hCEF; Section 3.2.1) to drive ABCA3 expression in the lung could result in overexpression of ABCA3 in non-target cell populations, leading to the activation of cellular stress and aggregation pathways [207]. A potential solution might be to utilise a promoter more specific to ATII cells; for example, the endogenous promoter for the human *SFTPB* gene restricts SP-B expression to club and ATII cells in the lung [208]. Therefore, a minimal *SFTPB* promoter (mSPB; bases 4925-5554 of GenBank NCBI *SFTPB* reference sequence NG_016967.1) together with three new promoters based on mSPB but with additional enhancer sequences (Alv1, Alv2, Alv3; Figure 3.10B) developed by Dr Altar Munis (unpublished) were evaluated for ATII cell specificity.

The mSPB promoter was found to drive GFP expression at a level comparable to hCEF in a SALI culture following transduction of titre-matched rLV (Figure 3.10A) (Altar Munis, Personal communication). To test the expression profile *in vivo*, lentiviral genome plasmids were generated expressing fLux under the control of the mSPB promoter to generate pmSPB.fLux (pGM1131) as well as variants incorporating the Alv1 (pGM1132), Alv2 (pGM1133) and Alv3 (pGM1134) promoters (Table 2.2). The rLV genome plasmids pmSPB.fLux, pAlv1.fLux and pAlv2.fLux (pGM1131, pGM1132, pGM1133 respectively) were made by digesting pCMV-.fLux (pGM350), pmSPB.GFP (pGM769), pAlv1.GFP (pGM119) and pAlv2.GFP (pGM1120) with NheI (5') and PspOMI (3') before ligation and sequencing (GM16-1075; Table 2.1) to confirm successful insertion of the fLux transgene into backbone plasmids. Plasmid pAlv3.fLux (pGM1134) was cloned using NheI (5') and HindIII (3') for digestion of pCMV-.fLux (pGM350) and pAlv3.GFP (pGM1121) due to the presence of a PspOMI restriction site within the Alv3 promoter sequence. The rSIV.F/HN vectors (vGM396 – vGM399; Table 2.3) were subsequently produced and titrated by the Virus Production team at LS (Section 2.4) for testing *in vivo*.

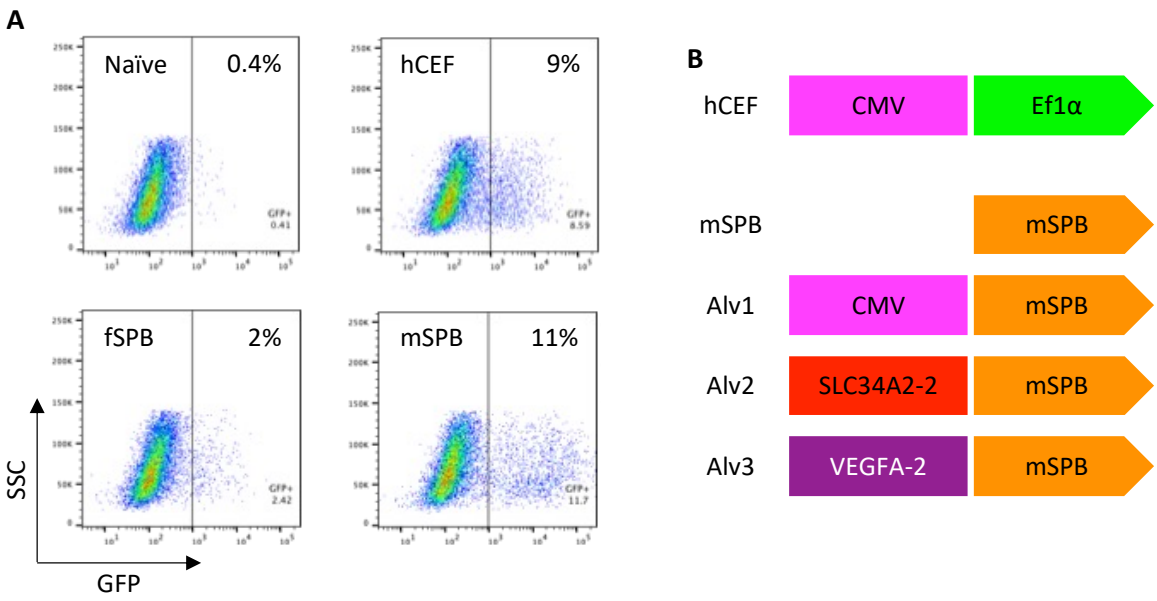


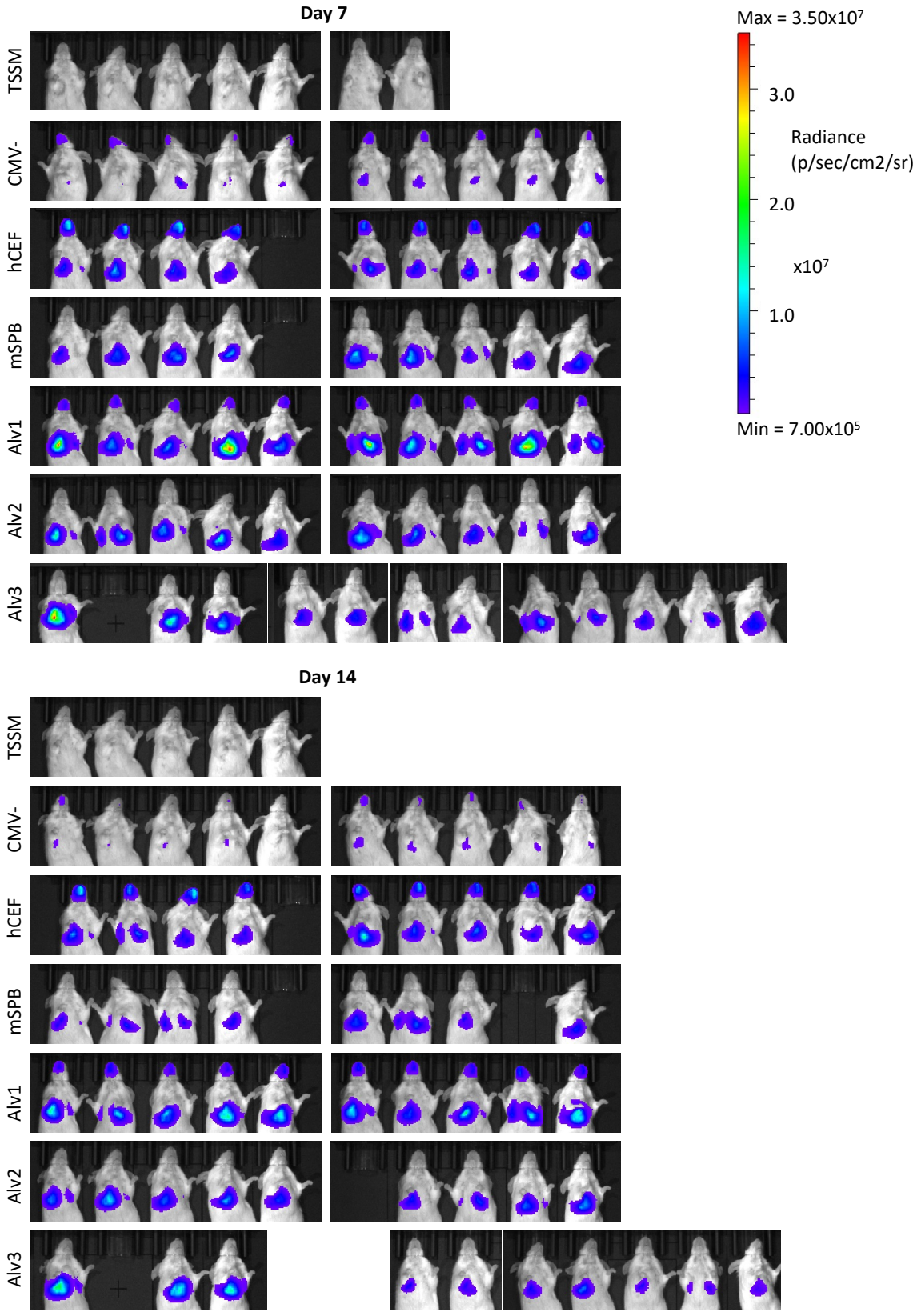
Figure 3.10 Design of novel promoters based on endogenous human *SFTPB* promoter (Altar Munis)

Three promoters were designed, constructed and evaluated by Dr Altar Munis (unpublished). **(A)** H441 cells cultured in an air-liquid interface model (SALI; Section 2.3.2.1) were transduced with $1x10^6$ TU rSIV.F/HN vectors encoding GFP (n=3). Activity of full length (f) and minimal (m) versions of the endogenous alveolar specific promoter surfactant protein B (SPB) were compared with the hCEF promoter by flow cytometry. After gating for single cells, the percentage of GFP positive cells was calculated. Naïve H441 cells were used as a negative control for setting the GFP gate. **(B)** Novel promoters were generated using mSPB coupled with different enhancer sequences (Alv1 = CMV; Alv2 = SLC34A2-2; Alv3 = VEGFA-2). SSC, side scatter.

3.2.4.1 Assessment of alveolar specificity by in vivo BLI

BALB/c mice were dosed via the intranasal route (Section 2.7.5) with $2.5x10^8$ TTU of rLV encoding fLux (Table 2.3) and luciferase expression assessed by *in vivo* BLI (Section 2.7.6) for up to 3 months to compare expression from the various promoters with the ubiquitous hCEF promoter (Figure 3.11). As an indicator of alveolar specificity, the bioluminescence produced from fLux expression was captured in both the lung (containing both airway and alveolar epithelial cells) and nasal areas (containing only airway epithelial cells) (Figure 3.11). In the lung, bioluminescence was detected for all promoters, albeit to various degrees, however bioluminescence in the nose was only apparent from CMV-, hCEF and Alv1 promoter-driven fLux expression. Of note, the bioluminescence produced from CMV-driven fLux in the lung appeared much weaker than from all other promoter groups. As anticipated, no bioluminescence was apparent in images of control mice dosed with formulation buffer (TSSM) only (Figure 3.11).

When bioluminescence in the nose and lung was quantified at day 28, a high lung/nose ratio was confirmed for fLux expression driven by mSPB, Alv2 and Alv3 promoters (Figure 3.12). A high lung/nose ratio indicates alveolar specificity since alveolar epithelial cells are localised to the lung parenchyma whereas airway epithelial cells are distributed throughout the nose and lungs of mice. Statistical analysis confirmed a significant different in promoter specificity between groups ([H (6)=55.53, $p<0.0001$]; Figure 3.12). Dunn's post hoc test confirmed a significant difference between animals dosed with rSIV.F/HN.hCEF.fLux (0.4892 ±0.0716) and those receiving rSIV.F/HN.mSPB.fLux (11.58±1.077; $p<0.0001$), rSIV.F/HN.Alv2.fLux (15.90±1.640, $p<0.0001$) or rSIV.F/HN.Alv3.fLux (13.63±1.350, $p<0.0001$). Interestingly, there was no difference between rSIV.F/HN.hCEF.fLux and rSIV.F/HN.Alv1.fLux treated animals ($p=0.1658$), suggesting the addition of a CMV enhancer to the mSPB promoter (Alv1; Figure 3.10B) causes the loss of parenchymal specificity.


Day 7
TSSM
CMV-
hCEF
mSPB
Alv1
Alv2
Alv3
Max = 3.50x10^7
3.0
Radiance
(p/sec/cm2/sr)
2.0
x10^7
1.0
Min = 7.00x10^5
Day 14
TSSM
CMV-
hCEF
mSPB
Alv1
Alv2
Alv3

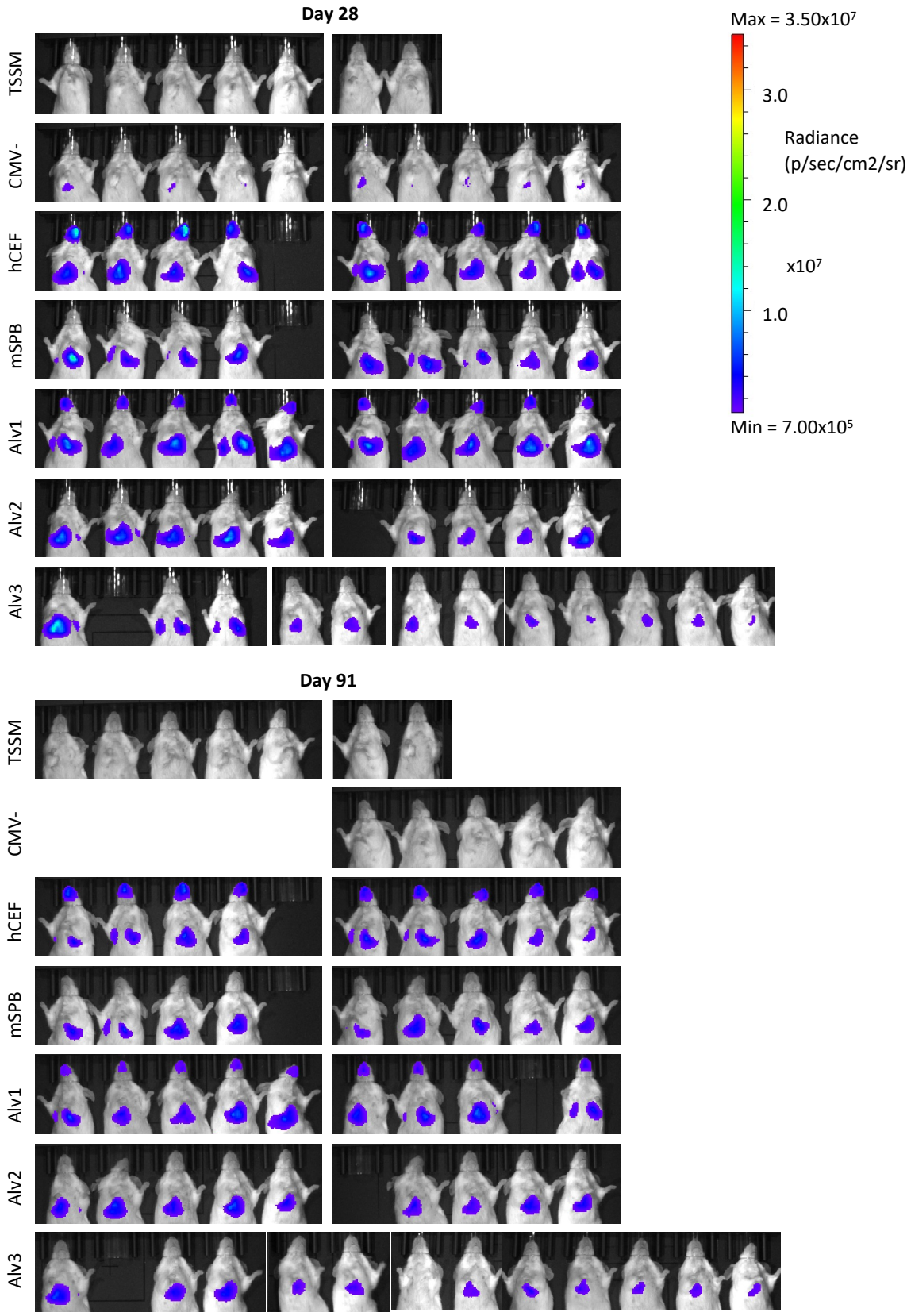


Figure 3.11 *In vivo* bioluminescence imaging of BALB/c mice dosed with recombinant lentivirus expressing *firefly* luciferase

Figure 3.11 *In vivo* bioluminescence imaging of BALB/c mice dosed with recombinant lentivirus expressing *firefly* luciferase

Female BALB/c mice (n=7-12) were administered $2.5x10^8$ TU of rSIV.F/HN lentivirus encoding *firefly* luciferase (fLux) under the control of CMV-, hCEF, mSPB, Alv1, Alv2 or Alv3 promoters, or formulation buffer (TSSM) as a control, via intranasal administration (Section 2.7.5). At timepoints 7, 14, 28 days and 3 months after transduction, mice were administered D-luciferin and imaged using an IVIS Lumina LT Imaging System (Section 2.7.6) to visualise the distribution of fLux activity in the lung and nasal cavity. Activity of fLux is measured in photons/second/cm^2/sr and all images are taken on the same scale ($7.0x10^5$ - $3.5x10^7$).

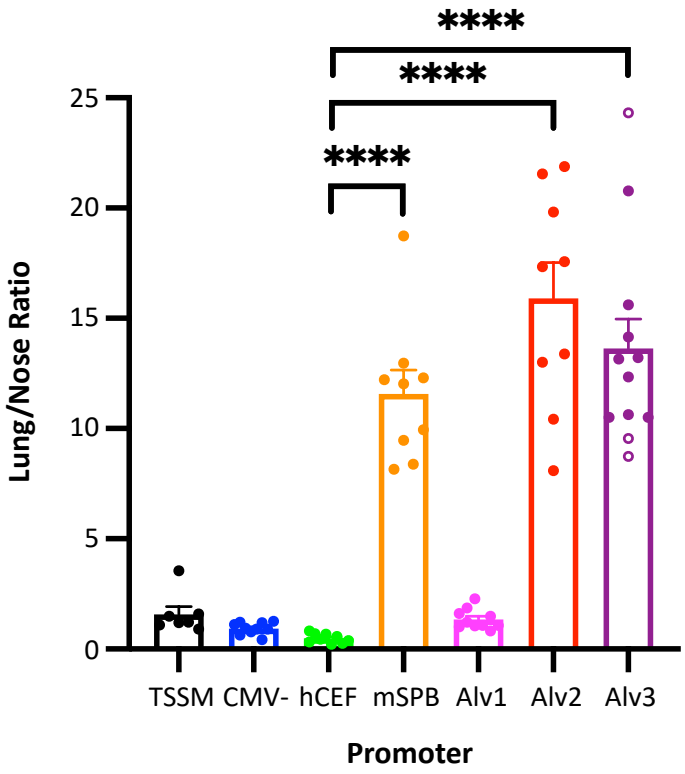


Figure 3.12 Alveolar specificity of novel promoters

Female BALB/c mice (n=7-12) were administered 2.5x10^8 TU of rSIV.F/HN lentivirus encoding *firefly* luciferase (fLux) under the control of CMV-, hCEF, mSPB, Alv1, Alv2 or Alv3 promoters, or formulation buffer (TSSM), via intranasal administration (Section 2.7.5). At 28 days after transduction, mice were administered D-luciferin and imaged using an IVIS Lumina LT Imaging System (Section 2.7.6) to visualise the distribution of fLux activity in the lung and nasal cavity. Regions of interest (ROIs) were drawn over the chest and nasal cavities and the signal within these regions measured in photons/second/cm^2/sr. Alveolar specificity was determined by calculating the ratio of signal in the lung to signal in the nasal cavity (Section 2.7.6). Data are shown as individual values and mean ± SEM. Significant differences between the values were determined by Kruskall-Wallis [H (6)=55.53, p<0.0001] with Dunn's post hoc multiple comparison test comparing vector groups with hCEF. A calculated p value of <0.05 was deemed significant (p<0.0001 ,****). Alv3: open circles, W345; closed circles W362 (Table 2.3).

3.2.4.2 Long-term expression profile of newly developed promoters

The ability of new promoters to drive sustained transgene expression in the lung over three months was assessed by calculating the area under the curve (AUC) for fLux expression for each animal (Figure 3.13). Expression of fLux in the lung was consistently sustained across the study period for all promoter evaluated (Figure 3.13A), although expression driven by the CMV- promoter was at noticeably lower levels and not apparent at day 91 (three months) of the experiment (Figure 3.11). Furthermore, although a Kruskal-Wallis test comparing animals dosed with rSIV.F/HN.mSPB.fLux with other rLV treatment groups identified a significant difference between the groups [$H(5)=39.23$, $p<0.0001$], Dunn's post hoc analysis only identified a statistically significant difference between mSPB-driven ($4.25x10^7 \pm 3.04x10^6$) and CMV-driven ($5.85x10^6 \pm 3.58x10^5$) fLux expression ($p=0.0053$). The level of bioluminescence (AUC) produced from Alv1, Alv2 and Alv3 promoters were all within 1.6-fold of that produced by mSPB, of rSIV.F/HN.mSPB.fLux, indicating that the addition of enhancer sequences to the mSPB promoter did not provide any significant improvement to transgene expression levels in the lung parenchyma. Perhaps, most importantly, there was no difference in fLux expression levels driven by mSPB and hCEF promoters (Figure 3.13B).

In conclusion, the use of luciferase bioluminescence to compare promoters showed that mSPB, Alv2 and Alv3 promoters can drive expression of fLux in the murine lung equivalent to hCEF but have expression profiles that are more restricted to the alveolar region.

3.3 Discussion

The aim of the studies described in this chapter was to develop an rLV expressing ABCA3 which could be used for proof-of-principle of ABCA3 gene addition therapy.

3.3.1 Improving lentiviral vector yields

The yield and titre to which an rLV can be produced is a crucial factor in the successful translation to *in vivo* clinical applications, as it fundamentally impacts manufacturing capabilities and restricts

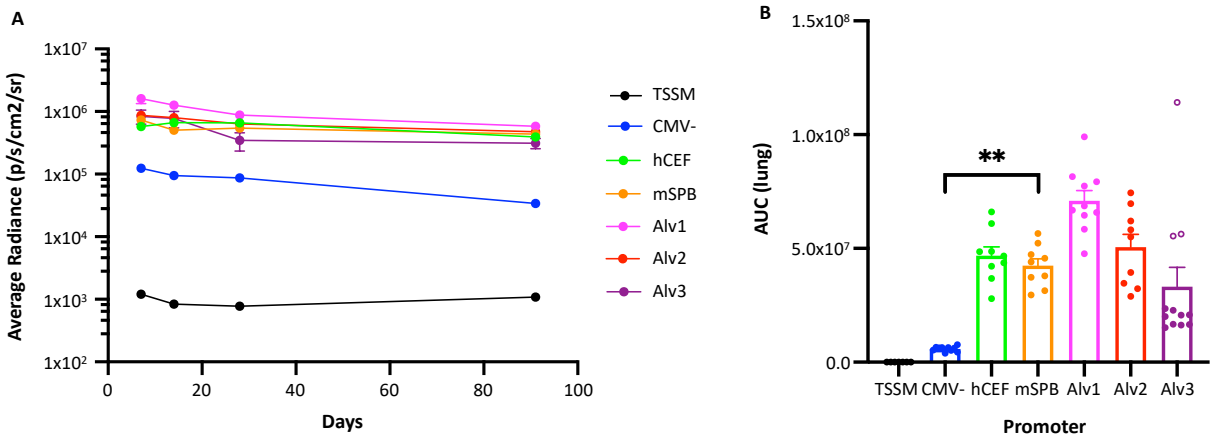


Figure 3.13 Expression profile of novel promoters in murine lung

Female BALB/c mice (n=7-12) were administered 2.5x10^8 TU of rSIV.F/HN lentivirus encoding *firefly* luciferase (fLux) under the control of CMV-, hCEF, mSPB, Alv1, Alv2 or Alv3 promoters, or formulation buffer (TSSM), via intranasal administration (Section 2.7.5). At 7, 14, 28 days and 3 months after transduction, mice were administered D-luciferin and imaged using an IVIS Lumina LT Imaging System to quantify fLux activity in the lung (Section 2.7.6). Regions of interest (ROIs) were drawn over the chest cavity and the signal within these regions measured in photons/second/cm^2/sr. **(A)** Average radiance and **(B)** area under the curve (AUC) analysis were plotted to understand the expression profile induced by promoters. Data are shown as individual values and/or mean ± SEM. **(B)** Significant differences between vector treated groups were determined by Kruskall-Wallis [H (5)=39.23, p<0.0001] with Dunn's post hoc multiple comparison test comparing vector groups with mSPB. A calculated p value of <0.05 was deemed significant (p<0.01,**). Alv3: open circles, W345; closed circles, W362 (Table 2.3).

the upper dose limit when given to patients. The relative yield achieved for the LS production of rSIV.F/HN.hCEF.ABCA3 (Figure 3.8) was $2.23x10^9$ TTU/l, which was 12.7-fold lower than equivalent productions of rLV encoding the reporter GFP transgene ($2.84x10^{10}$ TTU/l). This was unsurprising because the ABCA3 transgene is 4.4kb larger than GFP and a decrease in vector packaging efficiency is expected as cargo size (the distance from LTR to LTR) increases [197, 198]. However, the yield was similar to that achieved for rSIV.F/HN.hCEF.CFTR ($1.04x10^9$ TTU/l; Figure 3.8), an rLV which has been successfully transitioned to 200l scale, GMP-compliant manufacturing for clinical trials (Steve Hyde, Personal communication). Reducing the size of the *ABCA3* expression cassette to 8.1kb, keeping it within recommended guidelines [198], is expected to have contributed to the successful production of ABCA3 rLV with a viable manufacturing yield. By only including the protein-coding regions of *ABCA3* (Figure 3.1) the transgene was made 1.5kb smaller. For recombinant HIV-1 production, increasing vector genome length is generally thought to result in a semi-logarithmic decrease in functional viral titre [197, 198] when the production scale and final volume of purified rLV obtained are the same (all factors taken into consideration by the relative manufacturing yield).

If required, additional improvements to relative yield for ABCA3 rLV may be attainable by further truncating the transgene cDNA sequence [199]. The western blot in Figure 3.4 highlights two ABCA3 isoforms (of 170kDa and 190kDa). The apparent 190kDa protein corresponds to the full-length protein, whilst it is thought that the 170kDa truncated ABCA3 protein is produced following cleavage by cathepsin L at Lys174 (Section 1.2.3.1) [42, 209]. The purpose of this N-terminal truncation is not completely understood but has been suggested to regulate ABCA3 activity. Indeed, certain *ABCA3* mutations appear to alter the ratio of 170/190kDa protein [210]. If the 170kDa isoform of ABCA3 is identified as the active protein, then it may be possible to remove approximately 500bp from the *ABCA3* transgene which correspond to the 20kDa cleavage product.

The rLV production method is another area of exploration for improving vector yield. As seen in Figure 3.7, using the commercially available LV-MAX cell system for production offered no advantage over our established in-house platform which uses suspension adapted HEK293T cells. The ratio of plasmids used at transfection for our in-house platform is based on plasmid mass. This means that the larger the genome plasmid the fewer copies of genome plasmid that are available during virus production. Altering the plasmid ratios to normalise genome plasmid copy number, by accounting for genome size, may improve vector yields for larger transgenes, and is something that will be explored in Chapter 5.

3.3.2 Lentiviral vector design for clinical translation

To have the most benefit in patients and limit off-target effects, the virus produced should drive efficient protein expression specifically in the target cell population. The target cell population for ABCA3 gene therapy is ATII cells in the parenchyma of the lung. Obtaining transgene expression in these cells, along with limiting expression in other cell types can be achieved by a combination of lentiviral pseudotyping and promoter selection. Pseudotyping dictates which cell types will be transduced, whilst the transgene promoter is then responsible driving transgene expression in those cells. It is known that the rSIV.F/HN vector platform used in this research (Section 3.2.2), can transduce ATII cells in mice, however it also transduces other predominant cell types in the airway [139]. An alternative pseudotype, such as VSV-G, may therefore be a better choice for ABCA3 rLV. Unlike F/HN, transduction via VSV-G is poor in airway epithelial cells *in vivo* as the VSV-G receptors are located predominantly on the basolateral surface [211, 212]. Research by Borok *et al.* [212] suggests this same access limitation is not observed with alveolar epithelial cells. To test if the VSV-G pseudotype would transduce ATII cells whilst restricting transduction of other cell types in the respiratory system, lungs harvested from mice treated with a VSV-G pseudotyped rLV, containing a ubiquitous mammalian promoter (e.g. Ef1α or hCEF) and a fluorescent transgene (e.g. GFP), could be immunostained for cell co-localisation markers (e.g. pro-SPC for ATII cells). Whilst

overexpression of ABCA3 in the airways may cause adverse reactions, these cells physiologically express ABCA3 at very low levels and only in ATII cells has its role been found to be non-redundant [31, 92, 213]. Overexpression of ABCA3 in the airways may result in cellular stress and immune response activation if the cells are unable to correctly process the protein in large quantities.

Following transduction, selective transgene expression is regulated by promoter selection. The rLV genome plasmids were generated with three ubiquitous promoters (Section 3.2.1.2). Of these, hCEF was selected due to its ability to drive sustained, high level protein expression in the lung (Figure 3.13) and due to previous use in clinical trials [137, 139]. However, hCEF is also highly active in the airways (as confirmed by nasal bioluminescence, Figure 3.11; low lung/nose ratio, Figure 3.12) having first been developed for treatment of an airway-affected disease (CF). The promoters evaluated in Section 3.2.3 were originally designed to be ATII cell-specific, however, the *in vivo* BLI used for analysis could only provide an indication of alveolar specificity (Figure 3.12). To acquire more detailed knowledge of the individual cell types transduced an immunohistochemistry (IHC) co-staining approach would be necessary – for example, immunostaining for the transgene and pro-SPC would identify ATII cells.

Within all the new promoters assessed by BLI, a mSPB promoter was used for transcription initiation. Based on results in Section 3.2.4, mSPB, Alv2 or Alv3 promoters all appeared to drive alveolar restricted expression (Figure 3.11 and 3.12) at a level analogous to hCEF. Whilst this indicates that one of these three promoters should be considered if ATII cell-specific transgene expression is required, SP-B is also expressed to some degree in club cells [208] and thus promoters developed based on the SP-B promoter are unlikely to be completely specific for ATII cells. In contrast, SP-C is thought to only be expressed in ATII cells [208], thus novel promoters designed around the endogenous SP-C promoter sequence may be more fully ATII cell-specific. Using an SP-C promoter, Degiulio *et al*. [214] have demonstrated ATII cell-specific expression of YFP in mouse lung after intratracheal delivery of DNA. However, preliminary work by Dr Altar Munis

(unpublished), transducing an ATII SALI culture model [103] with rLV expressing GFP under the control of full and minimal SP-C promoters, indicates SP-C based promoters drive very low levels of expression with >10-fold reduction in GFP-positive cells compared with hCEF and mSPB promoters.

3.3.3 Conclusion

In summary, genome plasmids containing a truncated ABCA3 transgene were engineered and confirmed to express ABCA3 protein in a relevant lung cell line, when packaged into rLV pseudotyped with F/HN. Using a luciferase transgene the alveolar specificity and long-term expression profile of several new promoters were evaluated to identify alternative vector configurations with improved alveolar specificity. For the current preferred vector configuration (rSIV.F/HN.hCEF.ABCA3; vGM336), viral titres obtained from LS in-house production indicate this is a viable gene therapy product for transitioning to GMP-compliant manufacturing methods for clinical use. This candidate ABCA3 gene therapy vector should now be tested in *in vitro* and *in vivo* models of ABCA3-mediated surfactant dysfunction to confirm functionality of the therapeutic product and to understand its effectiveness at rescuing a disease phenotype.

Chapter 4 Development of cell models for ABCA3-mediated surfactant dysfunction

4.1 Introduction

In chapter 3, the generation of rLV vectors encoding ABCA3 was described. To evaluate these, a cellular model of ABCA3-mediated surfactant dysfunction was required. For a model to be representative of the target disease it was important that it recapitulated multiple features of ATII cells (Section 1.4.1). Patient-derived primary ATII cells should, in theory, be the ideal model for evaluating ABCA3 gene therapy strategies. Cells derived from patients with ABCA3-mediated surfactant dysfunction are difficult to obtain due to the rarity of the disease. In addition, culturing primary human ATII cells can present challenges due to their progenitor properties (Section 1.4.1.1). Most attempts at culturing human-derived primary ATII cells have failed as the cells either transdifferentiate into cells with an ATI phenotype or the cells become senescent and stop proliferating after two generations [215]. Consequently, the field has frequently turned to immortalised cell lines to model the lung parenchyma. Whilst immortalised cell lines resolve problems surrounding maintenance in cell culture, there is ongoing debate about which cell lines best mimic the properties of primary ATII cells. Swain *et al.* [216] concludes that the biomolecular fingerprint of the A549 cell line, used as a mainstay model of ATII cells since the 1970s, is statistically more like ATI cells than ATII cells. Cooper *et al.* [217] attempts to explain the phenotypic discrepancies found in A549 cells, suggesting that time in culture prior to use has a large influence on the observed phenotype. Most notably, immortalised cells tend not to secrete surfactant – a key function of ATII cells in the lung [218]. Furthermore, the cancer-derived origin of many immortalised cell lines means that the transcriptome is altered which, in turn, may influence their responses to potentially therapeutic agents.

The overall aim of the studies described in this chapter was to identify an ATII-like cell system amenable to both genetic manipulation for generation of an ABCA3 surfactant dysfunction disease model, and transduction with rSIV.F/HN for evaluation of potential gene therapy products.

4.2 Results

4.2.1 Primary alveolar pneumocytes as a model of alveolar type II cells

A mixed population of primary pneumocytes was obtained (EP71AL, Epithelix) and characterised for stability of the constituent ATII cell population and the presence of intracellular LB organelles as a defining feature of ATII cells (Section 1.2.1). Particular attention was paid to the progenitor nature of ATII cells (Section 1.1) [10, 114]. When grown in submerged culture (Section 2.3.1.2) cells were assessed for the presence of LBs by TEM (Section 2.6.5.2). At an early passage (passage 6) densely packed concentric circles with a diameter of about 400nm, characteristic of mature LBs [219], were abundant (Figure 4.1). Additional structures associated with LB biogenesis were also observed, such as potential MVBs with small ILVs (Figure 4.1B) and immature LBs (Figure 4.1C).

After confirming the presence of LBs in the pneumocyte mix, attempts were made to quantify the proportion of ATII cells present in the population and to determine if this proportion changed as cells were maintained in culture over a period of four weeks. Cells were cultured in 24-well plates and stained for the ATII cell marker pro-SP-C in triplicate (Section 2.6.4.1; Table 2.5). Two images per replicate were captured at random and the percentage of pro-SP-C positive (pro-SP-C+) cells calculated using Fiji (Section 2.6.5.1). ATII cells were present in cultures at passages 4 and 8 (four weeks; Figure 4.2), however, there was a 23.6% reduction ($p=0.0022$; Mann Whitney test) in the proportion of pro-SP-C+ cells by passage 8. This indicated a loss of ATII cells with continued cell culture, suggesting that the primary ATII cell population was unstable and that the development of a model of ABCA3 deficiency would not be feasible in these cells under the current culturing conditions. No further characterisation of these cultures was undertaken.

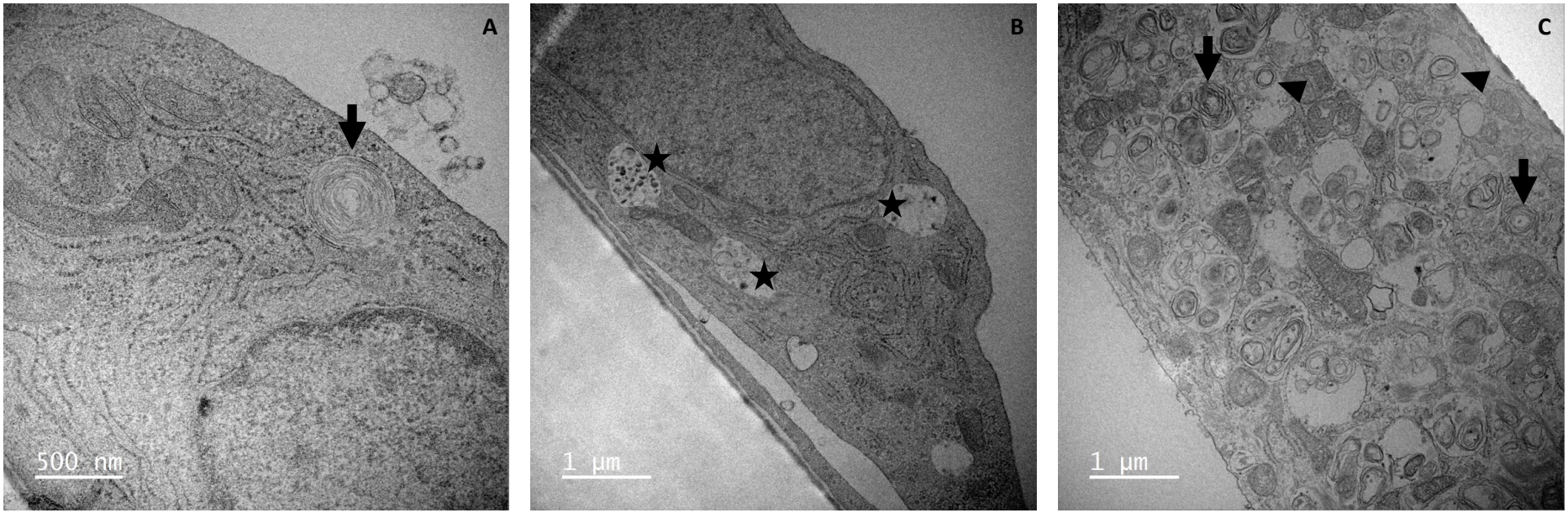


Figure 4.1 Electron microscopy of human primary alveolar pneumocytes showing lamellar body structures

A vial of primary alveolar pneumocytes (Epithelix) was thawed, cultured submerged to passage 6 (Section 2.3.1.2) and analysed by TEM (Section 2.6.5.2) for the presence of lamellar bodies (LBs). Approximately, 30 images captured at random were analysed. **(A)** A mature LB with a diameter of approximately 400nm. **(B)** Potential multivesicular bodies (MVBs) containing small intralumenal vesicles (ILVs). **(C)** A mix of mature and immature LBs. Mature LB: arrow. MVB: star. Immature LB: arrowhead.

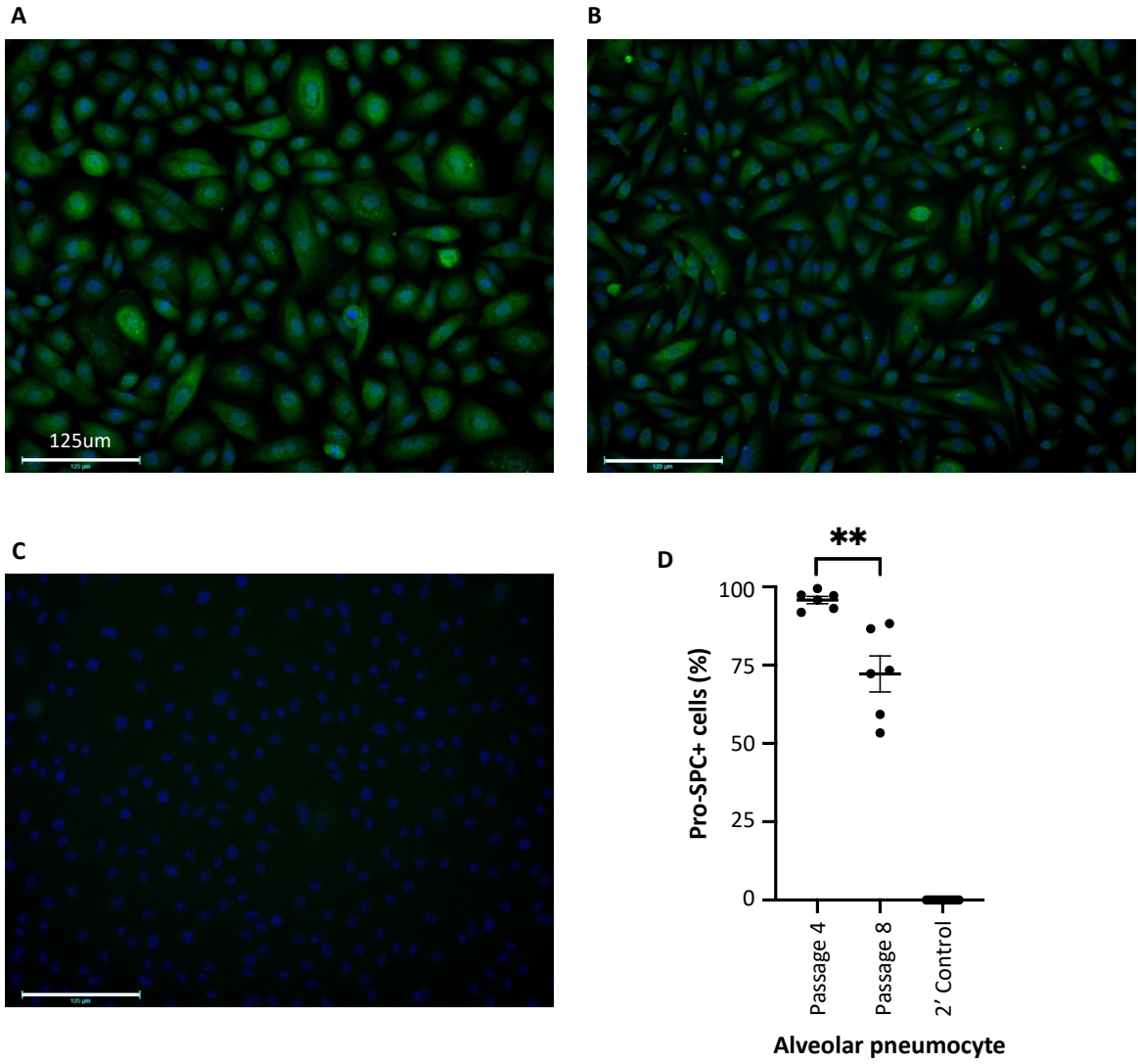


Figure 4.2 Immunofluorescence microscopy analysis of an ATII cell marker in alveolar pneumocytes

Primary human alveolar pneumocytes at passage 4 **(A)** and passage 8 **(B)** submerged cultures were stained for the ATII cell marker, pro-SPC (green), after permeabilisation with triton X-100. Staining for pro-SPC was detected at both passages. **(C)** Secondary antibody-only (AlexaFluor555) staining control of passage 8 alveolar pneumocytes. Nuclei counterstained with DAPI (blue); all scale bars indicate 125µm. Representative images from three replicates taken using the EVOS Auto 2 with fixed brightness settings. **(D)** The percentage of pro-SPC+ cells was quantified using Fiji to count the

number of blue (DAPI) and green (pro-SPC) particles. Mean ± SEM shown. A significant difference between the values was determined by a Mann-Whitney test [U (6, 6)=0, z=-23.64, p=0.0022]. A calculated p value of <0.01 (**) was deemed significant.

4.2.2 Evaluation of immortalised lung cell lines

In addition to replicating features of ATII cells (Section 1.1), a suitable cell model should also be amenable to genomic manipulation and lentiviral delivery; therefore immortalised lung cell lines of human (A549 [210], H441 [103]), murine (LA-4 [220], MLE-12 [221]) and rat (RLE-6TN [222]) origin were considered due to their prior use in studies of the alveolar epithelium. In this study, human-derived cells were selected for characterisation because of the increased relevance to human lung and improved translatability to patients. The human H441 and A549 adenocarcinoma cells were characterised with regards to the expression of endogenous surfactant-related genes and presence of LB organelles. The impact of culturing cells in an 'air-lifted' ALI system (Section 1.4.1.2, Section 2.3.2) was also investigated.

The mRNA expression from surfactant-related genes *ABCA3* and *SFTPB* was measured in H441 and A549 cells by relative qRT-PCR (Section 2.5.5; Bact_qRT_F, Bact_qRT_R, ABCA3_qRT_F_Ex4, ABCA3-qRT_R_Ex5, SPB_qRT_F, SPB_qRT_R; Table 2.1) and compared with expression levels in a non-lung, human cell line (HEK293T). *ABCA3* expression was detected in all three cell lines grown under submerged conditions (Section 2.3.1.1), with the highest level observed in HEK293T cells ([F (4, 24)=3.464, p=0.0227]; Figure 4.3A). When lung cell lines were grown air-lifted in a more physiologically relevant ALI culture system (Section 1.4.1.2; Section 2.3.2) RNA expression tended to increase in both H441 (3.2-fold) and A549 cells (1.9-fold), although these differences did not reach statistical significance (Figure 4.3A; $p>0.05$). When mRNA expression from *SFTPB* (Section 1.2.1.1) was measured, differences in expression between the cell lines and culture systems were more readily apparent – *SFTPB* expression tended to increase after cells were air-lifted (Figure 4.3B). SALI cultures expressed significantly higher levels of *SFTPB* compared with H441 cells grown in submerged culture or A549 cells in either culture system ([F (4, 10)=44.02, $p<0.0001$]; post hoc all $p<0.0001$).

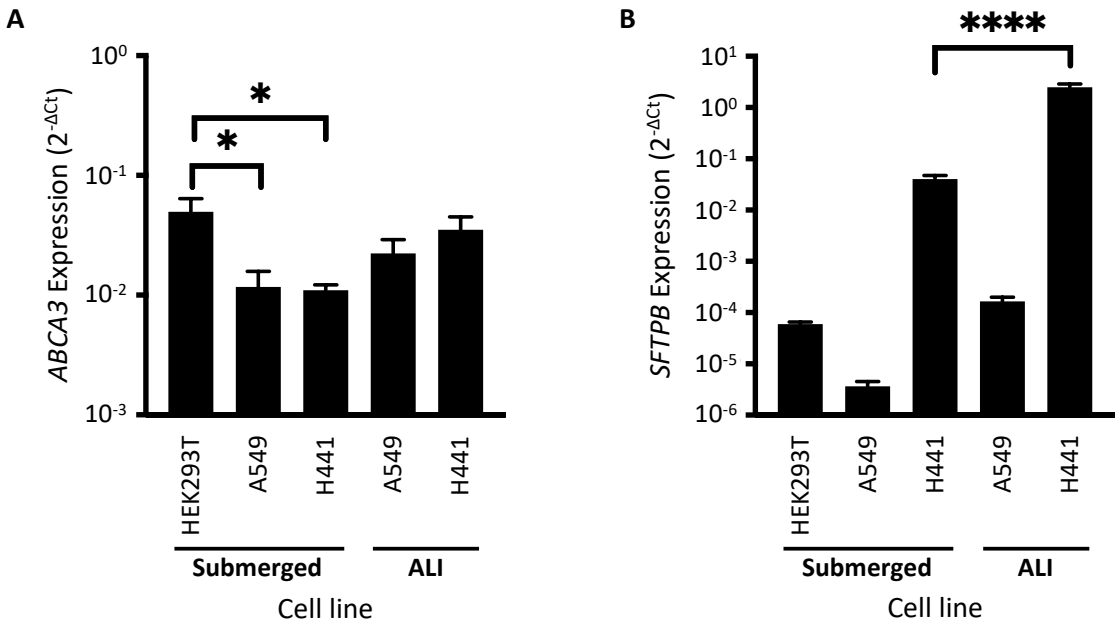


Figure 4.3 RNA expression of surfactant protein genes *ABCA3* and *SFTPB* in immortalised lung cell lines

RNA was extracted from HEK293T, A549 and H441 cells cultured in submerged conditions or an ALI system. **(A)** *ABCA3* and **(B)** *SFTPB* expression was determined via relative qRT-PCR and normalised to β-actin expression using $2^{-\Delta Ct}$ analysis. Data are shown as mean ± SEM of n=2 independent experiments, with three technical replicates per qRT-PCR assessment. Significant differences between the values were determined by one-way ANOVA with Tukey's all means post hoc analysis. A calculated p value of >0.05 was deemed non-significant (ns). Calculated p values of <0.05 (*) and <0.0001 (****) were deemed significant.

Differences in protein expression by western blotting (Section 2.5.6) were also observed, although interestingly these were not consistent with the qRT-PCR data (Figure 3.3A). Figure 4.4A shows that endogenous ABCA3 protein was only detectable in lysates of HEK293T cells, A549 cells cultured in submerged conditions and H441 SALI cultures. As an alternative analytical approach, ICC for ABCA3 protein (Section 2.6.4.1) was performed for H441 and A549 cells grown in submerged culture. Punctate ABCA3 staining around the nuclei was observed in both A549 (Figure 4.4D, E) and H441 (Figure 4.4F, G) permeabilised cultures. There appeared to be no obvious differences in expression pattern or level between the two submerged cell lines. In contrast, when exogenous ABCA3 was expressed from a plasmid (Figure 4.4C) greater levels of staining were observed.

Lamellar bodies have previously been identified in A549 cells [223, 224]. Analysis of A549 and H441 cells by TEM (Section 2.6.5.2) was performed on cells grown submerged or in an ALI system to look for LB structures (Figure 4.5). Under submerged conditions, LBs were abundant in A549 cells (Figure 4.5A, B), being observed in 12/23 images, but not present in H441 cells (24 images taken). Air-lifting H441 cells to generate SALI cultures resulted in the emergence of some LB structures (Figure 4.5C), although these were rare and lacked the highly structured LB organisation seen in primary cells (Figure 4.1). What was observed in H441 cells, both submerged but more abundantly in SALI cultures, were electron-lucent structures (Figure 4.5D, E). These could be regions of low electron density that are typically generated when lipid-dense structures are not successfully fixed by aldehyde-based fixatives (used here; Section 2.6.5.2) [225].

To investigate if extending the time in cell culture after air-lift could increase the number of LBs or induce LB maturation, SALI cultures were maintained for a maximum of 6 weeks after air-lift and samples fixed for TEM at regular intervals. As seen in Figure 4.6, prolonged culture did not result in a discernible difference in observed LB numbers or structure. However, the abundance of electron-lucent regions surrounded by dark, electron-dense rings (Figure 4.6, asterisks) increased, which may indicate an increase in lipid species with prolonged culture (Section 4.3.2).

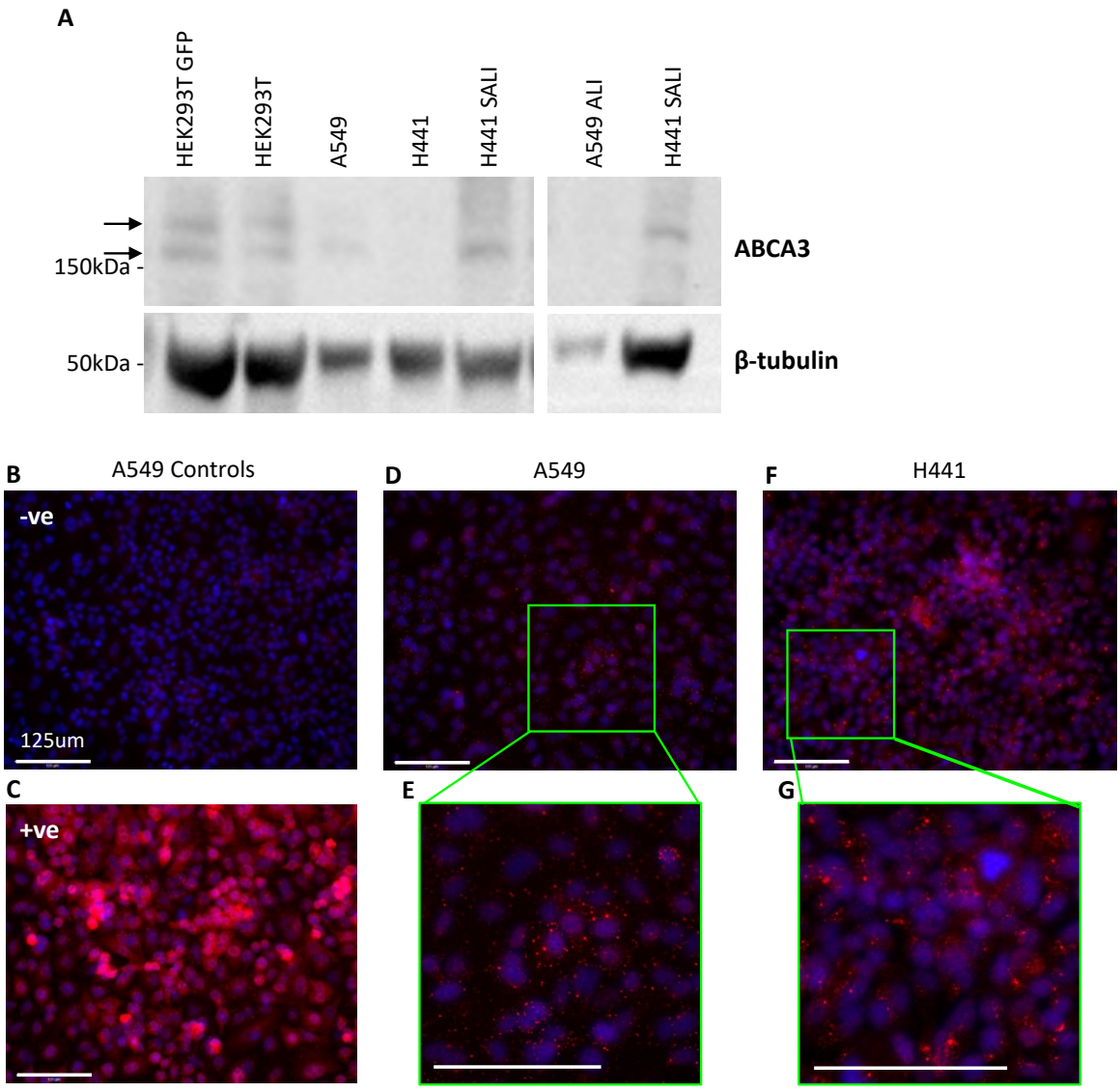


Figure 4.4 Endogenous ABCA3 protein expression in immortalised lung cell lines

(A) Representative western blot (n=3) of HEK293T, A549 and H441 cell protein lysates extracted from cells grown in submerged or ALI culture systems. 40µg of protein lysates were denatured under reducing conditions for western blot analysis using ABCA3 primary antibody PA5-52478. HEK293T cells transiently transfected with plasmid expressing GFP (pGM311; Table 2.2) were included as a positive (+ve) control (HEK293T GFP) having previously been demonstrated to express detectable levels of ABCA3 (Figure 3.3); β-tubulin was used as a protein loading control. The size of the closest protein markers to the target proteins are indicated. **(B-G)** Merged images of ABCA3

staining (PA5-52478) in **(B-E)** A549 and **(F-G)** H441 cells grown in submerged culture. In duplicate, cells were fixed, permeabilised and immunostained for ABCA3 (red) using AlexaFluor555-conjugated anti-rabbit secondary antibody. Nuclei were counterstained with DAPI (blue). Three images were captured at random for each replicate and representative images are shown. **(B)** Secondary antibody only stained control (-ve) and **(C)** ABCA3 staining in A549 cells transiently transfected with p.CMV-ABCA3 (pGM955; +ve). **(E, G**) Zoomed in view of green box regions indicated in **(D, F**). All scale bars indicate 125µm.

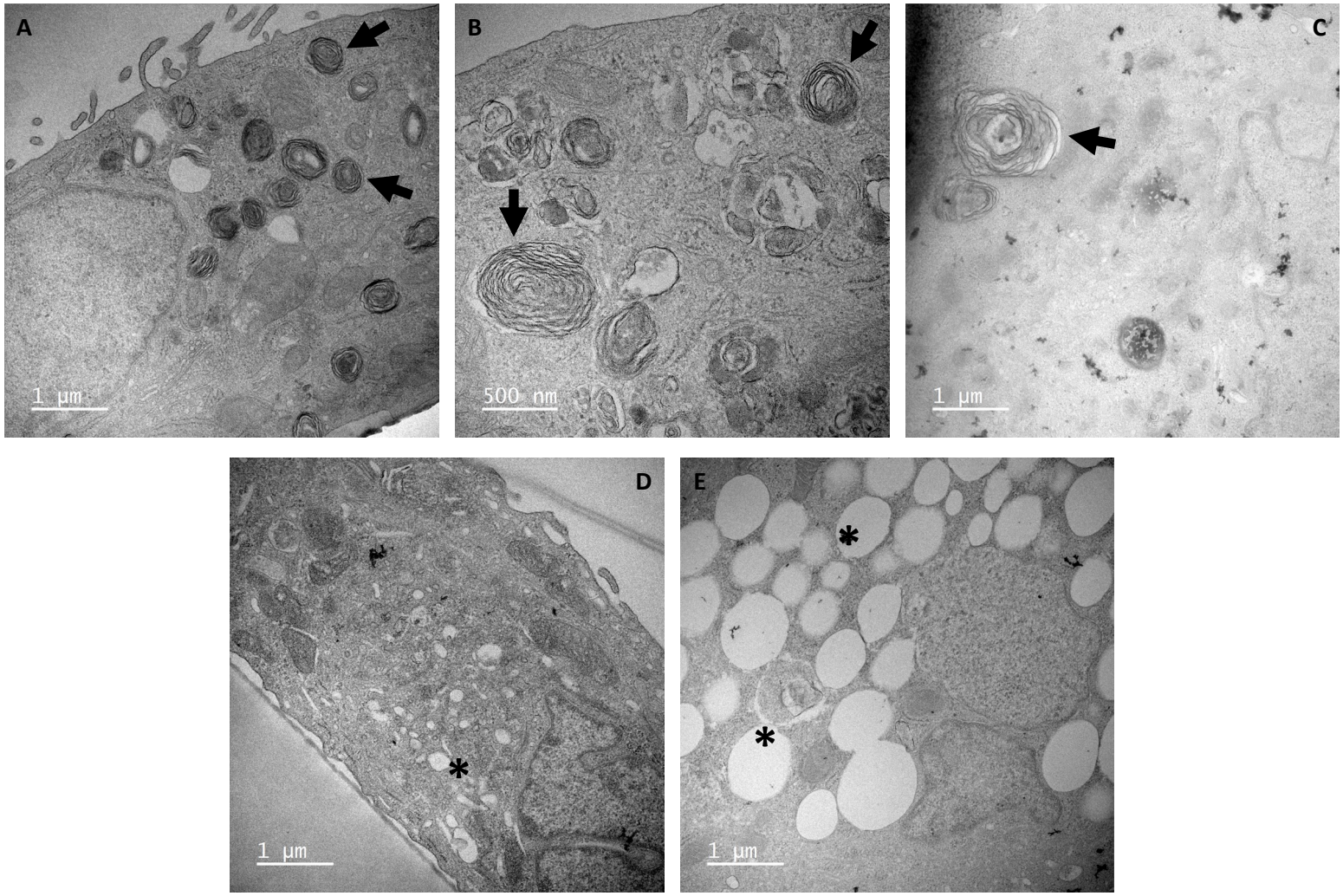


Figure 4.5 Electron microscopy of immortalised lung cell lines showing lamellar body structures

Figure 4.5 Electron microscopy of immortalised lung cell lines showing lamellar body structures

Transmission electron microscopy (Section 2.6.5.2) was performed to detect lamellar bodies (LBs) and approximately 25 images per sample, taken at random, were analysed. **(A-B)** Mature LBs with highly organised intra-organelle structure (arrow) observed in A549 cells grown in submerged culture. **(C)** LB structure (arrow) identified in SALI culture (air-lifted H441 cells). **(D-E)** Electron-lucent structures (asterisk) in H441 cells grown under **(D)** submerged conditions and **(E)** as a SALI culture. Scale bars as annotated.

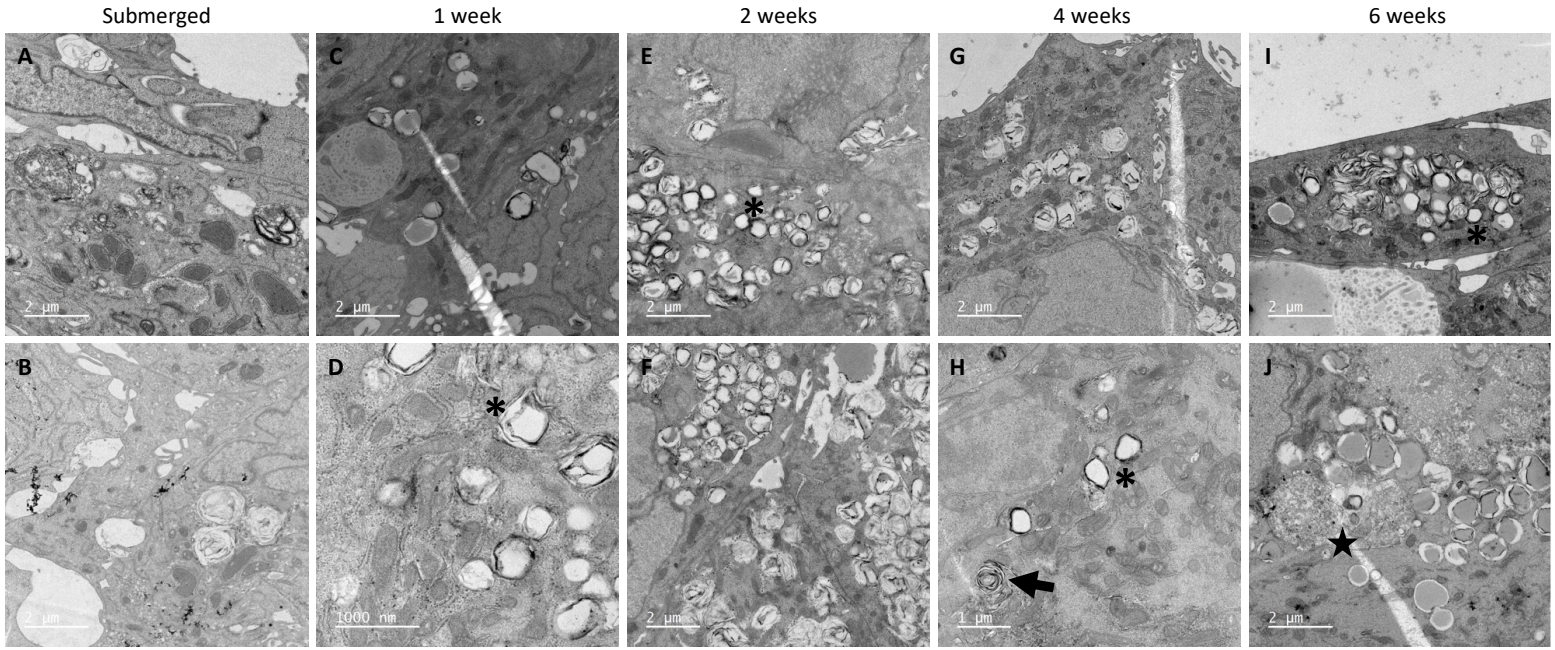


Figure 4.6 Electron microscopy of H441 SALI cultures over time

Representative TEM images (n=50 per replicate) of H441 cells grown under submerged **(A-B)** or air-lifted **(C-J)** culture conditions. Cells were seeded in duplicate on Transwell inserts and fixed for analysis at 1 week **(C-D)**, 2 weeks **(E-F)**, 4 weeks **(G-H)** or 6 weeks **(I-J)** post air-lift. **(A-B)** H441 cells cultured under submerged conditions for 3 days were included for comparison. **(C-J)** Electron-lucent regions (asterisk) were detected at multiple timepoints after air-lift. **(H)** LB (arrow) measuring approximately 500nm and three electron-lucent regions (asterisk) each measuring approximately 400nm in diameter. Scale bars as annotated. Potential multivesicular body: star.

4.2.3 Development of an *ABCA3* knockout cell line

The aim of this experiment was to ablate ABCA3 expression by CRISPR-mediated disruption of the *ABCA3* coding sequence. Two strategies for CRISPR-mediated KO of *ABCA3* are described below – first, the deletion of whole exons of *ABCA3* (Section 4.2.3.2) and second, the introduction of indels downstream of the ATG start codon in exon 4. The H441 cell line was initially selected as it had previously been used successfully in genome editing experiments targeting *SFTPB* [103] and another ABC protein, ABC subfamily C member 1 (*ABCC1*) [226] loci.

4.2.3.1 gRNA selection and cloning

To identify gRNA sequences for efficient SpCas9-mediated cleavage of the *ABCA3* locus, the CRISPOR open access web tool (GRCh37/hg19 reference genome assembly) [227] was used. This service provided a range of potential gRNAs (Table 4.1) which were ranked according to the criteria in Table 4.2 – a MIT specificity score above 90 and no off-target sequence complementarity within three mismatches were considered essential criteria for gRNAs. A high Doench efficiency score [228] was also desirable however predictive efficiency scores often correlate poorly to editing outcomes *in vitro* and thus a Doench score above 50 was accepted [229]. The highest-ranking gRNA sequences targeting regions upstream of the transcription and translation start sites (intron -1 and exon 4 respectively), in addition to two early translated exons (exon 5, exon 8) were selected for analysis (Figure 1.5). A further three gRNA sequences (Table 4.1) targeting the first full coding exon (exon 5) were subsequently designed by Synthego. The gRNAs were cloned into the SpCas9 backbone plasmid PX459 V2.0 (Table 2.2) which contains a puromycin mammalian selection marker (Sections 2.2.1 and 2.2.3; Table 2.1: SpCas9 gRNA plasmid cloning).

Table 4.1 Guide RNA sequences targeting *ABCA3*

Name	Target	Sequence	Source
Int-1_hABCA3	Intron -1	CGAACCGCGTACTCGCAGGG	CRISPOR
Ex4_hABCA3	Exon 4	CGTCCCTTTCAGCAAGACGA	CRISPOR
Ex5_hABCA3	Exon 5	CCGTCACTGAGACAGTGCGC	CRISPOR
Ex5.1_hABCA3	Exon 5	GCCCGGGTAGATGGTGGCGT	Synthego
Ex5.2_hABCA3	Exon 5	TGTTTTCTGGGATCCTCATC	Synthego
Ex5.3_hABCA3	Exon 5	CTTCCAGAAGCGGAAGGTCC	Synthego
Ex8_hABCA3	Exon 8	AAGGGGTCTGCGATGAACGG	CRISPOR

Table 4.2 Selection criteria for SpCas9 gRNAs targeting *ABCA3*

gRNA	MIT Specificity	Doench Efficiency	Off-target next to PAM (0-1-2-3-4 mismatches)
Int-1_hABCA3	97	71	0-0-0-0-1
Ex4_hABCA3	94	62	0-0-0-0-0
Ex5_hABCA3	93	57	0-0-0-0-1
Ex8_hABCA3	96	61	0-0-0-0-0

The editing efficiency of each gRNA in the absence of puromycin was assessed in H441 cells by ICE (Section 2.2.16). Cells were seeded in 24-well plates ($1.2x10^5$ cells/well) 24 hours before Lipofectamine-mediated transfection (Section 2.3.3) of SpCas9 plasmids or a GFP-expressing transfection control plasmid (pGM311, Table 2.2) in triplicate. At 72 hours after transfection, when GFP signal was detectable in transfection control cells, gDNA was extracted (Section 2.5.1) and the region surrounding the target edit sites PCR amplified (Section 2.2.2; Table 2.1: ICE (PCR amplification)). Single amplicons of the expected size (400-500bp) were obtained for gRNAs targeting exons 4, 5 and 8, however, the region surrounding intron -1 failed to amplify and no further analysis of this gRNA was pursued. Following Sanger sequencing (Table 2.1: ICE (sequencing)), and comparison of the resulting traces with non-transfected controls by ICE, no evidence for SpCas9-mediated genomic editing with any gRNA was observed (Table 4.3). A multi-sgRNA targeting the T cell receptor α constant (*TRAC*) loci was obtained from Synthego and used as a positive control for ICE analysis – an ICE score of 16% was recorded. These results indicated that the editing efficiency of the gRNAs was poor and could not be detected by ICE analysis without puromycin selection to enrich the edited cell population.

Table 4.3 Editing efficiency of gRNAs targeting *ABCA3* in H441 cells

ICE score (0-100) reflects overall editing efficiency. Knockout (KO) score (0-100) reflects frameshifts or >21bp indels.

gRNA	ICE Score	KO score
Ex4_hABCA3	0	0
Ex5_hABCA3	0	0
Ex5.1_hABCA3	0	0
Ex5.2_hABCA3	0	0
Ex5.3_hABCA3	0	0
Ex8_hABCA3	0	0
TRAC (+ve control)	16	16

4.2.3.2 Approach 1: Large fragment deletion strategy

Given the low editing efficiency achieved with the individual gRNAs (Section 4.2.3.1), a multi-gRNA strategy aimed at excising a fragment of *ABCA3* was used to increase the probability of gene disruption. H441 cells were transfected (Section 2.3.3) with a total of 2500ng DNA split equally between four gRNA-expressing plasmids (pGM1042 – pGM1045, Table 2.2; Int-1_hABCA3, Ex4_hABCA3, Ex5_hABCA3, Ex8_hABCA3; Table 4.1). After puromycin selection (Section 2.3.4), the number of viable cells in transfected wells was greatly reduced compared with non-puromycin treated controls; thus, transfected cells were serially diluted in three 96-well plates to a concentration of 0.8 cells/well in 50% conditioned media (0.2µm filter purified from parental H441 cells). Four weeks after seeding in 96-well plates, colonies had formed in ~10% of wells (29/288). Genomic DNA was extracted from these potential *ABCA3* KO cell clones and genotyped by endpoint PCR (Section 2.2.2; ABCA3KO_gt_In-1_F, ABCA3KO_gt_In+8_R; Table 2.1). Attempts to identify large fragment deletions in the *ABCA3* locus by amplification of a 22kb region (between intron -1 and intron 8) were unsuccessful. Instead, two sequential endpoint PCR reactions targeting exon 2 to exon 4 (ABCA3KO_gt_Ex2_F, ABCA3KO_gt_Ex4_R; Table 2.1) and exon 3 to exon 6 (ABCA3_KOgt_Ex3c_F, ABCA3_KOgt_Ex6c_R; Table 2.1) were performed (Figure 4.7). When visualised by agarose gel electrophoresis, amplicons of approximately 3kb (exons 2 – 4) and 4kb

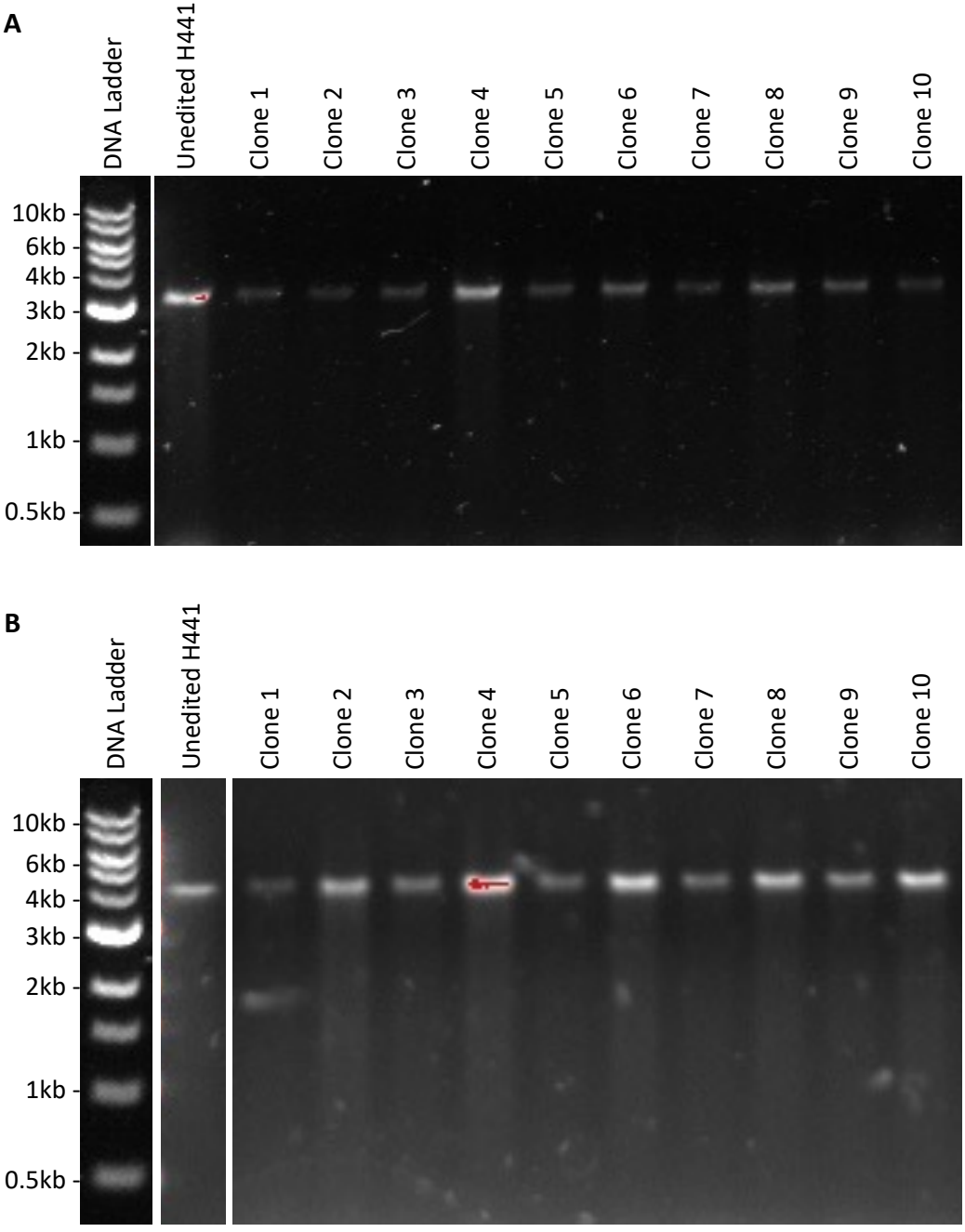


Figure 4.7 Genotyping potential H441 *ABCA3* knockout clones by endpoint PCR

Genomic DNA was extracted from naïve, unedited H441 cells together with putative H441 edited clones transfected with CRISPR plasmids targeting the *ABCA3* locus and was PCR amplified with primers targeting **(A)** exon 2 to exon 4 and **(B)** exon 3 to exon 6. After agarose gel electrophoresis, visualised amplicons of approximately 3kb (exon 2 to exon 4) and 4kb (exon 3 to exon 6) indicated no fragment deletion. Representative images are shown.

(exons 3 – 6) were observed in unedited control samples and potential *ABCA3* KO clones (Figure 4.7) indicating that substantial target fragment deletion was unsuccessful – although a 2kb amplicon suggestive of heterozygous editing was seen in Clone 1 (Figure 4.7B). Sanger sequencing of gel extracted amplicons also failed to identify a clone containing a disrupted *ABCA3* locus.

4.2.3.3 Approach 2: Disruption of exon 5

The absence of detectable editing when targeting a large region of *ABCA3* prompted a change in strategy of both editing and CRISPR delivery. In a preliminary study, three sgRNAs targeting exon 5 of *ABCA3* (Table 4.1; Synthego) were complexed with SpCas9 protein before reverse transfecting H441 cells using Lipofectamine CRISPRMAX (Section 2.3.3). Preliminary ICE analysis (Table 2.1: ICE) in H441 cells reported a 5% editing efficiency; this was lower than that achieved for a manufacturer-provided control of multi-sgRNA targeting the *TRAC* locus (9%), but higher than that observed previously with individual gRNAs and plasmid delivery (Section 4.2.3.1). Following preliminary analysis of editing efficiency, H441 cells were reverse transfected (Section 2.3.3) using a ratio of 1.3:1 sgRNA to SpCas9 for RNP formation. After four days in culture, gDNA was extracted from a proportion of the cells for editing analysis. PCR amplification of *ABCA3* exon 5 (SynthKO_ABCA3_F2, SynthKO_ABCA3_R2; Table 2.1), from cells transfected with the SpCas9-RNP complex targeting *ABCA3,* produced three detectable amplicons between 300-400bp upon gel electrophoresis (Figure 4.8A). This amplification pattern differed from that of control *TRAC*-edited cells and naïve, non-transfected H441 cells where a singular amplicon of approximately 400bp was detected. Sanger sequencing of the individual amplicons (Figure 4.8B) identified a 97bp deletion, indicating editing of the *ABCA3* locus had been successful – ICE analysis of the combined amplicons further supported this conclusion with an editing efficiency score of 10% and KO score of 6% recorded. According to the reference *ABCA3* gDNA sequence (ENSG00000167972), the 97bp deletion in exon 5 was expected to cause a frameshift resulting in KO of ABCA3 protein expression. Efforts to isolate an *ABCA3* KO clonal cell line by limiting dilution of some of the remaining cells were unsuccessful and

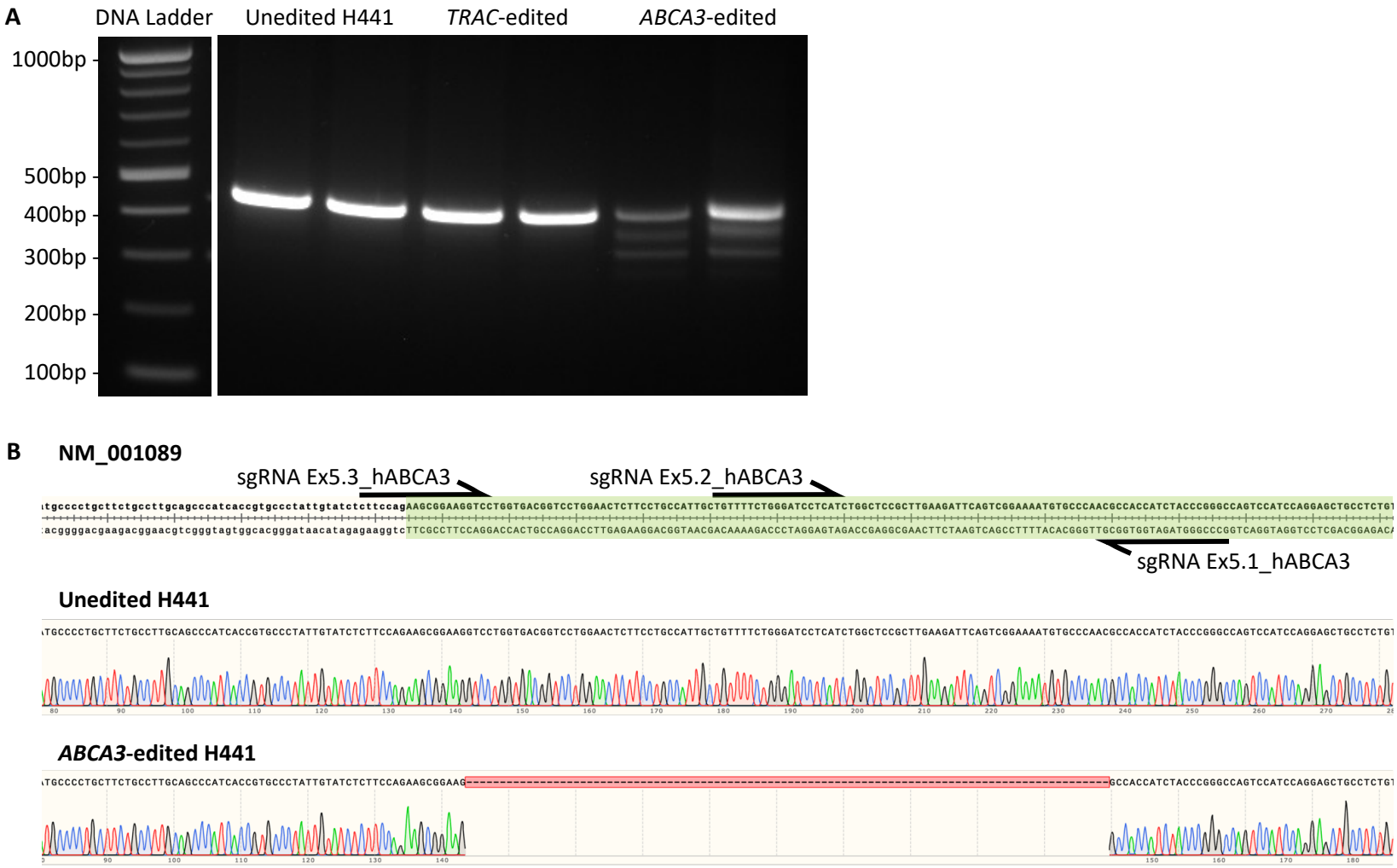


Figure 4.8 Transfection of H441 cells with sgRNA-spCas9 RNP complexes targeting *ABCA3* exon 5 produces indels

Figure 4.8 Transfection of H441 cells with sgRNA-spCas9 RNP complexes targeting *ABCA3* exon 5 produces indels

(A) Agarose gel electrophoresis of PCR-amplified gDNA extracted from naïve H441 cells (Unedited H441) and H441 cells transfected with an sgRNA-spCas9 RNP complex targeting the *ABCA3* locus (*ABCA3*-edited) or a control locus (*TRAC*-edited). PCR amplification of *ABCA3* exon 5 was expected to produce a single amplicon of 438bp without large indels. Three amplicons (approximately 300 – 400bp) were detected in *ABCA3*-edited H441 gDNA samples. **(B)** Sanger sequencing traces from gel purified amplicons showing a 97bp deletion in *ABCA3* exon 5 in *ABCA3*-edited H441 cells. The sgRNA binding locations are annotated on the reference sequence (NM_001089). *ABCA3* exon 5 is highlighted in green. *ABCA3*-edited H441 trace obtained from 300bp amplicon.

the proportion of cells which contained edits in the *ABCA3* locus (in the pool) was deemed too low to accurately model ABCA3 deficiency. Successful editing of the *ABCA3* locus was subsequently confirmed (Hamid Dolatshad, Personal communication; Figure 4.9) after isolation of edited H441 cell clones from cryopreserved H441 cells (Section 2.3.1) transfected with the SpCas9-RNP complex targeting *ABCA3*. Unfortunately, these edited *ABCA3* KO clones were available too late to be included in this study.

4.2.4 Expression of clinically relevant ABCA3 variants in H441 cells

As an alternative to generating an *ABCA3* KO cell line for disease modelling and proof-of-principle studies, the impact of overexpressing known pathogenic ABCA3 variants in H441 cells was explored. Clinically significant mutations in *ABCA3* were identified from the literature [88] and five mutations spanning the length of the gene and with differing associated disease severities (Section 1.3.2.3) were selected for modelling: ABCA3 variants L101P, E292V, N568D, T1114M, Y1515X (ClinVar variation ID: 8012, 203381, 8015, 2627240, 992982). In brief, H441 cells were transduced with rLV encoding each of the ABCA3 variants (Section 4.2.4.2) before the protein expression profile and functional activity of the pathogenic variants (Section 4.2.4.3) was investigated. Finally, the cell lines expressing ABCA3 variants were transduced with rSIV.F/HN.hCEF.ABCA3 (pGM336, Table 2.3) to demonstrate functionality of the vector (Section 4.2.5).

4.2.4.1 Construction of rLV genome plasmids expressing pathogenic ABCA3 variants

Mutations were first introduced into the pCMV-.ABCA3 plasmid pGM955 (Table 2.2) using SDM (Section 2.2.8). After confirmation of successful mutagenesis (Section 2.2.15; Table 2.1: site-directed mutagenesis), the *ABCA3* cDNAs present in the new derivatives of pGM955 (pGM992 – pGM996; Table 2.2) were amplified by PCR using Fw_20220510 and Rv_WT_20220510 primers (Table 2.1) and a low error-rate amplifying enzyme to isolate the *ABCA3* variant transgenes. For the Y1515X variant, which corresponds to an early stop codon, Rv_Y1515X_20220510 reverse primer

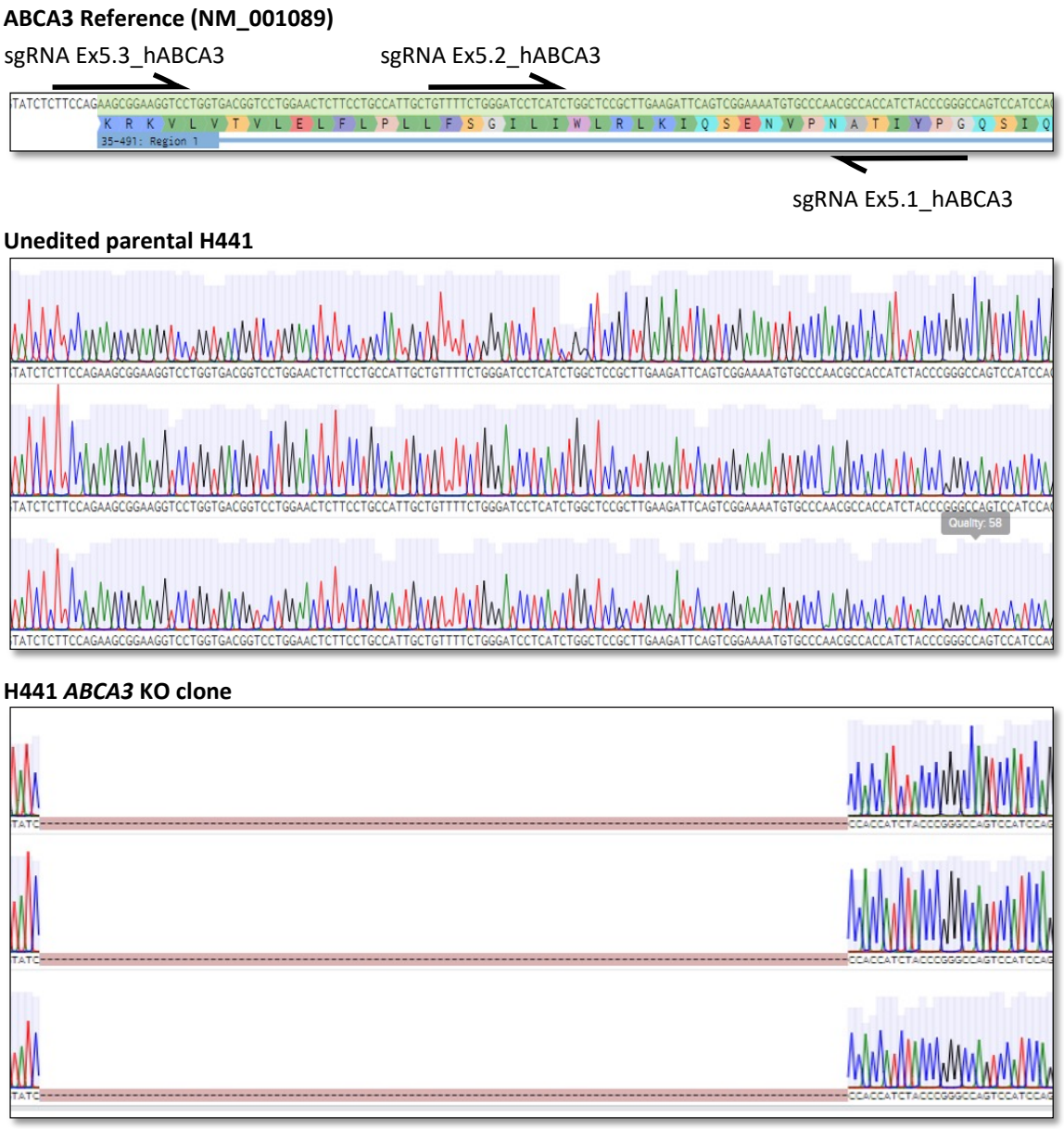


Figure 4.9 Evaluation of *ABCA3* KO in a H441 cell clone using Sanger sequencing (Hamid Dolatshad)

Sanger sequencing traces of gDNA obtained from unedited, parental H441 cells and a H441 *ABCA3* KO clone isolated from a pool of edited cells (Section 4.2.3.3, Figure 4.8). A deletion in *ABCA3* exon 5 causing a frameshift and premature stop codon was confirmed using three different sequencing primers. The binding locations of sgRNAs used are annotated on the reference sequence (NM_001089) with *ABCA3* exon 5 highlighted in green.

(Table 2.1) was used to prevent PCR amplification of *ABCA3* beyond position 1515. During amplification, the 5' NheI restriction sites were maintained whilst 3' stop codons were replaced with an XbaI restriction site (Figure 4.10). The removal of stop codons was necessary to allow expression of a downstream F2A linked puromycin resistance gene in the final plasmid backbone (Figure 4.10). Following PCR amplification of the *ABCA3* variants (including a WT *ABCA3* transgene), all PCR products and the backbone plasmid pGM943 (HIV-CMV-coACE2-F2A-PuroR; Table 2.2) were digested with both NheI and XbaI (Section 2.2.1) before ligation with T4 ligase (Section 2.2.6). In each final plasmid (pGM1122-1127; Table 2.2), the sequence of the regions of *ABCA3* cDNA which underwent SDM, in addition to those encompassing the NheI and XbaI restriction sites, were confirmed by Sanger sequencing (Section 2.2.15; GM20-1351, GM20-1361; Table 2.1).

4.2.4.2 Development and characterisation of ABCA3 variant cell lines

For each ABCA3 variant, rHIV vector pseudotyped with VSV-G was produced at SS (25ml culture volume) in suspension-adapted HEK293T-17 cells (Section 2.4.2.1) and purified by centrifugation (Section 2.4.3.1) to generate 500µl of purified vGM449 to vGM454 virus preparations (Table 2.3). H441 cells were seeded in 6-well plates ($3.5x10^5$ cells/well) and transduced after 24 hours with 100µl of non-titrated virus (Section 2.4.3.1). Approximately 4 weeks after puromycin selection was initiated (Section 2.3.4) the rate of puromycin-induced cell death appeared to stabilise, and the resulting ABCA3 variant bulk cell lines were characterised. PCR amplification of gDNA extracted from transduced cells indicated the successful insertion of *ABCA3* variants into the genome of each bulk line. The amplification primers targeted the CMV- promoter (GM19-1241; Table 2.1), and a region present in all *ABCA3* variants used (qRT-PCR_R_Ex5; Table 2.1) to prevent signal amplification of endogenous H441 *ABCA3*. As seen in Figure 4.11, an approximately 500bp (predicted 488bp) amplicon was detected in all transduced cell lysates, which was not present in parental, non-transduced H441 gDNA samples. Sanger sequencing of gel extracted amplicons confirmed the presence of the *ABCA3* variants in H441 cells (data not shown).

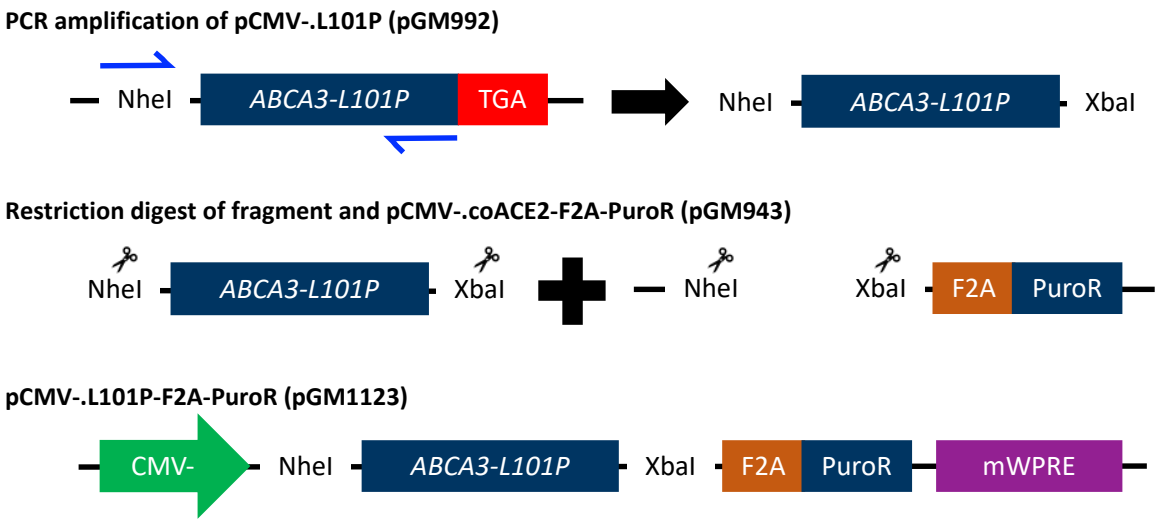


Figure 4.10 Generation of recombinant lentiviral vector genomes encoding pathogenic ABCA3 variants

Recombinant LV genome plasmids encoding ABCA3 variants and a puromycin selection cassette were constructed in two stages. First, SDM of pCMV-.ABCA3 (pGM955) introduced mutations (L101P, E292V, N569D, T1114M, Y1515X) into the cDNA (pGM992 – 996) before PCR amplification was used to isolate the transgenes, remove the stop codon and introduce an XbaI restriction site. Subsequently, the isolated transgene and lentiviral backbone plasmid pCMV-.coACE2-F2A-PuroR (pGM943) were ligated after double restriction digestion with NheI (5') and XbaI (3') to generate the final rLV genome plasmids (pGM1122– 1126; Table 2.2) for rHIV production.

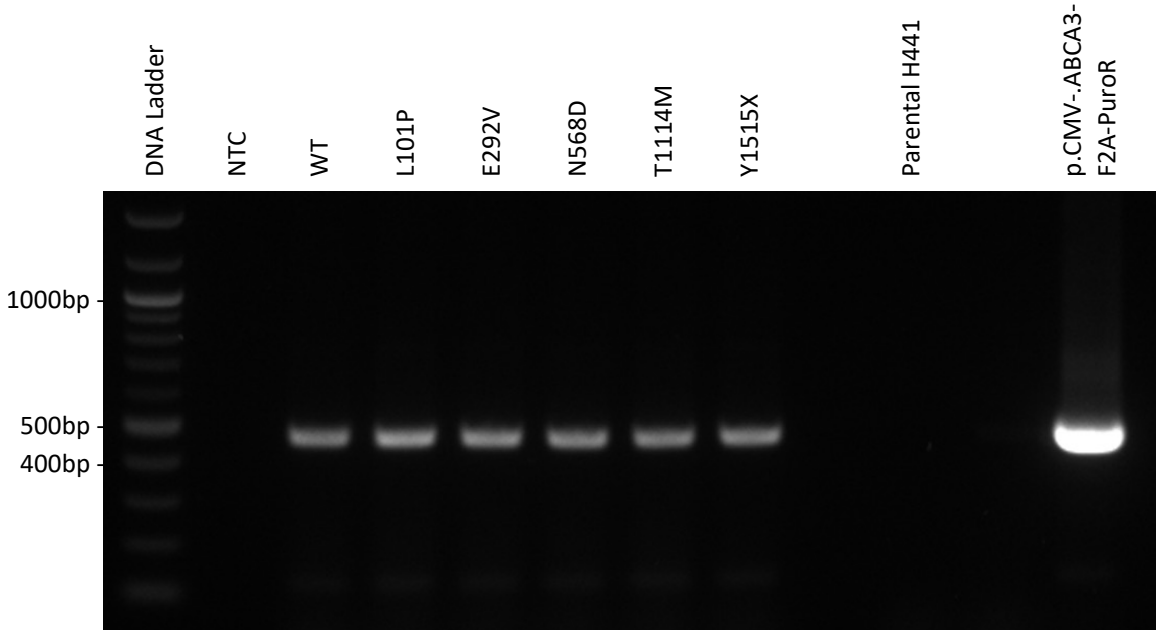


Figure 4.11 PCR amplification to confirm insertion of *ABCA3* variants into the H441 genome

H441 cells were transduced with rHIV.VSV-G.CMV-.ABCA3*-F2A-PuroR (vGM449 – 454; Table 2.3) and gDNA was extracted for PCR amplification with primers targeting CMV- and *ABCA3* (Table 2.1). Reactions containing primers and polymerase but without DNA acted as a no template control (NTC) and gDNA from parental, non-transduced H441 cells or H441 cells transfected with p.CMV-.ABCA3-F2A-PuroR (pGM1122; Table 2.2) were included as negative and positive controls. Amplicons were visualised by agarose gel electrophoresis with predicted 488bp amplicon. *: ABCA3 variant.

To assess how many copies of the rHIV expressing the *ABCA3* variants had integrated into the H441 cell genome, ddPCR with copy number analysis was completed (Section 2.5.6; Table 4.4). The number of copies of WPRE (mWPRE_dd_01_SH; Table 2.1), present in the lentiviral plasmid backbone, relative to that of endogenous *CFTR* (hCFTRint1; Table 2.1) in the reference genome, was calculated for triplicate samples of each new *ABCA3* variant-expressing cell line, assuming a diploid copy number of *CFTR* in H441 cells. As indicated in Table 4.4, fewer than 1% of parental, non-transduced H441 cells were positive for WPRE compared to an average of 25% across all newly generated *ABCA3* variant cell lines. The L101P variant cell line had the lowest percentage of WPRE-positive cells (19.7%) whilst the Y1515X variant had the highest (29.1%).

Table 4.4 Copy number analysis of H441 ABCA3 variant cell lines

Vector copy number (VCN)

H441 Cell line	VCN per cell	Percentage WPRE+ cells
Naïve, non-transduced	0.008	0.8%
ABCA3-WT	0.261	26.1%
ABCA3-L101P	0.197	19.7%
ABCA3-E292V	0.237	23.7%
ABCA3-N568D	0.245	24.5%
ABCA3-T1114M	0.242	24.2%
ABCA3-Y1515X	0.291	29.1%

4.2.4.3 Impact of mutations on ABCA3 expression and function

As shown in Chapter 3 (Figure 3.3B), the predominant ABCA3 protein product observed by immunoblotting is approximately 190kDa with an additional 170kDa product due to proteolytic cleavage after Lys174 [42]. To investigate if the presence of the selected mutations in *ABCA3* altered the protein expression profile, western blot analysis was performed on lysates from the bulk *ABCA3* variant cell lines (Section 4.2.4.2). As previously observed (Section 4.4.2, Figure 4.4A), endogenous ABCA3 was not detectable in protein lysates from parental H441 cells grown in submerged culture, but when transiently transfected with a plasmid expressing ABCA3 (pGM955; pCMV-.ABCA3; Table 2.2) the characteristic 190kDa and 170kDa bands for ABCA3 were observed

(Figure 4.12). Lysates from the ABCA3 WT variant H441 bulk cell line showed a similar 190kDa/170kDa banding pattern to plasmid transfected cells. The 190kDa and 170kDa bands were also present at similar ratios and intensities in the H441 bulk cell lines transduced with ABCA3 variants E292V, N568D and T1114M, although additional smaller bands (e.g. at approximately 150kDa and 130kDa) were also detected. Interestingly, despite similar detection levels of the housekeeping control, β-tubulin, the ABCA3 L101P variant showed a reduced intensity of the 190kDa product as well as diminished overall detection of ABCA3. It is of note, that of the bulk cell lines created for the various ABCA3 variants, the L101P bulk cell line had the lowest VCN (Table 4.4), which may have affected ABCA3 expression levels. In lysates obtained from the Y1515X premature stop codon variant, the largest ABCA3 protein product detected was approximately 170kDa with a secondary band present at 150kDa (Figure 4.12). It is also interesting to note that expression level of the Y151X variant appeared lower than the other missense mutations, which might be a result of NMD or protein instability affecting the expression profile of this variant.

Subsequently, the impact of the pathogenic mutations on potential ABCA3 transporter function was considered. In addition to transporting surfactant lipids, the ABCA3 transporter protein can sequester and detoxify cytotoxic agents such as doxorubicin [230, 231]. This function of ABCA3 was assessed using an MTT viability assay after the ABCA3 variant bulk cell lines were exposed to 10µM doxorubicin (Section 2.3.5). As seen in Figure 4.13, when treated with doxorubicin, H441 cells stably expressing ABCA3 variants had a significantly reduced viability compared with cells expressing WT ABCA3 (one-way ANOVA, [F (5, 12)=4.772, p=0.0124]). These findings show that the transporter function of ABCA3 is negatively impacted by each of the *ABCA3* mutations and that the doxorubicin assay described above can be used to indirectly monitor this function.

4.2.5 Expression of functional ABCA3 from rSIV.F/HN.hCEF.ABCA3 in cell culture

To demonstrate expression of functional ABCA3 from the rLV vector generated in Chapter 3, three ABCA3 variant cell line pools (L101P, E292V and Y1515X) were each transduced with

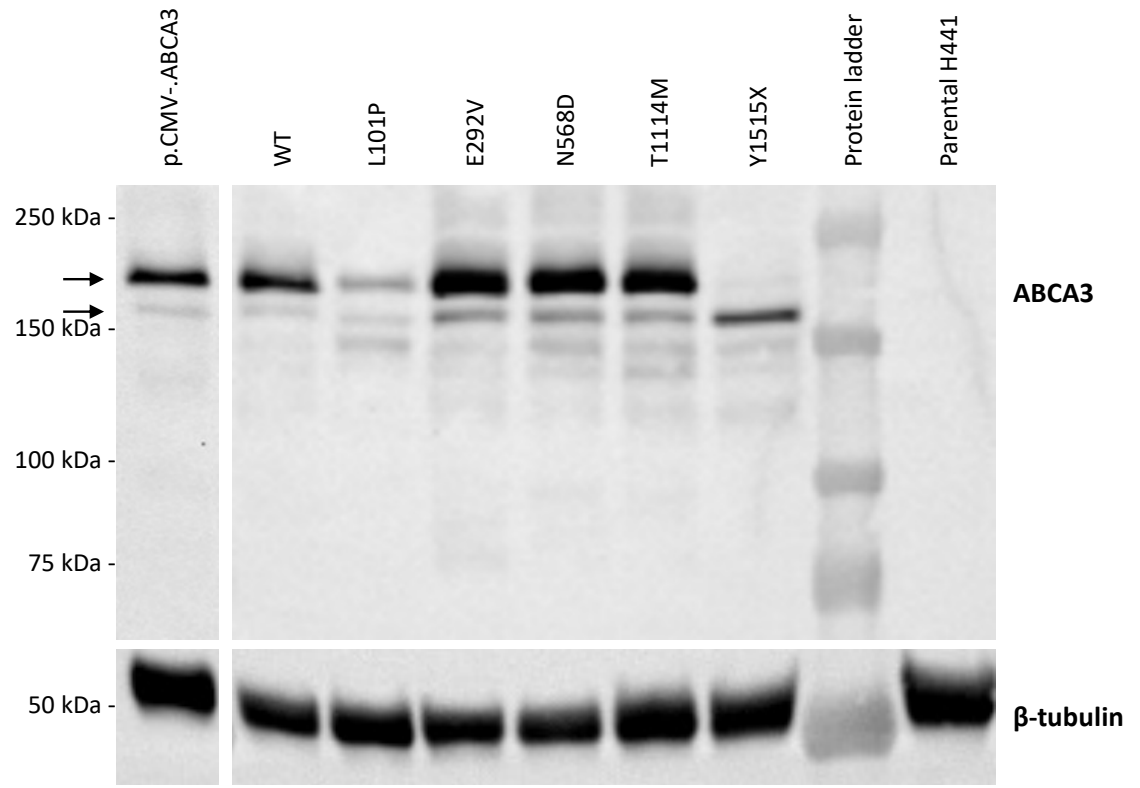


Figure 4.12 Immunoblotting of H441 cells transduced with rHIV virus expressing ABCA3 variants

H441 cells ($3.5x10^5$ cells/well) were transduced with 100µl crude rHIV expressing wild-type ABCA3 (WT) or pathogenic variants (L101P, E292V, N568D, T1114M, Y1515X). Following protein extraction, 20µg of total protein lysate was denatured under reducing conditions for SDS-PAGE (8% gel) and western blotting completed using ABCA3 primary antibody PA5-52478 (Section 2.5.6). Non-transduced H441 cells (Parental H441) were used as a control for background staining. Protein lysate from HEK293T cells transfected with plasmid expressing ABCA3 (p.CMV-.ABCA3; pGM955; Table 2.2) was included as a positive control. β-tubulin was used as a protein loading control. The size of the closest protein markers to the target proteins are labelled.

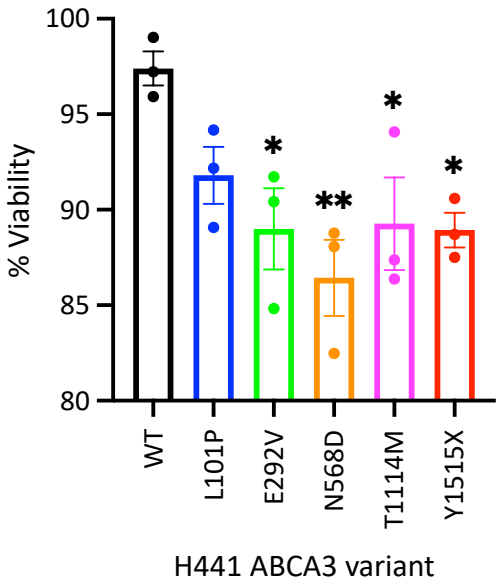


Figure 4.13 Cell viability of H441 ABCA3 variant cell lines in response to doxorubicin

The H441 ABCA3 variant cell lines were seeded in triplicate in 96-well plates ($1.4x10^4$ cells/well) and, after 72 hours, 10µM doxorubicin was added to the culture medium for 3 hours before replacing with complete media minus doxorubicin (Section 2.3.5). 24 hours after doxorubicin exposure, cell viability was assessed by MTT assay (Section 2.3.5) and the percentage viability of doxorubicin-treated cells calculated compared with untreated cells for each ABCA3 variant cell line. Significant differences between the groups were determined by one-way ANOVA [F (5, 12)=4.772, p=0.0124] with Dunnet's multiple comparisons post hoc analysis comparing to H441 ABCA3 WT variant cell line. Calculated p values of <0.05 (*) and <0.01 (**) were deemed significant. Data are shown as individual values and mean ± SEM.

rSIV.F/HN.hCEF.ABCA3 (vGM336; Table 2.3, MOI 30). After passaging twice, the cells were exposed to doxorubicin and cell viability (Section 2.3.5) measured compared with non-transduced controls. As shown in Figure 4.14, the three control (non-transduced) ABCA3 variant bulk cell lines showed significantly reduced cell viability on exposure to doxorubicin compared with the WT ABCA3 bulk cell line ($p<0.05$). When the ABCA3 variant cells that had been transduced with vector expressing ABCA3 (rSIV.F/HN.hCEF.ABCA3; vGM336) were exposed to doxorubicin cell viability was not significantly different from the non-transduced WT ABCA3 bulk cell line ($p>0.05$; Figure 4.14). This suggests that ABCA3 expressed from rSIV.F/HN.hCEF.ABCA3 was able to detoxify the doxorubicin and restore cell viability – confirmation that this vector expresses functional ABCA3 protein.

4.3 Discussion

The studies described in this chapter were used to explore possible cell models of ABCA3-mediated surfactant dysfunction that could be used to provide 'proof-of-principle' for rSIV.F/HN.hCEF.ABCA3 expressing functional ABCA3.

4.3.1 Maintaining an ATII cell phenotype in primary culture

Primary alveolar pneumocytes were initially considered, but as described in Section 4.2.1, these were deemed unsuitable for generating an ABCA3 disease model due to the gradual loss of ATII cell identity over time in culture (Figure 4.2). This is in line with published reports where 2D culturing of human ATII pneumocytes causes transdifferentiate into ATI cells [100]. In one report, sustained expression of ATII cells markers in culture was achieved by co-culture with fibroblast feeder cells plus inclusion of a Rho kinase (ROCK) inhibitor in the media [215]. Fibroblasts are a known source of keratinocyte growth factor (KGF), reported to help restrict ATII cell transdifferentiation, whilst ROCK inhibitors promote cell proliferation. Inclusion of these factors in the culture medium may have limited transdifferentiation of the ATII cell population but would not remove ATI cells already present in the primary pneumocyte population. Fluorescence-activated cell sorting (FACS) could

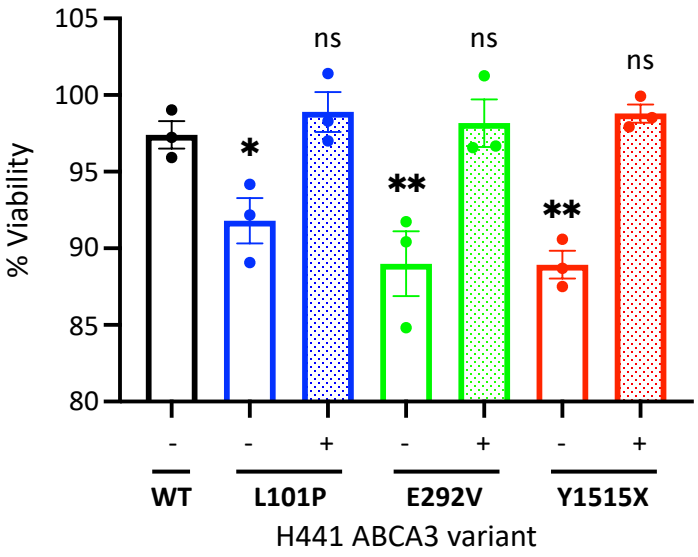


Figure 4.14 Cell viability of H441 ABCA3 variant cell lines exposed to cytotoxic doxorubicin can be rescued by transduction with rSIV.F/HN.hCEF.ABCA3

H441 ABCA3 variant cell lines (as annotated) were transduced with rSIV.F/HN.hCEF.ABCA3 (vGM336; Table 2.3; +) at MOI 30, 9 days (2 passages) before incubation with 10µM doxorubicin (three hours). At 24 hours after doxorubicin incubation, cell viability was assessed by MTT assay (Section 2.3.5) and the percentage viability of doxorubicin-treated cells calculated compared with no doxorubicin incubation. Data are shown as individual values and mean ± SEM with n=3 per condition. For H441 ABCA3 variant cells significant differences between the values were determined by one-way ANOVA [F (6, 14)=11.69, $p<0.0001$] with Dunnet's multiple comparisons post hoc analysis comparing to non-transduced (-) H441 ABCA3 WT variant cell line. A calculated p value of <0.05 was deemed significant ($p<0.05$, *; $p<0.01$, **). A calculated p value >0.05 was deemed non-significant (ns).

have been used to isolate ATII cells, however the ATII cell specific marker SP-C is expressed intracellularly making it unsuitable for FACS of live cells. HTII-280 has been used as a surrogate, extracellularly expressed marker of ATII cells [232] although there is debate in the field about how stable HTII-280 expression is in *ex vivo* samples (like alveolar pneumocytes) [233]. Using reprogrammed iPSCs (Section 1.4.1.1) [108] should resolve some of the issues encountered with primary alveolar pneumocytes and this technology is now being explored in our laboratory. An additional advantage of iPSC-derived cells is their amenability to gene editing [234] which generally necessitates cells in 2D culture; iPSCs are cultured in 2D prior to redifferentiation into iPSC-derived ATII cells [108].

4.3.2 Selection of H441 cells as an ATII cell model

The H441 lung cell line was investigated as a potential ABCA3 disease model due to the ability to form SALI cultures and the surfactant related gene *SFTPB* to be gene edited (Section 1.4.1.2) [103, 235]. In addition, the resultant SP-B KO SALI cultures displayed physiologically relevant disease phenotypes, such as a reduction in barrier integrity (TEER), which could be corrected by transduction with rLV expressing the SP-B protein [103]. Importantly, although both H441 and A549 cell lines expressed endogenous *ABCA3* (Figure 4.3), in the more physiologically relevant ALI cell culture models (Section 1.4.1.2) ABCA3 protein could only be detected, by immunoblotting, in H441 cells grown as SALI cultures (Figure 4.4A). Therefore, there was a strong argument to utilise the H441 SALI model as it exhibits physiologically relevant ATII cell characteristics.

One disadvantage of using H441 cells was the lack of mature LB organelles visible in cells grown in submerged or SALI culture (Figures 4.5 and 4.6). It was hoped that the shift in H441 transcriptome with air-lift [235] would facilitate the emergence of visible LBs. The TEM images of the SALI cultures showed unidentified electron-lucent regions (Figure 4.5E) which increased in abundance the longer H441 cells were maintained in SALI cultures (Figure 4.6). The electron-lucent regions are thought to be associated with lipid-rich structures. Aldehyde-based fixatives (e.g. formaldehyde,

glutaraldehyde) commonly used for TEM react with amino groups of proteins and amino-containing phospholipids (e.g. phosphatidylserine), stabilising the cellular structure by cross-linking [225]. Unfortunately, many lipids, including the major lipid component of surfactant phosphatidylcholine, lack amino groups and are thus poorly fixed by aldehyde-based fixatives and susceptible to extraction during subsequent TEM sample preparation steps. It is possible that further optimisation of the protocol for sample preparation, to promote fixation of lipid structures, could increase detection of LB and MVBs in the H441 cultures. This seemed to be a specific limitation of the H441 cells, however, as LB organelles were detected with ease in A549 cells (Figure 4.5A, B) and alveolar pneumocytes (Figure 4.1) with a routine TEM protocol. Alternatively, if the LB protein marker lysosomal associated membrane protein 3 (LAMP3) was expressed in H441 cells [235] immunostaining, or immunogold labelling prior to TEM, might have provided evidence of LBs in these cells.

In addition to data shown here (Figure 4.4A), where ABCA3 was only detectable by western blot in SALI cultures, Raman micro-spectroscopy by Swain *et al.* [216] suggests the A549 cell phenotype more closely aligns with ATI cells rather than the desired ATII cells. Furthermore, Munis *et al.* [103] and Ren *et al*. [236] have demonstrated A549 cells after air-lift do not have functional barrier properties, as measured be TEER and transepithelial potential difference (TEPD) – epithelial cell characteristics which only develop in SALI culture. On balance, H441 cells were the preferred option for generating a model of ABCA3-mediated surfactant dysfunction.

4.3.3 Gene editing in H441 cells

Attempts to edit the *ABCA3* locus in H441 cells were initially unsuccessful (Section 4.2.3.2), however transitioning from plasmid delivery of SpCas9 (Section 4.2.3.2) to direct delivery of SpCas9 protein in RNP complexes (Section 4.2.3.3) resulted in low-level editing (Figure 4.8). Genome editing after plasmid delivery of CRISPR-SpCas9 machinery is partially dependent on promoter activity in recipient cells. In the PX459 V2.0 plasmid backbone (Table 2.2) SpCas9 is expressed under

the constitutive hybrid (CBh) promoter and there is little information about its activity in H441 cells, or lung cells more generally; but as a strong, ubiquitous promoter [237] it is unlikely that low promoter activity was a problem. Indeed, low levels of editing were observed in preliminary characterisation of gRNAs using the PX459 plasmid backbone, and Munis *et al.* [103] successfully used PX459-derived plasmids to generate a H441 SP-B KO cell line. The two main differences in the approach described in Section 4.2.3.2 and the successful attempt by Munis *et al.* [103] are gRNA placement and clonal isolation technique. The *SFTPB*-targeting gRNAs span the first four exons over a region of 1.6kb. In contrast, although three of the *ABCA3* gRNAs target the first four coding exons (exon 4 to 8) the fourth was designed to anneal upstream of the assumed transcription start site in exon 1. Even excluding the uncharacterised UpEx1 *Abca3* gRNA (Table 4.1), the remaining gRNAs span 6.8kb, over 5kb larger than for *SFTPB.* Despite concerns that excision of the large *ABCA3* fragment would be less efficient than the much smaller *SFTPB* fragment, this approach was initially selected because deletion of whole exons was assumed to increase the likelihood of achieving complete gene ablation. Single gRNA-mediated gene disruption relies on the introduction of frameshift-inducing indels to break the reading frame downstream of the cut site and prevent protein expression. However, depending on the location of the indel introduced, expression of functional, C-terminally truncated proteins is possible [238] – although in many cases these truncated proteins will be unstable and rapidly degraded.

The approach used in Section 4.2.3.3 aimed to retain the potential benefits of a multi-guide approach whilst reducing the size of the target DNA fragment to a 100bp region within exon 5 of *ABCA3*. Figure 4.8 and subsequent sequencing confirmed gene editing via fragment deletion had occurred, however clonal isolation was not attempted due to time constraints and the lack of a selection marker to isolate edited cells. Combing RNP delivery of CRISPR-SpCas9 with gene trap technology [238] would have allowed for inclusion of a selection marker, whilst retaining cell line independent SpCas9 expression, but also would have added complexity. Importantly, several H441 ABCA3 KO clones have since been isolated from the heterogeneous cell pool obtained in Section

4.2.3.3. This relied on automated seeding of single cells into 96-well plates after FACS for cell viability and subsequent sequencing analysis of the *ABCA3* locus in colony-forming cells (Hamid Dolatshad, Personal communication).

4.3.4 Protein processing of pathogenic ABCA3 variants

When H441 cells were transduced to express pathogenic variants of ABCA3 (Section 4.2.4), western blotting (Figure 4.12) showed that E292V, N568D and T1114M variants expressed characteristic 190kDa/170kDa forms of ABCA3 at similar ratios and intensities to that observed with the WT *ABCA3* sequence. This observation is in accordance with Wambach *et al.* [239] and Kinting *et al.* [157], suggesting that type II *ABCA3* mutations do not disrupt intracellular ABCA3 protein trafficking or processing (Section 1.3.2.3) but are trafficked to post-Golgi compartments correctly where they are cleaved by cathepsin L [42], resulting in the 170kDa form similar to WT ABCA3 protein. The additional smaller bands detected for these three ABCA3 variants (Figure 4.12) are currently unexplained but may be non-specific or ABCA3 proteolytic breakdown products. In contrast, the banding pattern observed for the type I L101P *ABCA3* mutation differs from published western blots [109, 239] where others have stated that only the non-cleaved 190kDa protein is observed. Here (Figure 4.12), both non-cleaved and cleaved ABCA3 was detected, in addition to a denser band at approximately 150kDa, although the overall ease of detection was diminished. Typically, type I *ABCA3* mutations (including L101P; Section 1.3.2.3) are retained in the ER and subsequently degraded, although H441 cells may have an altered ER quality control system (ERQC) [240] for ABCA3 where variants could be trafficked to post-Golgi compartments for processing. Western blotting of the Y1515X ABCA3 variant also suggest the ERQC system in H441 cells may be unusual; similar to the L101P ABCA3 mutation, the Y1515X type I mutation should also be retained in the ER (Section 1.3.2.3) [90] but the immunoblot showed a predominant, albeit reduced in intensity compared to most other ABCA3 variants, band at approximately 170kDa (Figure 4.12). The 20kDa reduction from 190kDa (WT ABCA3) to 170kDa (Y1515X variant) as the largest ABCA3 protein

product detected can be explained by the premature stop codon which truncates the ABCA3 C-terminal by 189 amino acids. Comparing protein processing (western blot) and intracellular localisation (ICC) of ABCA3 L101P-expressing H441 cells with A549 cells transduced to express the L101P variant may help elucidate if the differences in protein processing observed in Figure 4.12 are a H441 cell-specific effect.

4.3.5 Measuring ABCA3 transporter function

In addition to the physiological role of ABCA3 in transporting surfactant lipids, it is frequently upregulated in cancers where cytotoxic drugs are sequestered and detoxified as part of cellular multi-drug resistance [231]. An assay to assess detoxification of the cytotoxic drug doxorubicin (Section 2.3.5) [241] was used to confirm functional activity of ABCA3 expressed from rSIV.F/HN.hCEF.ABCA3 (Section 4.2.4.3) and to investigate the effect of *ABCA3* mutations on protein function. As anticipated, expression of pathogenic ABCA3 variants was associated with reduced cell viability after exposure to doxorubicin compared with WT ABCA3 (Figure 4.13). Furthermore, transduction of H441 cells expressing variant ABCA3 protein with rSIV.F/HN.hCEF.ABCA3 (vGM336, Table 2.3) improved cell viability following doxorubicin exposure (Figure 4.14) – suggesting the ABCA3 protein expressed from the vector is functional in this assay.

The use of heterogenous cell pools instead of clonal cell lines may have contributed to variation observed between replicates in the doxorubicin assay (Figure 4.13, 4.14). The effect of any potential variability introduced would have been diminished if the working range of the doxorubicin assay was larger. In cells exposed to doxorubicin, the biggest reduction recorded in cell viability was approximately 15% whilst the variability between sample replicates was generally around 5%. Without an *ABCA3* KO cell line (not available at the time) the full working range of the assay could not be determined, so it is unknown how much of the variation in viability can be attributed to 'noise' within the assay and whether a 15% reduction represents some retention or complete loss of ABCA3 function. Additionally, only approximately 25% of cells in the bulk cell lines

contained the *ABCA3* variant transgenes (Table 4.4); isolating clonal lines may have made differences in response to doxorubicin between WT ABCA3 and ABCA3 variants more apparent, thus increasing the working range of the assay. An alternative, more direct, assay of ABCA3 function could be to track intracellular localisation of fluorescently tagged lipid molecules [157, 242], although fluorescence microscopy was considered less quantitative.

4.3.6 Conclusion

In summary, gene editing was used to ablate the *ABCA3* locus in H441 cells and an *ABCA3* KO clone has been obtained (Figure 4.9). An *in vitro* doxorubicin detoxification assay was established and used to show that *ABCA3* pathogenic mutations led to compromised cell viability. This assay was also used to confirm that rSIV.F/HN.hCEF.ABCA3 (vGM336; Table 2.3) could express functional ABCA3 protein sufficient to restore cell viability to levels seen in cells expressing WT ABCA3. These results supported progression to testing of rSIV.F/HN.hCEF.ABCA3 *in vivo*, described in Chapter 5.

Chapter 5 *In vivo* lentiviral gene transfer for ABCA3-mediated surfactant dysfunction

5.1 Introduction

In previous chapters, the generation of rLV encoding ABCA3 was described along with confirmation of expression of functional ABCA3 in human cell models of ABCA3-mediated surfactant dysfunction. To further investigate the therapeutic potential of these vectors they were tested in mice, which are frequently used to model genetic diseases and support drug development. Intranasal delivery of rLV to mice is routinely used to investigate vector delivery to the lung airways and parenchyma [155]. This route of administration delivers into the luminal space of the murine lung such that the vector is directly exposed to the polarised lung epithelium similar to lung delivery in patients. The shared homology of murine and human genes also makes mice excellent disease models.

Constitutive KO of *Abca3* in mice is neonatal lethal (Section 1.4.2.1.1) [92], therefore a conditional KO model (Section 1.4.2.1.2) was used to evaluate the protective effect of gene therapies given after birth. In this conditional model (referred to as *Abca3*-iKO; Section 5.2.2) tamoxifen activates the Cre-loxP system (Figure 1.8) resulting in KO of *Abca3* (and expression of EYFP). In the absence of tamoxifen, Cre recombinase is expressed under the endogenous *SFTPC* promoter but held in the cytoplasm due to fusion with a mutated oestrogen receptor ligand binding domain ($CreER^{T2}$; Figure 1.8B, Table 2.7). Tamoxifen is an oestrogen receptor antagonist and, as such, binds to ER^{T2} inducing the translocation of Cre recombinase to the nucleus, where it catalyses recombination between loxP sites located in intron 3 and intron 7 of *Abca3* (Figure 1.8C) and either side of a transcriptional stop sequence within an *EYFP* expression cassette in the *Rosa26* locus.

The aim of the studies described in this Chapter was to test rSIV.F/HN.hCEF.ABCA3 in a mouse disease model of ABCA3-mediated surfactant dysfunction.

5.2 Results

5.2.1 Evaluation of rSIV.F/HN.hCEF.ABCA3 in BALB/c mice

Dosing with rLV was first undertaken in WT BALB/c mice before progressing to the conditional *Abca3* KO disease model (Section 2.7.3, Section 5.2.2). Recombinant LV expressing ABCA3 (rSIV.F/HN.hCEF.ABCA3) was intranasally administered to BALB/c mice at a total dose of $5.8x10^8$ TTU (vGM336, Wave 321) or $8x10^8$ TTU (vGM438, Wave 372; Table 2.3; Section 2.7.5), or formulation buffer (TSSM) as a negative control. A dose of $8x10^8$ TTU was selected since this dose of rLV expressing SP-B was already known to be effective in a conditional murine model of SP-B surfactant deficiency (Kamran Miah, Personal communication). However, the previously characterised batch of vGM336 (W321; Chapters 3 and 4) was insufficiently concentrated to allow delivery of this dose due to dosing volume restrictions (100µl) for intranasal delivery to mice, leading to a maximum feasible dose of $5.8x10^8$ TTU. Further optimisation of the plasmid ratios during rSIV.F/HN.hCEF.ABCA3 production (Section 3.3.1) resulted in a batch (W372) with an increased viral titre, allowing a total dose of $8x10^8$ TTU (vGM438) to be administered.

The WT BALB/c mice dosed with rLV expressing ABCA3 showed no outward signs of distress, with weights increasing in parallel with control TSSM-dosed animals until all mice were electively culled 14 days after receiving the final vector dose (data not shown). When lungs were harvested (Section 2.7.7), the gross pathology of all organs appeared normal and H&E staining (Section 2.6.2) of lung cryosections (Section 2.6.1) showed no pathological changes or differences between the TSSM-dosed and rLV-dosed mice (Figure 5.1). The alveoli were free from visible inflammatory cell infiltrates and the alveolar walls appeared normal and not thickened [243]. Together, these observations support the notion that administration of a relatively high dose ($8x10^8$ TTU) of rSIV.F/HN.hCEF.ABCA3 should not induce acute toxicity.

To investigate ABCA3 expression in the murine lung after transduction, multiple assays targeting human ABCA3 were performed. The relative expression levels of murine *Abca3*

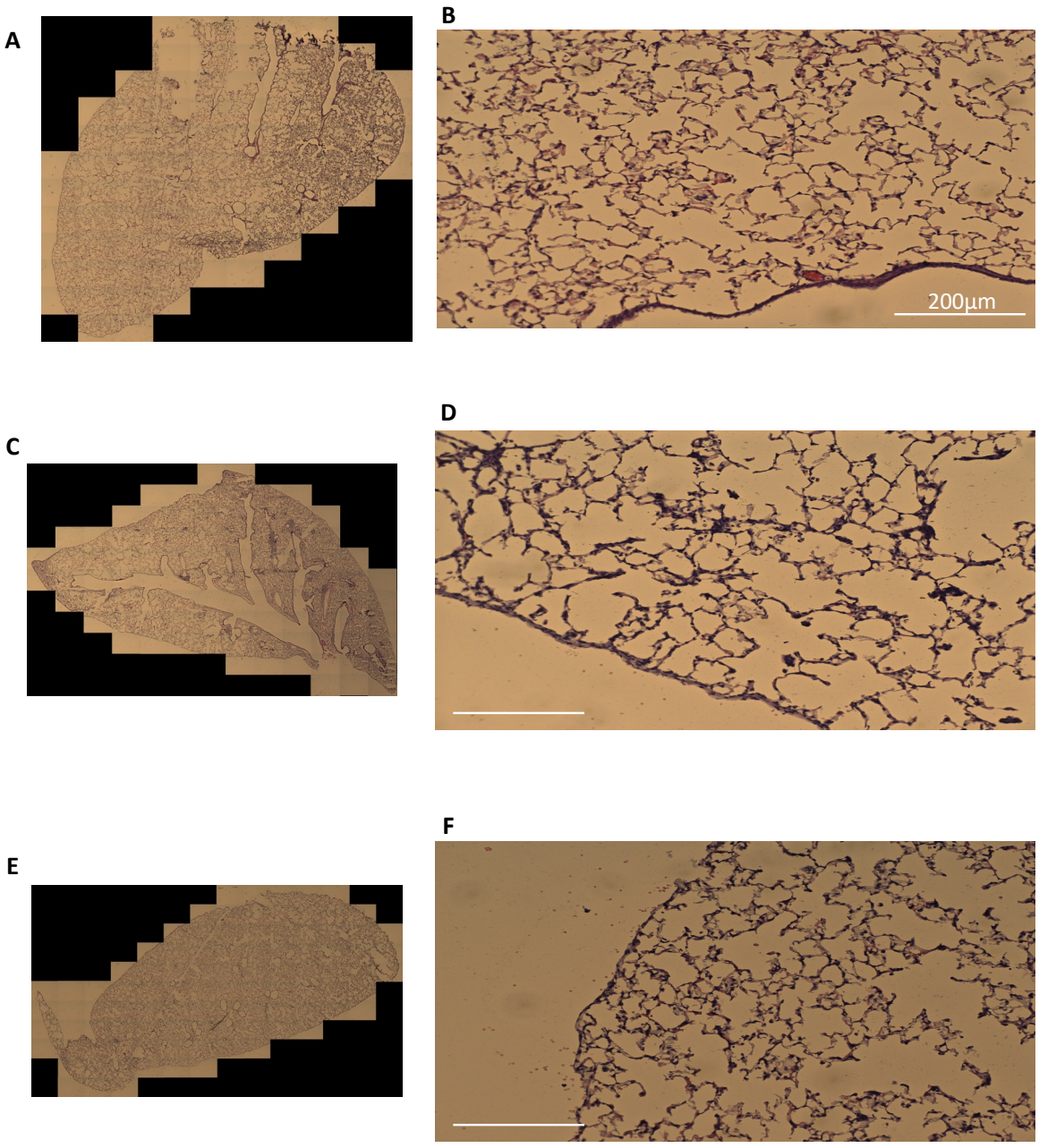


Figure 5.1 H&E staining of WT BALB/c murine lungs transduced with recombinant lentivirus rSIV.F/HN.hCEF.ABCA3 expressing human ABCA3

Lung cryosections were obtained from mice intranasally administered **(A, B)** formulation buffer (TSSM), **(C, D)** 5.8x10^8 TTU rSIV.F/HN.hCEF.ABCA3 (vGM336, W321) or **(E, F)** 8x10^8 TTU rSIV.F/HN.hCEF.ABCA3 (vGM438, W372; Table 2.3) and stained with H&E. No histopathological abnormalities were identified. Representative images from n=3 mice per treatment (n=3 cryosections/mouse); all scale bars represent 200µm.

(mABCA3_qRT_Fex10_202011, mABCA3_qRT_Rex11_202011; Table 2.1) and human *ABCA3* (ABCA3_qRT_F_Ex4, ABCA3_qRT_R_Ex5; Table 2.1) were measured by qRT-PCR (Section 2.5.4; Bact_qPCR_F1_2102, Bact_qPCR_R1_2102; Table 2.1). As expected, endogenous murine *Abca3* expression was not significantly different between mice treated with formulation buffer, or mice treated with rLV expressing human ABCA3 (Figure 5.2A; Kruskal-Wallis [χ^2 (2)=6.507, p=0.0262] with Dunn's post hoc tests; all p>0.05). In contrast, there was a significant increase in human *ABCA3* expression [χ^2 (2)=6.667, p=0.0231] in mice treated with $8x10^8$ TTU rSIV.F/HN.hCEF.ABCA3 (vGM438; p=0.04) compared with control mice dosed with TSSM (Figure 5.2B). Unexpectedly, evidence of vector-induced ABCA3 protein expression was difficult to obtain – western blotting of murine lung homogenates (Section 2.5.6) was uninformative as neither human ABCA3 nor murine Abca3 was detected (data not shown), whilst IHC (Section 2.6.4.2) was unable to distinguish between murine and human protein expression (Figure 5.3). By IHC, positive staining with an anti-ABCA3 antibody (Table 2.6) was observed in left lung cryosections from all mice (Figure 5.3). In the absence of a known ABCA3-negative control lung sample, the anti-ABCA3 antibody was assumed to be ABCA3/Abca3 cross-reactive (the manufacturer reports human reactivity with 82% antigen sequence identity to the mouse ortholog). Importantly, no positive staining was observed in samples incubated with only the primary (data not shown) or secondary antibody (A21428, Table 2.6; Figure 5.3A). Thus, ultimately there appeared to be no clear differences in the staining intensity, or number of ABCA3-positive cells, between control (Figure 5.3B) and vector-treated (Figure 5.3C, 5.3D) mice. This might indicate that the doses of rSIV.F/HN.hCEF.ABCA3 tested were insufficient given the efficacy of the anti-ABCA3/Abca3 antibody utilised.

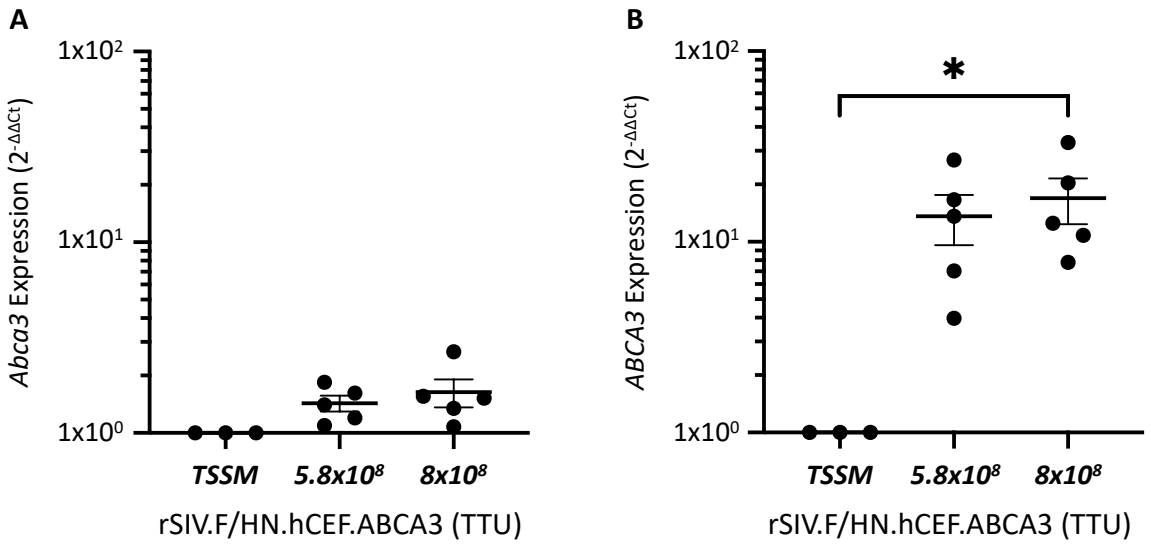


Figure 5.2 Expression of murine *Abca3* and human *ABCA3* mRNA in mouse lungs transduced with rSIV.F/HN.hCEF.ABCA3

RNA was extracted from the right lung lobes of BALB/c mice intranasally administered $5.8x10^8$ TTU (vGM336; n=5) or $8x10^8$ TTU (vGM438; n=5) rSIV.F/HN.hCEF.ABCA3 or formulation buffer (TSSM; n=3). **(A)** *Abca3* and **(B)** *ABCA3* expression was determined via relative qRT-PCR and normalised to murine β-actin expression using $2^{-\Delta\Delta Ct}$ analysis. Data are shown as mean ± SEM with three technical replicates. Significant differences between the values were determined by Kruskal-Wallis with Dunn's post hoc analysis. A calculated p value of <0.05 (*) was deemed significant.

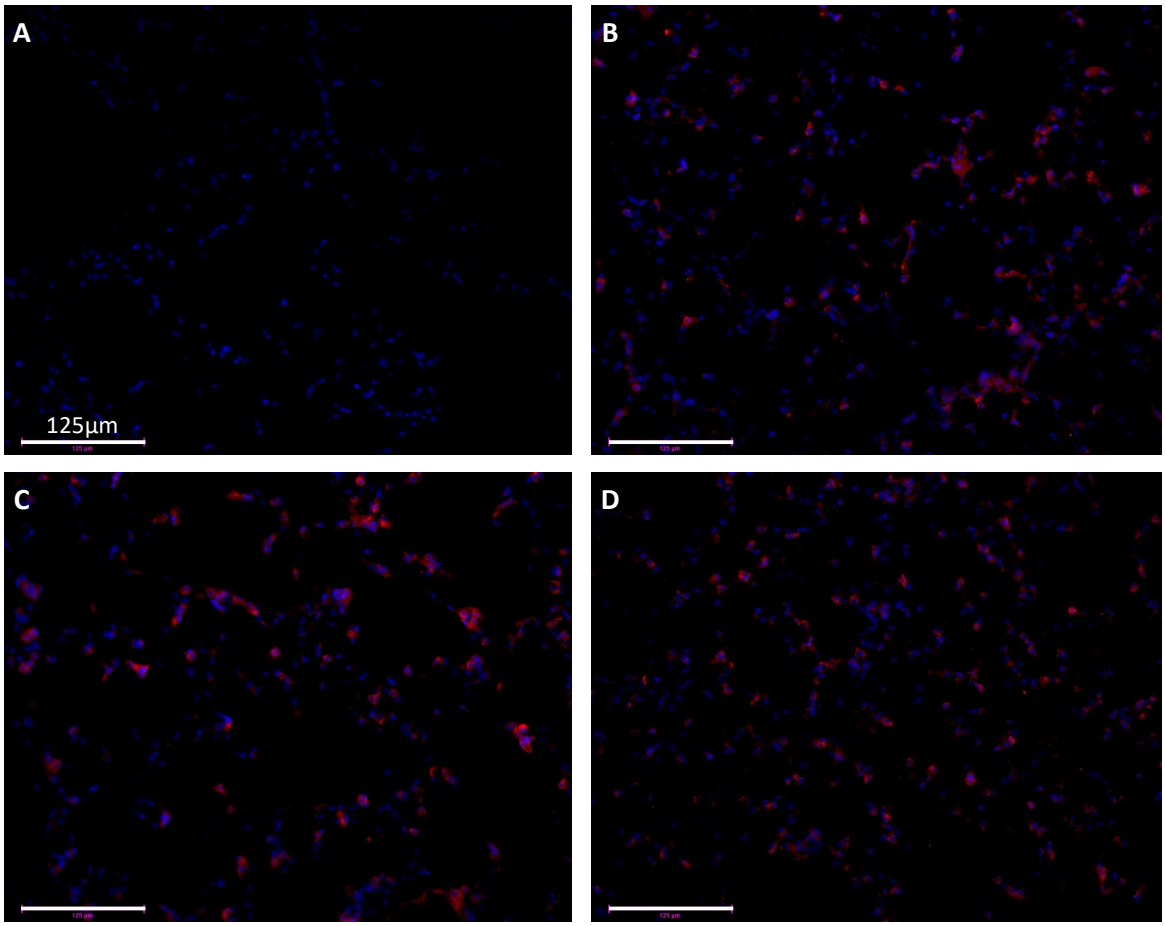


Figure 5.3 Expression of ABCA3 protein in BALB/c mice dosed with rSIV.F/HN.hCEF.ABCA3

Lung cryosections (Section 2.6.1) obtained from mice intranasally administered **(A, B)** formulation buffer (TSSM), **(C)** $5.8x10^8$ TTU rSIV.F/HN.hCEF.ABCA3 (vGM336, W321) or **(D)** $8x10^8$ TTU rSIV.F/HN.hCEF.ABCA3 (vGM438, W372; Table 2.3) were stained for ABCA3 expression (red) using the primary antibody WRAB-70565 and a secondary antibody conjugated to AlexaFluor555 (Table 2.6). Cryosections were counterstained with DAPI (blue) before images (n=3/sample) were taken using an EVOS FL Auto 2 Imaging System with associated RFP and DAPI light cubes (Section 2.6.5.1). **(A)** Secondary antibody only control. Representative images shown. Capture settings were kept consistent between samples. All scale bars indicate 125µm.

5.2.2 Establishment of a conditional *Abca3* knockout mouse model

5.2.2.1 Breeding mice containing Sftpc-CreER^{T2}, floxed Abca3 and EYFP cassettes to generate a colony of Abca3-iKO mice

As discussed in Section 1.4.2, constitutive KO of *Abca3* is neonatal lethal and thus cannot be used to assess gene therapy strategies delivered after birth. Mice homozygous for a floxed *Abca3* allele (*Abca3$^{flox/flox}$*), in addition to mice with an *Abca3$^{flox/WT}$, Sftpc-CreER^{T2+}, EYFP^{+}* genotype (Table 2.7), were acquired from Jeffrey Whitsett (Cincinnati Children's Hospital, USA) and used to establish a homozygous *Abca3*-iKO (*Abca3$^{flox/flox}$, Sftpc-CreER$^{T2+/+}$, EYFP$^{+/+}$*) mouse colony over a period of 10 months (Section 2.7.3). As described in Figure 5.4A, *Abca3$^{flox/flox}$* and *Abca3$^{flox/WT}$, Sftpc-CreER^{T2+}, EYFP^{+}* mice were first crossed to produce pups with an *Abca3$^{flox/flox}$, Sftpc-Cre^{+}, EYFP^{+}* genotype before further breeding of *Abca3$^{flox/flox}$, Sftpc-Cre^{+}, EYFP^{+}* mice until fully homozygous mice could be established as breeders (Section 2.7.3.1).

The genotype of each animal was first assessed after recommended endpoint PCR genotyping assays were established (Section 2.7.3.2). The zygosity of the floxed *Abca3* locus was determined with a single PCR reaction identified to span the 5' loxP site in intron 3 of *Abca3* (mAbca3-F, mAbca3-R; Table 2.1). As seen in Figure 5.4B, amplification of gDNA containing a WT *Abca3* locus, produced a single band of approximately 300bp when visualised by agarose gel electrophoresis, whereas gDNA from confirmed *Abca3$^{flox/flox}$* mice amplified a 400bp product due to the inserted loxP sequences. Accordingly, mice containing one floxed *Abca3* allele and one WT *Abca3* allele were identified by two PCR products of 300bp and 400bp. However, the recommended primers (Whitsett, Cincinnati Children's Hospital) to amplify the *Sftpc-CreER^{T2}* (CRE comm-F (Cre-A), CRE comm-R (Cre-B); Table 2.1) and *EYFP* (EYFP MUT-F, EYFP MUT-R; Table 2.1) transgenic cassettes could not distinguish between heterozygous and homozygous mice, as an amplicon was only produced by the transgenic allele (Figure 5.4B). It was hypothesised that the zygosity of the *Sftpc-CreER^{T2}* cassette would likely affect the efficiency of recombination at the *Abca3* and *EYFP* loci;

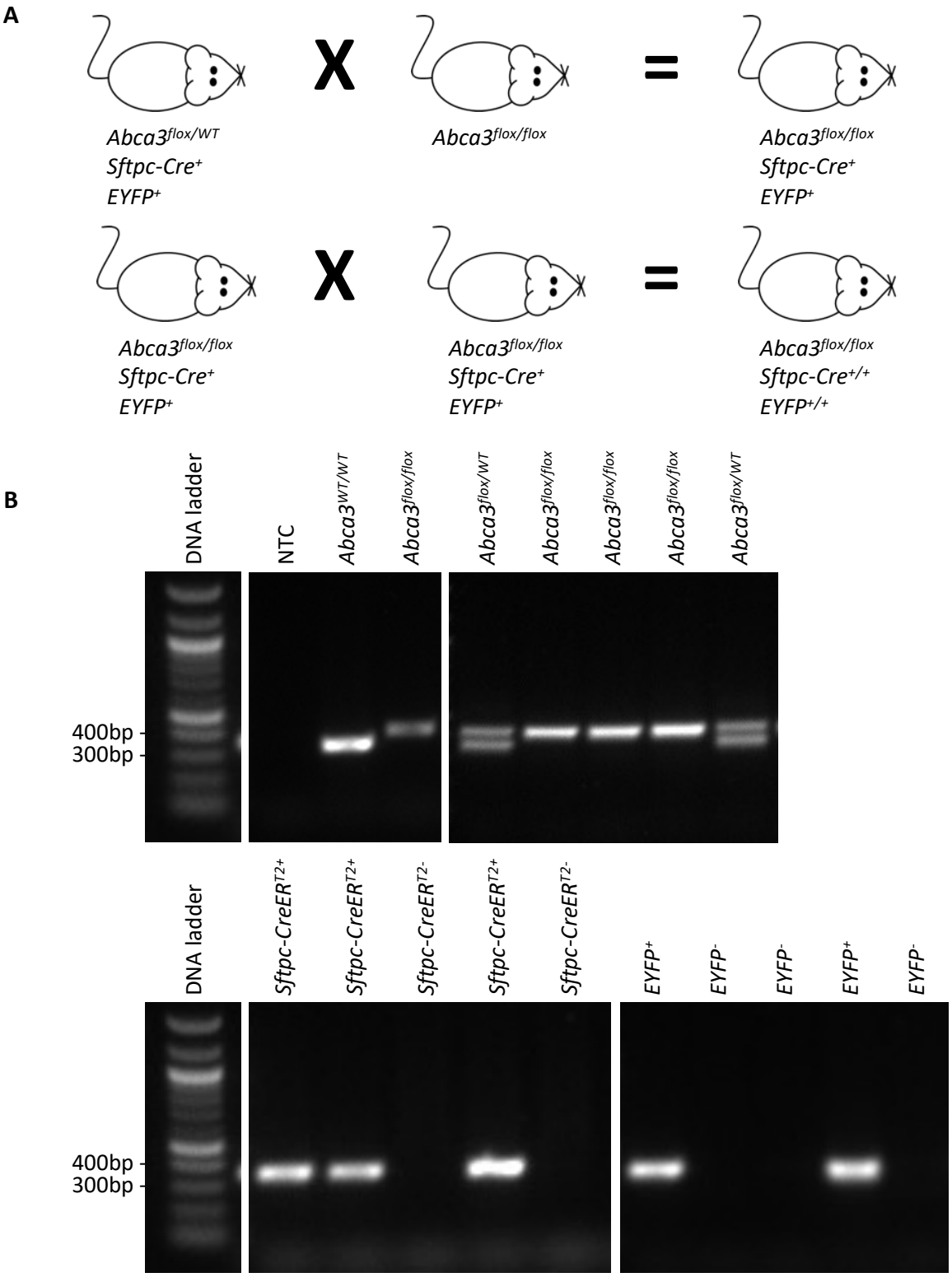


Figure 5.4 Breeding strategy to generate a conditional *Abca3* knockout mouse model

Figure 5.4. Breeding strategy to generate a conditional *Abca3* knockout mouse model

(A) Mice homozygous for a floxed *Abca3* allele (*Abca3*$^{flox/flox}$) were bred with mice heterozygous for floxed *Abca3* (*Abca3*$^{flox/WT}$), positive for a Cre recombinase oestrogen receptor fusion protein (CreER^{T2}) under the transcriptional control of the endogenous *Sftpc* promoter (*Sftpc-CreER*$^{T2+}$) and positive for an EYFP cassette with an upstream floxed stop sequence (*EYFP*$^{+}$). Breeding produced mice homozygous for floxed *Abca3* and positive for *Sftpc-CreER*^{T2} and *EYFP* cassettes. These mice were then bred until mice homozygous for each transgenic cassette were identified. **(B)** The genotype of the mice was assessed by PCR amplification of gDNA. The zygosity of the floxed *Abca3* allele was determined by the size and number of amplicons (300bp, *Abca3*^{WT}; 400bp, *Abca3*^{flox}). *Sftpc-CreER*^{T2} and *EYFP* cassettes were identified as positive (amplicon detected) or negative (no amplicon). NTC: no template control; *Abca3*$^{WT/WT}$ DNA obtained from BALB/c mice.

therefore, additional commercial (Transnetyx) real-time PCR analyses were performed using proprietary primer sequences to independently confirm the result prior to study initiation. This confirmed the ability to distinguish between *Sftpc-CreER*^{T2} and *EYFP* heterozygous and homozygous mice to ensure only mice homozygous for all three transgenic elements were included.

5.2.2.2 Characterisation of Abca3-iKO transgenic cassettes

To confirm amplicon specificity for the inherited genotyping primers, PCR amplicons were gel extracted (Section 2.2.5), cloned into TOPO plasmids (Section 2.2.7; Table 2.2, pGM1326 – 1328) and analysed by Sanger sequencing (Section 2.2.15; Table 2.1: M13F, M13R). Comparing the 400bp *Abca3* amplicon (Section 5.2.2.1) with an *Abca3* reference transcript (ENSMUST00000039013.15, Ensembl) revealed a 60bp insertion containing the canonical 34bp loxP sequence (Figure 5.5A) [244]. The remaining sequence aligned with intron 3 of *Abca3* indicating the recommended genotyping primers targeted the 5' loxP site. Alignment of the *Sftpc-CreER*^{T2} amplicon with the *CreER*^{T2} sequence (1029bp; obtained from plasmid 44580, Addgene) was a 100% match with bases 311-784 (data not shown). Interestingly, when attempts were made to align the *EYFP* amplicon to the EYFP sequence (obtained from Snapgene) a small number of bases in the middle of the 384bp sequence consistently failed to align to EYFP but aligned perfectly with bases 11-395 of Topaz YFP (Figure 5.5B). This suggested that the EYFP cassette had been misannotated and contained Topaz YFP, although full sequencing of the cassette would be needed for confirmation.

To ensure that recombination could occur between the loxP sites in the *Abca3* locus, an *in vitro* model of the system was developed. The location of the 5' loxP site had already been elucidated (Figure 5.5A) and the location of the 3' loxP site was identified after amplification of intron 7 of *Abca3* (Figure 5.6A). Direct sequencing of amplicons obtained from *Abca3*$^{flox/flox}$ and BALB/c mice (Table 2.1: F1-8_mAbca3_in7sq), with subsequent alignment to an *Abca3* reference sequence (ENSMUST00000039013.15) identified a 176bp insertion, including a loxP site, 1728bp downstream of exon 7 in *Abca3*$^{flox/flox}$ mice (Figure 5.6A). With the location of the two loxP sites now known, a

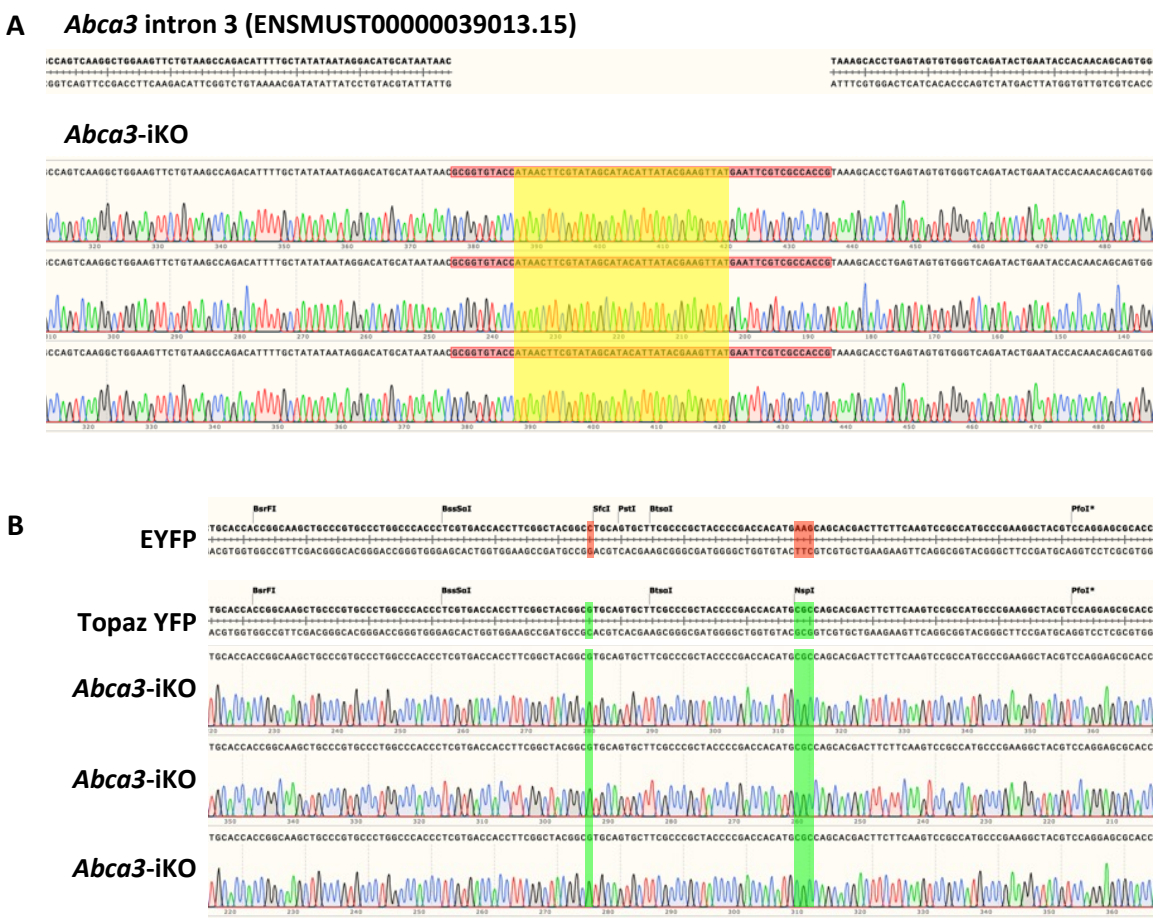


Figure 5.5 Sanger sequencing traces of *Abca3*-iKO transgenic cassettes

PCR amplicons obtained from genotyping *Abca3*-iKO mice were cloned into TOPO plasmids and sequenced. **(A)** Alignment of *Abca3*^{flox} cassette with an *Abca3* reference sequence (ENSMUST00000039013.15) identified a 60bp insertion (red) in intron 3 (450bp upstream of exon 4) containing a loxP site (yellow). **(B)** Alignment of *EYFP* cassette with EYFP and Topaz YFP reference sequences (Snapgene). The *EYFP* cassette contained four base pair mismatches (red) when aligned with EYFP. No mismatches (green) between the *EYFP* cassette and Topaz YFP reference sequence were found.

fragment of DNA encompassing either the floxed *Abca3* sequence (*Abca3*$^{flox/flox}$ mice) or WT *Abca3* sequence (BALB/c mice) was isolated from gDNA using mABCA3_in3_F and mABCA3_in7_R6 primers (Table 2.1; Section 2.2.2) and cloned into TOPO plasmids to generate pSP6.Abca3in3-7 (BALB/c) (pGM1240; WT *Abca3*) and pSP6.Abca3in3-7(*Abca3*-iKO) (pGM1241; floxed *Abca3*; Table 2.2; Section 2.2.7). After linearising the plasmids by restriction digest (NheI; Section 2.2.1), plasmids were incubated with Cre recombinase protein (M0298, NEB) before products were visualised by agarose gel electrophoresis (Figure 5.6B; Section 2.2.17). In pGM1241, containing floxed *Abca3,* two products were detected (Figure 5.6C), whereas only the full-length linearised plasmid (9.2kb) was identified in pGM1240 containing WT *Abca3*. This suggested that recombination between the two loxP sites of the *Abca3* should be possible *in vivo*.

5.2.3 Natural history of uninduced *Abca3*-iKO mice

Once a breeding colony of *Abca3*-iKO mice was established, the probability of survival after weaning (four weeks) was assessed in uninduced animals (no tamoxifen; Figure 5.7). The oldest animal electively culled was at 49 weeks although most healthy animals were electively culled earlier to reduce husbandry costs. The survival probability beyond 45 weeks appeared similar between males and females (98.7% and 97.6% respectively; Figure 5.7). Three out of the 267 male mice included in the analysis were found dead or were culled due to sickness between 8-10 weeks old. For females, three mice were found dead or culled due to sickness between 7-9 weeks old; one mouse went missing at 18 weeks and 220 females were still alive or had been electively culled when the data were analysed. These results are in accordance with other non-harmful transgenic lines maintained under identical husbandry conditions.

It has been reported that some Cre-loxP systems can be leaky [245] resulting in Cre recombinase activity in the absence of tamoxifen. To assess if Cre was active in uninduced *Abca3*-iKO mice, EYFP expression was assessed. Ideally, isogenic control *Abca3*$^{flox/flox}$*, Sftpc-CreER*$^{T2-}$*, EYFP*$^{+}$ mice would have been used to assess baseline recombination activity at the primary locus of interest, *Abca3*,

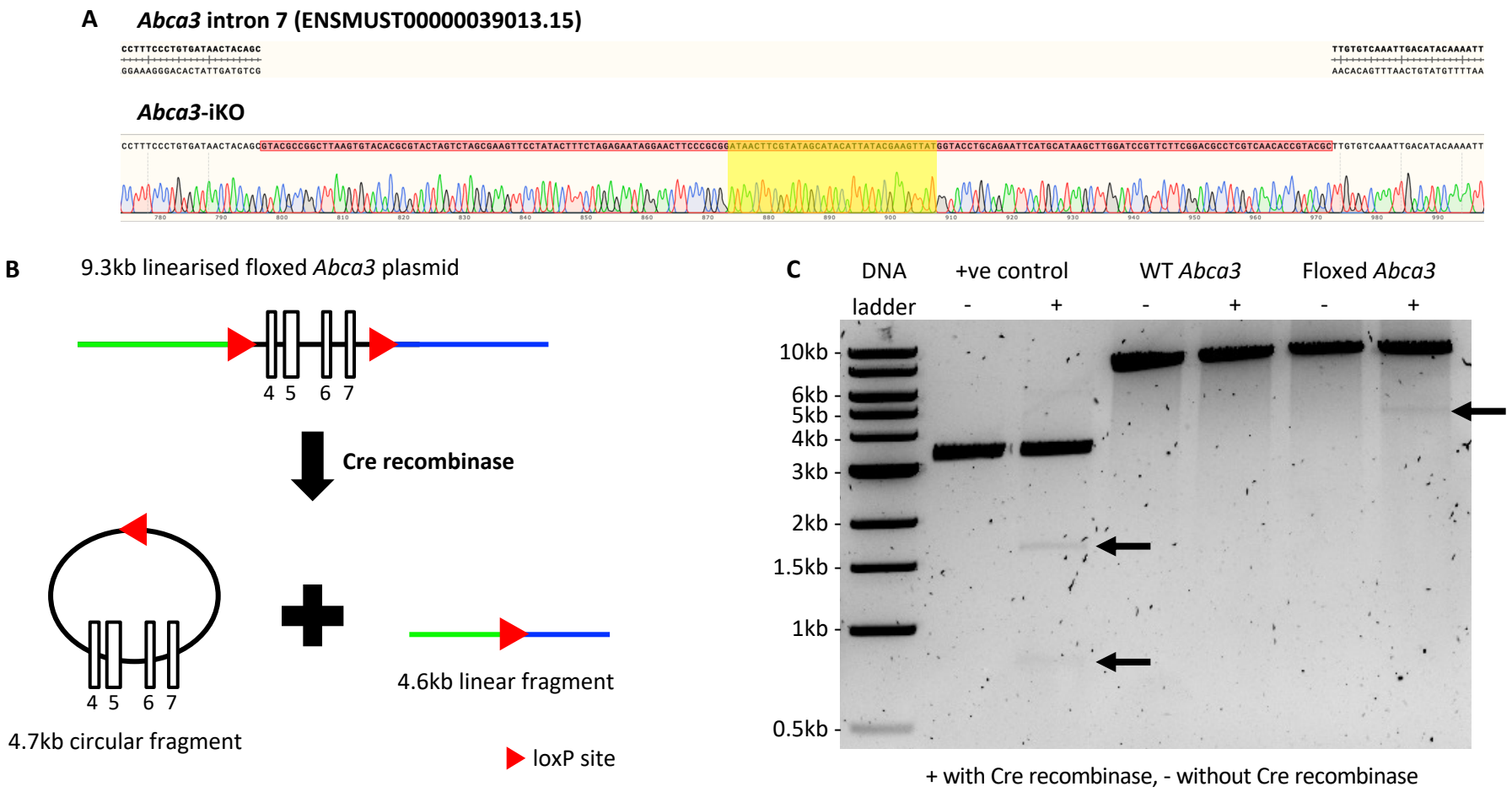


Figure 5.6 Recombination of floxed *Abca3* pSP6.Abca3in3-7(*Abca3*-iKO) *in vitro*

Figure 5.6 Recombination of floxed *Abca3* pSP6.Abca3in3-7(*Abca3*-iKO) *in vitro*

The Cre recombinase/loxP system present in *Abca3*-iKO mice was recreated *in vitro* by cloning the floxed *Abca3* sequence (or WT BALB/c control) into a TOPO plasmid and incubating with Cre recombinase. **(A)** The 3' loxP site (yellow) in the $Abca3^{flox}$ allele was located 1.7kb downstream of exon 7. The 5' loxP site was previously identified 450bp upstream of exon 4 (Figure 5.5A). **(B)** Diagram of pSP6.Abca3in3-7(*Abca3*-iKO) (floxed *Abca3*; pGM1241; Table 2.2) linearised with NheI. Treatment of linearised plasmid with Cre recombinase should result in recombination between the loxP sites (red arrowheads) to produce a 4.7kb circular fragment and 4.6kb linear fragment of DNA. **(C)** Agarose gel electrophoresis of linearised pSP6.Abca3in3-7(BALB/c) (pGM1240), pSP6.Abca3in3-7(*Abca3*-iKO) (pGM1241) and a positive recombination control (+ve) plasmid (supplied by manufacturer) incubated with (+) and without (-) Cre recombinase. An approximately 4.6kb fragment was detected in pGM1241 treated with Cre recombinase but not in pGM1240, nor in samples not treated with Cre recombinase. Black arrows indicate recombination products.

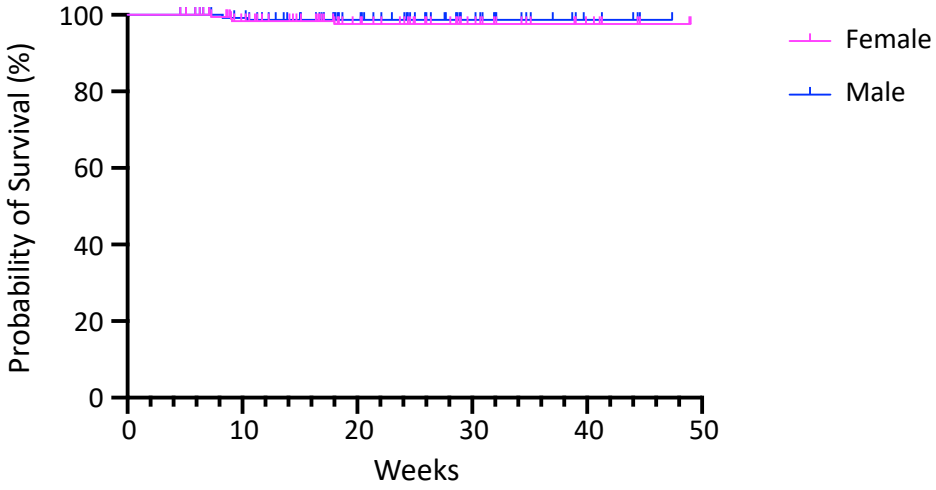


Figure 5.7 Kaplan-Meier survival curve for uninduced *Abca3*-iKO mice

The survival probability of *Abca3*-iKO mice without disease induction was estimated from 224 females and 267 males who did not receive experimental intervention. Mice found dead (n=3), culled due to sickness (n=3) or that went missing (n=1) were allocated a value of 1. Mice culled electively whilst healthy (n=332) or still alive when the data was captured (n=72) were allocated a value of 0. Males and females had a 98.7% and 97.6% probability respectively of surviving beyond 45 weeks. All mice <4 weeks old (pre-wean) were excluded from analysis.

because recombination activity may vary at different loci. However, as these mice were not available, gain of EYFP expression was used as a surrogate marker. Upstream of the *EYFP* sequence in the *Rosa26* loci is a floxed tpA sequence (Section 2.7.3) [191] which should prevent EYFP transcription in uninduced *Abca3*-iKO mice. Lung cryosections and PCLS from BALB/c and uninduced *Abca3*-iKO mice were compared for native EYFP expression (Section 2.6.3). Samples from BALB/c mice were used as a negative control for background tissue autofluorescence. In cryosections from *Abca3*-iKO mice, signal in the YFP channel, above that present in BALB/c mice, was readily detected in a modest number of cells (Figure 5.8A); this observation suggested that Cre was active to some degree at the *EYFP* locus prior to tamoxifen administration. Similarly, when PCLS were generated and cultured from two *Abca3*-iKO mice, EYFP expression was also visible (Figure 5.8B). Together with the observed survival of uninduced *Abca3*-iKO mice (Figure 5.7), these data suggest that the Cre-loxP system present in *Abca3*-iKO mice is slightly leaky but does not appear to induce disease with a physiologically detrimental impact on animal wellbeing.

5.2.4 Induction of a disease phenotype in *Abca3*-iKO mice

A tamoxifen dosing regimen that could predictably and consistently induce high levels of Cre recombination at the *Abca3* locus in the *Abca3*-iKO mice was required, before they could be used to assess the efficacy of dosing with rSIV.F/HN.hCEF.ABCA3. It was anticipated that deletion of *Abca3* exons 4 – 7 would lead to a measurable change in the *in vivo* phenotype such that disease protection studies using rSIV.F/HN.hCEF.ABCA3 would not be solely reliant on molecular analysis of tissue samples obtained after death. In addition to looking for a respiratory disease phenotype, the weight and temperature of animals were closely monitored (Section 2.7.3.4) for animal wellbeing and as a potential indicator of a disease phenotype, as these parameters have proven useful when monitoring SP-B deficient animals (Kamran Miah, Personal communication) [122]. In accordance with the UK Home Office project licence under which these studies were performed

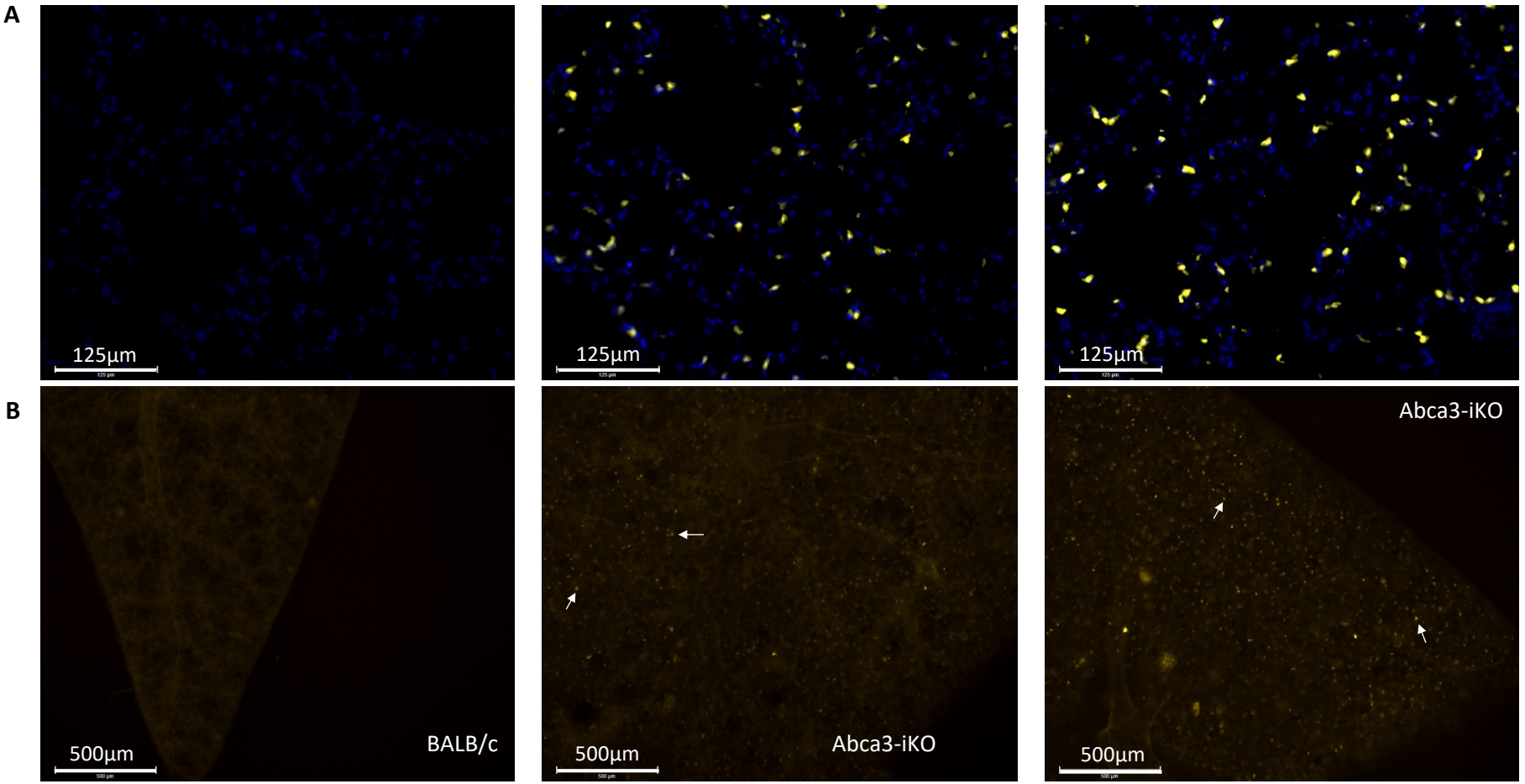


Figure 5.8 *Sftpc-CreER*^{T2} expression is leaky in uninduced *Abca3*-iKO mice

Figure 5.8 *Sftpc-CreER^{T2}* expression is leaky in uninduced *Abca3*-iKO mice

(A) Lung cryosections and **(B)** precision cut lung slices (PCLS) from *Abca3*-iKO and BALB/c mice were imaged for expression of EYFP to assess if *Sftpc-CreER^{T2}* is active in the uninduced (baseline) state. Cryosections were counterstained with DAPI (blue; Section 2.6.3) before images (n=3 per sample) were taken using an EVOS FL Auto 2 Imaging System with associated YFP and DAPI light cubes (Section 2.6.5.1). PCLS were imaged live *in situ* for EYFP (yellow). Representative images shown. Capture settings were kept consistent between samples. White arrows indicate EYFP expression.

animals were humanely culled if a weight loss of 15% or more was recorded more than once or the mice appeared to be in severe distress.

5.2.4.1 Approach 1: Intraperitoneal injection of tamoxifen

Intraperitoneal delivery of tamoxifen is a commonly used method of activating Cre-loxP systems *in vivo* [246] due to its accommodation of weight-based dosing, to ensure that each animal is dosed with the correct amount of tamoxifen. An initial dosing regimen of 100mg/kg IP tamoxifen (20mg/ml solution; Section 2.7.3.3.1) dosed for 5 days consecutively was selected based on similar doses being reported in the literature to induce Cre-loxP systems in the lungs [247]. The *Abca3*-iKO mice (n=4 male and n=4 female) were dosed with tamoxifen and the weight, external body temperature and respiratory phenotype [122] monitored (Section 2.7.3.4) 2-4 times a day for a month after the final dose of tamoxifen was administered (day 0; Figure 5.9) before lungs were harvested for analysis. During the 5 days of tamoxifen dosing, and the proceeding 4 days, body weight steadily declined in all animals (Figure 5.9A and 5.9B). Two males (FLSL 16.1A, FLSL 16.1B) continued to lose weight and were sacrificed 4 and 10 days after receiving the final dose of tamoxifen, but the weights of the remaining six animals stabilised after the initial period of loss and these animals were electively culled at the end of the study. In general, and in contrast to the weight loss observed during dosing (Figure 5.9A), the external body temperature of the mice did not decline during tamoxifen dosing (Figure 5.9C). However, a small decline in temperature was observed in FLSL 16.1A prior to being culled. There were no outward signs of respiratory distress (Section 2.7.3.4) in any of the animals (data not shown) but when lungs of FLSL 16.1B (sacrificed on day 10) were harvested, dark red regions which may indicate lung collapse or consolidation (when air in the lung is replaced with something else) were apparent (Figure 5.10A). In general, the lungs of mice harvested at the end of the study appeared pink and healthy (Figure 5.10B). Despite the differences in gross lung pathology between FLSL 16.1B (sacrificed on day 10) and FLSL 15.1E

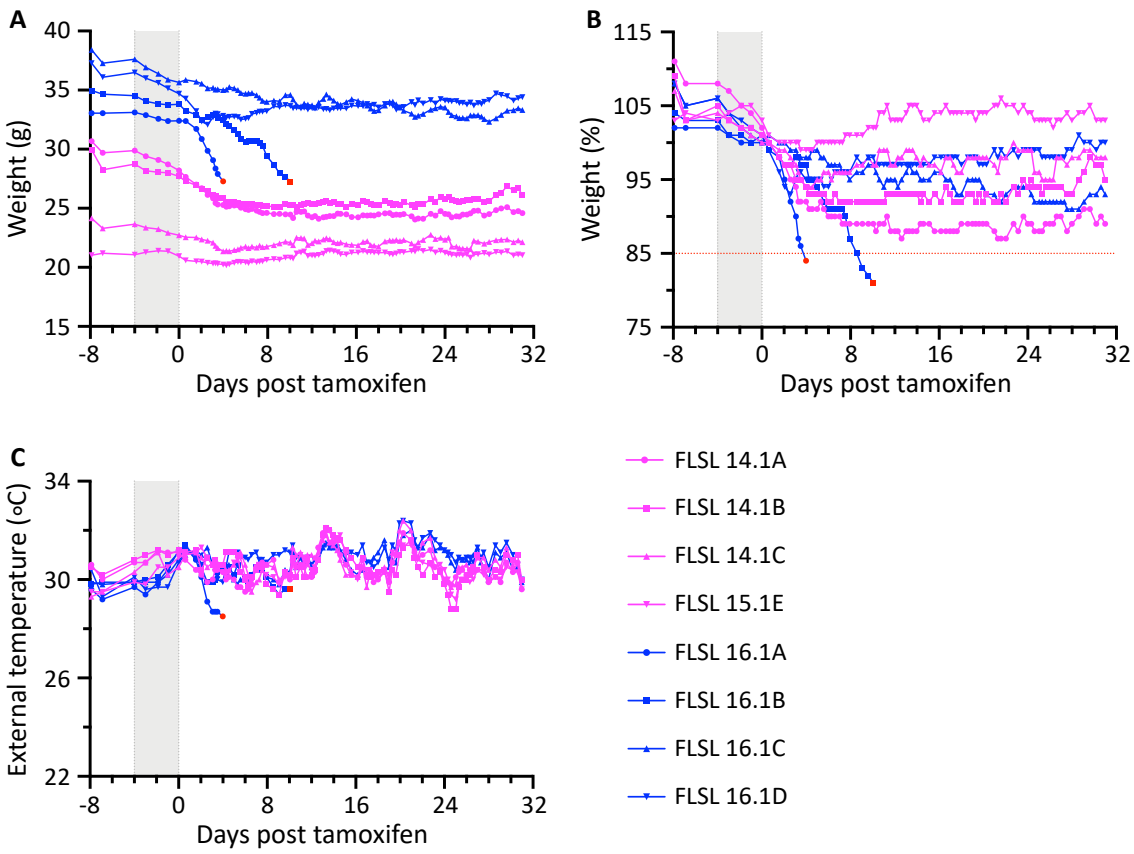


Figure 5.9 Weight and temperature of *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection

The *Abca3*-iKO mice were treated on 5 days consecutively with 100mg/kg IP tamoxifen (20mg/ml; Section 2.7.3.3.1). **(A, B)** Weight and **(C)** external body temperature were monitored 2-4 times a day for 1 month. Mice losing more than 15% of body weight compared to the final day of tamoxifen dosing (day 0) were culled and lungs harvested. Data are plotted as the 3-point rolling average. Grey bars indicate period of tamoxifen dosing. Pink and blue data points correspond to female and male mice respectively. Mice culled prior to the study termination are indicated with a red dot for the final measurements taken.

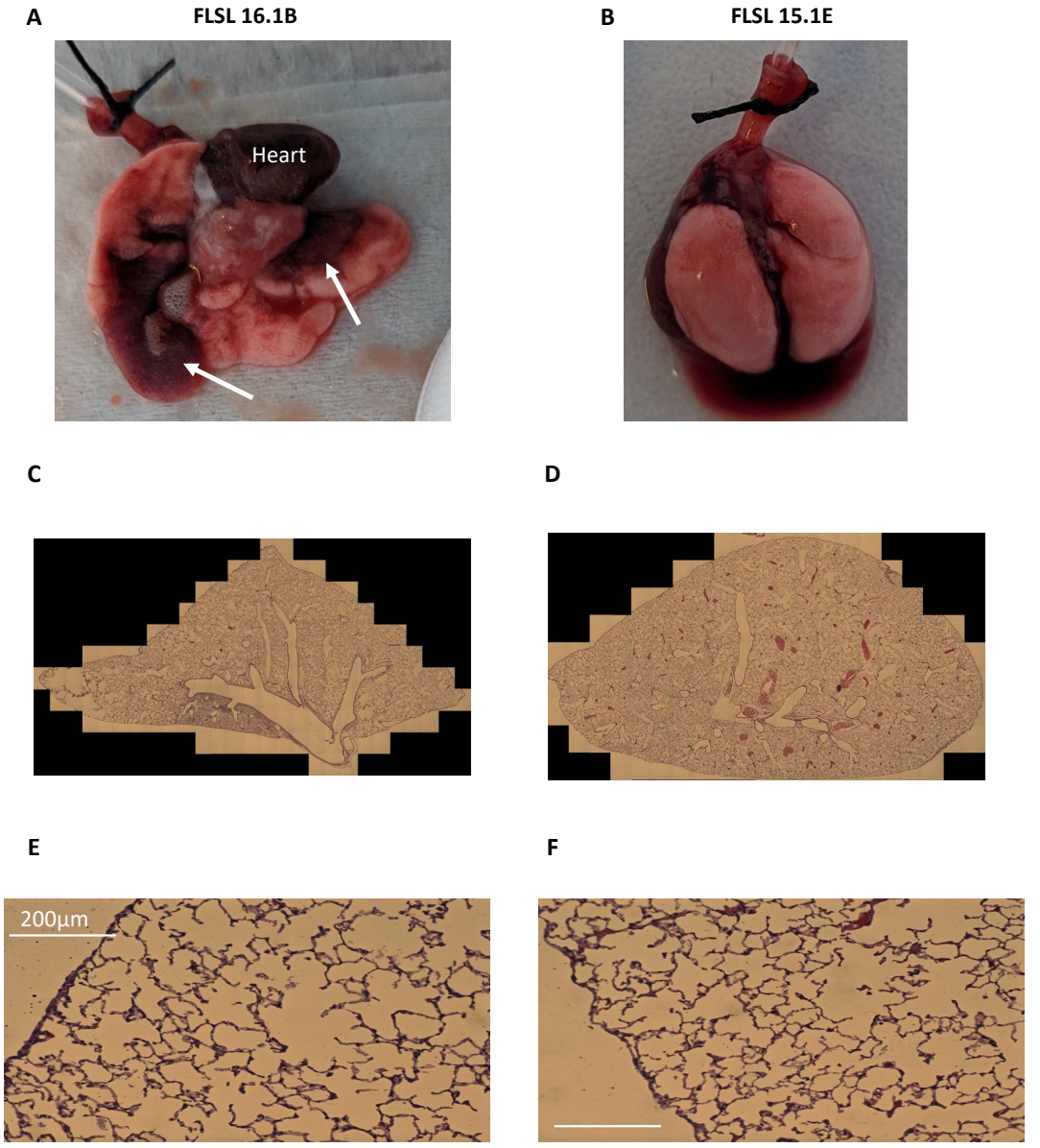


Figure 5.10 Evidence of pathology in lungs harvested from *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection

Organs were harvested from *Abca3*-iKO mice treated on 5 days consecutively with 100mg/kg IP tamoxifen (20mg/ml; Section 2.7.3.3.1; Figure 5.9). **(A, B)** Photos of the heart and lungs obtained when tissue was collected from **(A)** FLSL 16.1B, culled 10 days after the final dose of tamoxifen, and **(B)** FLSL 15.1E, culled at the end of the study. Dark red regions of suspected lung pathology are indicated by arrows. **(C-F)** H&E staining of cryosections obtained from the lungs (**C, E**: FLSL 16.1B; **D, F**: FLSL 16.1C). All scale bars indicate 200µm, representative H&E images shown.

(sacrificed electively at the end of the study), no differences were identified by H&E staining for all animals in the study (Figure 5.10 C-F).

DNA and RNA extracted from the lungs (Section 2.7.4) were analysed to assess the level of recombination achieved at the *Abca3* locus and how this impacted *Abca3* mRNA transcript levels (Figure 5.11). It was hypothesised that an increase in Cre recombination activity would correspond to decreased *Abca3*. Following endpoint PCR amplification of intron 3 – 7 of *Abca3* (Figure 5.11A; Table 2.1: mABCA3_in3_F, mABCA3_in7_R6) an approximately 5kb amplicon (contains exons 4-7) was detected in all samples whereas a secondary 1kb amplicon (exon 4-7 deletion) was only detected in samples from tamoxifen-treated animals, suggesting that recombination had occurred in all animals treated with tamoxifen. To quantify the extent of recombination at the *Abca3* locus, the percentage recombination was calculated using ddPCR (Section 2.5.5). The ddPCR assay utilised was designed such that one primer/probe specifically amplified the *Abca3* locus after recombination had occurred (Table 2.1: loxP3_X_ddPCR) whilst the other specifically amplified non-recombined floxed *Abca3* (Figure 5.11B; Table 2.1: loxP1_X_ddPCR). As seen in Figure 5.11C, ~10% recombination was achieved on average for tamoxifen-treated mice, however the variability was smaller for male *Abca3*-iKO mice. Interestingly, the two animals culled prematurely (FLSL 16.1A, FLSL 16.1B) were not associated with the highest percentage of Cre recombination. When the relative *Abca3* expression was subsequently assessed by qRT-PCR (Section 2.5.4) using primers targeting the deleted fragment of *Abca3* (Table 2.1: mABCA3_qRT_Fex6_202011, mABCA3_qRT_Rex7_202011), 100mg/kg IP tamoxifen dosing was associated with an apparent 10-fold reduction in *Abca3* expression compared with uninduced *Abca3*-iKO mice (Figure 5.11D).

To investigate if recombination at the *Abca3* locus could be increased above the 10% achieved with five IP doses of 100mg/kg tamoxifen (Figure 5.11C), and perhaps also induce a consistent respiratory or weight loss disease phenotype, the daily IP tamoxifen dose was increased to

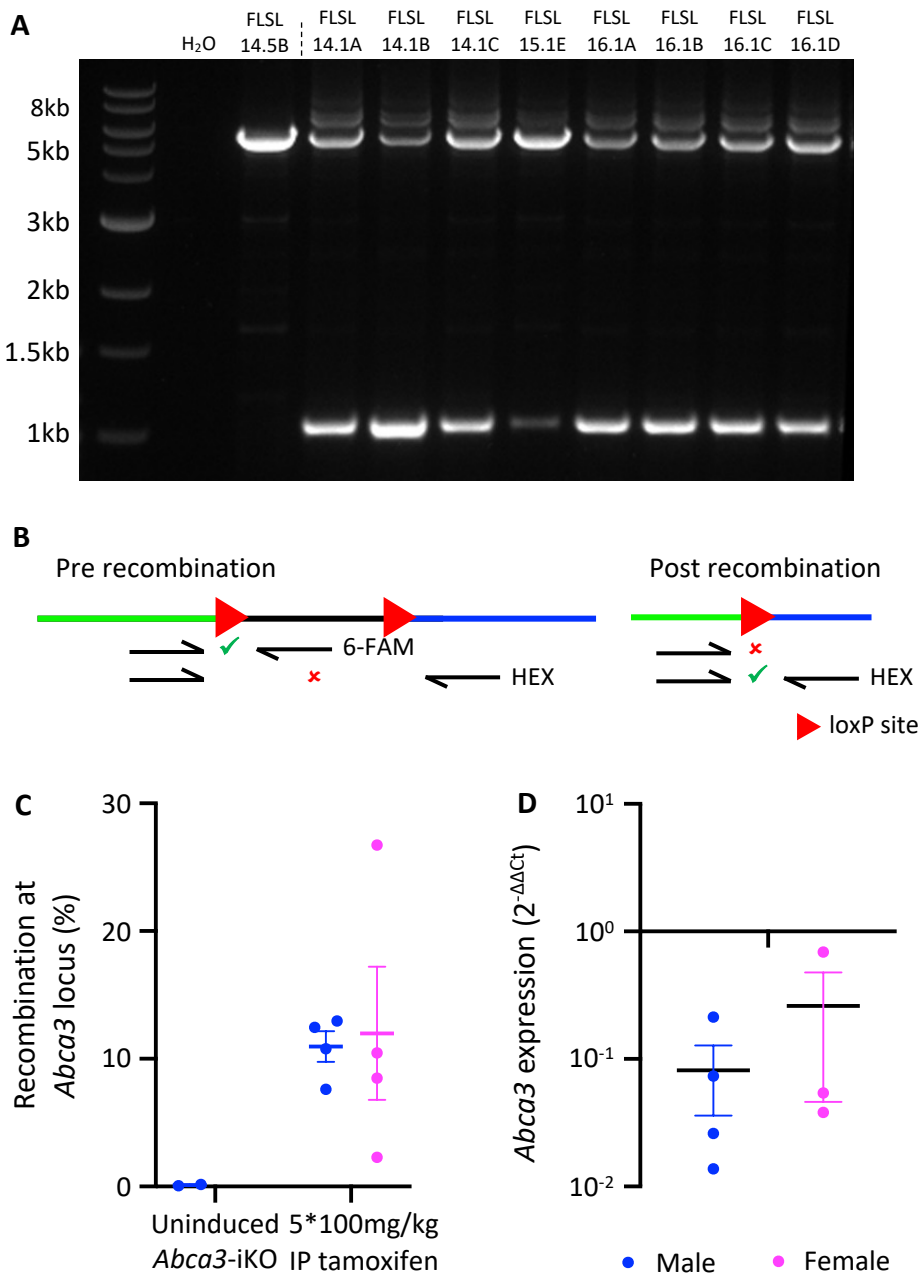


Figure 5.11 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered 100mg/kg tamoxifen by intraperitoneal injection

Organs were harvested from *Abca3*-iKO mice treated on 5 days consecutively with 100mg/kg IP tamoxifen (20mg/ml; Section 2.7.3.3.1; Figure 5.9) and DNA and RNA extracted from the harvested lungs to assess induction of the Cre-loxP system. **(A)** Genomic DNA samples from an uninduced *Abca3*-iKO mouse (FLSL 14.5B) and mice treated with 100mg/kg IP tamoxifen for 5 days underwent PCR amplification of intron 3 – 7 of *Abca3* to assess recombination of floxed *Abca3*. An

approximately 5kb amplicon was expected when recombination had not occurred, whilst deletion of exons 4 – 7 should result in amplification of a 1kb product. Both 5kb and 1kb amplicons were detected for all tamoxifen-dosed animals, but in the uninduced *Abca3*-iKO control only a 5kb product was detected. **(B)** Primer binding sites for ddPCR Cre recombination. **(C)** Quantification of Cre recombination at the *Abca3* locus by ddPCR. Results are separated by gender (Male, blue; Female, pink). **(D)** qRT-PCR analysis of exons 6 – 7 of *Abca3* in RNA extracted from uninduced and tamoxifen-treated *Abca3*-iKO mice. Relative expression of *Abca3* was calculated compared to uninduced *Abca3*-iKO mice. Data are shown as mean ± SEM with three technical replicates.

200mg/kg. An internal temperature probe was implanted into *Abca3*-iKO mice for monitoring purposes (Section 2.7.3.4), in case it was more sensitive to physiological changes than external body temperature. Two BALB/c mice also received 200mg/kg IP tamoxifen injections. The results showed rapid weight loss in all mice after 200mg/kg tamoxifen IP dosing was started (Figure 5.12A); *Abca3*-iKO mice lost approximately 10% of their body weight between the first and last tamoxifen dose (Figure 5.12B). A similar trend in weight loss was observed in BALB/c mice. In addition, internal and external body temperatures dropped by approximately 3°C in multiple animals (Figure 5.12C, 5.12D) and mice appeared lethargic. As a consequence, tamoxifen dosing was stopped after 4 days as the similar responses in WT BALB/c mice indicated that the tamoxifen dosing rather than development of an *Abca3* deficiency disease phenotype was the cause. Within 4 days of the final dose of 200mg/kg IP tamoxifen nine of the ten study mice, including both BALB/c control mice, had reached the protocol severity limits and were humanely culled (Figure 5.12A). Gross examination of abdominal organs showed that the liver and intestines of these nine animals were abnormal whilst the lungs of most animals appeared pink and healthy (Figure 5.13A, 5.13B). Staining of the lungs with H&E did not identify signs of immune cell infiltration (Figure 5.13C, 5.13D). Together these observations suggested that IP injection of 200mg/kg tamoxifen generated unacceptable non-ABCA3 disease related physiological changes. Furthermore, when the impact of the higher dose of tamoxifen on recombination activity at the *Abca3* locus and mRNA expression in *Abca3*-iKO mice was assessed, the percentage of Cre recombination was not increased (Figure 5.14A). Similarly, expression of *Abca3* exons 6-7 was not further reduced by increasing the tamoxifen dose (Figure 5.14B).

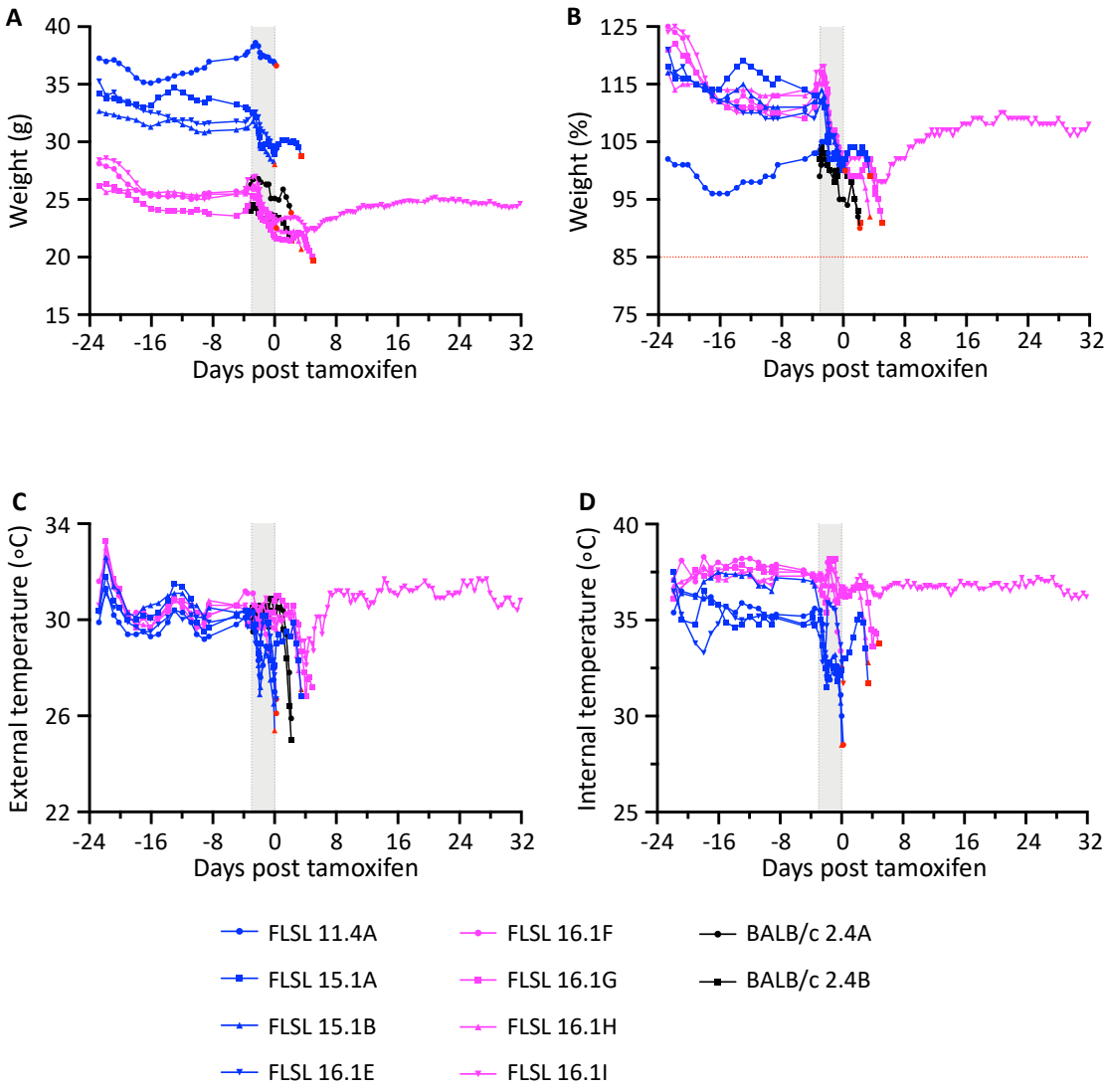


Figure 5.12 Weight and temperature of *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection

The *Abca3*-iKO mice and BALB/c mice were treated on 4 days consecutively with 200mg/kg IP tamoxifen (20mg/ml; Section 2.7.3.3.1). **(A, B)** Weight, **(C)** external and **(D)** internal body temperature were monitored 2-4 times a day for 1 month. Mice losing more than 15% of body weight compared to the final day of tamoxifen dosing (day 0) were culled and lungs harvested. Data are plotted as the 3-point rolling average. Grey bars indicate period of tamoxifen dosing. Pink and blue data points correspond to female and male *Abca3*-iKO mice respectively. Black data points correspond to BALB/c mice. Mice culled prior to the study termination are indicated with a red dot for the final measurements taken.

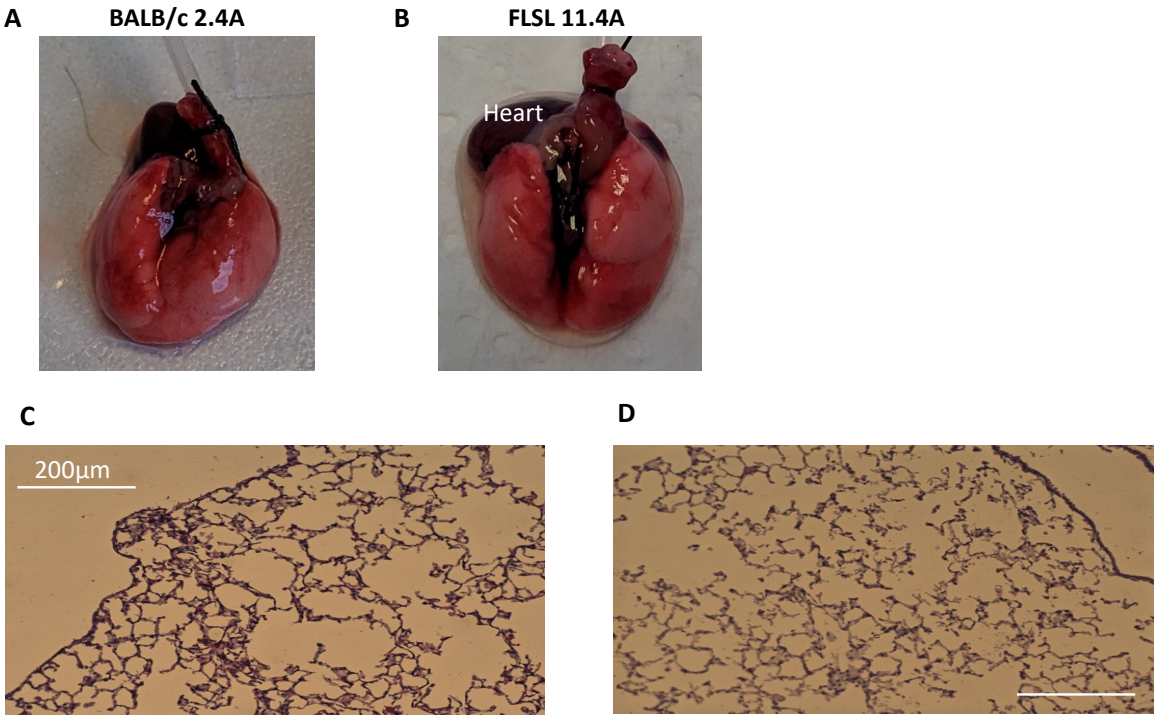


Figure 5.13 No evidence of lung pathology in *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection

(A, B) Photos of the heart and lungs obtained when tissue was collected from (A) BALB/c 2.4A and (B) FLSL 11.4A after administration of 200mg/kg IP tamoxifen for 4 days consecutively. Lungs appeared pink and healthy. **(C, D)** H&E staining of cryosections obtained from lungs (**C:** BALB/c 2.4A; **D:** FLSL 15.1B). All scare bars indicate 200µm, representative H&E images shown.

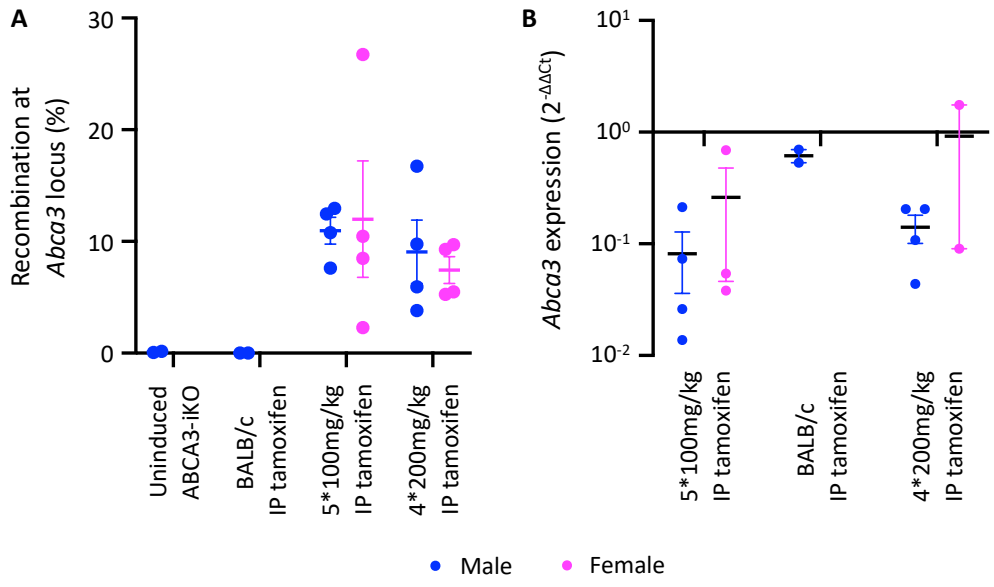


Figure 5.14 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered 200mg/kg tamoxifen by intraperitoneal injection

Organs were harvested from *Abca3*-iKO mice and BALB/c mice treated for 4 days consecutively with 200mg/kg IP tamoxifen (20mg/ml; Section 2.7.3.3.1; Figure 5.12) and DNA and RNA extracted from the harvested lungs to assess induction of the Cre-loxP system. **(A)** Quantification of Cre recombination at the *Abca3* locus by ddPCR. Results are separated by gender (Male, blue; Female, pink). **(B)** qRT-PCR analysis of exons 6 – 7 of *Abca3* in RNA extracted from uninduced and tamoxifen-treated *Abca3*-iKO mice. Relative expression of *Abca3* was calculated compared to uninduced *Abca3*-iKO mice. Data are shown as mean ± SEM with three technical replicates.

5.2.4.2 Approach 2: Delivery of tamoxifen in diet

A tamoxifen diet has been used successfully to induce disease in a very similar conditional *Abca3* KO model (Section 1.4.2.1.2) [128] but is not the preferred route of administration because the dosage of tamoxifen is dependent on how much the mouse eats, which is difficult to assess, and introduces another source of variability into studies. Additionally, some experienced animal workers suggest that mice may choose to starve instead of eating a tamoxifen diet due to an aversion to the taste (Jordan Tanner, Personal communication). Mice were administered a tamoxifen diet (0.25g/kg tamoxifen, 5% sugar) continuously for 8 or 14 days (Section 2.7.3.3.2) and monitored 2-4 times a day for up to 1 month. There was no change in respiratory phenotype and body temperature remained stable (data not shown) but, as anticipated, mice appeared reluctant to eat the tamoxifen diet and began to lose weight immediately (Figure 5.15). One female mouse (FLSL 15.6C) with a low starting bodyweight (approximately 17g) was culled 3 days after the tamoxifen diet was introduced, having lost 15% of body weight (Figure 5.15B). Upon gross examination lungs appeared healthy and the death was not considered to be related to the development of an *Abca3* KO phenotype. Remaining animals (n=9) appeared to adjust to the tamoxifen diet after 3 days with weights plateauing or beginning to increase (Figure 5.15). After 8 days eating tamoxifen-containing food, animals which then returned to eating standard chow (n=4) gained weight for the remaining study duration with one exception. One male (FLSL 18.1A) increased in weight in the first 6 days when standard chow was returned but then rapidly began losing weight, resulting in study termination 17 days after the tamoxifen diet was first introduced (Figure 5.15A). Gross examination of the lungs upon harvest revealed small dark red patches of suspected pathology (Figure 5.16A), however no changes in the lungs were visible by H&E staining (Figure 5.16B, 5.16C).

In contrast to mice who received a tamoxifen diet for 8 days, male mice (n=4) kept on the diet for an additional 6 days (14 days total) continued to lose weight and were culled 12-17 days after first

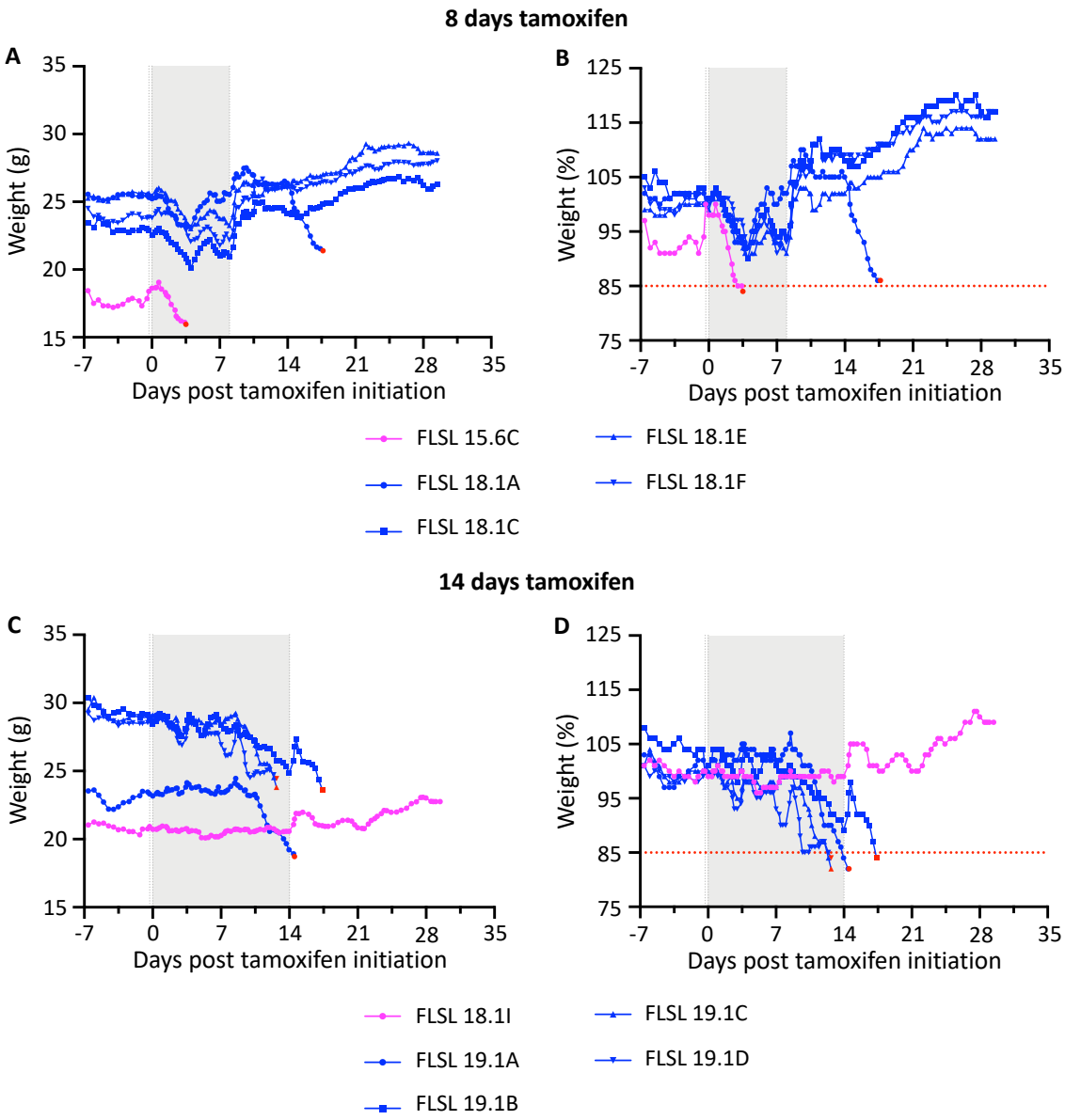


Figure 5.15 Weight of *Abca3*-iKO mice administered a tamoxifen diet (0.25g/kg) for 8 or 14 days

The *Abca3*-iKO mice were given a tamoxifen diet (0.25g/kg; Section 2.7.3.3.2) for **(A, B)** 8 or **(C, D)** 14 days continuously. Weight was monitored 2-4 times a day for 1 month. Mice losing more than 15% of body weight compared to the first day of tamoxifen dosing (day 0) were culled and lungs harvested. Data are plotted as the 3-point rolling average. Grey bars indicate period of tamoxifen dosing. Pink and blue data points correspond to female and male mice respectively. Mice culled prior to the study termination are indicated with a red dot for the final measurements taken.

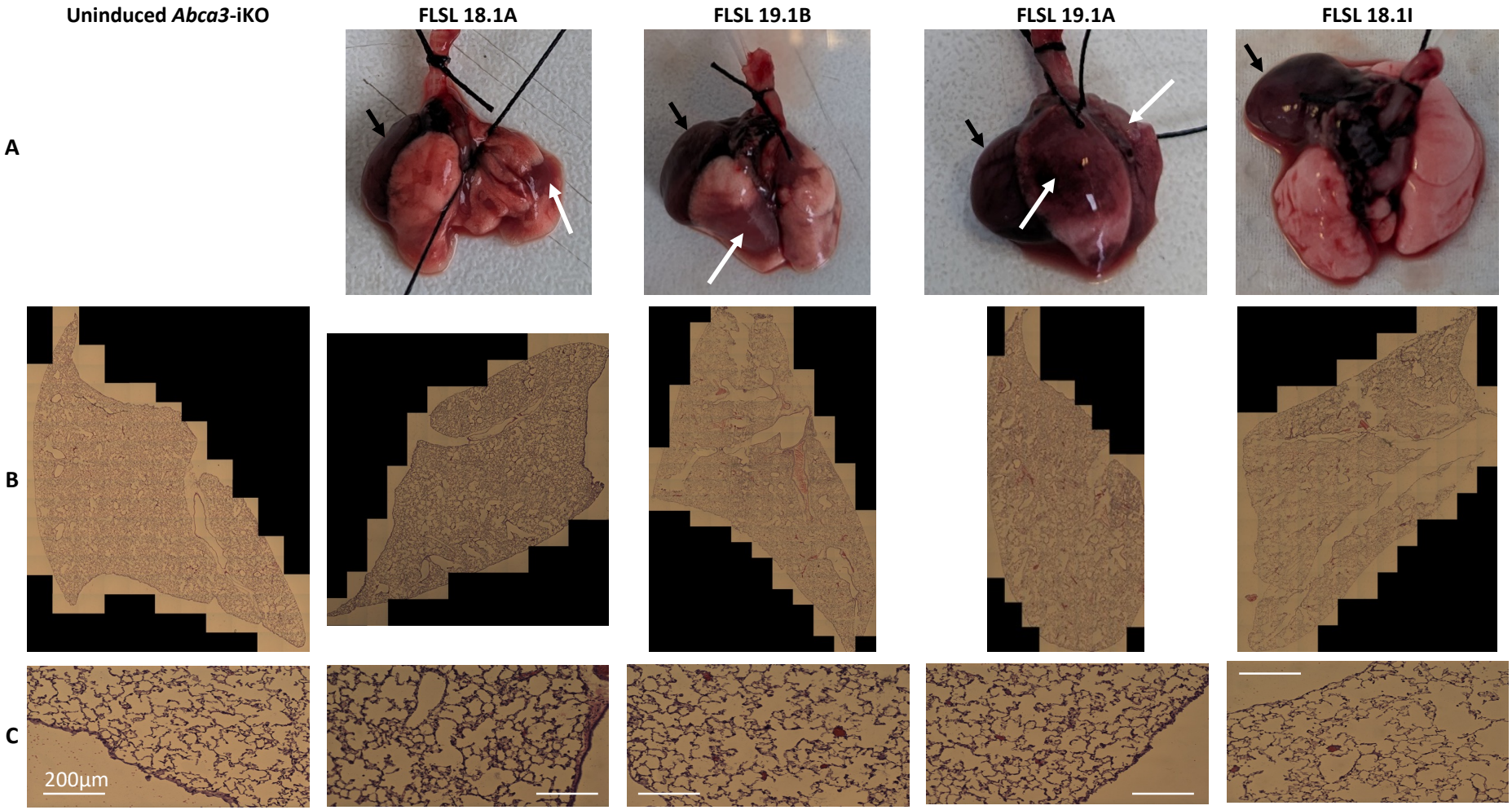


Figure 5.16 Evidence of pathology in lungs harvested from male *Abca3*-iKO mice administered a tamoxifen diet for 8 or 14 days

Figure 5.16 Evidence of pathology in lungs harvested from male *Abca3*-iKO mice administered a tamoxifen diet for 8 or 14 days

Organs were harvested from *Abca3*-iKO mice administered a tamoxifen diet (0.25g/kg; Section 2.7.3.3.2) for 8 or 14 days continuously (Figure 5.15). **(A)** Photos of the heart and lungs obtained when tissue was collected from FLSL 19.2A (uninduced *Abca3*-iKO), FLSL 18.1A (8 days tamoxifen diet) and FLSL 9.1B, FLSL 19.1A and FLSL 18.1I (14 days tamoxifen). Dark red regions of suspected lung pathology are indicated by white arrows, heart indicated by black arrows. **(B, C)** H&E staining of cryosections obtained from the photographed lungs. All scale bars indicate 200µm, representative H&E images shown.

receiving tamoxifen (Figure 5.15C). Interestingly, the weight of the female mouse (FLSL 18.1I; n=1) remained stable on tamoxifen before steadily increasing when returned to standard chow (Figure 5.15C). As previously seen in two animals sacrificed prematurely after dosing with 100mg/kg IP tamoxifen (Figure 5.10A) or 8 days of tamoxifen diet (Figure 5.16A), the lungs of male mice culled prematurely whilst receiving a 14-day tamoxifen diet contained dark regions of suspected lung pathology (Figure 5.16A). The extent of pathology varied between animals but were in stark contrast to lungs harvested from the female mouse (FLSL 18.1I) at the study endpoint, which outwardly appeared pink and healthy (Figure 5.16A). Staining with H&E (Figure 5.16B, 5.16C) did not reveal noticeable differences in the histopathology of the lungs of uninduced *Abca3*-iKO mice or mice treated with tamoxifen.

Analysis of Cre recombination at the *Abca3* locus suggested that, in males, recombination was increased compared with earlier levels after IP dosing of tamoxifen (Figure 5.17A; 8-day tamoxifen diet: mean recombination=13.6%, 14-day tamoxifen diet: mean recombination=17.2%). Furthermore, the highest levels of Cre recombination were observed in animals culled prematurely – approximately 23% of *Abca3* transcripts were recombined in the male culled at day 17 after receiving a tamoxifen diet for 8 days. Surprisingly, increased recombination at the *Abca3* locus did not appear to correlate with a greater reduction in *Abca3* mRNA levels. The mean lowest levels of *Abca3* exons 6-7 amplification (Figure 5.17B) were observed in mice dosed with 100mg/kg IP tamoxifen. However, in-line with the ddPCR recombination data (Figure 5.17A), *Abca3* levels appeared slightly lower in male mice that received a 14-day tamoxifen diet (mean=0.35) compared with those who received a tamoxifen diet for 8 days (mean=0.64; Figure 5.17B).

Regardless of tamoxifen dosing strategy, Cre recombination at the *Abca3* locus in female mice was generally lower than in male mice (Figure 5.17A). Correspondingly, excluding animals dosed with IP 200mg/kg tamoxifen, where tamoxifen toxicity was considered a major contributing factor, male mice proportionally accounted for more of the animals culled prematurely than females. It was

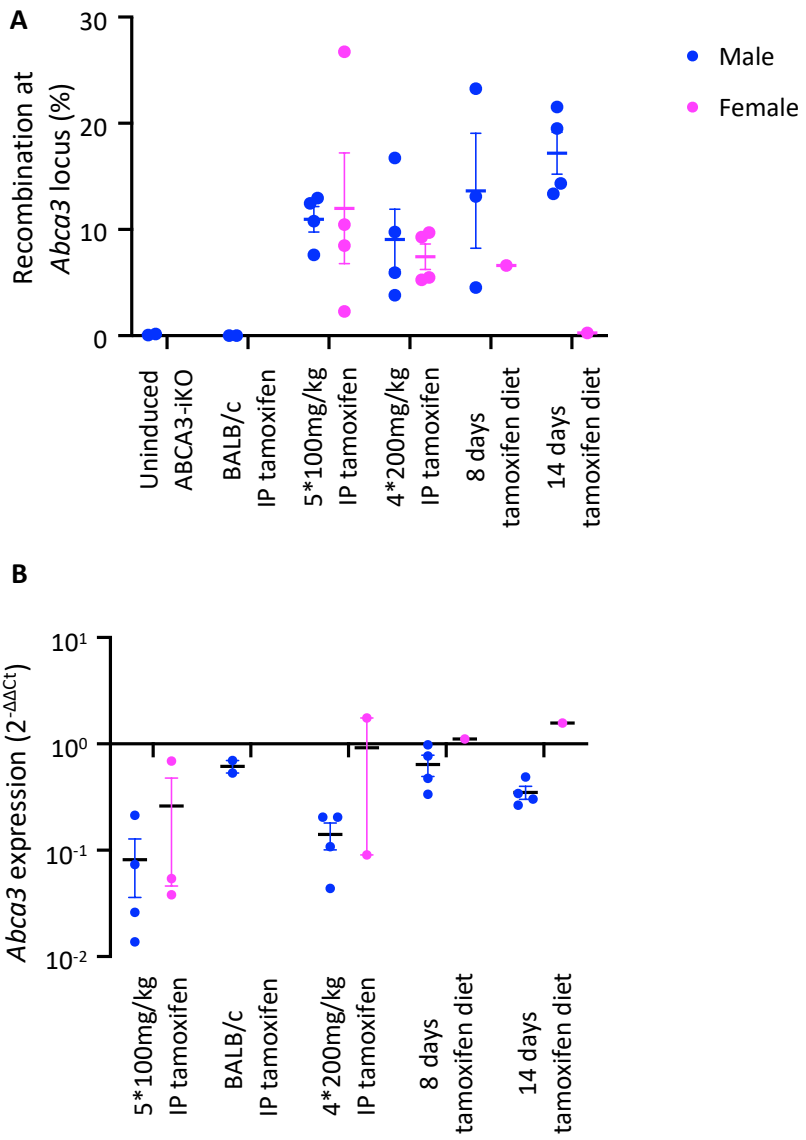


Figure 5.17 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered a continuous tamoxifen diet

Organs were harvested from *Abca3*-iKO mice administered a tamoxifen diet (0.25g/kg; Section 2.7.3.3.2; Figure 5.15) for 8 or 14 consecutive days and DNA and RNA extracted from the harvested lungs to assess induction of the Cre-loxP system. **(A)** Quantification of Cre recombination at the *Abca3* locus by ddPCR. Results are separated by gender (Male, blue; Female, pink). **(B)** qRT-PCR analysis of exons 6 – 7 of *Abca3* in RNA extracted from uninduced and tamoxifen-treated *Abca3*-iKO mice. Relative expression of *Abca3* was calculated compared to uninduced *Abca3*-iKO mice. Data are shown as mean ± SEM with three technical replicates.

hypothesised that male mice were more sensitive to tamoxifen (Section 5.3.3), and only male mice were included in subsequent studies.

Alternative tamoxifen diet strategies were also investigated; *Abca3*-iKO mice (n=10) and BALB/c mice (n=3) were given a tamoxifen diet for 5 days, followed by standard chow for 2 days, until they had received a tamoxifen diet for a total of 15 days (Section 2.7.3.3.2). All mice survived until the elected study endpoint with no visible signs, *in vivo*, of disease development and showed a repeating cycle of weight loss followed by weight gain when tamoxifen food and then standard chow was provided (Figure 5.18). This supported the hypothesis that tamoxifen diet avoidance was a contributing factor to weight loss in previous studies (Figure 5.15), and thus bodyweight may be an unreliable indicator of the development of an *Abca3* deficiency phenotype in mice when tamoxifen was given orally *ad libitum*. Given the dependence of tamoxifen dose on individual mouse food consumption, molecular analyses (ddPCR, qRT-PCR) of lung tissue obtained from animals pulsed with a tamoxifen diet revealed a large variability in Cre recombination at the *Abca3* locus (Figure 5.19A) and expression of exons 6 – 7 of *Abca3* (Figure 5.19B). On average, the extent of Cre recombination with the pulsed tamoxifen diet (16%) was similar to that achieved in males receiving a continuous tamoxifen diet for 14 days (17%). Likewise, in contrast to IP administration of tamoxifen, both continuous provision and pulsing of a tamoxifen diet did not appear to reduce *Abca3* mRNA expression. Of note, the general discrepancy between qRT-PCR (Figure 5.19B) and ddPCR (Figure 5.19A) data obtained is hypothesised to be a result of high Ct values and technical replicate variability for qRT-PCR. Given that no visible change in respiratory phenotype was observed in tamoxifen-treated *Abca3*-iKO mice, the ddPCR Cre recombination assay data is assumed to be more reflective of the *in vivo* picture. Had time allowed, efforts would have been made to optimise the qRT-PCR assay and investigate the impact of Cre recombination on Abca3 protein expression.

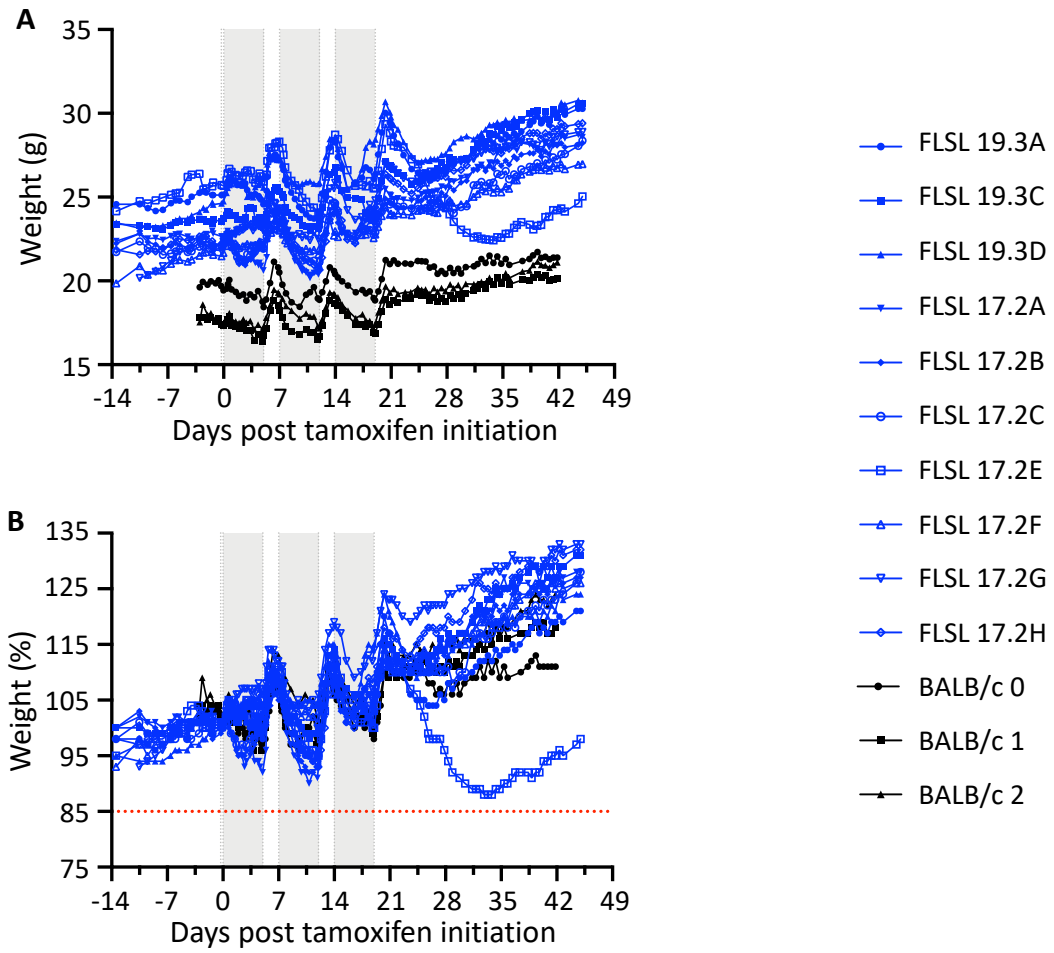


Figure 5.18 Weight of *Abca3*-iKO mice administered a pulsed tamoxifen diet (0.25g/kg) for a total of 15 days

The *Abca3*-iKO mice were given a tamoxifen diet (0.25g/kg; Section 2.7.3.3.2) for a total of 15 days (5 days on tamoxifen diet, 2 days on standard chow). Weight was monitored 2-4 times a day for about 1 month. Mice losing more than 15% of body weight compared to the first day of tamoxifen dosing (day 0) were culled and lungs harvested. Data are plotted as the 3-point rolling average. Grey bars indicate periods of tamoxifen dosing. Blue data points correspond to male *Abca3*-iKO mice and black data points correspond to BALB/c mice. Mice culled prior to the study termination are indicated with a red dot for the final measurements taken.

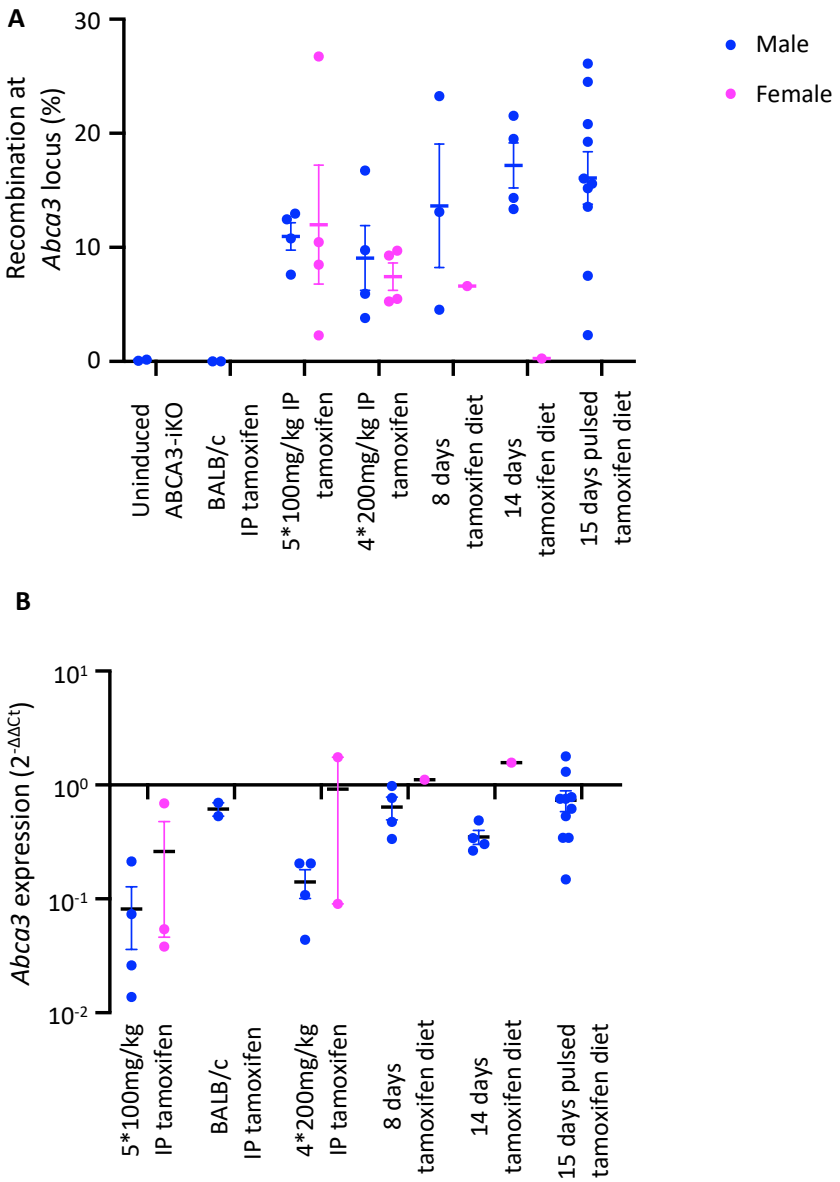


Figure 5.19 Cre recombination and *Abca3* expression in the lungs of *Abca3*-iKO mice administered a pulsed tamoxifen diet for a total of 15 days

Organs were harvested from *Abca3*-iKO mice administered a pulsed tamoxifen diet (0.25g/kg; Section 2.7.3.3.2; Figure 5.18) for a total of 15 days and DNA and RNA extracted from the harvested lungs to assess induction of the Cre-loxP system. **(A)** Quantification of Cre recombination at the *Abca3* locus by ddPCR. Results are separated by gender (Male, blue; Female, pink). **(B)** qRT-PCR analysis of exons 6 – 7 of *Abca3* in RNA extracted from uninduced and tamoxifen-treated *Abca3*-iKO mice. Relative expression of *Abca3* was calculated compared to uninduced *Abca3*-iKO mice. Data are shown as mean ± SEM with three technical replicates.

5.3 Discussion

Studies described in this Chapter aimed to evaluate the performance of rSIV.F/HN.hCEF.ABCA3 *in vivo* and to specifically develop a murine disease model of ABCA3 deficiency which could be used to further evaluate potential efficacy of this vector.

5.3.1 Molecular analysis of murine lung tissue

Overall, the results in this chapter indicated that the molecular assays and reagents used did not provide sufficient sensitivity or working range for the depth of analysis desired. For molecular analyses of lung tissue, DNA, RNA and protein were routinely harvested from the right lung lobes of mice whilst the left lobe was fixed and sectioned for histological analysis (Section 2.7.7). This approach reduced the number of animals required and maximised the information obtained for any given experiment, supporting the principles of the 3Rs of animal research [129]; this approach should also have minimised variability. In *Abca3*-iKO mice, however, an additional source of variability was the level of disease induction achieved. In Figures 5.10 and 5.16, observed lung pathology was patchy and unevenly distributed between different lung lobes (Figure 1.7), with visible pathology restricted to only one or two lobes. This uneven distribution might also be reflective of what was happening on a cellular level and could account for variability between gross organ pathology, histology and molecular assays. For example, one mouse (FLSL 16.1B) receiving 100mg/kg IP tamoxifen was culled prematurely on day 10 with regions of visible lung pathology upon tissue harvest (Figure 5.10A), but Cre recombination (Figure 5.11) in this mouse was lower than the average for male mice dosed with the 100mg/kg IP tamoxifen regimen. Homogenising all lung lobes for extraction of RNA, DNA and protein and/or fixing all lobes for histological analysis should theoretically reduce inter-lobe variability and discrepancies between *in vivo* observations and molecular analyses but the number of animals required to complete all necessary analyses would not be practical.

Many of the assays used in this chapter were dependent on the extraction of DNA, RNA or protein from lung homogenates, and whilst isolation of genetic material from lung homogenates is arguably the simplest sample preparation approach, the percentage of ATII cells (the cell type that highly expresses ABCA3) is likely to represent a very small proportion of the total lysate. In Rindler *et al.* [128] the extent of *Abca3* deletion induced in an ATII cell-restricted conditional *Abca3* KO model (*Abca3*-cKO; Section 1.4.2.1.2) changed from 60% in whole lung samples to 90% when purified ATII cells were used. In Section 5.2.1, qRT-PCR confirmed expression of ABCA3 from rSIV.F/HN.hCEF.ABCA3 (Figures 5.2) but this was not detected by immunoblotting – nor was endogenous murine Abca3 detected. This is likely because endogenous Abca3 expression from ATII cells became 'diluted' by the greater number of non-ATII cell types, which have much lower (or non-existent) Abca3 expression levels (Figure 5.3). Isolation of murine ATII cells [248] prior to the extraction of genetic material should resolve this issue, allowing for easier detection of Abca3 and other ATII cell markers. Similarly, the much greater contribution of genetic material from 'non-ATII' cells in the molecular analysis of *Abca3*-iKO mice may have been a contributing factor to the low levels of recombination (Figure 5.19A) observed after tamoxifen dosing.

5.3.2 Tamoxifen route of administration

Whilst the use of Cre-loxP in animal models is a widely used technology there remains a lack of consensus on the optimal route and dosage of tamoxifen administration [249]. As a prodrug, a decision must also be made whether to administer tamoxifen, metabolised by cytochrome P450 enzymes in the liver [250], or the active metabolite 4-hydroxytamoxifen. Tamoxifen is more commonly used because it is less expensive but, as it must be metabolised to become active, it can result in a delayed effect. Here, when IP injection of tamoxifen (Section 5.2.4.1) and consumption of a tamoxifen diet (Section 5.2.4.2) were each evaluated for the ability to induce *Abca3* deletion in *Abca3*-iKO mice, neither route of administration resulted in an altered *in vivo* respiratory phenotype and Cre recombination at the *Abca3* locus was only between 10-20% for either strategy

(Figure 5.19). For a disease phenotype to be observed in the conditional *Abca3* KO mice, it is anticipated that *Abca3* expression must be reduced by more than 50%, as mice heterozygous for constitutive *Abca3* KO have no recognisable disease phenotype despite confirmation that *Abca3* levels are halved [92].

Intraperitoneal administration is a preferred method of administering tamoxifen because the dosage can be adjusted for animal weight and it can be ensured that the full dose is given to each animal. On the assumption that the tamoxifen dose would directly control the extent of Cre recombination, and thus Abca3 knockdown in ATII cells, it was hypothesised that by limiting the variability in tamoxifen dose received the level of *Abca3* deletion would be more consistent between animals. Whilst this appeared somewhat true (Figure 5.19A), since Cre recombination was less variable in mice administered tamoxifen IP rather than by diet, an IP dosing regimen which induced a respiratory disease phenotype (as anticipated with KO of *Abca3*) was not found. When the daily dose was increased from 100mg/kg to 200mg/kg undesirable non-Abca3 related physiologic changes ensued (Figure 5.12, 5.13).

In contrast to IP administration of tamoxifen, delivery through diet is harder to control as it is dependent on each animal eating a specific quantity of the diet. This issue is further compounded by the bitter taste of tamoxifen, which is difficult to disguise and can result in diet avoidance [251]. However, the diet approach to tamoxifen delivery was attempted here since Rindler *et al.* [128] reported its successful use to induce high levels of *Abca3* deletion in *Abca3*-cKO mice (Section 1.4.2.1.2). When mice were continuously sustained on a tamoxifen diet body weights declined (Figure 5.15) and manifestation of a disease phenotype was difficult to identify. On a molecular level, responses to a tamoxifen diet were more heterogeneous than from IP delivery (Figure 5.17). Pulsing the tamoxifen diet with standard chow is recommended by Inotiv (Indiana, USA, [252]) and Queen's University (Ontario, Canada, [253]) to manage aversion to the tamoxifen diet and accompanying weight loss. This strategy reduced excessive weight loss, but mice tended not to eat

when given a tamoxifen diet and then gorged on standard chow (Figure 5.18). Mixing tamoxifen powder in gel supplements (ClearH_2O, Maine, USA) as an alternative to tamoxifen food pellets may be better tolerated and could be attempted in future studies, however this approach would require the purchase of specialised feeders.

Oral gavage is an alternative administration method used to deliver tamoxifen to murine Cre-loxP models [254]. Similar to IP dosing, the drug is delivered directly to the stomach such that the dose can be well controlled. However, oral gavage is technically challenging, more stressful for the mice and when done incorrectly tamoxifen can be delivered into the lungs. Recently, oral gavage has been tested in *Abca3*-iKO mouse model and a change in respiratory phenotype was observed *in vivo* (Kamran Miah, Personal communication). The extent of recombination at the *Abca3* locus in these mice has not yet been measured. It is anticipated that this approach could be utilised in future studies to address the correction of Abca3 deficiency by administration of rSIV.F/HN.hCEF.ABCA3.

5.3.3 Disease phenotype in *Abca3*-iKO mice

In the *Abca3*-iKO mice, a consistent respiratory disease phenotype was rarely observed after treatment with tamoxifen and, generally, individual mice were only culled prematurely due to a drop in weight of >15% (Figure 5.9B, 5.15). In constitutive *Abca3* KO models (Section 1.4.2) differences in the rate and depth of breathing compared with WT mice are readily apparent [92]. Furthermore, in the most similar conditional *Abca3* KO model published (*Abca3*-cKO) [128] signs of respiratory distress, indicated by dehydration, dyspnoea, a hunched appearance and lethargy, were noted in those mice where *Abca3* KO had been induced. In the studies described here (Section 5.2.4), the apparent lack of a disease phenotype after tamoxifen administration may be explained by the low levels of recombined *Abca3* observed – the greatest level of recombination observed was approximately 25% (Figure 5.19). In constitutive heterozygous *Abca3* KO mice (*Abca3*$^{+/-}$), where Abca3 expression is thought to be reduced by 50%, the *in vivo* phenotype was unchanged compared with WT mice [92]. This suggests that Abca3 expression must be reduced by >50% for

potential changes in physiology to be observed. Likewise, in the context of gene addition therapy, this also implies that restoration of ABCA3 levels to 50% should correct a disease phenotype. Additionally, due to the demanding monitoring schedule established for *Abca3*-iKO mice during disease induction studies, monitoring was completed by a team of individuals. Mice were assigned a phenotype score of M0-M3 [122], however the differences between the scoring criteria were subtle and thus the scores assigned somewhat subjective.

These studies were not powered to detect differences in disease induction between sexes as this was not hypothesised to be an important factor in Cre recombination events. However, one important observation from the *Abca3*-iKO induction studies provided an early indication that a mouse sex difference in response to tamoxifen dosing might exist (Figure 5.19). In mice dosed with 100mg/kg IP tamoxifen half of the males were culled prematurely due to weight loss, whilst all the females were electively culled at the end of the study (Figure 5.9). Similarly, when tamoxifen was provided in food for 14 days all male mice had to be culled before the end of the study due to weight loss, whilst the female appeared unaffected (Figure 5.15D). Suitably powered studies would be required to investigate whether sex is a contributing factor to successful Cre recombination and disease induction. Furthermore, the impact of other potentially confounding factors (e.g. age at tamoxifen administration) should be considered.

5.3.4 Conclusion

For the first time, rLV expressing ABCA3 has been administered to mice leading to an increase in ABCA3 expression in the lungs of these animals. An important goal is to establish whether the quantity of ABCA3 produced in the lungs from the rLV is sufficient to correct a disease phenotype. So far, correction of ABCA3 deficiency has not been reported for any vector in any mouse model of this disease. Significant progress has been made in investigating the induction of an ABCA3 disease phenotype in a conditional KO model (*Abca3*-iKO) and its use to evaluate the efficacy of potential ABCA3 gene therapy products, including rSIV.F/HN.hCEF.ABCA3.

Chapter 6 Discussion

Genetic therapies have the potential to treat or even cure a variety of disease indications with high unmet clinical needs. The aim of this thesis was to investigate the development of a lentiviral gene therapy for the treatment of severe ABCA3-mediated surfactant dysfunction in babies for whom there is currently no treatment. In Chapter 3, rLV vectors expressing ABCA3 were generated and the efficacy of one of the candidate vectors, rSIV.F/HN.hCEF.ABCA3, was subsequently evaluated in models of ATII cells (Chapter 4) before *in vivo* administration to BALB/c mice in Chapter 5. A conditional KO murine model of ABCA3 deficiency was also investigated.

6.1 Utility of a conditional *Abca3* knockout mouse for evaluating new therapeutics

Mouse disease models can be helpful in the development of new gene therapeutics in order to demonstrate proof-of-principle *in vivo*. A review of the literature (Section 1.4.2) indicated that a conditional *Abca3* KO model was most suitable for modelling lethal ABCA3-mediated surfactant dysfunction, whilst still allowing for evaluation of potential gene therapies. Whereas constitutive *Abca3* KO mouse models appear to closely mimic the ABCA3-deficient patient population that present with RDS at birth (Section 1.4.2.1.1), in mice there is an accelerated timeline for disease progression from birth to death (hours versus weeks [91]) which makes constitutive *Abca3*-KO mice impractical for evaluating therapeutics which are not of immediate benefit.

Studies described in Chapter 5 aimed to induce lung disease in a conditional *Abca3* KO murine model of ABCA3-mediated surfactant dysfunction similar to that generated by Rindler *et al.* [128]. In general, the parenchymal region of murine lungs is considered to be a good model of human alveoli as it contains a similar distribution of ATI and ATII cells in comparison with the more proximal airways where the distribution of submucosal glands, basal cells, cartilage and pseudostratified epithelium varies significantly between the two species [255]. Importantly for this thesis, ATII cells in both mouse and human are considered comparable in terms of appearance, gene expression

patterns and response to injury [256]. In addition, for the F/HN pseudotyped rLV vector used in this study, ATII cells in humans and mice both express alpha2,3-sialyated receptors [257] which is the preferred receptor subtype for HN-mediated vector cell entry [257].

One key difference between mice and humans is the response to respiratory distress. In babies experiencing RDS, supplemental oxygen can be provided via nasal cannulas or, more invasively, by intubation. Although intubation of mice is also possible [258], this is typically only done for brief periods (e.g. to support ventilation during surgery) as the anatomically smaller size of the respiratory system makes this challenging and raises ethical concerns regarding animal wellbeing if used for prolonged periods. As prolonged intubation of constitutive *Abca3*-KO mice would likely be required to allow therapeutics sufficient time to take effect, the ability to temporally control disease induction in conditional models of *Abca3*-KO can be viewed as akin to human intervention (oxygen therapy) in RDS patients prior to administration of therapeutics.

A significant challenge to working with the tamoxifen-inducible *Abca3*-iKO mouse colony (Section 5.2.2) was the consistent activation of the Cre recombinase system to ablate *Abca*3. Despite testing multiple dosing regimens for delivery of tamoxifen to the mice (Section 5.2.4), a noticeable change in respiratory phenotype was rarely observed and only modest levels of Cre recombination could be demonstrated. Attempts to increase the tamoxifen dosage resulted in drug-induced toxicity instead of enhanced disease induction (Figure 5.14), emphasising that in conditional models the type and bioavailability of the drug used for disease induction can be an important confounding factor in subsequent phenotypic observations. The results from the disease induction studies also suggested that male mice might undergo more efficient tamoxifen-induced Cre recombination (Section 5.2.4). This was unexpected although, recently, sex-specific effects of Cre recombinase expression on mouse behaviour have been reported in neuronal mouse lines [259, 260] and Besnard *et al.* [127] noted that postnatal survival of doxycycline-induced *Abca3*$^{\Delta/\Delta}$ mice was improved in females. Furthermore, sex has been described as an influencing factor in the responses

of animals to tamoxifen [261, 262]. Of note, however, is that in patients, such sex differences (e.g. in survival or lung disease outcomes) have not been correlated in those with ABCA3-mediated surfactant dysfunction. Based on the findings described in this thesis, a small study to investigate this further was performed by Dr Kamran Miah (Personal communication), using oral gavage for tamoxifen delivery to *Abca3*-iKO male and female mice (n=5 per group). The results showed that all male mice in the study required humane culling prior to the planned end of the experiment, whereas only two females reached that endpoint (Figure 6.1A); these preliminary results support the observations in Section 5.2.4. Analysis of harvested lung tissue from this additional study were suggestive of greater Cre-mediated recombination in *Abca3* in males based on the intensity of the products from endpoint PCR amplification of the floxed *Abca3* region (Figure 6.1B). Sex differences have not been reported previously in studies using the *Abca3*-cKO model [128] although it is unclear whether both sexes were used in the experiments. For tamoxifen-inducible models utilising a modified oestrogen ligand binding domain ($CreER^{T2}$) to regulate Cre recombinase activity, it could be hypothesised that male mice contain fewer endogenous oestrogen receptors to sequester tamoxifen, thus making it more available to bind $CreER^{T2}$, translocate to the nucleus and mediate higher levels of *Abca3* KO. In future studies, it might be informative to measure the serum levels of tamoxifen achieved by alternate murine delivery routes and to correlate this with the degree of Cre-mediated recombination. Although the reason for the observed sex differences in *Abca3*-iKO mice has not yet been confirmed, the preliminary findings (Figure 6.1A) suggest it might be possible to use male mice (in larger group sizes) as a model in future studies.

In addition to the chemical dependence of conditional KO models for disease induction, the stability of the disease state, once induced, is an important factor when evaluating the efficacy of therapeutics. In the *Abca3*-cKO model [128], incomplete KO of *Abca3* in ATII cells resulted in a self-resolving disease phenotype that was supported by subsequent molecular analysis showing restoration of *Abca3* expression after an initial decrease. Given the role ATII cells play in repopulating the alveoli after damage, it is possible that unedited ATII cells have a selective

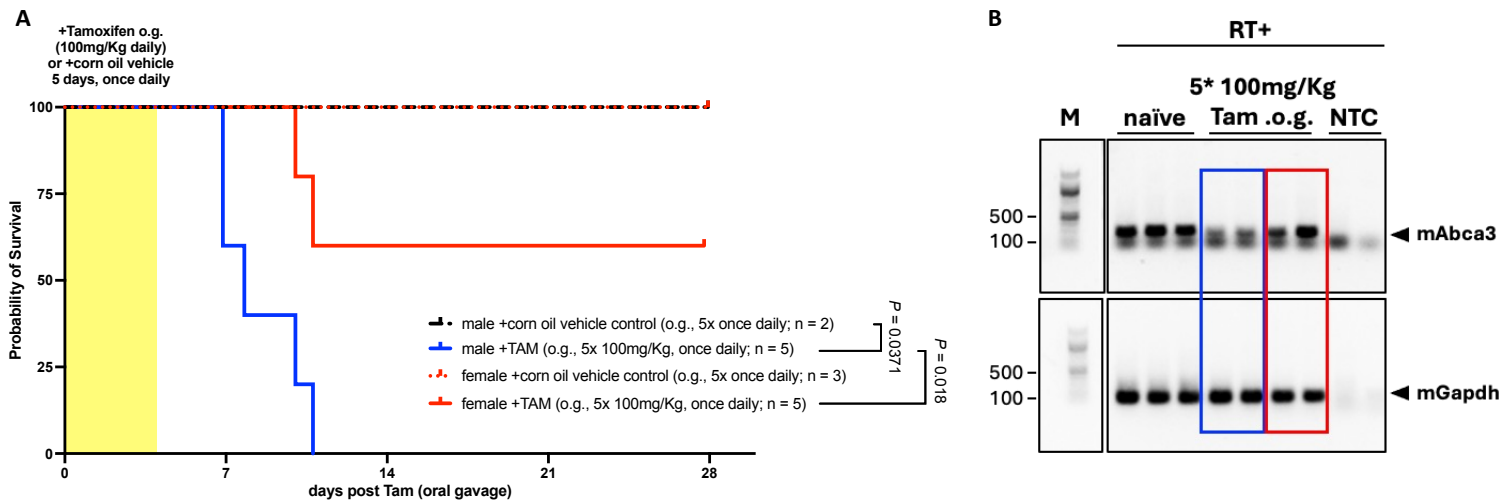


Figure 6.1 *Abca3*-iKO male mice are more susceptible to disease induction with tamoxifen delivered by oral gavage compared with females (Kamran Miah)

Five male and five female *Abca3*-iKO mice were administered tamoxifen by oral gavage (100mg/kg) once daily over 5 consecutive days. Weight, temperature and respiratory phenotype were monitored three times a day for 1 month. Mice losing more than 15% of body weight compared to the first day of tamoxifen dosing, and/or those in severe respiratory distress were culled and lungs harvested. **(A)** Kaplan-Meier curve showing median survival of ~8 days for male *Abca3*-iKO mice, compared with an undefined median survival for female mice. Results were compared with male and female mice administered corn oil (vehicle control). **(B)** RT-PCR analysis of *Abca3* exon 6-7 in RNA extracted from uninduced and tamoxifen-treated *Abca3*-iKO mice (male, blue; female, red). Murine *Gapdh* (mGapdh) was used as a loading control. RNA obtained from mice with a non-transgenic *Abca3* locus (uninduced *Sftpb* KO) was used as a baseline control.

advantage and proceed to replace those ATII cells which do not express *Abca3*. Consequently, if a conditional *Abca3*-KO model was used to demonstrate proof-of-principle for potential new therapies, close attention would need to be paid as to the extent of initial gene KO and self-resolution of the disease to prevent incorrect conclusions about the drug efficacy. In such studies, judicious selection of control treatment groups will be critical – sham treatment and a control vector expressing an irrelevant transgene are encouraged. Together, these factors indicate that conditional *Abca3*-KO mice remain a complex model in which to attempt to demonstrate *in vivo* proof-of-principle for ABCA3 gene therapy.

6.2 The ideal murine model of ABCA3-mediated surfactant dysfunction

Murine models of ABCA3-mediated surfactant dysfunction have so far only been used for improving understanding of disease mechanisms and aetiology. To evaluate potential therapeutics, a disease model should, ideally, not require chemical induction, nor should it have the ability to 'self-resolve' in the absence of intervention (Section 6.1). In addition, a disease phenotype should develop which can be monitored in live animals so that studies are not solely reliant on the molecular analysis of harvested tissues. Murine *Abca3* KO models do not meet these criteria – constitutive *Abca3*-KO mice die too quickly for effective intervention strategies to be used [92] and conditional models require complex regimens of chemical induction of disease phenotypes [127, 128].

A model in which a clinically relevant *ABCA3* mutation is inserted into the murine *Abca3* gene would not only be more representative of disease aetiology in humans (in which gene mutation, rather than gene KO occurs) but may result in a more consistent and persistent respiratory phenotype amenable to reversal by gene therapeutics. Furthermore, if in addition to the loss of functional ABCA3 protein the existence/intracellular accumulation of mutated ABCA3 protein plays a part in disease presentation, KI models of specific *ABCA3* mutations rather than complete gene KO would be more relevant. Currently, the only published *Abca3* KI mouse model incorporates the E292V

ABCA3 missense allele [131] (described in Section 1.4.2.2); although E292V is of clinical importance as one of the most common *ABCA3* mutations [88], its introduction into mice did not appear to result in a respiratory disease phenotype *in vivo*. Small, electron dense LBs and age-dependent lung inflammation could be seen by molecular analysis, but the mice had a normal life span and did not present with phenotypically altered behaviour [131]. In patients, E292V is classified as a type II (functional deficit) mutation and is generally associated with chILD (milder disease) [70]. It could be hypothesised that introducing a type I *ABCA3* mutation, more commonly associated with RDS in patients, would result in a respiratory phenotype in mice. The L101P mutation was one of the first *ABCA3* mutations to be associated with surfactant dysfunction, having been described in the seminal paper of Shulenin *et al.* in 2004 [69]. Since then, it has been characterised in multiple *in vitro* models where it has been associated with misformed LBs [210] and ER stress [102] – making it a potentially good candidate for generating a murine model. However, although there is significant protein homology between murine Abca3 and human ABCA3, L101P falls within a region of relatively poor homology and the leucine residue is not conserved in rodents (phenylalanine) [69]. As a consequence, mutation of leucine residue 101 to a proline in mice may not impact protein function in the same way it does in humans. Modelling a type I mutation which lies in a conserved region of ABCA3 may improve the translation of a disease phenotype and efficacy data for therapeutics between murine models and humans.

One concern with modelling type I *ABCA3* mutations is the associated disease severity in mice. Whilst a recognisable respiratory disease phenotype in mice is desirable, a balance must be found between a modifiable phenotype and survivability. For rLV (F/HN) gene therapy in particular, work with reporter transgenes suggests that expression peaks one week after administration before plateauing [139] – indicating that any mouse model must be viable for at least a week after birth to allow expression from the rLV transgene to be established. Furthermore, whilst it is possible to intranasally administer substances to mice before weaning (approximately four weeks) this carries higher risks such as rejection and/or cannibalisation of the pup by the mother after human

handling. Finally, for a transgenic line to be sustainable the mice need to be capable of reproduction. If homozygous mutant *Abca3* mice are not alive long enough to reproduce then pups must be bred from heterozygous parents, which introduces additional complexities to colony management. If pups are born according to Mendelian frequency, only 25% of the offspring generated by the breeding scheme would be useful in survival studies and up to 75% of the offspring may have to be culled without providing scientifically useful information.

An alternative to using mice homozygous for type I *Abca3* mutations for the evaluation of therapeutics may be the development of a type I/type II compound heterozygous mouse model. The residual *Abca3* expression and function arising from the type II mutant allele may allow mice to survive the neonatal period whilst, coupled with a type I mutation on the second allele, should result in <50% expression of Abca3 compared with WT mice. Careful selection of the specific *ABCA3* mutations to model would be critical to the success of this approach. Identification of compound heterozygous mutations in *ABCA3,* that have been identified in patients and lie within conserved sequences of the murine gene, would be necessary and functional characterisation of each mutation using *in vitro* models would likely be needed prior to generating to transgenic mice.

6.3 Improved animal models for respiratory diseases

Modelling diseases in rodents, and subsequently demonstrating efficacy and safety for novel therapeutics, has for many years been a requirement of the drug development pipeline. Up until December 2022, the FDA required all potential new medicines to be tested for safety and efficacy in animals, by law, before a clinical trial in humans could be approved [263]. Likewise, in the UK, the Medicine and Healthcare products Regulatory Agency (MHRA) generally require animal testing to provide essential data about the safety of a potential new medicine before it is used in human clinical trials. Typically, for small molecule drugs, data are obtained from one rodent and one non-rodent species. However, for Advanced Therapy Medicinal Products (ATMPs) such as cell and gene therapies, one animal species may be sufficient provided an acceptable risk assessment can be

provided. With the advancement of alternative, multi-cellular models of human origin, such as 'organ-on-a-chip' and organoid models (Section 1.4.1), there is renewed debate on the necessity of animal testing. In particular, the utility of rodent models has been brought into question [264] due to the discrepancies between human and rodent physiology. In the development of therapeutics for CF, it was initially assumed that murine models would be suitable (and a requirement) for modelling CF lung disease as mice have a *Cftr* gene with 78% amino acid homology to human *CFTR* [265]. However, whilst the multiple KO and mutation models that were developed recapitulated some features of CF, key characteristics, namely spontaneous lung disease and mucus obstruction, are absent from mouse models [265]. These differences are attributed, in part, to a variation in the distribution and abundance of submucosal glands in mice and, also, the presence of alternative, compensatory mechanisms for Cftr which do not exist in humans. Due to these differences between human disease and murine models, many gene therapy trials for CF have been approved with the assistance of only single species *in vivo* safety data and in the absence of murine proof-of-principle efficacy data. In addition to demonstrating that a murine disease model is not essential for progression of new therapeutics, patient data have also demonstrated the differential effect drugs can have in mice and humans – the CF modulator drug ivacaftor is ineffective in a murine model of CF but is effective in humans [266].

In contrast to rodents, large animals better reflect the anatomy and physiology of humans and as such can offer more relevant information for clinical translation [267]. For respiratory research, sheep and young pigs are favoured due to the similarities in lung structure, size and cell types compared to humans [268]. In addition, the breathing patterns (respiratory frequency and breathing volume) in these animals allow for evaluation of drug deposition when therapeutics are administered either by inhalation or breathing tube, as would most commonly be used clinically. The distribution of a drug following direct administration into the lungs can vary with changes in breathing patterns, and is itself influenced greatly by body size; thus the pattern of drug deposition in rodent lungs is unlikely to reflect that of patients [269]. Mice are also preferential nose-breathers

however it has been shown that inhalation through the mouth increases the proportion of drug which reach the airways [270]. For the treatment of surfactant dysfunction that presents with refractory RDS at birth, it is envisioned that a gene therapeutic could be delivered directly into the lungs of newborn babies via an intubation tube. This delivery route has been mimicked in WT piglets, to improve understanding of how vector dose and delivery volume can affect transduction efficiency within the lung (Gerry McLachlan, Personal communication).

Large animals have also been used to directly model respiratory disease, either by exposure to exogenous substances (e.g. to induce asthma [271]) or by genome manipulation to generate transgenic animals [272]. After KO of *Cftr* in mice failed to recapitulate key features of CF disease, models were generated in sheep [272] and pigs [273]. Unlike murine models of CF, in which the CFTR potentiator ivacaftor had no effect [266], lambs mutated to express the common F508del mutation respond to CFTR correctors and potentiators [272] supporting the view that large animal models can have a role to play in drug development. Compared with rodents, large animals are more costly, more difficult to handle and require specialised facilities for husbandry [267] - in a porcine model of CF an intensive care unit had to be established as the manifestation of CF disease was very severe [274]. However, large animal models could be considered for the development of therapeutics for ABCA3-mediated surfactant dysfunction, as an additional advantage over rodent models would be the ability to intubate and ventilate the animal in a similar manner to humans, thereby modelling the drug delivery route in patients.

6.4 Alternatives to animal models

Ethically, the use of 'farm animals' like sheep and pigs in medical research is more widely accepted than using non-human primates and companion animals such as dogs. However, the principle of the 3Rs to replace, reduce and refine the use of animals in research [129] dictates that alternatives to animal models must be explored. Conventional, *in vitro* (cellular) models are a good starting place. However, many cell models utilise immortalised cell lines which are considered a poor replica

for the human body response due to the often-perturbed nature of cancer-derived cells, differences in growth (culture) conditions and the unicellular format which prevents paracrine signalling between cell types (Section 1.4.1). The A549 and H441 human lung cell lines explored in Chapter 4 highlight one of the strategies which can be attempted to increase the relevance of immortalised cell lines. Culturing cells at the ALI can mimic the exposure of lung cells to air *in vivo* and is reflected in changes to the transcriptome [235].

Primary cells obtained from humans are usually considered a good alternative to animal models, however there are limitations in sourcing and maintaining primary ATII cells in culture (Section 4.3.1) as demonstrated in Figure 4.2. Interestingly, the commercial source of primary alveolar pneumocytes (Epithelix) now also sell a primary alveolar ALI model (AlveolAir) which they claim maintains a stable ratio of ATI and ATII cells for at least a two week period [275]. AlveolAir may be worth investigating as a model, although two weeks is still a very short timeframe for experimentation. The inclusion of endothelial cells seeded on the basolateral surface might release factors to support the ATII phenotype - a similar strategy to that employed for maintenance of iPSC-derived ATII cells [114, 115] before feeder-free culture conditions were established . This could be tested by the addition of the drug cocktail used to maintain iATII cells [108] to the AlveolAir media. In the context of rLV development for ABCA3-mediated surfactant dysfunction, AlveolAir cultures could help in optimising vector design for improved transduction and transgene expression. For example, does alteration of the producer plasmid ratios during rLV vector production improve the performance of the final rLV, perhaps by altering expression of the F and HN envelope proteins and the transgene (vGM336 vs vGM438), affecting the number of transduced cells? Likewise, transduction of AlveolAir cultures with rLV vectors containing the alveolar promoters (evaluated in mice in Section 3.2.4) could be used to investigate whether these elements are equally active in human ATI and ATII cells.

Cultures of the AlveolAir model, or similar primary cells, are unlikely to be useful for modelling surfactant dysfunction disorders due to the rarity of patient samples and the requirement of cells to be in submerged culture for most gene editing approaches. In contrast, iPSC-derived ATII cells (Section 1.4.1) have a greater potential for disease modelling as the cells can be continuously cultured (under appropriate conditions) and gene edited prior to differentiation to an ATII cell lineage. Furthermore, if patients carrying mutations of interest can be identified, obtaining fibroblasts for iATII cell derivation is simpler and less invasive than collecting primary ATII cells directly. A recent publication by Sun *et al.* [109] describes the generation and characterisation of iATII cells obtained from patients with *ABCA3* mutations and by gene editing WT iPSC lines. Lipidomic analysis of apical supernatant demonstrated a reduction in PC species, including DPPC, in the ABCA3 mutant iATII cells which was partially reversed in CRISPR-corrected iATII cells. A reduction in PC species is also observed in patients with ABCA3-mediated surfactant dysfunction – demonstrating the translatability of the iATII model and furthermore, the feasibility of restoring surfactant lipid transport by increasing expression of WT ABCA3. In ABCA3 mutant iATII cells, smaller LBs, upregulation of the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB) inflammatory pathway and a reduced proliferation capacity were also observed [109]. Whilst changes in LB structure and size have previously been reported in patients [69], and can be replicated in some immortalised cell models [210], the reduction in proliferation capacity has not been reported in humans and has only recently been suggested in the *Abca3*-cKO mouse model developed by Rindler *et al.* [128]. Likewise, the link between *ABCA3* mutations and increased canonical NF-kB signalling has not been previously appreciated [109]. If this arises from a mutant ABCA3 protein with a 'gain-of-function' then this may have important implications for gene therapy approaches to treating ABCA3-mediated surfactant dysfunction. Expression of the mutant ABCA3 protein may need to be reduced to limit downstream effects of prolonged inflammatory signalling, which can include the development of lung fibrosis [276]. As such, it is likely that iATII cells

expressing mutant ABCA3 variants, maintained in a 2D ALI culture, represent a highly relevant model for investigating efficacy of gene therapeutic strategies targeting *ABCA3*.

6.5 Assays to assess the expression and function of ABCA3

One challenge experienced in the studies reported here was the availability of high quality anti-ABCA3 antibody reagents (Table 2.5), which were polyclonal and could not distinguish between murine and human orthologs. Moreover, there appeared to be a trade-off between antibody sensitivity and non-specific binding that was noticeable by western blotting (Section 3.2.2.1). The use of monoclonal antibodies would likely reduce non-specific binding, however, there are none currently available. Aside from the time-consuming generation of a hybridoma cell line [277], adding a protein tag, such as FLAG (DYKDDDDK) or GFP, to the C-terminal of ABCA3 should allow for more specific protein detection. Although tagging membrane bound proteins can alter protein processing and localisation, and is best avoided when using potential clinical vectors in pre-clinical studies, tagged ABCA3 has been used successfully [29, 42]. An alternative method to increase specific detection of ABCA3 would be to use digital proximity ligation assay (dPLA) [278] where two primary antibodies (raised in different species and targeting different epitopes) are used with two secondary antibodies tagged with oligonucleotide dPLA probes. This approach has been used successfully with antibodies targeting CFTR (Anthony Sinadinos, Personal communication) but remains problematic for ABCA3 as the majority of anti-ABCA3 antibodies are raised in rabbits and target the first ECD (Figure 1.5C).

In addition to demonstrating expression of ABCA3, an assay for functional ABCA3 was required. ABCA3 acts predominantly as an intracellular lipid transporter (Section 1.2.3), but in a pathophysiologic setting (e.g. cancer) it can mediate multi-drug resistance due to its ability to transport cytotoxic drugs [231]. In Chapter 4, doxorubicin detoxification (Section 2.3.5), similar to that reported by Wittmann *et al*. [230], was used as an indirect measure of ABCA3 function in H441 cells expressing ABCA3 variants. Cells with known pathogenic mutations in *ABCA3* were less viable

after doxorubicin exposure, indicating attenuated ABCA3 function compared with WT control. As hypothesised, significant differences were observed between the ABCA3 variant cell pools (Figure 4.13), and cells transduced with rSIV.F/HN.hCEF.ABCA3 (Figure 4.14). However, the changes were modest and also likely affected by other extracellular transporters and cellular drug evasion mechanisms. The sequestration of cytotoxic drugs may not be a physiological role of ABCA3 but has also been observed for other ABC type proteins [279].

An alternative, though also indirect, method for assessing ABCA3 function is imaging of LB morphology by TEM. When *ABCA3* is mutated, smaller LBs and/or LBs reminiscent of a 'fried egg' have been observed (Figure 1.6). This was described in Chapter 4, when considering which cells would be appropriate for the generation of an *in vitro* model of ABCA3-mediated surfactant dysfunction. Lamellar bodies can also be identified with the lipophilic dye Nile-Red which stains intracellular lipid droplets [280], but this is not specific for LBs and does not provide the same level of detail obtained by TEM analysis. Interestingly, whilst LB images are frequently included in patient case reports [69] and characterisation of disease models [109], they are rarely quantifiable and the morphology is highly subjective to user interpretation. A high degree of natural variation in LB morphology already exists in the absence of disease and can be observed in Figure 4.1 and 4.5. Moreover, as perturbations in LB morphology are not observed with all *ABCA3* mutations, and the organisation of lipids within LBs is largely dependent on SP-B [27], correction of LB morphology by a potential therapeutic may not be representative of fully functional ABCA3.

Alternative assays of ABCA3 function focus on transport of phospholipids, [242, 281], which are the most abundant component in surfactant (Figure 1.2) and a known ABCA3 substrate. Labelling of phospholipid species has been used to track their intracellular localisation and quantify changes associated with mutations in *ABCA3* or with potential therapeutic interventions. Radiolabelled-PC has been used [282] with safer fluorescent [281] and metabolic [242] labelling approaches now emerging. TopFluor-PC is a fluorescently labelled PC analogue which can be incorporated into

surfactant-like liposomes and has been demonstrated to be taken up into ABCA3+ vesicles in A549 cells [281]. Using this approach Kinting *et al.* [157] and Wambach *et al.* [210] both demonstrated that expression of L101P and E292V *ABCA3* mutations in A549 cells reduced the amount of TopFluor-PC-labelled liposomes inside ABCA3+ vesicles. Alternatively, propargyl-choline, a choline analogue which fluoresces after a click chemistry reaction, can be used to localise endogenously synthesised PC and PC-lipid derivatives [242] in the cytosol of ATII cells [283, 284]. These approaches could be used to measure the efficacy of potential therapeutics, using *in vitro* disease models such as iPSC lines that express GFP-tagged ABCA3 variants from the endogenous *ABCA3* loci [109].

The assays described above, and the doxorubicin detoxification assay used in Chapter 4, are indirect measures of ABCA3 function. There is also an assumption that the 'correction' of LB morphology should reflect correction of dysregulated surfactant composition and homeostasis. Ideally, assays that can directly quantify surfactant release from LBs onto the cell surface should be developed; adaptation of a surfactant/lipid release assay using lipophilic fluorescent dye FM-64 might be helpful here [109]). In addition, lipidomic and proteomic analysis will be required to more fully understand the impact of novel therapeutics targeting ABCA3.

6.5 Conclusion and future directions

There are currently no targeted therapies for treating surfactant dysfunction disorders arising from mutations in *ABCA3*. Here, for the first time a viral vector encoding human ABCA3 has been generated (rSIV.F/HN.hCEF.ABCA3), exploiting the large LV packaging capacity to accommodate sequences for expression of full-length ABCA3. The utility of a conditional *Abca3* KO mouse was explored for its potential as an ABCA3 deficiency disease model but requires further work to refine disease induction and disease characterisation. Importantly, animal models that can usefully mimic human disease are not always available to provide the required pre-clinical data for clinical translation; thus, the requirement for pre-clinical efficacy studies in animal models needs more

consideration. Discussions with regulators will be required to debate the relative weight attributed to data obtained from animal models compared with human cell culture systems. Ultimately, to support progression of a gene therapeutic for ABCA3-mediated surfactant dysfunction, well-characterised *in vitro* cell models may have to be sufficient to provide proof-of-principle pre-clinical efficacy. Overall, rSIV.F/HN.hCEF.ABCA3 represents an exciting opportunity to meet a currently unmet medical need in newborns affected by this condition. A key question will be the demonstration that restoration of ABCA3 expression, and its function, can benefit surfactant homeostasis.

References

1. Knudsen, L. and M. Ochs, *The micromechanics of lung alveoli: structure and function of surfactant and tissue components.* Histochem Cell Biol, 2018. **150**(6): p. 661-676.
2. Ochs, M., et al., *The number of alveoli in the human lung.* Am J Respir Crit Care Med, 2004. **169**(1): p. 120-4.
3. Frohlich, E., et al., *Measurements of Deposition, Lung Surface Area and Lung Fluid for Simulation of Inhaled Compounds.* Front Pharmacol, 2016. **7**: p. 181.
4. Dunhill, M.S., *Postnatal Growth of the Lung.* Thorax, 1962. **17**(4): p. 329-333.
5. Herring, M.J., et al., *Growth of alveoli during postnatal development in humans based on stereological estimation.* Am J Physiol Lung Cell Mol Physiol, 2014. **307**(4): p. L338-44.
6. Low, F.N., *The pulmonary alveolar epithelium of laboratory mammals and man.* Anat Rec, 1953. **117**(2): p. 241-63.
7. Weibel, E.R., *On the tricks alveolar epithelial cells play to make a good lung.* Am J Respir Crit Care Med, 2015. **191**(5): p. 504-13.
8. Crapo, J.D., et al., *Cell number and cell characteristics of the normal human lung.* Am Rev Respir Dis, 1982. **126**(2): p. 332-7.
9. Fehrenbach, H., *Alveolar epithelial type II cell: defender of the alveolus revisited.* Respir Res, 2001. **2**(1): p. 33-46.
10. Evans, M.J., et al., *Transformation of alveolar type 2 cells to type 1 cells following exposure to NO2.* Exp Mol Pathol, 1975. **22**(1): p. 142-50.
11. Clements, J.A., *Lung surfactant: a personal perspective.* Annu Rev Physiol, 1997. **59**: p. 1-21.
12. Mason, R.J. and M.C. Williams, *Type II alveolar cell. Defender of the alveolus.* Am Rev Respir Dis, 1977. **115**(6 Pt 2): p. 81-91.
13. Bastacky, J., et al., *Alveolar lining layer is thin and continuous: low-temperature scanning electron microscopy of rat lung.* J Appl Physiol (1985), 1995. **79**(5): p. 1615-28.
14. Andreev, K., et al., *Structural Changes in Films of Pulmonary Surfactant Induced by Surfactant Vesicles.* Langmuir, 2020. **36**(45): p. 13439-13447.
15. Ma, C.C. and S. Ma, *The role of surfactant in respiratory distress syndrome.* Open Respir Med J, 2012. **6**: p. 44-53.
16. Avery, M.E. and J. Mead, *Surface properties in relation to atelectasis and hyaline membrane disease.* AMA J Dis Child, 1959. **97**(5, Part 1): p. 517-23.
17. Wright, J.R. and J.A. Clements, *Metabolism and turnover of lung surfactant.* Am Rev Respir Dis, 1987. **136**(2): p. 426-44.
18. Perez-Gil, J., *Structure of pulmonary surfactant membranes and films: the role of proteins and lipid-protein interactions.* Biochim Biophys Acta, 2008. **1778**(7-8): p. 1676-95.
19. Batenburg, J.J. and H.P. Haagsman, *The lipids of pulmonary surfactant: dynamics and interactions with proteins.* Prog Lipid Res, 1998. **37**(4): p. 235-76.
20. Weaver, T.E. and J.J. Conkright, *Function of surfactant proteins B and C.* Annu Rev Physiol, 2001. **63**: p. 555-78.
21. Weaver, T.E. and J.A. Whitsett, *Function and regulation of expression of pulmonary surfactant-associated proteins.* Biochem J, 1991. **273(Pt 2)**(Pt 2): p. 249-64.
22. Pastva, A.M., J.R. Wright, and K.L. Williams, *Immunomodulatory roles of surfactant proteins A and D: implications in lung disease.* Proc Am Thorac Soc, 2007. **4**(3): p. 252-7.
23. Wright, J.R., *Immunoregulatory functions of surfactant proteins.* Nat Rev Immunol, 2005. **5**(1): p. 58-68.
24. Motwani, M., et al., *Mouse surfactant protein-D. cDNA cloning, characterization, and gene localization to chromosome 14.* J Immunol, 1995. **155**(12): p. 5671-7.

25. Watson, A., J. Madsen, and H.W. Clark, *SP-A and SP-D: Dual Functioning Immune Molecules With Antiviral and Immunomodulatory Properties.* Front Immunol, 2020. **11**: p. 622598.
26. Weaver, T.E., C.L. Na, and M. Stahlman, *Biogenesis of lamellar bodies, lysosome-related organelles involved in storage and secretion of pulmonary surfactant.* Semin Cell Dev Biol, 2002. **13**(4): p. 263-70.
27. Hobi, N., et al., *A small key unlocks a heavy door: The essential function of the small hydrophobic proteins SP-B and SP-C to trigger adsorption of pulmonary surfactant lamellar bodies.* Biochim Biophys Acta, 2016. **1863**(8): p. 2124-34.
28. Wang, Z., et al., *Differential activity and lack of synergy of lung surfactant proteins SP-B and SP-C in interactions with phospholipids.* J Lipid Res, 1996. **37**(8): p. 1749-60.
29. Beers, M.F. and S. Mulugeta, *The biology of the ABCA3 lipid transporter in lung health and disease.* Cell Tissue Res, 2017. **367**(3): p. 481-493.
30. Veldhuizen, R., et al., *The role of lipids in pulmonary surfactant.* Biochim Biophys Acta, 1998. **1408**(2-3): p. 90-108.
31. Yamano, G., et al., *ABCA3 is a lamellar body membrane protein in human lung alveolar type II cells.* FEBS Lett, 2001. **508**(2): p. 221-5.
32. Conkright, J.J., et al., *Secretion of surfactant protein C, an integral membrane protein, requires the N-terminal propeptide.* J Biol Chem, 2001. **276**(18): p. 14658-64.
33. Stahlman, M.T., et al., *Lamellar body formation in normal and surfactant protein B-deficient fetal mice.* Lab Invest, 2000. **80**(3): p. 395-403.
34. Gower, W.A., et al., *Fatal familial lung disease caused by ABCA3 deficiency without identified ABCA3 mutations.* J Pediatr, 2010. **157**(1): p. 62-8.
35. Whitsett, J.A., S.E. Wert, and T.E. Weaver, *Alveolar surfactant homeostasis and the pathogenesis of pulmonary disease.* Annu Rev Med, 2010. **61**: p. 105-19.
36. Ryan, U.S., J.W. Ryan, and D.S. Smith, *Alveolar type II cells: studies on the mode of release of lamellar bodies.* Tissue Cell, 1975. **7**(3): p. 587-99.
37. Wirtz, H.R. and L.G. Dobbs, *Calcium mobilization and exocytosis after one mechanical stretch of lung epithelial cells.* Science, 1990. **250**(4985): p. 1266-9.
38. Han, S. and R.K. Mallampalli, *The Role of Surfactant in Lung Disease and Host Defense against Pulmonary Infections.* Ann Am Thorac Soc, 2015. **12**(5): p. 765-74.
39. Stahlman, M.T., et al., *Expression of ABCA3 in developing lung and other tissues.* J Histochem Cytochem, 2007. **55**(1): p. 71-83.
40. Connors, T.D., et al., *The cloning of a human ABC gene (ABC3) mapping to chromosome 16p13.3.* Genomics, 1997. **39**(2): p. 231-4.
41. Nagata, K., et al., *Human ABCA3, a product of a responsible gene for abca3 for fatal surfactant deficiency in newborns, exhibits unique ATP hydrolysis activity and generates intracellular multilamellar vesicles.* Biochem Biophys Res Commun, 2004. **324**(1): p. 262-8.
42. Hofmann, N., et al., *Analysis of the Proteolytic Processing of ABCA3: Identification of Cleavage Site and Involved Proteases.* PLoS One, 2016. **11**(3): p. e0152594.
43. Schaller-Bals, S., et al., *Surface-expressed lamellar body membrane is recycled to lamellar bodies.* Am J Physiol Lung Cell Mol Physiol, 2000. **279**(4): p. L631-40.
44. Onnee, M., et al., *Structure-Based Understanding of ABCA3 Variants.* Int J Mol Sci, 2021. **22**(19).
45. Kos, V. and R.C. Ford, *The ATP-binding cassette family: a structural perspective.* Cell Mol Life Sci, 2009. **66**(19): p. 3111-26.
46. Xie, T., et al., *Cryo-EM structures of the human surfactant lipid transporter ABCA3.* Sci Adv, 2022. **8**(14): p. eabn3727.
47. Walker, J.E., et al., *Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold.* EMBO J, 1982. **1**(8): p. 945-51.

48. Dean, M., Y. Hamon, and G. Chimini, *The human ATP-binding cassette (ABC) transporter superfamily.* J Lipid Res, 2001. **42**(7): p. 1007-17.
49. Zimmermann, L.J., et al., *Surfactant metabolism in the neonate.* Biol Neonate, 2005. **87**(4): p. 296-307.
50. Chakraborty, M., Kotecha, S., *Pulmonary surfactant in newborn infants and children.* Breathe, 2013. **9**(6): p. 476-488.
51. NHS. *Newborn respiratory distress syndrome*. 2021 [cited 2024 08/10/2024]; Available from: https://www.nhs.uk/conditions/neonatal-respiratory-distress-syndrome/.
52. Boyle, E.M., Roehr, C.C., *Optimising the management of respiratory distress in late preterm and early term babies.* Infant, 2022. **18**(4): p. 137-141.
53. Miao, J., et al., *Continuous positive pressure ventilation combined with pulmonary surfactant in the treatment of neonatal respiratory distress syndrome.* Pak J Med Sci, 2020. **36**(4): p. 647-651.
54. Farnaceutici, C. *Curosurf (poractant alfa) Intratracheal Suspension*. 2020 [cited 2024 08/10/2024]; Available from: https://curosurf.com/.
55. Sardesai, S., et al., *Evolution of surfactant therapy for respiratory distress syndrome: past, present, and future.* Pediatr Res, 2017. **81**(1-2): p. 240-248.
56. Nogee, L.M., *Genetic causes of surfactant protein abnormalities.* Curr Opin Pediatr, 2019. **31**(3): p. 330-339.
57. Hamvas, A., et al., *Surfactant protein B deficiency: antenatal diagnosis and prospective treatment with surfactant replacement.* J Pediatr, 1994. **125**(3): p. 356-61.
58. Hentschel, R., et al., *Surfactant replacement therapy: from biological basis to current clinical practice.* Pediatr Res, 2020. **88**(2): p. 176-183.
59. Kurath-Koller, S., et al., *Surfactant Protein B Deficiency Caused by Homozygous C248X Mutation-A Case Report and Review of the Literature.* AJP Rep, 2015. **5**(1): p. e53-9.
60. Benden, C., *Pediatric lung transplantation.* J Thorac Dis, 2017. **9**(8): p. 2675-2683.
61. Children, G.O.S.H.f., *Imogen's double lung transplant*. 2016, GOSH, NHS Foundation Trst: https://www.gosh.nhs.uk/press-releases/imogens-double-lung-transplant/#:~:text=Imogen%20was%20only%20five%20months,the%20operation%20in%20the%20UK.
62. Eldridge, W.B., et al., *Outcomes of Lung Transplantation for Infants and Children with Genetic Disorders of Surfactant Metabolism.* J Pediatr, 2017. **184**: p. 157-164 e2.
63. Kuo, C.S. and L.R. Young, *Interstitial lung disease in children.* Curr Opin Pediatr, 2014. **26**(3): p. 320-7.
64. Chen, F., et al., *Case Report: Report of Two Cases of Interstitial Lung Disease Caused by Novel Compound Heterozygous Variants in the ABCA3 Gene.* Front Genet, 2022. **13**: p. 875015.
65. Chua, W.C., et al., *Congenital Surfactant C Deficiency with Pulmonary Hypertension-A Case Report.* Children (Basel), 2022. **9**(10).
66. Salerno, T., et al., *Surfactant Protein C-associated interstitial lung disease; three different phenotypes of the same SFTPC mutation.* Ital J Pediatr, 2016. **42**: p. 23.
67. Avital, A., et al., *Natural history of five children with surfactant protein C mutations and interstitial lung disease.* Pediatr Pulmonol, 2014. **49**(11): p. 1097-105.
68. Nogee, L.M., et al., *Brief report: deficiency of pulmonary surfactant protein B in congenital alveolar proteinosis.* N Engl J Med, 1993. **328**(6): p. 406-10.
69. Shulenin, S., et al., *ABCA3 gene mutations in newborns with fatal surfactant deficiency.* N Engl J Med, 2004. **350**(13): p. 1296-303.
70. Bullard, J.E., et al., *ABCA3 mutations associated with pediatric interstitial lung disease.* Am J Respir Crit Care Med, 2005. **172**(8): p. 1026-31.
71. Nogee, L.M., et al., *A mutation in the surfactant protein C gene associated with familial interstitial lung disease.* N Engl J Med, 2001. **344**(8): p. 573-9.

72. Wert, S.E., J.A. Whitsett, and L.M. Nogee, *Genetic disorders of surfactant dysfunction.* Pediatr Dev Pathol, 2009. **12**(4): p. 253-74.
73. Cole, F.S., et al., *Population-based estimates of surfactant protein B deficiency.* Pediatrics, 2000. **105**(3 Pt 1): p. 538-41.
74. Nogee, L.M., et al., *Allelic heterogeneity in hereditary surfactant protein B (SP-B) deficiency.* Am J Respir Crit Care Med, 2000. **161**(3 Pt 1): p. 973-81.
75. Nogee, L.M., et al., *A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds.* J Clin Invest, 1994. **93**(4): p. 1860-3.
76. Beers, M.F., et al., *Pulmonary surfactant metabolism in infants lacking surfactant protein B.* Am J Respir Cell Mol Biol, 2000. **22**(3): p. 380-91.
77. Edwards, V., et al., *Ultrastructure of lamellar bodies in congenital surfactant deficiency.* Ultrastruct Pathol, 2005. **29**(6): p. 503-9.
78. Gower, W.A., S.E. Wert, and L.M. Nogee, *Inherited Surfactant Disorders.* NeoReviews, 2008. **9**(10): p. e458-e467.
79. Vorbroker, D.K., et al., *Aberrant processing of surfactant protein C in hereditary SP-B deficiency.* Am J Physiol, 1995. **268**(4 Pt 1): p. L647-56.
80. Nogee, L.M., *Alterations in SP-B and SP-C expression in neonatal lung disease.* Annu Rev Physiol, 2004. **66**: p. 601-23.
81. Mulugeta, S., et al., *A surfactant protein C precursor protein BRICHOS domain mutation causes endoplasmic reticulum stress, proteasome dysfunction, and caspase 3 activation.* Am J Respir Cell Mol Biol, 2005. **32**(6): p. 521-30.
82. Maguire, J.A., S. Mulugeta, and M.F. Beers, *Multiple ways to die: delineation of the unfolded protein response and apoptosis induced by Surfactant Protein C BRICHOS mutants.* Int J Biochem Cell Biol, 2012. **44**(1): p. 101-12.
83. Stewart, G.A., et al., *4-Phenylbutyric acid treatment rescues trafficking and processing of a mutant surfactant protein-C.* Am J Respir Cell Mol Biol, 2012. **47**(3): p. 324-31.
84. Hawkins, A., et al., *A non-BRICHOS SFTPC mutant (SP-CI73T) linked to interstitial lung disease promotes a late block in macroautophagy disrupting cellular proteostasis and mitophagy.* Am J Physiol Lung Cell Mol Physiol, 2015. **308**(1): p. L33-47.
85. Chibbar, R., et al., *Nonspecific interstitial pneumonia and usual interstitial pneumonia with mutation in surfactant protein C in familial pulmonary fibrosis.* Mod Pathol, 2004. **17**(8): p. 973-80.
86. Wambach, J.A., et al., *Single ABCA3 mutations increase risk for neonatal respiratory distress syndrome.* Pediatrics, 2012. **130**(6): p. e1575-82.
87. Onnee, M., et al., *Instability of Mature ABCA3 Protein: Toward a New Classification of ABCA3 Mutations?* Am J Respir Cell Mol Biol, 2022. **67**(5): p. 602-605.
88. Wambach, J.A., et al., *Genotype-phenotype correlations for infants and children with ABCA3 deficiency.* Am J Respir Crit Care Med, 2014. **189**(12): p. 1538-43.
89. Matsumura, Y., et al., *Characterization and classification of ATP-binding cassette transporter ABCA3 mutants in fatal surfactant deficiency.* J Biol Chem, 2006. **281**(45): p. 34503-14.
90. AlAnazi, A., et al., *The most frequent ABCA3 nonsense mutation -p.Tyr1515* (Y1515X) causing lethal neonatal respiratory failure in a term neonate.* Ann Thorac Med, 2017. **12**(3): p. 213-215.
91. Cheong, N., et al., *ABCA3 is critical for lamellar body biogenesis in vivo.* J Biol Chem, 2007. **282**(33): p. 23811-7.
92. Fitzgerald, M.L., et al., *ABCA3 inactivation in mice causes respiratory failure, loss of pulmonary surfactant, and depletion of lung phosphatidylglycerol.* J Lipid Res, 2007. **48**(3): p. 621-32.
93. Bruder, E., et al., *Ultrastructural and molecular analysis in fatal neonatal interstitial pneumonia caused by a novel ABCA3 mutation.* Mod Pathol, 2007. **20**(10): p. 1009-18.

94. Garmany, T.H., et al., *Surfactant composition and function in patients with ABCA3 mutations.* Pediatr Res, 2006. **59**(6): p. 801-5.
95. Brasch, F., et al., *Alteration of the pulmonary surfactant system in full-term infants with hereditary ABCA3 deficiency.* Am J Respir Crit Care Med, 2006. **174**(5): p. 571-80.
96. Krude, H., et al., *Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency.* J Clin Invest, 2002. **109**(4): p. 475-80.
97. Carre, A., et al., *Five new TTF1/NKX2.1 mutations in brain-lung-thyroid syndrome: rescue by PAX8 synergism in one case.* Hum Mol Genet, 2009. **18**(12): p. 2266-76.
98. Hamvas, A., et al., *Heterogeneous pulmonary phenotypes associated with mutations in the thyroid transcription factor gene NKX2-1.* Chest, 2013. **144**(3): p. 794-804.
99. Guillot, L., et al., *NKX2-1 mutations leading to surfactant protein promoter dysregulation cause interstitial lung disease in "Brain-Lung-Thyroid Syndrome".* Hum Mutat, 2010. **31**(2): p. E1146-62.
100. Dobbs, L.G., *Isolation and culture of alveolar type II cells.* Am J Physiol, 1990. **258**(4 Pt 1): p. L134-47.
101. Maguire, J.A., S. Mulugeta, and M.F. Beers, *Endoplasmic reticulum stress induced by surfactant protein C BRICHOS mutants promotes proinflammatory signaling by epithelial cells.* Am J Respir Cell Mol Biol, 2011. **44**(3): p. 404-14.
102. Weichert, N., et al., *Some ABCA3 mutations elevate ER stress and initiate apoptosis of lung epithelial cells.* Respir Res, 2011. **12**(1): p. 4.
103. Munis, A.M., S.C. Hyde, and D.R. Gill, *A human surfactant B deficiency air-liquid interface cell culture model suitable for gene therapy applications.* Mol Ther Methods Clin Dev, 2021. **20**: p. 237-246.
104. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.
105. Van Haute, L., et al., *Generation of lung epithelial-like tissue from human embryonic stem cells.* Respir Res, 2009. **10**(1): p. 105.
106. Wang, D., et al., *A pure population of lung alveolar epithelial type II cells derived from human embryonic stem cells.* Proc Natl Acad Sci U S A, 2007. **104**(11): p. 4449-54.
107. Jacob, A., et al., *Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells.* Nat Protoc, 2019. **14**(12): p. 3303-3332.
108. Jacob, A., et al., *Differentiation of Human Pluripotent Stem Cells into Functional Lung Alveolar Epithelial Cells.* Cell Stem Cell, 2017. **21**(4): p. 472-488 e10.
109. Sun, Y.L., et al., *Human pluripotent stem cell modeling of alveolar type 2 cell dysfunction caused by ABCA3 mutations.* J Clin Invest, 2024. **134**(2).
110. Silva, S., et al., *Air-liquid interface (ALI) impact on different respiratory cell cultures.* Eur J Pharm Biopharm, 2023. **184**: p. 62-82.
111. Abo, K.M., et al., *Air-liquid interface culture promotes maturation and allows environmental exposure of pluripotent stem cell-derived alveolar epithelium.* JCI Insight, 2022. **7**(6).
112. Ohlinger, K., et al., *Air-liquid interface culture changes surface properties of A549 cells.* Toxicol In Vitro, 2019. **60**: p. 369-382.
113. Loret, T., et al., *Air-liquid interface exposure to aerosols of poorly soluble nanomaterials induces different biological activation levels compared to exposure to suspensions.* Part Fibre Toxicol, 2016. **13**(1): p. 58.
114. Barkauskas, C.E., et al., *Type 2 alveolar cells are stem cells in adult lung.* J Clin Invest, 2013. **123**(7): p. 3025-36.
115. Gotoh, S., et al., *Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells.* Stem Cell Reports, 2014. **3**(3): p. 394-403.
116. Chen, Y.W., et al., *A three-dimensional model of human lung development and disease from pluripotent stem cells.* Nat Cell Biol, 2017. **19**(5): p. 542-549.

117. Meyer-Berg, H., et al., *Identification of AAV serotypes for lung gene therapy in human embryonic stem cell-derived lung organoids.* Stem Cell Res Ther, 2020. **11**(1): p. 448.
118. Meyer-Berg, H., *Gene Therapy for Surfactant Protein B Deficiency using Recombinant AAV Vectors*, in *Radcliffe Department of Medicine*. 2021, University of Oxford. p. 301.
119. Danopoulos, S., J. Shiosaki, and D. Al Alam, *FGF Signaling in Lung Development and Disease: Human Versus Mouse.* Front Genet, 2019. **10**: p. 170.
120. Melton, K.R., et al., *SP-B deficiency causes respiratory failure in adult mice.* Am J Physiol Lung Cell Mol Physiol, 2003. **285**(3): p. L543-9.
121. Glasser, S.W., et al., *Altered stability of pulmonary surfactant in SP-C-deficient mice.* Proc Natl Acad Sci U S A, 2001. **98**(11): p. 6366-71.
122. Kang, M.H., et al., *A lung tropic AAV vector improves survival in a mouse model of surfactant B deficiency.* Nat Commun, 2020. **11**(1): p. 3929.
123. Cook, M.J., *The Anatomy of the Laboratory Mouse*, M.G. Informatics, Editor. 1965, Elsevier: https://www.informatics.jax.org/cookbook/. p. 143.
124. Ban, N., et al., *ABCA3 as a lipid transporter in pulmonary surfactant biogenesis.* J Biol Chem, 2007. **282**(13): p. 9628-9634.
125. Hammel, M., et al., *Targeted inactivation of the murine Abca3 gene leads to respiratory failure in newborns with defective lamellar bodies.* Biochem Biophys Res Commun, 2007. **359**(4): p. 947-51.
126. Herber-Jonat, S., et al., *Abca3 haploinsufficiency is a risk factor for lung injury induced by hyperoxia or mechanical ventilation in a murine model.* Pediatr Res, 2013. **74**(4): p. 384-92.
127. Besnard, V., et al., *Conditional deletion of Abca3 in alveolar type II cells alters surfactant homeostasis in newborn and adult mice.* Am J Physiol Lung Cell Mol Physiol, 2010. **298**(5): p. L646-59.
128. Rindler, T.N., et al., *Alveolar injury and regeneration following deletion of ABCA3.* JCI Insight, 2017. **2**(24).
129. NC3Rs. *The 3Rs*. 2024 [cited 2024 08/10/2024]; Available from: https://nc3rs.org.uk/who-we-are/3rs.
130. Beers, M.F., et al., *Aberrant lung remodeling in a mouse model of surfactant dysregulation induced by modulation of the Abca3 gene.* Ann Anat, 2017. **210**: p. 135-146.
131. Tomer, Y., et al., *The common ABCA3(E292V) variant disrupts AT2 cell quality control and increases susceptibility to lung injury and aberrant remodeling.* Am J Physiol Lung Cell Mol Physiol, 2021. **321**(2): p. L291-L307.
132. Anderson, W.F., *September 14, 1990: the beginning.* Hum Gene Ther, 1990. **1**(4): p. 371-2.
133. Administration, U.F.a.D., *FDA approves innovative gene therapy to treat pediatric patients with spinal muscular atrophy, a rare disease and leading genetic cause of infant mortality*. 2019, FDA: https://www.fda.gov/news-events/press-announcements/fda-approves-innovative-gene-therapy-treat-pediatric-patients-spinal-muscular-atrophy-rare-disease.
134. Agency, E.M., *New gene therapy to treat spinal muscular atrophy (corrected)*. 2020, European Medicines Agency: https://www.ema.europa.eu/en/news/new-gene-therapy-treat-spinal-muscular-atrophy-corrected.
135. Ogbonmide, T., et al., *Gene Therapy for Spinal Muscular Atrophy (SMA): A Review of Current Challenges and Safety Considerations for Onasemnogene Abeparvovec (Zolgensma).* Cureus, 2023. **15**(3): p. e36197.
136. Sun, D., et al., *CAR-T cell therapy: A breakthrough in traditional cancer treatment strategies (Review).* Mol Med Rep, 2024. **29**(3).
137. Alton, E., et al., *Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial.* Lancet Respir Med, 2015. **3**(9): p. 684-691.

138. Strayer, M.S., S.H. Guttentag, and P.L. Ballard, *Targeting type II and Clara cells for adenovirus-mediated gene transfer using the surfactant protein B promoter.* Am J Respir Cell Mol Biol, 1998. **18**(1): p. 1-11.
139. Alton, E.W., et al., *Preparation for a first-in-man lentivirus trial in patients with cystic fibrosis.* Thorax, 2017. **72**(2): p. 137-147.
140. Buck, J., et al., *Lipid-Based DNA Therapeutics: Hallmarks of Non-Viral Gene Delivery.* ACS Nano, 2019. **13**(4): p. 3754-3782.
141. Rosewell, A., Hiatt, P., McConnell, R., Dang, D., Vetrini, F., Grove, N., Palmer, D., Finegold, M., Beadet, A., Ng, P., *277. Efficient Transduction of Alveolar Type II Epithelial Cells in Nonhuman Primates by Helper-Dependent Adenovirus 5-Fiber 35 Chimeric Vectors.* Molecular Therapy, 2011. **19**: p. S108.
142. Bulcha, J.T., et al., *Viral vector platforms within the gene therapy landscape.* Signal Transduct Target Ther, 2021. **6**(1): p. 53.
143. Dong, B., H. Nakai, and W. Xiao, *Characterization of genome integrity for oversized recombinant AAV vector.* Mol Ther, 2010. **18**(1): p. 87-92.
144. Marrone, L., P.M. Marchi, and M. Azzouz, *Circumventing the packaging limit of AAV-mediated gene replacement therapy for neurological disorders.* Expert Opin Biol Ther, 2022. **22**(9): p. 1163-1176.
145. Sun, L., J. Li, and X. Xiao, *Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization.* Nat Med, 2000. **6**(5): p. 599-602.
146. Odom, G.L., et al., *Gene therapy of mdx mice with large truncated dystrophins generated by recombination using rAAV6.* Mol Ther, 2011. **19**(1): p. 36-45.
147. Tornabene, P., et al., *Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina.* Sci Transl Med, 2019. **11**(492).
148. Cooney, A.L., et al., *Gene Therapy Potential for Genetic Disorders of Surfactant Dysfunction.* Front Genome Ed, 2021. **3**: p. 785829.
149. Kormann, M.S., et al., *Expression of therapeutic proteins after delivery of chemically modified mRNA in mice.* Nat Biotechnol, 2011. **29**(2): p. 154-7.
150. Barnett, R.C., et al., *Featured Article: Electroporation-mediated gene delivery of surfactant protein B (SP-B) restores expression and improves survival in mouse model of SP-B deficiency.* Exp Biol Med (Maywood), 2017. **242**(13): p. 1345-1354.
151. Mahiny, A.J., et al., *In vivo genome editing using nuclease-encoding mRNA corrects SP-B deficiency.* Nat Biotechnol, 2015. **33**(6): p. 584-6.
152. Earley, J., et al., *Evading and overcoming AAV neutralization in gene therapy.* Trends Biotechnol, 2023. **41**(6): p. 836-845.
153. Miah, K., Ruffle, A., Viegas, M.A., Castells, E., Dolatshad, H., Sumner-Jones, S., Dean, R., Du, Y., McStea, A., Munday, R.J., Munis, A.M., Hyde, S.C., Gill, D.R., *74. Rescue of Lethal SP-B Deficiency in a Murine Model Using Lentiviral Vector-Mediated Gene Therapy, ASGCT 27th Annual Meeting Abstracts.* Molecular Therapy, 2024. **32**(4): p. 40.
154. Donnelley, M., et al., *Repeat or single-dose lentiviral vector administration to mouse lungs? It's all about the timing.* Gene Ther, 2023. **30**(9): p. 698-705.
155. Griesenbach, U., et al., *Assessment of F/HN-pseudotyped lentivirus as a clinically relevant vector for lung gene therapy.* Am J Respir Crit Care Med, 2012. **186**(9): p. 846-56.
156. Alapati, D., et al., *In utero gene editing for monogenic lung disease.* Sci Transl Med, 2019. **11**(488).
157. Kinting, S., et al., *Potentiation of ABCA3 lipid transport function by ivacaftor and genistein.* J Cell Mol Med, 2019. **23**(8): p. 5225-5234.
158. Domm, J.M., *Adeno-Associated Virus-Based Gene Therapy for Treating Lung Surfactant Deficiencies*, in *Department of Pathobiology*. 2018, University of Guelph. p. 134.

159. van Lieshout, L.P., et al., *A Novel Triple-Mutant AAV6 Capsid Induces Rapid and Potent Transgene Expression in the Muscle and Respiratory Tract of Mice.* Mol Ther Methods Clin Dev, 2018. **9**: p. 323-329.
160. Durand, S. and A. Cimarelli, *The inside out of lentiviral vectors.* Viruses, 2011. **3**(2): p. 132-159.
161. VandeWoude, S. and C. Apetrei, *Going wild: lessons from naturally occurring T-lymphotropic lentiviruses.* Clin Microbiol Rev, 2006. **19**(4): p. 728-62.
162. Sharp, P.M. and B.H. Hahn, *The evolution of HIV-1 and the origin of AIDS.* Philos Trans R Soc Lond B Biol Sci, 2010. **365**(1552): p. 2487-94.
163. Hirsch, V.M., et al., *An African primate lentivirus (SIVsm) closely related to HIV-2.* Nature, 1989. **339**(6223): p. 389-92.
164. Milone, M.C. and U. O'Doherty, *Clinical use of lentiviral vectors.* Leukemia, 2018. **32**(7): p. 1529-1541.
165. German Advisory Committee Blood, S.A.o.P.T.b.B., *Human Immunodeficiency Virus (HIV).* Transfus Med Hemother, 2016. **43**(3): p. 203-22.
166. Wolff, J.H. and J.G. Mikkelsen, *Delivering genes with human immunodeficiency virus-derived vehicles: still state-of-the-art after 25 years.* J Biomed Sci, 2022. **29**(1): p. 79.
167. Vogt, V.M., *Retroviral Virions and Genomes*, in *Retroviruses*, J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997: Cold Spring Harbor (NY).
168. Munis, A.M., *Gene Therapy Applications of Non-Human Lentiviral Vectors.* Viruses, 2020. **12**(10).
169. Naldini, L., et al., *In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector.* Science, 1996. **272**(5259): p. 263-7.
170. Nikolic, J., et al., *Structural basis for the recognition of LDL-receptor family members by VSV glycoprotein.* Nat Commun, 2018. **9**(1): p. 1029.
171. Zufferey, R., et al., *Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo.* Nat Biotechnol, 1997. **15**(9): p. 871-5.
172. Zufferey, R., et al., *Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery.* J Virol, 1998. **72**(12): p. 9873-80.
173. Dull, T., et al., *A third-generation lentivirus vector with a conditional packaging system.* J Virol, 1998. **72**(11): p. 8463-71.
174. Follenzi, A., et al., *Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences.* Nat Genet, 2000. **25**(2): p. 217-22.
175. Zufferey, R., et al., *Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors.* J Virol, 1999. **73**(4): p. 2886-92.
176. Shao, L., et al., *Genome-wide profiling of retroviral DNA integration and its effect on clinical pre-infusion CAR T-cell products.* J Transl Med, 2022. **20**(1): p. 514.
177. Yoder, K.E., et al., *Strategies for Targeting Retroviral Integration for Safer Gene Therapy: Advances and Challenges.* Front Mol Biosci, 2021. **8**: p. 662331.
178. Sandrin, V., S.J. Russell, and F.L. Cosset, *Targeting retroviral and lentiviral vectors.* Curr Top Microbiol Immunol, 2003. **281**: p. 137-78.
179. Duverge, A. and M. Negroni, *Pseudotyping Lentiviral Vectors: When the Clothes Make the Virus.* Viruses, 2020. **12**(11).
180. Gutierrez-Guerrero, A., F.L. Cosset, and E. Verhoeyen, *Lentiviral Vector Pseudotypes: Precious Tools to Improve Gene Modification of Hematopoietic Cells for Research and Gene Therapy.* Viruses, 2020. **12**(9).
181. Milani, M., et al., *GP64-pseudotyped lentiviral vectors target liver endothelial cells and correct hemophilia A mice.* EMBO Mol Med, 2024. **16**(6): p. 1427-1450.
182. Oakland, M., et al., *Intrapulmonary Versus Nasal Transduction of Murine Airways With GP64-pseudotyped Viral Vectors.* Mol Ther Nucleic Acids, 2013. **2**(1): p. e69.

183. Ichim, C.V. and R.A. Wells, *Generation of high-titer viral preparations by concentration using successive rounds of ultracentrifugation.* J Transl Med, 2011. **9**: p. 137.
184. Mazarakis, N.D., et al., *Rabies virus glycoprotein pseudotyping of lentiviral vectors enables retrograde axonal transport and access to the nervous system after peripheral delivery.* Hum Mol Genet, 2001. **10**(19): p. 2109-21.
185. Kong, J., et al., *Correction of the disease phenotype in the mouse model of Stargardt disease by lentiviral gene therapy.* Gene Ther, 2008. **15**(19): p. 1311-20.
186. Al-Khuzaei, S., et al., *An Overview of the Genetics of ABCA4 Retinopathies, an Evolving Story.* Genes (Basel), 2021. **12**(8).
187. Weng, J., et al., *Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in abcr knockout mice.* Cell, 1999. **98**(1): p. 13-23.
188. Cooney, A.L., et al., *Lentiviral-mediated phenotypic correction of cystic fibrosis pigs.* JCI Insight, 2016. **1**(14).
189. Trust, C.F. *Lenticlair (TM) 1: A study to assess how well a new genetic therapy is tolerated and whether it improves lung function in people with CF*. 2024 [cited 2024 08/10/2024]; Available from: https://www.cysticfibrosis.org.uk/get-involved/clinical-trials/trialstracker/tt014335.
190. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
191. Srinivas, S., et al., *Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus.* BMC Dev Biol, 2001. **1**: p. 4.
192. Chapman, H.A., et al., *Integrin alpha6beta4 identifies an adult distal lung epithelial population with regenerative potential in mice.* J Clin Invest, 2011. **121**(7): p. 2855-62.
193. Vaughan, A.E., et al., *Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury.* Nature, 2015. **517**(7536): p. 621-5.
194. Xenariou, S., et al., *Use of ultrasound to enhance nonviral lung gene transfer in vivo.* Gene Ther, 2007. **14**(9): p. 768-74.
195. Kobayashi, M., et al., *Pseudotyped lentivirus vectors derived from simian immunodeficiency virus SIVagm with envelope glycoproteins from paramyxovirus.* J Virol, 2003. **77**(4): p. 2607-14.
196. Mitomo, K., et al., *Toward gene therapy for cystic fibrosis using a lentivirus pseudotyped with Sendai virus envelopes.* Mol Ther, 2010. **18**(6): p. 1173-82.
197. Sweeney, N.P. and C.A. Vink, *The impact of lentiviral vector genome size and producer cell genomic to gag-pol mRNA ratios on packaging efficiency and titre.* Mol Ther Methods Clin Dev, 2021. **21**: p. 574-584.
198. Kumar, M., et al., *Systematic determination of the packaging limit of lentiviral vectors.* Hum Gene Ther, 2001. **12**(15): p. 1893-905.
199. Harper, S.Q., et al., *Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy.* Nat Med, 2002. **8**(3): p. 253-61.
200. Au, H.K.E., M. Isalan, and M. Mielcarek, *Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings.* Front Med (Lausanne), 2021. **8**: p. 809118.
201. Ho, J.Y., et al., *Promoter usage regulating the surface density of CAR molecules may modulate the kinetics of CAR-T cells in vivo.* Mol Ther Methods Clin Dev, 2021. **21**: p. 237-246.
202. Hyde, S.C., et al., *CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression.* Nat Biotechnol, 2008. **26**(5): p. 549-51.
203. Moore, L.D., T. Le, and G. Fan, *DNA methylation and its basic function.* Neuropsychopharmacology, 2013. **38**(1): p. 23-38.
204. Farinha, C.M., et al., *Biochemical methods to assess CFTR expression and membrane localization.* J Cyst Fibros, 2004. **3 Suppl 2**: p. 73-7.

205. Tsuji, Y., *Transmembrane protein western blotting: Impact of sample preparation on detection of SLC11A2 (DMT1) and SLC40A1 (ferroportin).* PLoS One, 2020. **15**(7): p. e0235563.
206. Capurro, V., et al., *Partial Rescue of F508del-CFTR Stability and Trafficking Defects by Double Corrector Treatment.* Int J Mol Sci, 2021. **22**(10).
207. Pan, D., H. Buning, and C. Ling, *Rational Design of Gene Therapy Vectors.* Mol Ther Methods Clin Dev, 2019. **12**: p. 246-247.
208. Weaver, T.E., *Synthesis, processing and secretion of surfactant proteins B and C.* Biochim Biophys Acta, 1998. **1408**(2-3): p. 173-9.
209. Engelbrecht, S., et al., *The surfactant lipid transporter ABCA3 is N-terminally cleaved inside LAMP3-positive vesicles.* FEBS Lett, 2010. **584**(20): p. 4306-12.
210. Wambach, J.A., et al., *Functional Genomics of ABCA3 Variants.* Am J Respir Cell Mol Biol, 2020. **63**(4): p. 436-443.
211. Johnson, L.G., et al., *Pseudotyped human lentiviral vector-mediated gene transfer to airway epithelia in vivo.* Gene Ther, 2000. **7**(7): p. 568-74.
212. Borok, Z., et al., *Vesicular stomatitis virus G-pseudotyped lentivirus vectors mediate efficient apical transduction of polarized quiescent primary alveolar epithelial cells.* J Virol, 2001. **75**(23): p. 11747-54.
213. Mulugeta, S., et al., *Identification of LBM180, a lamellar body limiting membrane protein of alveolar type II cells, as the ABC transporter protein ABCA3.* J Biol Chem, 2002. **277**(25): p. 22147-55.
214. Degiulio, J.V., C.D. Kaufman, and D.A. Dean, *The SP-C promoter facilitates alveolar type II epithelial cell-specific plasmid nuclear import and gene expression.* Gene Ther, 2010. **17**(4): p. 541-9.
215. Bove, P.F., et al., *Breaking the in vitro alveolar type II cell proliferation barrier while retaining ion transport properties.* Am J Respir Cell Mol Biol, 2014. **50**(4): p. 767-76.
216. Swain, R.J., et al., *Assessment of cell line models of primary human cells by Raman spectral phenotyping.* Biophys J, 2010. **98**(8): p. 1703-11.
217. Cooper, J.R., et al., *Long Term Culture of the A549 Cancer Cell Line Promotes Multilamellar Body Formation and Differentiation towards an Alveolar Type II Pneumocyte Phenotype.* PLoS One, 2016. **11**(10): p. e0164438.
218. Grek, C.L., et al., *Characterization of alveolar epithelial cells cultured in semipermeable hollow fibers.* Exp Lung Res, 2009. **35**(2): p. 155-74.
219. Schmitz, G. and G. Müller, *Structure and function of lamellar bodies, lipid-protein complexes involved in storage and secretion of cellular lipids.* J Lipid Res, 1991. **32**(10): p. 1539-70.
220. Kim, K., et al., *Loss of SP-A in the Lung Exacerbates Pulmonary Fibrosis.* Int J Mol Sci, 2022. **23**(10).
221. Xu, W., et al., *GCN5L1 regulates pulmonary surfactant production by modulating lamellar body biogenesis and trafficking in mouse alveolar epithelial cells.* Cell Mol Biol Lett, 2023. **28**(1): p. 90.
222. Tanjore, H., et al., *Alveolar epithelial cells undergo epithelial-to-mesenchymal transition in response to endoplasmic reticulum stress.* J Biol Chem, 2011. **286**(35): p. 30972-30980.
223. Lieber, M., et al., *A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells.* Int J Cancer, 1976. **17**(1): p. 62-70.
224. Hermanns, M.I., et al., *Lung epithelial cell lines in coculture with human pulmonary microvascular endothelial cells: development of an alveolo-capillary barrier in vitro.* Lab Invest, 2004. **84**(6): p. 736-52.
225. Fujimoto, T., et al., *Imaging lipid droplets by electron microscopy.* Methods Cell Biol, 2013. **116**: p. 227-51.

226. Sake, J.A., et al., *Knockout of ABCC1 in NCI-H441 cells reveals CF to be a suboptimal substrate to study MRP1 activity in organotypic in vitro models.* Eur J Pharm Sci, 2023. **181**: p. 106364.
227. Concordet, J.P. and M. Haeussler, *CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens.* Nucleic Acids Res, 2018. **46**(W1): p. W242-W245.
228. Doench, J.G., et al., *Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9.* Nat Biotechnol, 2016. **34**(2): p. 184-191.
229. Haeussler, M., et al., *Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR.* Genome Biol, 2016. **17**(1): p. 148.
230. Wittmann, T., et al., *Increased Risk of Interstitial Lung Disease in Children with a Single R288K Variant of ABCA3.* Mol Med, 2016. **22**: p. 183-191.
231. Chapuy, B., et al., *Intracellular ABC transporter A3 confers multidrug resistance in leukemia cells by lysosomal drug sequestration.* Leukemia, 2008. **22**(8): p. 1576-86.
232. Gonzalez, R.F., et al., *HTII-280, a biomarker specific to the apical plasma membrane of human lung alveolar type II cells.* J Histochem Cytochem, 2010. **58**(10): p. 891-901.
233. Evans, K.V. and J.H. Lee, *Alveolar wars: The rise of in vitro models to understand human lung alveolar maintenance, regeneration, and disease.* Stem Cells Transl Med, 2020. **9**(8): p. 867-881.
234. Korogi, Y., et al., *In Vitro Disease Modeling of Hermansky-Pudlak Syndrome Type 2 Using Human Induced Pluripotent Stem Cell-Derived Alveolar Organoids.* Stem Cell Reports, 2019. **12**(3): p. 431-440.
235. Munis, A.M., et al., *RNA-seq analysis of the human surfactant air-liquid interface culture reveals alveolar type II cell-like transcriptome.* Mol Ther Methods Clin Dev, 2022. **24**: p. 62-70.
236. Ren, H., N.P. Birch, and V. Suresh, *An Optimised Human Cell Culture Model for Alveolar Epithelial Transport.* PLoS One, 2016. **11**(10): p. e0165225.
237. Gray, S.J., et al., *Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors.* Hum Gene Ther, 2011. **22**(9): p. 1143-53.
238. Reber, S., et al., *CRISPR-Trap: a clean approach for the generation of gene knockouts and gene replacements in human cells.* Mol Biol Cell, 2018. **29**(2): p. 75-83.
239. Wambach, J.A., et al., *Functional Characterization of ATP-Binding Cassette Transporter A3 Mutations from Infants with Respiratory Distress Syndrome.* Am J Respir Cell Mol Biol, 2016. **55**(5): p. 716-721.
240. Araki, K. and K. Nagata, *Protein folding and quality control in the ER.* Cold Spring Harb Perspect Biol, 2011. **3**(11): p. a007526.
241. Steinbach, D., et al., *ABCA3 as a possible cause of drug resistance in childhood acute myeloid leukemia.* Clin Cancer Res, 2006. **12**(14 Pt 1): p. 4357-63.
242. Li, Y., et al., *Metabolic labelling of choline phospholipids probes ABCA3 transport in lamellar bodies.* Biochim Biophys Acta Mol Cell Biol Lipids, 2019. **1864**(12): p. 158516.
243. Chen, H., et al., *Pulmonary permeability assessed by fluorescent-labeled dextran instilled intranasally into mice with LPS-induced acute lung injury.* PLoS One, 2014. **9**(7): p. e101925.
244. Shaw, D., L. Serrano, and M. Lluch-Senar, *Lox'd in translation: contradictions in the nomenclature surrounding common lox-site mutants and their implications in experiments.* Microbiology (Reading), 2021. **167**(1).
245. Kristianto, J., et al., *Spontaneous recombinase activity of Cre-ERT2 in vivo.* Transgenic Res, 2017. **26**(3): p. 411-417.
246. Pyo, K.H., et al., *Establishment of a Conditional Transgenic Mouse Model Recapitulating EML4-ALK-Positive Human Non-Small Cell Lung Cancer.* J Thorac Oncol, 2017. **12**(3): p. 491-500.

247. Gui, Y.S., et al., *SPC-Cre-ERT2 transgenic mouse for temporal gene deletion in alveolar epithelial cells.* PLoS One, 2012. **7**(9): p. e46076.
248. Corti, M., A.R. Brody, and J.H. Harrison, *Isolation and primary culture of murine alveolar type II cells.* Am J Respir Cell Mol Biol, 1996. **14**(4): p. 309-15.
249. Donocoff, R.S., et al., *Optimization of tamoxifen-induced Cre activity and its effect on immune cell populations.* Sci Rep, 2020. **10**(1): p. 15244.
250. White, I.N., et al., *Species differences in the covalent binding of [14C]tamoxifen to liver microsomes and the forms of cytochrome P450 involved.* Biochem Pharmacol, 1995. **49**(8): p. 1035-42.
251. Smith, B.M., et al., *Oral and Injected Tamoxifen Alter Adult Hippocampal Neurogenesis in Female and Male Mice.* eNeuro, 2022. **9**(2).
252. Inotiv. *Tamoxifen diets for inducible Cre-LoxP systems*. 2024 [cited 2024 08/10/2024]; Available from: https://insights.inotivco.com/hubfs/resources/data-sheets/10337-envigo-r227-tamoxifen-info-sheet-eu-final-web.pdf.
253. University, Q.s. *SOP 7.31 - Tamoxifen Administration (Mice)*. 2024 [cited 2024 08/10/2024]; Available from: https://www.queensu.ca/animals-in-science/policies-procedures/sop/mice/7-31#:~:text=Option%201%3A%20Gradually%20acclimate%20mice,diet%20on%20weekends%20and%20holidays.
254. Whitfield, J., T. Littlewood, and L. Soucek, *Tamoxifen administration to mice.* Cold Spring Harb Protoc, 2015. **2015**(3): p. 269-71.
255. Basil, M.C. and E.E. Morrisey, *Lung regeneration: a tale of mice and men.* Semin Cell Dev Biol, 2020. **100**: p. 88-100.
256. Beers, M.F. and Y. Moodley, *When Is an Alveolar Type 2 Cell an Alveolar Type 2 Cell? A Conundrum for Lung Stem Cell Biology and Regenerative Medicine.* Am J Respir Cell Mol Biol, 2017. **57**(1): p. 18-27.
257. Munday, R.J., et al., *Sendai F/HN pseudotyped lentiviral vector transduces human ciliated and non-ciliated airway cells using α 2,3 sialylated receptors.* Mol Ther Methods Clin Dev, 2022. **26**: p. 239-252.
258. Das, S., et al., *A simple method of mouse lung intubation.* J Vis Exp, 2013(73): p. e50318.
259. Baghdadi, M., et al., *Sex-specific effects of Cre expression in Syn1Cre mice.* Sci Rep, 2023. **13**(1): p. 10037.
260. Costa, K.M., D. Schenkel, and J. Roeper, *Sex-dependent alterations in behavior, drug responses and dopamine transporter expression in heterozygous DAT-Cre mice.* Sci Rep, 2021. **11**(1): p. 3334.
261. Hemieda, F.A., *Influence of gender on tamoxifen-induced biochemical changes in serum of rats.* Mol Cell Biochem, 2007. **301**(1-2): p. 137-42.
262. Stout, M.B., R.L. Scalzo, and E.A. Wellberg, *Persistent Metabolic Effects of Tamoxifen: Considerations for an Experimental Tool and Clinical Breast Cancer Treatment.* Endocrinology, 2021. **162**(9).
263. Han, J.J., *FDA Modernization Act 2.0 allows for alternatives to animal testing.* Artif Organs, 2023. **47**(3): p. 449-450.
264. Rydell-Tormanen, K. and J.R. Johnson, *The Applicability of Mouse Models to the Study of Human Disease.* Methods Mol Biol, 2019. **1940**: p. 3-22.
265. Lavelle, G.M., et al., *Animal Models of Cystic Fibrosis Pathology: Phenotypic Parallels and Divergences.* Biomed Res Int, 2016. **2016**: p. 5258727.
266. Bose, S.J., et al., *Differential thermostability and response to cystic fibrosis transmembrane conductance regulator potentiators of human and mouse F508del-CFTR.* Am J Physiol Lung Cell Mol Physiol, 2019. **317**(1): p. L71-L86.
267. Van der Velden, J. and K.J. Snibson, *Airway disease: the use of large animal models for drug discovery.* Pulm Pharmacol Ther, 2011. **24**(5): p. 525-32.

268. Frohlich, E., *Animals in Respiratory Research.* Int J Mol Sci, 2024. **25**(5).
269. Meeusen, E.N., et al., *Sheep as a model species for the study and treatment of human asthma and other respiratory diseases.* Drug Discovery Today: Disease Models, 2009. **6**(4): p. 101-106.
270. Raeburn, D., S.L. Underwood, and M.E. Villamil, *Techniques for drug delivery to the airways, and the assessment of lung function in animal models.* Journal of Pharmacological and Toxicological Methods, 1992. **27**(3): p. 143-159.
271. Bischof, R.J., et al., *Induction of allergic inflammation in the lungs of sensitized sheep after local challenge with house dust mite.* Clin Exp Allergy, 2003. **33**(3): p. 367-75.
272. Viotti Perisse, I., et al., *Sheep models of F508del and G542X cystic fibrosis mutations show cellular responses to human therapeutics.* FASEB Bioadv, 2021. **3**(10): p. 841-854.
273. Welsh, M.J., et al., *Development of a porcine model of cystic fibrosis.* Trans Am Clin Climatol Assoc, 2009. **120**: p. 149-62.
274. Rogers, C.S., et al., *Disruption of the CFTR gene produces a model of cystic fibrosis in newborn pigs.* Science, 2008. **321**(5897): p. 1837-41.
275. Epithelix. *AlveolAir (AA)*. 2024 [cited 2024 09/10/2024]; Available from: https://epithelix.com/products/alveolair.
276. Wallace, W.A., et al., *Inflammation-associated remodelling and fibrosis in the lung - a process and an end point.* Int J Exp Pathol, 2007. **88**(2): p. 103-10.
277. Lipman, N.S., et al., *Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources.* ILAR J, 2005. **46**(3): p. 258-68.
278. Lin, J., et al., *Ultra-sensitive digital quantification of proteins and mRNA in single cells.* Nat Commun, 2019. **10**(1): p. 3544.
279. Nanayakkara, A.K., et al., *Targeted inhibitors of P-glycoprotein increase chemotherapeutic-induced mortality of multidrug resistant tumor cells.* Sci Rep, 2018. **8**(1): p. 967.
280. Cheong, N., et al., *Functional and trafficking defects in ATP binding cassette A3 mutants associated with respiratory distress syndrome.* J Biol Chem, 2006. **281**(14): p. 9791-800.
281. Höppner, S., et al., *Quantification of volume and lipid filling of intracellular vesicles carrying the ABCA3 transporter.* Biochim Biophys Acta Mol Cell Res, 2017. **1864**(12): p. 2330-2335.
282. Gilliard, N., et al., *Radiolabeling of the hydrophobic components of lung surfactant with 3-(trifluoromethyl)-3-(m-[125I]iodophenyl)diazirine.* Anal Biochem, 1991. **193**(2): p. 310-5.
283. Batenburg, J.J., *Surfactant phospholipids: synthesis and storage.* Am J Physiol, 1992. **262**(4 Pt 1): p. L367-85.
284. Gault, C.R., L.M. Obeid, and Y.A. Hannun, *An overview of sphingolipid metabolism: from synthesis to breakdown.* Adv Exp Med Biol, 2010. **688**: p. 1-23.